METHODOLOGICAL APPROACHES TO THE AMBULATORY ASSESSMENT OF ANXIETY DURING SITUATIONAL EXPOSURE

ANDREW J. WHITE

Inaugural dissertation submitted in partial fulfillment of the requirements for the degree Doctor of Social Sciences in the Graduate School of Economic and Social Sciences at the University of Mannheim

Department of Psychology
School of Social Sciences
University of Mannheim
July 2014

School of Social Sciences, University of Mannheim

Dean: Prof. Dr. Michael Diehl
First reviewer: Prof. Dr. Georg W. Alpers (Main supervisor)
Second reviewer: Prof. Dr. Michael Witthöft (Second supervisor)
Third reviewer: Prof. Dr. Sabine Sonnentag

Date of oral defence: 22nd October, 2014

Andrew J. White: *Methodological Approaches to the Ambulatory Assessment of Anxiety During Situational Exposure*

Dedicated to the loving memory of John C. White

1946 – 2011

PUBLICATIONS AND DECLARATION OF SUPPORT RECEIVED

The studies presented in current thesis were supported by a scholarship from the Center for Doctoral Studies in Social and Behavioral Sciences (CDSS) at the University of Mannheim, and financial support received through the PanikNetz project, which was funded by the German Federal Ministry of Education and Research (project number 01GV0617).

The first study presented in this dissertation was recently published in a peer-reviewed journal:

White, A. J., Umpfenbach, K., & Alpers, G. W. (2014). Where have they gone? Tracking movement patterns to document the process of situational exposure in agoraphobia. *Professional Psychology: Research and Practice*, 45, 171-179.

In White, Umpfenbach & Alpers (2014), I was responsible for the study design, training the clinician (the second author) to use the ambulatory monitoring device, data analysis and manuscript preparation. Dr. Katja Umpfenbach delivered therapy to the patient, obtained informed consent, and administered clinical scales throughout treatment. Discussions I had with Prof. Dr. Georg Alpers were instrumental in the conception of the study and he also provided helpful advice during my preparation of the manuscript.

The remaining studies presented in this dissertation are based on data collected as part of a PanikNetz project, a multicentre clinical treatment project. The principal investigators were Prof. Dr. Alfons Hamm, Prof. Dr. Alexander L. Gerlach, Dr. Thomas Lang, and Prof. Dr. Georg W. Alpers. The principal investigators were responsible for securing the initial funding, designing of the overarching study, coordinating patient recruitment and overseeing the delivery of treatment.

In Studies 2, 3, and 4, I was responsible for data preprocessing, data analysis and chapter preparation. Prof. Dr. Dieter Kleinböhl offered statistical advice related to data analysis. I obtained some feedback on earlier drafts from Prof. Dr. Dieter Kleinböhl, Dr. Andre Pittig, Dr. Florian Bublatzky, Josepha Zimmer, and Prof. Dr. Georg Alpers.

Because each of the presented studies were the result of collaboration and either have been, or will be, published with several co-authors, I have opted to use the first person plural pronoun, "we", when presenting arguments and results, which is in line with American Psychological Association's publishing guidelines.

ACKNOWLEDGEMENTS

I am indebted to my supervisor, Georg Alpers for his ongoing support and encouragement. For all our stimulating exchanges which helped me to see the forest when I began to fixate on the trees, and the opportunities he has provided for me to explore my interests, I am most appreciative.

I would like to thank Dieter Kleinböhl for his sage methodological advice, and for helping me to work through conceptual challenges that arose throughout my doctoral studies. I profited greatly from his down-to-earth, good-humoured approach.

A big thank you also goes out to my friends and colleagues for their support and feedback. In particular, my colleagues at the Chair and at OSI helped me develop ideas (and to discard flawed ones), and I have them to thank for helping to create a such a pleasant work environment.

My experience at the CDSS has been very rewarding. I am grateful for the contact I have had with staff and doctoral students, and for the financial support which I received.

Numerous research assistants assisted me to correspond with and collect data from the PanikNetz centres. They, the therapists, and patients who took part in the treatment, helped to make these studies possible.

My family and friends back home in Australia have always given me their unwavering support, and for this I am grateful. The encouragement of my mother and sister deserves a special mention.

And finally, I would like to thank to my partner, Frauke for her steadfast support over the past four years—she helped me get through some of life's darker moments and to celebrate the good times.

CONTENTS

i GENERAL INTRODUCTION 1

1 GENERAL INTRODUCTION 2

1.1 Ambulatory Assessment During Psychological Treatment . . . 2
1.2 Panic Disorder with Agoraphobia . . . 7
1.3 Theories on the Aetiology of Panic Disorder . . . 8
1.4 Open Questions Related to Exposure Therapy . . . 10
1.5 Theories of Exposure . . . 10
1.6 Psychophysiological Indicators of Fear . . . 12
1.7 Novel Contributions to the Field of Ambulatory Assessment 15
1.8 Methodological Approaches . . . 19

ii STUDIES 21

2 MOVEMENT PATTERNS DURING SITUATIONAL EXPOSURE 22

2.1 Introduction . . . 22
2.2 Method . . . 25
2.3 Results . . . 28
2.4 Discussion . . . 32

3 CLASSIFYING INDIVIDUAL HEART RATE RESPONSES OBTAINED DURING SITUATIONAL EXPOSURE 37

3.1 Introduction . . . 37
3.2 Method . . . 42
3.3 Results . . . 47
3.4 Discussion . . . 57

4 DOES HABITUATION MATTER? AN ASSESSMENT OF ADHERENCE TO END-OF-EXPOSURE INSTRUCTIONS AND EMOTIONAL PROCESSING THEORY PREDICTIONS 62

4.1 Introduction . . . 62
4.2 Method . . . 67
4.3 Results . . . 74
4.4 Discussion . . . 82

5 MIX IT UP: THE ADVANTAGES OF VARYING SITUATIONAL EXPOSURE CONTEXTS 88

5.1 Introduction . . . 88
5.2 Method . . . 91
5.3 Results . . . 95
5.4 Discussion . . . 99

iii GENERAL CONCLUSION 104

6 GENERAL CONCLUSION 105

6.1 Main Findings . . . 105
6.2 Discussion and Implications . . . 109

6.3 Conclusion 119

IV APPENDIX 121

A GPS-RELATED SUPPLEMENTS 122
A.1 A collection of GPS trajectories from a single patient 122

B HEART RATE SUPPLEMENTS 123
B.1 Spaghetti plot: Unsorted heart rate 123
B.2 Diceplot: Cluster assignment patterns 124
B.3 Operationalisation of emotional processing theory components 125
B.4 Distributional plots of EPT components 126
B.5 EPT predictors across binary treatment response 129
B.6 EPT component comparison 130

C ECOLOGICAL MOMENTARY ASSESSMENT SUPPLEMENTS 132
C.1 Empirical change plots for a subsample of patients 132
C.2 Sample ecological momentary assessment questions 134

REFERENCES 135

LIST OF FIGURES

Figure 1 The ambulatory device used throughout studies. 16
Figure 2 First driving exposure 30
Figure 3 Second driving exposure 31
Figure 4 Mean HR responses across clusters sorted by increasing absolute level 49
Figure 5 Mean HR by epochs across session number and treatment group 52
Figure 6 Self-reported anxiety across epoch by exposure session 54
Figure 7 Average self-reported anxiety across epoch by HR clusters 55
Figure 8 Mean exposure duration across exposure session 75
Figure 9 Changes in expected anxiety and fear of losing control as a function of mean exposure duration . 76
Figure 10 The relationship between exposure duration and consistency of adherence to end-of-exposure instructions 78
Figure 11 The relationship between consistency of adherence to end-of-exposure instructions and expected anxiety 80
Figure 12 The relationship between total number of exposures undertaken and changes in agoraphobic cognitions questionnaire scores 98
Figure 13 The association between multiple exposure contexts and maintenance of treatment gains 99
Figure 14 A collection of GPS trajectories from a single patient 122
Figure 15 Spaghetti plot of unsorted, unstandardised heart rates for the circumscribed segment of interest . . 123
Figure 16 Diceplot of cluster assignment and transitions . . 124
Figure 17 Operationalisation of within- and between-session habituation 125
Figure 18 Distributional plots for raw and winsorised initial fear activation 126
Figure 19 Distributional plots for raw and winsorised within-session habituation 127
Figure 20 Distributional plots for raw and winsorised between-session habituation 128
Figure 21 Emotional processing theory predictors across treatment non-responders and responder 129
Figure 22 Empirical growth plots for expected anxiety . . . 132
Figure 23 Empirical growth plots for fear of losing control . 133

LIST OF TABLES

Table 1 Clinical scale scores 29
Table 2 Mean HR of clusters and trial breakdown by session number and treatment 51
Table 3 Results of the repeated measures analysis of HR for a priori groupings 53
Table 4 Concordance: Pearson correlation between HR and self-reported anxiety across treatment, session, and cluster 56
Table 5 Descriptive statistics for emotional processing theory predictors 71
Table 6 Adherence to end-of-exposure instructions throughout exposure 77
Table 7 Average treatment progress measures and estimated rates of change across exposure 81
Table 8 Average treatment outcome and estimated rates of change throughout treatment derived from fixed effects parameters 96
Table 9 Results of fitting several multilevel models for change to expected anxiety. 130
Table 10 Results of fitting several multilevel models for change to fear of losing control. 131
Table 11 Sample questions presented on the ecological momentary assessment device 134

Part I

GENERAL INTRODUCTION

1 GENERAL INTRODUCTION

1.1 AMBULATORY ASSESSMENT DURING PSYCHOLOGICAL TREATMENT

It is an exciting time for ambulatory assessment researchers. Modern ambulatory devices can unobtrusively monitor the self-reports, physiology, and behaviours of individuals in their natural environment. For those who seek to unlock the mysteries of psychopathology, this also presents a significant opportunity—to use increasingly compact, precise, affordable instruments to validate and extend theories predominantly developed under controlled laboratory conditions. Ambulatory assessment offers the possibility of looking beyond the confines of the laboratory and studying psychological processes in natural settings where people live and spend the majority of their time. However, access to increasingly precise ambulatory devices alone will not be enough to address complex questions about mental disorders—chronic and disabling conditions, which are treated with trial-and-error (Insel, 2012). An understanding of the methodological issues which arise when conducting field research is thus required.

Developing more nuanced understandings of psychological processes involved in mental illness helps to further establish empirically supported treatments. Exposure-based therapy is a very common, evidence-based treatment for anxiety disorders (Chambless et al., 1998; Rachman, 2009), however is in need of further refinement as it is not effective for all patients, and relapse following treatment is quite common (Peter et al., 2008; Craske et al., 2014). Understanding how this already effective therapy exerts an influence on patients is therefore of considerable importance.

Against this backdrop, the overarching aim I address in this dissertation is how assessing anxiety symptoms in the field can help clarify who responds to exposure therapy. To understand *who* responds necessitates that we understand *how* individuals respond and under what circumstances. The methods used to examine patients therefore require close scrutiny, because the mapping between various indicators, psychological constructs and treatment outcomes is highly complex (Cacioppo et al., 1991, 2007). To this end, I examine the exposure procedures (*how* patients undertake exposure) and analytic approaches (data handling, operationalisation of constructs, statistical procedures) used to make sense of physiological, verbal and behavioural measures of anxiety.

Because methods are now available which allow psychological constructs to be acquired in increasingly relevant contexts (Wilhelm et al., 2012a), I argue for the adoption of novel assessment approaches. For instance, tracking movement with Global Positioning System (GPS) devices permits objective assessment of patient's compliance with exposure instructions (White et al., 2014). Correctly applied, these techniques have great potential to identify specific factors that, when tailored to individuals, promote enhanced treatment outcomes.

Although some treatments (e.g., cognitive restructuring) largely take place within therapist's offices, in this dissertation, exclusive attention will be given to patients undertaking situational exposure. This particular form of exposure therapy is an essential component of treatment for panic disorder with agoraphobia (PD/A), and involves confronting feared situations that appear in everyday, natural environments. Examining patients as they undertake actual situational exposure therefore involves taking some established assessment methods outside the laboratory (Alpers, 2009; Wilhelm & Grossman, 2010).

Within the current context, ambulatory assessment refers to both the monitoring of symptoms as they unfold in multifaceted everyday situations, and the subsequent assessment of data (Mehl & Conner, 2012). To address the central aim of this thesis, I seek to demonstrate that ambulatory assessments of clinical symptoms (i) are well suited to assess whether specific factors and mechanisms are associated with positive treatment response, and (ii) can build on, and promote an exchange with, laboratory-based research.

Relevance of Assessing Psychological Phenomena in Everyday Situations

Environmental influences on behaviour

Studying how psychopathological symptoms vary in natural settings offers several distinct advantages over controlled laboratory-based experiments. A commonly cited advantage is that studying people in their everyday environments helps to increase the ecological validity of findings. The concept of ecological validity derives from Brunswik's (1955) notion of representative experimental design though it originally had a somewhat different meaning. It encompasses the probabilistic relationship between an organism and its environment. In this sense, behaviours and psychological processes represent functional adaptations (in a Darwinian sense) to environments (Brunswik, 1955; Wilhelm et al., 2012b). Brunswik therefore took issue with methods that overlooked environmental influences. Current use of the term "ecological validity" was originally referred to as "ecological generalisability" by Brunswik (Brunswik, 1955, p.202) and could be achieved by adopting representative designs. This approach involves sampling participants *and* environments that are represent-

ative of the broader population to which results intend to be generalised (Brunswik, 1956; Hammond, 1996). Brunswik's conceptualisation of representative design and his formal criticism of controlled laboratory-based studies employing systematic experimental designs, helped set the stage for ambulatory research.

Systematic experimental designs involve the selection of systematically varied (manipulated) independent variables while all other relevant conditions are held constant or allowed to vary randomly (Brunswik, 1956; Dhami et al., 2004). Changes in the resulting dependent variables are then observed. The factorial design is a commonly employed extension of this design, whereby variables are artificially divided into several levels and then exhaustively combined (Dhami et al., 2004). Systematic designs strongly emphasise internal validity and reflect efforts to demonstrate causal relationships between several variables, but have two substantial drawbacks. First, controlling all variables not under investigation removes, or at least alters, the natural covariation among variables (Brunswik, 1955; Fiedler & Juslin, 2006). Second, the tasks used to elicit emotions in laboratories, although high in internal consistency and reliability, may be low in construct validity. That is, their mode of presentation may be somewhat contrived, which begs the question whether emotions evoked in laboratories are as authentic as those experienced in everyday life (Coan & Allen, 2007). Experiments designed to maximise internal validity therefore emphasise uncovering "what *can* happen" rather than "what *does* happen" (Reis, 2012, p.7).

Generalisation of results

In order to better understand what *does* happen, Brunswik proposed that studies should be based on a "representative design" (Brunswik, 1955). To some extent, a reliance on undergraduate students in psychological experiments has called into question whether the findings derived from this population generalise (Henrich et al., 2010). Less appreciated is the role of sampling representative environments—experimental conditions should represent those where phenomena of interest naturally occur to support claims of generalisability (Hammond, 2001). Although laboratory conditions may include elements that make stimuli or conditions more externally valid, behaviour studied within its natural context reflects "a correspondence between the conditions of a study and the conclusions that are drawn from it" (Reis, 2012, p. 7)—although this may not guarantee ecological validity, it is an important prerequisite. In sum, ambulatory research that samples individuals in a variety of everyday environments represents an approach that fosters ecological generalisability.

In basic laboratory-based research, randomised experiments and quasi-experiments are used to develop understandings of cause-and-effect relationships, and the internal validity of studies is

emphasised (Reis, 2012; Shadish et al., 2002). In such settings, it is feasible to manipulate several variables of interest and tightly control the temporal sequence of events. But this is more challenging in ambulatory studies and greater effort is made to demonstrate the external validity of the design and generalisability of findings. In such research, manipulating specific variables (e.g., the kind of treatment) can allow some degree of causal description (Shadish et al., 2002). However, continuous sampling of multiple streams of data means that ambulatory studies are better suited to discovering patterns of association between non-manipulated variables. As such, the data yielded from such ambulatory studies are suited to discovery-oriented research (Reis, 2012). In a similar vein, Wittmann's (2012) call to examine the correspondence between predictors and criterion measures in their level of generality supports the study of non-manipulated predictors when studying treatment outcomes. As the level of generality is different when research is conducted in real-life settings compared to stationary laboratories, it stands to reason that the degree of symmetry between predictors and treatment outcomes in ambulatory studies may also differ (Wittmann & Klumb, 2006; Wittmann, 2012). Examining associations between psychophysiological predictors (collected under ambulatory conditions) and treatment outcomes offers the opportunity to compare results with those collected under more constrained settings. In summary, gathering representative samples of participants and settings, and applying principles of symmetry are two methods to promote greater generalisability of results. This helps justify an ambulatory assessment approach when studying psychological processes and suggests that cooperation with basic research is needed to comprehensively account for organism-environment relations.

Retrospective recall bias and methodological reactivity

Another advantage of ambulatory studies is that they help circumvent concerns about the validity of retrospective or generalised responses (Wilhelm et al., 2012b). For example, it is possible to concurrently assess self-reports and a range of physiology signals during tasks of interest. Several studies have demonstrated the utility of monitoring symptoms central to mental illness, such as self-reported symptoms (e.g., anticipatory and current anxiety, Helbig-Lang et al., 2012) and physiological indicators (e.g., heart rate, respiration, Wilhelm & Grossman, 2010; Wilhelm et al., 2012a). This, in turn, can allow researchers to inspect intraindividual changes in symptoms throughout treatment (Hamaker, 2012).

Given that long, intensive sampling is not uncommon, researchers also examined whether the onerousness of sampling or discomfort associated with wearing certain devices was a source of bias. It has been shown that even long, intensive recordings do not appear to

bias outcome measures (Alpers, 2009; Stone et al., 2003). However, adopting sampling strategies that minimise the frequency, total number and length of each enquiry helps to reduce assessment burden, which is related to methodological reactivity and acceptance (for a review, see Santangelo et al., 2013). Thus, there are many good reasons to support the adoption of ambulatory assessment approaches to examine psychopathology, both as an adjunct to laboratory studies, and in and of their own right.

Using Ambulatory Methods to Probe Anxiety Disorders

Employing ambulatory methods to study emotional processes central to anxiety disorders has some precedence. Ambulatory assessment has been applied to study naturally manifesting symptoms from a wide range of anxiety disorders—driving phobia (Alpers et al., 2005), flying phobias (Wilhelm & Roth, 1998), claustrophobia (Alpers & Sell, 2008), and panic disorder with agoraphobia (Meuret et al., 2012). In these studies, psychophysiological measures were applied to reveal patterns of anxious responses during exposure therapy. As well as demonstrating the technical feasibility of multichannel psychophysiological assessment in a variety of settings, studies have also demonstrated the discriminant validity of specific ambulatory measures to distinguish between healthy and anxious individuals. For example, Wilhelm and Roth (1998) showed that during flights, flying phobics experienced greater fluctuations in skin conductance, higher HR, lower heart rate variability, more inspiratory pauses and greater self-reported anxiety, relative to healthy controls. Similarly, compared with health controls, cortisol levels during, and in anticipation of driving exposure, were elevated in a group of driving phobics (Alpers et al., 2003). In sum, demonstrating that specific physiological responses collected in the field discriminate between phobic and non-phobic individuals, helps elucidate the maladaptive processes central to anxiety disorders.

As many theories of anxiety have stemmed from clinical impressions and laboratory studies, the degree to which these ideas account for anxiety as it unfolds in everyday situations requires examination. A low correspondence between laboratory and ambulatory assessment findings has, however, been found in several studies (for a review, see Wilhelm et al., 2012b). For example, the effect of psychosocial demands on blood pressure and HR was examined under laboratory and during everyday life (Kamarck et al., 2003). Findings suggested only a moderate association between cardiovascular activity collected across several standardised laboratory tasks and during analogous, naturally-occurring events in daily settings. One difficulty in determining laboratory-field correspondence is finding equivalent operationalisations of outcome variables in both recording settings

(Kamarck et al., 2003)—a challenge, given that several extraneous factors influence the extent to which spontaneously-occurring events in daily life are discriminable. However, this problem is mitigated when ambulatory methods are applied to study circumscribed events, such as when phobic individuals are exposed to feared situations (e.g., Baker et al., 2010; Meuret et al., 2006; Wilhelm & Roth, 1998), and when event markers are clearly defined (Wilhelm & Grossman, 2010).

In summary, there are a range of reasons to look beyond the constrained laboratory settings when studying anxiety. Ambulatory measures of psychophysiological responses can reveal processes central to anxiety disorders, and can help to highlight inconsistencies with laboratory-based disorder models. In the current dissertation, there is a focus on processes involved in situational exposure, which is an essential component of modern treatment for panic disorder with accompanying agoraphobia (National Institute for Clinical Excellence, 2011).

1.2 PANIC DISORDER WITH AGORAPHOBIA

Panic disorder (PD) is an anxiety disorder in which individuals experience recurrent, uncued panic attacks (American Psychiatric Association, 2000, 2013), and follows a chronic clinical course. Panic attacks are sudden episodes of intense fear accompanied by several physical and mental symptoms such as heart palpitations, chest pain and sensations of shortness of breath. They are defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013) as a short period of intense fear or discomfort in the absence of actual danger in which four or more of a list of 13 symptoms emerge (American Psychiatric Association, 2000). Symptoms usually reach maximum intensity within 10 minutes. In addition to the immediate experience of symptoms during a panic attack, to meet criteria for a PD diagnosis individuals must also have developed substantial anticipatory anxiety over the possibility of re-experiencing a panic attack or over the implications of the attack or its consequences (e.g., belief that the attack will lead to a heart attack). Avoidance behaviours are sometimes exhibited by individuals with PD in an effort to prevent the reoccurrence of panic attacks. Agoraphobia has recently been classed as a standalone disorder in the DSM-5, in line with evidence that it is not unique to PD (Wittchen et al., 2010). The central feature of agoraphobia is the "marked, or intense, fear or anxiety triggered by the real or anticipated exposure to a wide range of situations" (American Psychiatric Association, 2013, p.218). Agoraphobic situations are avoided, endured with marked distress or require the presence of a companion, and escape from such

situations is not uncommon (Gloster et al., 2009; Richter et al., 2012). A diagnosis of agoraphobia requires that two or more of the following situations are actively avoided: Travelling on public transportation, being in open places (e.g., marketplaces), being in enclosed spaces (e.g., shops), standing in line or being in a crowd, and being outside the home alone (American Psychiatric Association, 2013). Depending on the pervasiveness of such situational avoidance, PD is classified as occurring with (PD/A) or without agoraphobia. Although the distinctness of agoraphobia as a diagnostic criteria cones to be explored (for a review, see Wittchen et al., 2010), my focus here concerns agoraphobia that accompanies PD.

It is also noteworthy that the experience of panic attacks is not restricted to those with panic disorder; they can also occur in the context of other mental disorders such as mood and substance disorders, and some general medical conditions such as asthma (Perna et al., 1997). Epidemiological research has shown that the estimated lifetime prevalence is 22.7% for isolated panic attacks, 3.7% for panic disorder without agoraphobia, and 1.1% for PD/A (Kessler et al., 2006). While panic attacks may be isolated experiences, in clinical settings they are significant risk markers for the onset and development of primary disorders and comorbidities, with anxiety disorders, major depression and alcohol use disorders being particularly common (Craske et al., 2010). For example, in one longitudinal birth cohort study involving 1265 individuals, experiencing a PA increased the risk of current major depression in the three years following the attack, after controlling for other early risk factors (Goodwin et al., 2004). Kessler and colleagues (2006) found that persistence and lifetime number of uncued attacks is significantly higher among those with PD/A, compared to other groups. Further, the clinical severity, as assessed by the clinician-administered version Panic Disorder Severity Scale (PDSS) confirmed that overall clinical severity was highest among those with PD/A: 86.3% had moderate or severe ratings, compared to 46.1% for PD without agoraphobia, and 6.7% for those who experienced isolated panic attacks. In summary, these findings demonstrate the chronic and disabling effects of experiencing the panic and phobic avoidance that characterise PD/A.

1.3 THEORIES ON THE AETIOLOGY OF PANIC DISORDER

Theories concerning the origins of PD have most commonly been derived from either cognitive or associative learning frameworks. According to cognitive theories, a person's "catastrophic misinterpretation" of somatic and auxiliary sensations are seen as central to the onset and maintenance of PD (Beck & Emery, 1985; Bouton et al., 2001; Clark, 1986). Accordingly, when the severity of normal sensa-

tions is misjudged (e.g., when heart palpitations are interpreted as evidence of an impending heart attack), these catastrophic thoughts give way to further unpleasant bodily sensations, which in turn results in additional catastrophic thoughts (Clark, 1986). As there is no disconfirmation of catastrophic thoughts, a vicious cycle ensues which can ultimately result in a panic attack. Over time, this leads to a heightened vigilance and sensitivity to normal physical sensations. Cognitive therapy, which is based on this model, is therefore premised on the idea that learning to disconfirm feared catastrophes facilitates reduction of panic symptoms (Hoffart et al., 2008; Salkovskis et al., 2007). In this sense, exposure therapy provides opportunities to disconfirm firmly held maladaptive beliefs. An alternate view of anxiety is outlined in the anxiety sensitivity (AS) theory. Here, it is proposed that panic stems from a trait-like belief that experiencing anxiety and its associated symptoms has far-reaching, negative implications (McNally, 2002; Reiss et al., 1986; Reiss, 1987). Individuals with elevated AS are therefore motivated to avoid cues that trigger dreaded symptoms. Rather than emphasising the immediate consequences of the experienced sensation, AS theory proposes that the physical, social and psychological harm from experiencing anxious symptoms accumulates over time.

In contrast to cognitive theories, conditioning theories of PD highlight disruptions in associative learning processes. According to this view, PD develops, "when stimuli, events or situations (conditioned stimuli, CSs) of anxiety are paired with a panic attack, the learning that may occur can allow the CSs to trigger panic and anxiety when they are encountered again" (Bouton et al., 2001, p.7). Given that panic-like symptoms are quite commonly experienced, what causes a minority of people to develop panic disorder? Goldstein and Chambless (1978) implicated interoceptive conditioning in the development of panic disorder—a process whereby internal sensations, normally present during moments of anxious arousal, become the conditioned stimuli that are capable of triggering higher levels of anxiety that culminate in panic attacks. They termed this process a "fear of fear", and proposed that this vicious cycle could generalise to external situations. Persistent avoidance of such situations could ultimately result in a person becoming agoraphobic (Chambless, 1985). More recently, converging evidence from animal and human conditioning studies has focussed attention on fear extinction. Individuals with anxiety disorders such as PD are thus characterised as having an impaired ability to extinguish fear—specifically a failure to learn that the CS no longer predicts the threat (for a historical review, see Milad & Quirk, 2012). This has helped generate considerable interest in factors that may help to strengthen extinction learning (Bouton, 2002; Rescorla, 2006; Craske et al., 2008). In summary, cognitive and conditioning

theories of panic have provided convincing explanations of factors involved in the development and maintenance of panic symptoms. Conditioning theories, in particular, allow findings to be translated from animals to humans (Bouton, 1994) and contribute to modern understandings of how fear can be both acquired and inhibited (Bouton & Swartzentruber, 1991; Bouton et al., 2001; Bouton, 2004).

1.4 OPEN QUESTIONS RELATED TO EXPOSURE THERAPY

A core component of treatment for panic disorder with agoraphobia (PD/A) is situational exposure, as indicated by the results of several randomised controlled clinical trials (Marks, 1987; Rachman, 2009). Exposure involves the repeated, systematic confrontation of feared cues. In situational (or in vivo) exposure, these cues are the feared situations that individuals associate with panic. Interoceptive exposure, also a common treatment for PD, is the deliberate provocation of feared physical symptoms that are experienced during panic attacks (Goldstein & Chambless, 1978). For example, when symptoms associated with low blood concentrations of carbon dioxide are thought to be involved in the onset of a panic attack, patients may be asked to hyperventilate or hold their breath (de Beurs et al., 1995).

One aspect of exposure that appears to divide both researchers and practitioners is the requirement for fear to habituate during and across exposure sessions. Habituation is a form of non-associative learning which refers to decreased behavioural response to repeated stimulation (Groves & Thompson, 1970). Guidelines for exposure therapy differ in their explicitness concerning the need to remain in feared situations until fear habituates: Some are non-specific, and instead emphasise the importance of reducing safety behaviours, and paying attention to the order in which various situations are confronted (Antony & Barlow, 2010). Others recommend that patients remain in feared contexts until they experience a clinically significant decline in anxiety (fear habituation) (Abramowitz et al., 2011; Clark & Beck, 2010; O'Donohue & Fisher, 2012; Puri & Treasaden, 2011). If habituation of fear is not a core component of exposure, as some suggest (Meuret et al., 2012; Craske et al., 2008, 2014), then alternative factors, grounded in other compelling theories of exposure, should be examined. In sum, although situational exposure is an evidence-based treatment for PD/A, the mechanisms that promote change, and thus the essential therapy components, remain unclear.

1.5 THEORIES OF EXPOSURE

Being one of the most well-supported therapy components, theorists have long sought to distil the factors that contribute to effective exposure (Bouton, 1994; Clark, 1986; Lang, 1968; Wolpe,

1958). Two influential learning processes that attempt to account for successful learning during exposure therapy are habituation and inhibitory learning.

In laboratory settings, the learning processes thought to underlie exposure are based on fear extinction (Vervliet et al., 2012). A useful framework for examining these processes is fear conditioning, where fear acquisition and extinction are treated as separate phases. Fear acquisition is a process whereby repeated pairing of a neutral stimulus with an unconditioned stimulus (US, e.g., electric shock) results in the US becoming "conditioned" as a predictor of the aversive stimulus. This newly conditioned stimulus (CS) has the capacity to elicit the cognitions, physiological and behavioural responses in anticipation of the aversive stimulus (Vervliet et al., 2012). Self-report and psychophysiological measures (e.g., skin conductance, startle reflex) are typically used to gauge the magnitude of the conditioned fear response (see Lipp, 2006). Fear extinction is seen as the gradual attenuation of anticipatory fear reactions when CS are consistently presented in the absence of aversive stimuli (US). Several brain regions have also been implicated in fear extinction, including the amygdala, medial frontal cortex (ventromedial prefrontal cortex, anterior cingulate cortex) and hippocampus (for a review, see Shin & Liberzon, 2010). Since return of fear (renewal, reinstatement, and spontaneous recovery) is a common post-extinction phenomenon (Bouton et al., 2001), this suggests that the original CS-US association is not erased, as previously assumed, but only inhibited. Within this context, a central goal is to identify factors that help inhibit the original CS-US association. In sum, laboratory paradigms continue to provide useful models to understand the mechanisms of extinction that have implications for exposure.

Often cited, the emotional processing theory (EPT) of Foa and Kozak (1986) is used to explain the beneficial effects of exposure treatment. EPT combines the concepts of habituation with corrective learning. Accordingly, exposure therapy is argued to be effective when exposure to feared situations activates a "fear structure" and allows for the integration of incompatible information (Foa & Kozak, 1986). The fear structure is conceptualised as a set of propositions (true/false statements) about a stimulus (e.g., open spaces), response (e.g., racing heart) and their meaning (e.g., "I am going crazy") that are stored in memory (Lang, 1971). Integration of new information into the fear structure is posited to result in the development of a non-fear structure that replaces or competes with the original one. The non-fear structure is also thought to emerge as individual's physiological fear response habituates within- and between-sessions. Subjective self-report and physiological (e.g., heart rate) measures have been found to be reliable indices of fear structure activation

(Barlow et al., 1994; Telch et al., 2000). Results from several studies have supported EPT, with higher initial heart rate during exposure found to be highly predictive of better treatment outcomes (Alpers & Sell, 2008) and fear reduction among non-clinical samples (Kozak et al., 1988; Telch et al., 2000).

There is, however, evidence that fear activation and subsequent habituation are unreliable indices of extinction learning (Meuret et al., 2012). Instead, it is argued that strengthening inhibitory learning is a more effective way of reducing PD/A symptoms (Craske et al., 2008). For example, exposure undertaken in multiple contexts may help strengthen inhibitory associations through generalisation and thus help prevent renewal of fear (Lang & Craske, 2000; Rowe & Craske, 1998). In addition, superior inhibitory learning has been demonstrated when multiple excitatory conditioned stimuli (compound stimuli) are used during extinction (Rescorla, 2006). In comparison to extinction trials which include separate conditioned stimuli, presenting two previously extinguished excitatory stimuli in compound has been found to attenuate subsequent return of fear. In summary, compelling findings from laboratory paradigms have helped identify some promising targets, but whether they translate to patients undertaking exposure in natural settings remains to be seen.

1.6 PSYCHOPHYSIOLOGICAL INDICATORS OF FEAR

A well-supported view of emotions is that they reflect engagement of appetitive or defensive motive systems (Lang et al., 1997; Lang & McTeague, 2009), which have evolved to manage physical reactions to environments that promote or threaten survival (Bradley et al., 2001). Contexts involved in the promotion of survival such as sustenance, procreation, and nurturance, activate the appetitive system, whereas contexts that threaten survival engage the defensive system (Bradley et al., 2001; Lang & McTeague, 2009). These motivational systems are posited to organise affective states (e.g., Obrist, 1981; Lang et al., 1997). For example, fearful organisms motivated to reduce or remove threat typically exhibit defensive reactions—physiological changes (e.g., autonomic reflexes such as heart rate changes), verbal responses (e.g., evaluative judgements), and behaviours (e.g., freezing, startle, fainting, escape, and attack) (Lang et al., 1997).

Within this context, the pathological fear and anxiety that characterise panic disorder can be understood as a collection of maladaptive defence responses. Compared to other anxiety disorders, panic patients exhibit greater physiological arousal and higher negative affect when actual danger or threat is absent, which is thought to reflect a compromised defence response (Cuthbert et al., 2003). Since defensive responses result in organism-wide changes, it

follows that examination of phenomena central to anxiety disorders benefits from a consideration of multiple response systems. To this end, Lang's multiple response theory (Lang, 1971, 1993, 1979) provides a unifying framework for empirical studies of anxiety disorders. Accordingly, emotions are expressed on three levels—verbal, physiological, and behavioural. A common finding across a range of laboratory procedures has been that response dimensions are often only loosely-coupled, such as when patients are presented with fear imagery (Lang & McTeague, 2009), or when claustrophobic patients are asked to remain in confined spaces (Alpers & Sell, 2008). Ambulatory studies have also replicated this finding, suggesting that concordance between physiology and self-reports is typically quite low (Barlow et al., 1980; Sievert, 2013). Yet other findings have suggested that as treatment advances, alignment across response dimensions is associated with better symptom reduction (Hodgson & Rachman, 1974; Liddell et al., 1987; Ning & Liddell, 1991). In sum, these results support the assessment of multiple responses when documenting emotional processes. Determining the rules which govern interactions between these systems holds promise for further characterising anxiety disorders and elucidating patterns of symptom variation within a single diagnostic category (Cuthbert et al., 2003; Cuthbert & Insel, 2013).

Self-report Measures

Self-reported measures provide insights into how anxiety symptoms are interpreted and perceived. Real-time assessment of self-reports form the basis of the experience sampling method (ESM), which has been successfully employed in ambulatory studies using clinical samples (Ebner-Priemer et al., 2009; Trull & Ebner-Priemer, 2009). ESMs have been used in a wide range of settings—for example, to study affective instability among patients with Borderline Personality Disorder (Ebner-Priemer & Trull, 2009), work-family conflict (Shockley & Allen, 2013), happiness (Zuzanek, 2013), and flow experiences among individuals who report dysregulated behaviours (Ceja & Navarro, 2012). There have also been successful applications among individuals with anxiety disorders, including the assessment of intrusions and flashbacks among patients with PTSD (Priebe et al., 2013).

Aside from to real-time measures, intervention studies typically assess symptoms such as the severity of anxious thoughts and agoraphobic avoidance using clinical scales, which are also based on self-reports. For example, the Mobility Inventory for Agoraphobia (Chambless et al., 1985), Panic and Agoraphobia Scale (Bandelow, 1995), and Agoraphobic Cognitions Questionnaire (Chambless et al., 1984), have been used to assess patient's avoidance, as well as the

severity of negative cognitions that occur when anxiety manifests. Routinely administered during clinical trials, these scales help to quantify pre- and post-treatment changes in avoidance. Although clinical scales often boast good internal consistency and test-retest reliability, data to support their external or ecological validity is lacking (Chambless et al., 1985). It is therefore uncertain the extent to which these measures can be generalised to different populations and settings.

Physiological Measures

In the absence of actual threat, cardiac defence responses (e.g., heart rate accelerations to facilitate escape from threat, Vila et al., 2007), are hallmark features of panic attacks (American Psychiatric Association, 2013). Abrupt heart rate (HR) acceleration is commonly endorsed by patients (Cox et al., 1994) and observed by researchers (Taylor et al., 1986). Pronounced changes in autonomic nervous system (ANS) activity have been linked to several anxiety disorders, such as panic disorder, post-traumatic stress disorder, and generalised anxiety disorder (Cohen et al., 2000; Friedman & Thayer, 1998a; Thayer et al., 1996). Being under autonomic control, the cardiovascular system reflects the interaction between the parasympathetic (associated with HR deceleration) and sympathetic (associated with HR acceleration) branches of the ANS (Thayer & Lane, 2009). Barring some exceptions (e.g., Margraf, 1990), the onset of panic symptoms coincides with a range of cardiovascular parameters under ANS control, including heart rate increases (Taylor et al., 1986), increases in blood pressure (Shear et al., 1992), and decreased heart rate variability (Friedman & Thayer, 1998a,b).

Ambulatory monitoring of HR among PD/A patients has some precedent and is well-supported by a range of studies (Alpers, 2009; Margraf et al., 1987; Roth et al., 1986; Taylor et al., 1986). Decisions to monitor patient's HR changes during treatment are grounded in findings that it reliably indexes sympathetic arousal (Saul et al., 1990)—a characteristic feature of PD and PD/A. As theories of exposure make predictions about habituation of fear across repeated confrontation of phobic stimuli, assessing heart rate provides a window into documenting changes in sympathetic arousal.

Behavioural Measures

Behavioural disturbances are also conspicuous symptoms among patients with anxiety disorders. In the case of PD/A avoidance of feared situations represents the most prominent symptom. A common method used to assess agoraphobic avoidance and accompanying physiological responses is the Behavioural Avoidance

Task (BAT). BATs were developed to assess clinical symptoms associated with behaviour disorders and have been used as part of intervention studies to demonstrate treatment effects (e.g., Baker et al., 2010; Mavissakalian, 1988). In BATs, individuals are asked to confront a feared stimulus or situation under controlled settings and their approach/avoidance is then measured. Physiological and self-report measures are also commonly recorded during approach. Approach is commonly measured along predefined dimensions such as steps taken towards a feared object; duration in a fear-relevant situation; or proximity to the feared situation (Antony & Barlow, 2010). BATs provide a more ecologically-valid index of avoidance compared to clinical scales such as the Mobility Inventory (Chambless et al., 1985). However, an important reason to be critical of the generalisability of BAT findings is that they are subject to demand characteristics. Experimenter demand characteristics and mode of instruction (personal vs. impersonal) have been consistently found to influence the extent to which participants confront feared stimuli (Bernstein & Nietzel, 1974; Speltz & Bernstein, 1976; Trudel, 1979). This means that the types of demands placed on participants during BATs represents a possible source of bias. In summary, behavioural measures of anxiety is an area that could benefit from new methods which could help characterise both the avoidance behaviours themselves, and the contexts in which they occur.

1.7 NOVEL CONTRIBUTIONS TO THE FIELD OF AMBULATORY ASSESSMENT

The studies presented in this thesis are united by a common disorder, treatment and assessment tools. To this end, patients with panic disorder with agoraphobia were assessed using a commercially-available sports monitor (see Figure 1) while undertaking repeated situational exposures. Specifically, physiological activation (heart rate, HR) and location derived from global positioning system (GPS) coordinates, were collected during each exposure using a commercial sports monitor. Using these indices, a variety of analytic methods were applied with a view to understanding who responds to exposure therapy. A guiding principle in each of the current studies was to retain a clear view of the individual patient.

Study 1 was a based on a single-case, where data was collected from a patient who received treatment at the Otto Selz Institute's outpatient treatment unit. The remaining studies were based on data collected from the PanikNetz project, a multicentre clinical treatment trial. We were interested in a subsample of patients who met criteria for panic disorder with accompanying agoraphobia (PD/A). Treatment consisted of 12 sessions with a psychotherapist, during which time patients completed several standardised situational exposures

Figure 1. A commercial sports monitor (Garmin Forerunner 310XT), seen here on the left arm, was used in all studies to record HR (with the aid of a chest strap, not seen in photo) and GPS-derived position. An ecological momentary assessment device, (Apple iPod Touch, with customised software, iDialogpad), seen here in the right hand, was used in Studies 2, 3, and 4, to assess a range of self-report measures before, during and after bus exposure.

(bus exposure), both with and without therapists. In addition to HR and location, various self-report measures were obtained (e.g., expected anxiety and current anxiety) before and during situational exposure using an ecological momentary assessment (EMA) device. We were particularly interested in examining expectancies about upcoming exposure tasks, as anticipatory anxiety is centrally involved in the maintenance of behavioural avoidance seen in PD/A (Craske & Barlow, 1988; Helbig-Lang et al., 2012; Rachman & Lopatka, 1986a).

Treatment was designed to compare two exposure variants: standard situational exposure and fear augmented exposure, in which patients focussed attention towards fear-inducing sensations, (e.g., bodily symptoms) or specific situational fear cues, and sometimes performed interoceptive exposure exercises if fear did not occur spontaneously. Apart from this contrast, post hoc analyses were designed to examine relationships between constructs derived from clinical scales and psychophysiological measures obtained throughout exposure therapy.

Study 1: A New Tool for Assessing How Patients Undertake Exposure

In this single case, we presented a novel approach to assess how a patient undertook situational exposure. Heart rate and GPS-derived position were collected to objectively document movement paths and accompanying physiological arousal during driving exposure. Tracking where exposure was conducted and the level of arousal

experienced at specific locations allowed us to gauge therapeutic compliance, inform the design of subsequent, suitably-challenging exposure tasks, and to track progress. We assessed the suitability of this novel approach for use within clinical settings. Specifically, the value of this approach as a means providing feedback to patients and therapists was assessed; depictions of the patient's movement and accompanying heart rate during exposure were presented for use within sessions. Further, we considered the extent to which person-specific factors contributed to the positive reception to this new approach.

This study was also motivated by the observation that the external validity of clinical scales used to measure agoraphobic avoidance is not well-established. For example, the Mobility Inventory (Chambless et al., 1985) is still heavily relied on to assess both the level of, and changes in, agoraphobic avoidance. The approach we present in this study is a first step towards expanding the range of methods that objectively document avoidance.

Study 2: Classifying Highly Variable Individual Responses During Exposure

As PD/A is a heterogeneous disorder (Andor et al., 2008; Roberson-Nay & Kendler, 2011) and not all patients respond to treatment (Boschen et al., 2009; Peter et al., 2008), this poses difficulties for obtaining valid, objective indicators of treatment response (Bandelow et al., 1995; Barlow et al., 1994; Whittal et al., 1996). Further, as physiological responses can be expected to be highly variable under complex field conditions, it was necessary to explore the extent to which intrinsic and extrinsic factors shaped physiological responses during exposure.

An overarching goal of the study was to clarify whether individual response typologies existed. Specifically, we tested whether individual's heart rate responses from separate exposures followed specific patterns, and whether different types of responses were apparent across individuals. This constituted a novel approach to assessing whether PD/A subtypes exist; to date, support for panic subtypes has been mixed (Kircanski et al., 2009). Identifying subtypes is important because a major, yet untested, assumption of treatment is that it uniformly applies to a diagnostic entity such as agoraphobia.

Using heart rate segments taken from a circumscribed segment at the beginning of bus exposure, the aim of the study was to sort both individuals and the HR responses into meaningful groups. The extent to which responses were systematically assigned was reasoned to be evidence of disorder subtypes. To this end, we conducted latent cluster analysis, a flexible model-based sorting procedure in which objects are assigned one of several latent, unobserved

subgroups (Vermunt & Magidson, 2002). Several reasons motivated the selection of bus boarding segment. First, bus boarding spanned both anticipatory and initial confrontation with the salient cue, which is a commonly feared situation (American Psychiatric Association, 2013; Wittchen et al., 2010). Second, we expected inter-individual response variance would increase after a few minutes of direct confrontation with the stimulus and could easily be driven by extraneous and complex context factors beyond our control. Finally, the initial segment is of importance to theories of exposure based on the principle of habituation, which assumes that fear level is initially elevated. It therefore made sense to restrict analysis to a peri-boarding interval.

Validation of clusters was also important and presented the opportunity to examine factors that influence concordance between response dimensions (in our case self-reported anxiety and HR). It has been proposed that fear and avoidance behaviours are not always synchronous (Rachman, 1984b; Rachman & Hodgson, 1974) throughout treatment.

Study 3: Examining the Importance of Fear Habituation During Exposure

In this study, we addressed an idea that has guided exposure therapy for the last several decades—that habituation of fear during exposure is necessary to promote reductions in anxiety (Foa & Kozak, 1986; Kozak et al., 1988). To this end, we examined central predictions from emotional processing theory (EPT). Specifically, we sought to determine whether the magnitude of initial fear activation, within-or between-session habituation predicted treatment outcomes (cf. Baker et al., 2010; Foa et al., 1995; Meuret et al., 2012). These predictors were derived from heart rate collected during exposure. The results of our previous study, however, led us to believe that the idiosyncratic HR response patterns would pose difficulties for the operationalisation of these constructs. An additional indicator of habituation was therefore based on the exposure procedure. Specifically, the therapist's exposure instructions (i.e., to remain in exposure until fear had reduced) provided an assessment context which was posited to interact with patient's behaviours (e.g., the duration of bus exposure) and psychological responses (e.g., progressive reduction in anxiety) (Stemmler, 1996). As such, we examined the degree to which apparent adherence to end-of-exposure instructions provided a window into habituation and exposure duration.

Study 4: The Effect of Multiple Exposure Contexts and Compound Stimuli on Maintenance of Treatment Gains

As return of fear is a common phenomena among patients who undertake exposure, we examined whether variables derived from models of fear extinction literature (Bouton & Swartzentruber, 1991; Bouton, 1994; Craske et al., 2014) could explain maintenance of treatment gains following treatment. Specifically, we examined the benefits of exposure conducted in multiple contexts, which was found to be beneficial in laboratory-based fear extinction studies (Balooch et al., 2012; Neumann, 2006; Shiban et al., 2013), applied therapy studies (Lang & Craske, 2000; Rowe & Craske, 1998). Since patients were free to choose where they would undertake unaccompanied exposure, we expected that those who chose more variable, and thus non-overlapping, exposure paths, would, through a process of generalisation, experience less return of fear following treatment. In addition, we explored whether exposure conducted in a mixture of rural and urban settings was superior to that conducted in urban settings alone. Here, we expected more variable exposure contexts to be associated with better maintenance of treatment gains after the end of exposure therapy. Finally, we examined whether the effects of compound stimuli (i.e., combining situational exposure with fear-provocation exercises such as interoceptive exposure; the treatment condition the PanikNetz project) would promote maintenance of gains. Support for this derives from several laboratory studies (Janak et al., 2012; Rescorla, 2006). In summary, using a GPS-derived measure of bus exposure path, and an a priori experimental condition, I aimed to identify factors associated with better maintenance of treatment gains.

1.8 METHODOLOGICAL APPROACHES

In the following studies, I outline several methodological approaches which elucidate how individuals change during and across repeated situational exposure. To reiterate, understanding *who* responds to psychological treatment necessitates that we understand *how* individuals respond. I therefore evaluate the capacity of ambulatory assessment of self-reported fear, physiology, and behaviour to clarify who responds to treatment. In each of the subsequent studies, the overarching aim is to demonstrate the value of exploring predominantly non-manipulated variables derived from psychophysiological responses that manifest during exposure.

Demonstrating the relevance of what does occur during exposure requires that one is mindful of both existing findings (that are often based on what *can* happen under more controlled settings) as well as the methodological issues central to ambulatory assessment. I

therefore present discovery-oriented approaches which are geared towards identifying factors that shape responses to exposure therapy under complex field conditions.

Part II

STUDIES

2

WHERE HAVE THEY GONE? TRACKING MOVEMENT PATTERNS TO DOCUMENT THE PROCESS OF SITUATIONAL EXPOSURE IN AGORAPHOBIA[1]

Therapists typically have limited information about how unaccompanied situational exposure is undertaken. To address this issue, we present a method of assessing movement patterns and concurrent arousal collected during situational exposure. We illustrate how this provides both objective and useful accounts of this important treatment component. In this case study, recordings of global position system-derived (GPS) position and heart rate were obtained from a 47-year-old female patient suffering from panic disorder with agoraphobia who received treatment through an outpatient clinic. Ambulatory assessment of movement and accompanying physiology (heart rate) during situational exposure is described. Visualizations of positional and physiological data recorded during exposure sessions revealed (1) that the patient actually confronted feared environmental cues, (2) that she experienced elevated physiological arousal, and (3) good therapeutic compliance. These depictions were used to plan subsequent exposure sessions and we discuss how this information provided unique insights into the process of exposure. Assessment of movement patterns using commercially available technology can yield clinically relevant information about treatment progress. We conclude that this method could extend traditional self-report measures of agoraphobic avoidance. Future directions, such as the possibility of using movement information to refine follow-up assessment, and the limitations of this approach are discussed.

2.1 INTRODUCTION

It is common knowledge that movement can be accurately tracked with cheap, reliable GPS (Global Positioning System) devices. To date, however, these tools have neither been integrated into routine clinical practice nor into research concerning the core features of mental illness, or the mechanisms underlying therapeutic treatments. Although potentially useful for many clinical disorders, GPS technology seems particularly well suited to track agoraphobic avoidance. Avoidance behaviour is a central feature of panic disorder

[1]APA copyright notice: This article may not exactly replicate the final version published in the APA journal. It is not the copy of record. http://www.apa.org/pubs/journals/pro

with agoraphobia (PD/A) and to varying degrees characterizes all anxiety disorders (American Psychiatric Association, 2000). Although avoidance serves a protective function when individuals encounter threats, among individuals with anxiety disorders, this disruptive behaviour markedly interferes with daily functioning (Hofmann et al., 2009). Not only is it associated with high costs for individuals (Pittig et al., 2014), their family and friends, but according to learning accounts of fear, avoidance serves to maintain fear and prevent extinction of fear responses (Lovibond et al., 2009).

There is some precedence for using technology to enhance clinical practice. Although hitherto not widely adopted within clinical psychology, GPS has been used by health researchers to measure physical activity (for a review, see Maddison & Ni Mhurchu, 2009), exposure to environmental contaminants (Elgethun et al., 2002), to document adolescent travel patterns (Wiehe et al., 2008), and to assess driving style (Porter & Whitton, 2002), to name a few uses. Within psychology, a growing body of literature also supports the use of simple technology to provide therapy-relevant information to clinicians (Boschen & Casey, 2008; Clough & Casey, 2011; Eonta et al., 2011; Morris & Aguilera, 2012). Eonta et al. (2011) showed, for example, how digital photos of a patient's living room could be used to objectively track hoarding behaviour.

Although situational exposure is a well-established treatment for agoraphobic avoidance (Chambless et al., 1998; Deacon & Abramowitz, 2004), objective depictions of how exposure is undertaken and how reductions in agoraphobic avoidance manifest are lacking. This is somewhat surprising since within the field of ambulatory assessment, there has been an increase in the number of targets (e.g., behavioural, cognitive) that can be precisely monitored (Trull & Ebner-Priemer, 2009). For example, within- and between-session changes in anxious responding during exposure have been well documented using physiological and self-reported measures (Alpers & Sell, 2008; Alpers et al., 2005; Bornas et al., 2011; Wilhelm & Roth, 1998). Nonetheless, quantification of behavioural avoidance has typically centred around the length of time participants remain in exposure contexts (e.g., Baker et al., 2010). A more fine-grained assessment of behaviour, however, is generally missing and would help expand current conceptions about how environmental conditions affect patients as they approach feared situations.

Among clinicians and researchers, there is a heavy reliance on self-report measures to provide insights into the personal experience of anxiety and to quantify pre- and post-treatment changes in avoidance—e.g., Mobility Inventory for Agoraphobia (Chambless et al., 1985). Although these measures are quick to administer, and boast high internal consistency and test-retest reliability, there is a paucity of evidence supporting their external or ecological validity.

Measurement of physiology offers useful insights into the processes of anxiety and has been widely employed to test hypotheses about the mechanisms underlying situational exposure. Physiological measures, such as heart rate, are used to evaluate whether initial fear reactivity is predictive of enhanced treatment outcome (Alpers et al., 2005; Meuret et al., 2012), and are often collected in conjunction with self-reported variables during BATs (e.g., Baker et al., 2010). In the case of panic disorder, heart rate (HR) recordings are typically justified on the grounds that tachycardia is the most common and most severe panic attack symptom reported by patients (Cox et al., 1994). Further, there is strong evidence that HR increases, in conjunction with negative thoughts, coincide with exposure to feared situations (Kenardy et al., 1993). However, to relate the often divergent findings from the physiological response domain with meaningful outcome measures (e.g., treatment progress), it remains important to clarify the often complex relationships between psychophysiological, behavioural and subjective reports (Cacioppo et al., 1991).

Behavioural Avoidance Tasks (BATs) are commonly employed to examine agoraphobic avoidance on a behavioural level. BATs assess clinical symptoms associated with behaviour disorders and involve individuals confronting a feared situation under controlled settings and measurement of their approach/avoidance along predefined dimensions, e.g., steps taken towards feared object; duration in a fear-relevant situation; or proximity to feared situation (Antony & Barlow, 2010). Although they provide a more ecologically valid index of avoidance than clinical scales, the behaviours targeted in BATs have, until now, been examined along a limited number of dimensions. This is particularly problematic when conclusions are drawn about avoidant behaviours since it is rare that the various facets of avoidant responding (e.g., hesitation, reliance on safety signals) are comprehensively examined. To avoid narrow conceptualisations of avoidance, we argue that GPS can be used to accurately assess a range of behavioural dimensions as individuals confront feared situations under complex environmental conditions. For example, GPS technology offers a means of objectively and unobtrusively tracking the location of an individual while outdoors (Kerr et al., 2011), and has been found to be more accurate than self-reported activity diaries (Badland et al., 2010). We propose that monitoring outdoor movement with GPS devices can extend our understanding of avoidant behaviour, and can help researchers to explore the interactions between other emotional response domains (physiological, self-report) as participants confront standardised fear-inducing situations.

In sum, among studies in which the mechanisms of situational exposure are examined, equal priority has not been given to these

three response domains—physiological and self-report responses dominate, and when behavioural measures are included, they are commonly based on BAT performance. We argue that more objective accounts of change within and between exposure sessions can be obtained by including measures of movement behaviour from actual exposure embedded in a therapeutic context.

In the current case study, we documented a patient's movement patterns and accompanying heart rate during driving exposure to primarily assess the benefits of this approach within therapeutic settings. Monitoring physiology during driving had the additional advantage that measures were not so strongly affected by exercise activation (Alpers et al., 2003, 2005). We explored whether collection and analysis of physiological and movement data yielded more detailed assessments of homework compliance. Compliance is typically rated by calculating the percentage of homework completed, or number of assignments completed (for a review, see Mausbach et al., 2010). Recent research suggests, however, that assessing the quality rather than quantity of homework completed is a better index of compliance (Cammin-Nowak et al., 2013). Further, guided by current models of feedback that stress the benefits of feedback that is objective, specific, and linked to personal goals (Archer, 2010), we examined the extent to which graphical depictions of movement and accompanying physiology, provided as feedback to the patient, were of therapeutic benefit.

2.2 METHOD

Participant

A 47-year-old German woman with a 12-year history of moderate agoraphobic avoidance who met DSM-IV criteria for panic disorder with agoraphobia (300.21) agreed to participate in the study. The patient left school without matriculating at the age of 16 and then trained for 2 years to become an office clerk. She met her husband at the age of 16, and moved into an apartment with him the following year. After her training, she worked in a firm for five years before taking leave to give birth to her daughter. She then joined a company where her husband also worked and has since remained in this position. She indicated that her presenting problem reduced her capacity to work and placed a strain on the relationship with her husband. She revealed that she began to avoid shopping centres in 2001, and felt increasingly hesitant about driving unaccompanied in her car, travelling on public transport (especially buses and trains), and spending time in crowded places. She received psychological and psychopharmacological (Opipramol) treatment for this problem in 2001, when she initially received her diagnosis, based on a German

version of the SCID interview (First et al., 1996). We confirmed this diagnosis using a SCID-I and II interview in the second and third sessions. During anamnesis, the patient reported that although the earlier therapy had helped to attenuate her symptoms, following this, she periodically felt weak, dizzy, and stressed while travelling on public transport or driving. She claimed that these physiological symptoms led her to develop a fear of losing consciousness. As these beliefs became particularly embedded, she became increasingly reliant on the support of her husband, who was often required to drive her to shopping centres. During the current treatment, she reported taking 37.5 mg of Venlafaxine daily.

Materials

A Garmin 310XT sports watch with an accompanying heart rate (HR) belt was used to capture physiological activation and positional data (GPS). HR, latitude and longitude were recorded at a sampling rate of 1Hz. A city map, downloaded from OpenStreetMap (`http://www.openstreetmap.org/`) was used to set goals for the patient—specifically, to plan the walking and driving paths that would be undertaken during the situational exposure sessions. Clinical symptoms were assessed using German versions of several standardised clinical scales: Body Sensation Questionnaire and Agoraphobic Cognitions Questionnaire (Chambless et al., 1984), Mobility Inventory (Chambless et al., 1985), Brief Symptom Inventory (Derogatis, 1993), and the Beck Depression Inventory-II (Beck et al., 1996).

Procedure

Prior to data collection, the patient provided informed consent and agreed to the publication of this clinical case. Psychological treatment was delivered by a therapist in her third year of psychotherapy training. As part of our institute's standard diagnostic procedures, clinical scales were administered on intake (T0), and after the tenth (T10), twentieth (T20), and thirtieth session (T30, the final session).

Following confirmation of diagnosis, psychoeducation was provided around the nature of PD/A according to Barlow and Craske's (1994) manual. One week before the day of each exposure, the therapist and patient set goals for the initial exposure which involved tracing her intended walking/driving path onto a city map. Prior to each recorded exposure session and while at the clinic, the heart rate belt and sports watch were attached just beneath the sternum and to the left wrist, respectively. The patient was provided with her previously completed city map and asked to refer to this during exposure. Before the driving exposure sessions on which we focus

in this study, the patient undertook four unaccompanied exposure tasks: two fixed-route train exposures in sessions 12 and 13 as part of the client's commute to and from the session, and two city-based walking exposures during sessions 13 and 15. Walking exposures were targeted at the patient's specific fears (walking unaccompanied through the city and entering a department store).

Movement depictions were presented to the patient in the session following each recorded exposure. Although not recorded in session 13, HR was recorded in session 15, which allowed us to present movement and accompanying HR plots. When discussing the results from the walking exposure sessions, the patient explained that entering large shopping centres was still anxiety-provoking. It was therefore agreed that as part of each driving exposure, she drive to and enter a shopping centre. Driving exposure was conducted in sessions 24 and 28 and in total, the patient attended 30 therapy sessions that were 50 minutes in duration.

Data Analysis

Following exposure, data were wirelessly transferred from the Garmin device to a computer (Garmin ANT agent™ software, Version 2.3.3) using a Garmin ANT+ USB stick that was provided with the watch. The resultant file contained time-stamped information about heart rate, longitude, latitude and altitude (TCX, Training Center XML). Garmin's online software tool, Garmin Connect was used to initially view the coordinate data. This web-based software allowed convenient visualisation of HR and movement trajectory[2]. This will suffice for those who want to obtain quick depictions of the route travelled and the accompanying HR response; however, we also outline an approach that affords more detailed analysis of movement patterns.

To more precisely examine the relationship between position and heart rate, we first developed an extraction script[3], developed in the statistical programming language R (R Development Core Team, 2013), to read TCX files and convert these to matrices composed of time-stamped vectors of positional coordinates. In order to more directly visualize the correspondence between physiological activation and position, we embedded HR (indicated by colour) into the path plot (see Figure 2) using R's ggmap (Kahle & Wickham, 2013) package.

[2] A desktop version of this software is also available

[3] This code can be found at: `https://github.com/shiroandy/garmin-tcx-parser`

2.3 RESULTS

Clinical Scales

BSQ, MI and BSI summary statistics indicated that between intake and the tenth session, the patient experienced initial increases in general and phobic anxiety, fear of somatic symptoms associated with anxiety and panic, as well as agoraphobic avoidance (see Table 1). In particular, the patient's MI scores indicated that she almost always avoided heights, shopping centres, crowded city sections, and driving in areas far from home, particularly when alone. A one-month gap in therapy and the patient's report that she felt pressured by her husband to demonstrate that therapy was helping her to change likely contributed to these initial increases. Moreover, large between-session mood fluctuations were a feature of the patient's presentation which partly explained these increases. Between sessions ten and twenty, these scale scores indicated symptom attenuation which confirmed the therapist's impression that the patient responded favourably to treatment, and particularly to the two walking exposures. Decreases in MI and BSI (anxiety scale) scores between T20 and T30 also indicated that further reductions in agoraphobic avoidance and general anxiety symptoms were achieved. Somewhat unexpected was that ACQ scores were considerably lower than in other studies of individuals with PD/A (e.g., Telch et al., 1989), and appeared to fall within subclinical range (Bibb, 1988). Compared to BSQ scores, this suggested that the patient's worries about her somatic symptoms were more severe than her fear-related cognitions. Finally, BDI-II scores indicated that the patient experienced minimal depression throughout treatment. In sum, these scale scores suggested that the extent to which the patient confronted feared situations unaccompanied, increased throughout therapy and that greatest reductions in phobic anxiety were achieved between sessions ten and twenty.

Driving Exposure

During the first driving exposure session for which we collected data, the patient drove for approximately 15 minutes away from the city and paused to undertake exposure in a small shopping centre for 2 hours. Thereafter, the driving exposure continued for another 30 minutes (Figure 2). This plot showed that the patient complied well with the therapist's instructions—the driving exposure was completed as planned and also appeared to allow enough time to undertake the exposure in the shopping mall. The accompanying heart rate suggested that the patient experienced quite high levels of arousal at the commencement of exposure which temporarily

Scale	Measurement Time			
	T0	T10	T20	T30
ACQ[a]	1.86	1.64	1.31	1.57
- Physical Concerns	1.71	1.42	1.50	1.57
- Loss of Control	2.00	1.86	1.71	1.57
BSQ[a]	2.53	3.53	2.41	2.76
MI				
- Accompanied[a]	1.42	1.54	1.25	1.07
- Alone[a]	2.23	3.30	2.37	1.78
BDI-II[b]	6	5	1	5
BSI				
- Anxiety [c]	59	76	50	37
- Phobic Anxiety[c]	60	80	44	44

Table 1. Clinical scale scores on intake, and after sessions 10, 20 and 30.

Note. ACQ = Agoraphobic Cognitions Questionnaire; BSQ = Body Sensations Questionnaire; MI = Mobility Inventory; BDI-II = Beck's Depression Inventory-II; BSI = Brief Symptom Inventory.
[a] Mean score. [b] Sum score. [c] T-score.

subsided and then increased to the initial level as she approached the shopping mall. On the second leg of this exposure, the patient's HR was noticeably lower and for most part was below 100 bpm. This, to some extent, was an expected finding since the patient reported experiencing fewer problems while driving in the vicinity of home.

The depiction of the second driving exposure (Figure 3) also confirmed that the patient undertook the task as planned. The accompanying plot of HR showed that her physiological arousal was particularly high at the commencement of exposure and the HR overlaid on the GPS track indicated that it marginally reduced on entering the car park of the large shopping centre. After completing the exposure within the shopping centre which lasted approximately 2 hours, the patient continued the driving exposure, during which time a further decline in HR was apparent.

Comparing the two driving exposure depictions allowed the therapist to determine that the patient experienced greater difficulties driving in unfamiliar (second driving exposure), compared to familiar (first driving exposure), areas. Although this challenged the idea that the gains would simply transfer from the first to the second driving exposure, an important feature, depicted in both figures, was that HR decreased throughout the exposure. Further, plots allowed the therapist and patient to gauge the difficulty of each exposure,

Figure 2. First driving exposure: GPS track with colour-coded HR embedded in path (above) with accompanying heart rate (beats per minute) plotted against time (below).

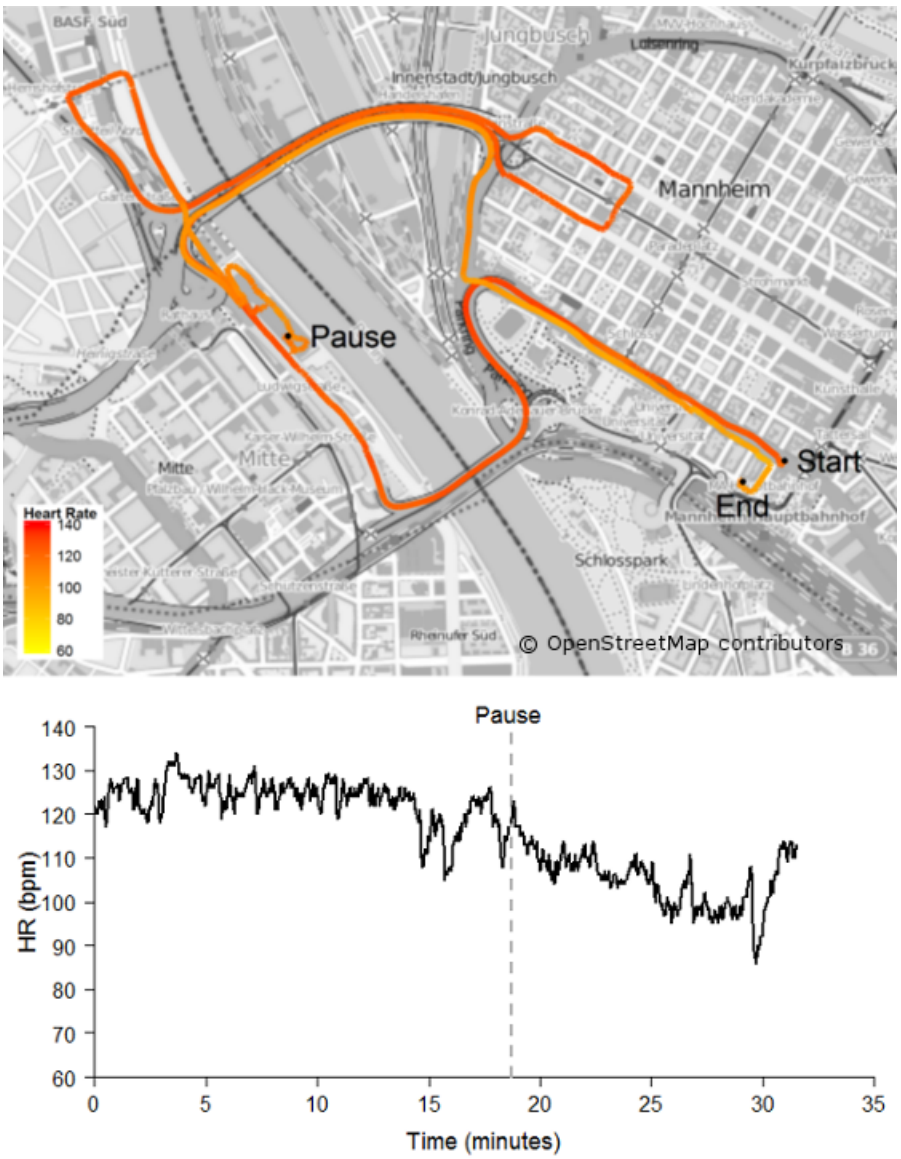

Figure 3. Second driving exposure: GPS track with colour-coded HR embedded in path (above) with accompanying heart rate (beats per minute) plotted against time (below).

which helped clarify the sorts of routes that would be most beneficial for subsequent driving exposure.

Patient Response to Feedback

When asked about the usefulness of the above plots, the patient indicated that seeing her exposure trajectory and accompanying HR reductions contributed to her sense of accomplishment and helped increase her motivation to practice confronting the feared situations. For example, towards the end of therapy the patient drove herself to the clinic more often, and reported that she had driven her

daughter to a neighbouring city. This provided some evidence that the depictions helped to reinforce the idea that despite experiencing strong anxiety symptoms during exposure, completing challenging driving tasks was possible.

The feedback also provided the therapist and patient with an opportunity to discuss the anxiety she experienced in considerable detail after each driving exposure. For both driving exposure sessions, the patient was able see that her HR decreased during the second leg of the journey after undertaking the shopping centre exposure. This, in turn, helped the therapist to demonstrate that given enough time anxiety symptoms could be expected to subside. Further, the therapist noted that HR plots helped the patient to quickly realise that she had overpredicted her fear and arousal prior to exposure.

The plots helped guide and enrich the therapeutic dialogue concerning the onset, character, and time course of emotions as she drove through unfamiliar areas. For example, the therapist was able to initiate a discussion about the experience of crossing a high bridge. Finally, the therapist also noted that the depictions facilitated comparison of different exposure sessions, which helped guide discussions about the patient's progress and helped in the planning of suitably challenging exposure tasks.

2.4 DISCUSSION

We outlined a method of documenting the movement and arousal of a patient with PD/A undertaking situational exposure to establish whether this yielded useful accounts of how this important treatment component was undertaken. Our findings indicated that the depictions provided the patient with therapeutically beneficial feedback, facilitated the planning of additional exposure tasks and allowed the therapist to gauge compliance.

We found that providing feedback to the patient about her movement and arousal patterns during exposure positively influenced her motivation to practice confronting additional feared situations. Inspecting changes in physiological parameters served to highlight the patient's achievements and to strengthen her sense of accomplishment. Discussing this personalised, detailed feedback also enriched follow-up discussions about the experience of completing exposure, and also helped plan additional, suitably-challenging exposure paths. We found that discussing the plots in therapy helped the patient to quickly reach the conclusion that her anticipatory anxiety (as indexed by her elevated HR) at the beginning of exposure was greater than what she experienced during the confrontation of the most feared situation. This, in turn, supported the rationale for situational exposure outlined during therapy.

Documenting movement and accompanying physiology can provide an accurate measure of homework compliance, an important factor associated with successful therapy (Mausbach et al., 2010). Our results indicated that the patient complied with the therapist's instructions as evidenced by a close correspondence between GPS paths and the planned routes. When this is not the case, the depictions can help the therapist to explore why this is so. More nuanced applications are also conceivable. For example, for patients who fear leaving their house, plots of GPS trajectories collected over successive, unaccompanied exposure sessions could objectively document the distance travelled from home throughout therapy. This information could help therapists determine dependence on safety signals, identify additional anxiety cues, and to plan future exposure sessions. Among patients who suggest that their physiological arousal is life threatening, heart rate plots might provide reassuring counterevidence. Further, combining self-report and physiological data with movement pattern depictions could serve as a powerful psychoeducational tool.

Practical Implications

As we present results from a single case, it is important to elaborate on several patient characteristics that may have contributed to the successful application of our method. Despite already having sought treatment in 2001, the patient's condition was chronic, and greatly limited the extent to which she could undertake daily tasks—a report confirmed by her elevated score on the alone subscale of the Mobility Inventory. The patient also appeared particularly motivated to change as she realised that her problem was a burden on the relationship with her husband. Another particularly relevant factor was that the patient was not concerned about having her position or HR monitored during exposure. We attribute this to the patient's positive attitude towards technology, which did not appear to hinge on her educational level and job experience, and also to the time we spent discussing the rationale for data collection. On this note, we cannot rule out that reactivity to ambulatory assessment may influence results—for some, the recording may elicit greater arousal, and thus distract the patient during exposure; for others, the device may be viewed as a safety signal and therefore reduce arousal and ostensibly enhance motivation (Alpers, 2009). More complicated psychophysiological recordings have been used in research studies without much indication of reactivity (Alpers et al., 2003, 2005; Wilhelm et al., 2001). Nonetheless, we recommend discussing patient expectations prior to recording and enquiring about the patient's experience of wearing devices that monitor GPS and physiology. Finally, as the patient took medication throughout treatment, it

could be argued that the anxiolytic effects of the drug were largely responsible for her willingness to engage in exposure and the gains she experienced. Given the low dose of Venlafaxine taken, and the chronic nature of the patient's avoidance, however, it is unlikely that medication substantially accounted for our findings. How well our method generalises to other patients with PD/A therefore requires further investigation.

Although we collected data from two exposure sessions, the quantity of data collected will affect the conclusions that can be drawn. We are confident that when patients are clear about the goals of exposure, movement data from a single session can provide useful insights into various facets of exposure (e.g., compliance, the extent to which challenging situations were confronted as planned, momentary physiological arousal). When changes in avoidance over time wish to be examined, particular attention should be paid to the instructions given to patients as these will affect the extent to which separate exposure sessions are comparable. At a more general level, discussing depictions of their movement and heart rate collected during exposure should bolster the patient's sense of self-efficacy in overcoming agoraphobic reactions. Particularly among patients with negative cognitive schemas who cannot attribute positive gains to their motivation during exposure, visualisation of progress is a tangible supplement to the therapist's encouragement and observations. To achieve valid recordings, we found it crucial to spend time configuring the device and training the therapist. We think that approximately 1 hour should suffice for novice users to learn how to configure a commercially-available device. In our case, we configured the device to minimise the chance that it would pose a distraction to the patient—we disabled all unwanted tone and vibration alarms, disabled visual display of HR, and enabled second-to-second sampling to obtain the most accurate recordings. In our case, a researcher configured the device and then spent 30 minutes informing the therapist how to attach the HR strap and check for a signal, and how to commence and end the recording. During therapy, the therapist spent 30 minutes discussing the details of the recording with the patient—the device was shown to the patient, it was made clear when the device would need to be attached, what would be recorded, and how the results would be of benefit. Our data management procedures were also clearly explained. On the day of exposure, it took approximately 5 minutes to both attach and detach the watch and HR belt. Currently, the device costs 230 Euros (US$335), and the patient received no payment for her participation.

Potential Research Uses

Tracking movement behaviour also holds promise for researchers. It may help inform theories of fear extinction and avoidance by offering detailed insights into the behavioural response domain. This would be timely given that agoraphobia has recently been reconceptualised as a standalone disorder in DSM-5 (American Psychiatric Association, 2013). For instance, various facets of movement could be classified and related to physiological activation during exposure: the order in which fear-relevant contexts are encountered; duration and total distance travelled during exposure; and speed at which feared situations are confronted. Tracking movement patterns might also elucidate how overprediction of anxiety, commonly seen in agoraphobic patients (de Beurs et al., 2002), affects the manner in which fear-evoking situations are confronted. In sum, documenting movement patterns during exposure could provide useful insights into the mechanisms underlying exposure, which remain hotly debated.

Limitations

We acknowledge that the feasibility of our method depends on user's familiarity with commercially-available sports watches. Setting up and undertaking recordings using modern devices should however, be within the grasp of novice users. Subsequent visualisation of GPS paths and accompanying HR is made quite simple with web applications that come bundled with most devices. Nonetheless, we anticipate that many full-time therapists may still hesitate at the time investment required to implement the simplest strategies outlined here. Most GPS devices have been developed for use as sports tools, and have not been tailored for clinical applications. However, we envision that the development of more tailored applications for smartphones will help to reduce the burden on therapists and patients alike.

Embedding HR into movement trajectories is more complex and geared for those with some programming experience. GPS drift sometimes occurs during overcast weather conditions or in built-up areas of cities, when GPS signals are occluded—this can occur on entering, or less frequently, walking near tall buildings. Drift can be corrected using algorithms such as Kalman filters (Jun et al., 2006), however, for the purposes of assessing how long patients have spent in various locations, small amounts of GPS drift should not overly hinder the clinical or experimental utility of plots. Because complex signal analysis is often challenging, becoming familiar with practical issues involved in the collection and processing of GPS data is important (for a review of these issues, see Kerr et al., 2011).

There were also some drawbacks associated with the administration of our clinical assessment scales (i.e., the ACQ, BSQ, MI, BDI-II and BSI). Assessing clinical symptoms directly before and after exposure sessions would have allowed us to more accurately gauge the impact of exposure sessions. As they stand, clinical scale scores only provide a broad impression of how symptoms varied throughout treatment.

Future Directions

In recent years, the use of movement and position sensors has become widespread, in part through less restricted access to devices such as mobile phones. With this technology, it is possible to design studies that help define what constitutes normal movement patterns for various groups of individuals. For example, conducting long-term ambulatory assessment of natural movement patterns among those with and without agoraphobia would yield rich data that could be used to determine the average levels of physiological activation associated with geographic locations throughout treatment. This, in turn, would enable more detailed tracking of symptom changes, and offer superior follow-up assessment possibilities that could allow therapists to detect, on the basis of movement patterns, when their patients are at risk of relapse. Furthermore, comparing the movement patterns of healthy with avoidant individuals could spur development of an objective measure of agoraphobic avoidance. These possibilities are only a sketch of some of the advantages of collecting and using movement patterns. Ultimately, we hope to inspire creative use of such devices (standalone or integrated in smart phones) in clinical practice. In the short term, however, we encourage refinement and extension of our analytic procedures, and anticipate critical evaluation of how this novel data source can be meaningfully used within the field of psychology.

3

CLASSIFYING INDIVIDUAL HEART RATE RESPONSES OBTAINED DURING SITUATIONAL EXPOSURE

Responses to situational exposure measured in the field produce large response variability, especially when heterogeneous clinical groups such as panic disorder with agoraphobia (PD/A) are studied. This calls for an analytic approach which puts the focus on individual responses. Identifying characteristic physiological profiles may contribute to the debate concerning the mechanisms of change underlying exposure, and the identification of clinical subgroups. Our sample consisted of 86 patients with PD/A drawn from the PanikNetz project who completed a total of 234 situational exposure exercises. Heart rate (HR), location (Global Positioning System), and self-report data were collected during a standardised exposure task (a bus ride) repeated on several occasions. Heart rate segments, 5 minutes long and centred around a salient event (bus boarding), were subject to latent class cluster analysis which assigned individual responses to one of a set of unknown classes. Cluster membership was explored post hoc by assessing systematic variation across two active treatment variants (standard exposure and fear augmented exposure) over four consecutive exposure sessions. The clustering procedure meaningfully sorted HR responses on the basis of form and level criteria and provided support for a response typology. Clusters with greater mean HR level were positively related with self-reported anxiety, and clusters with low absolute level and low variability tended to have lower levels of self-reported anxiety during the initial stages of exposure. We argue that traditional, nomothetic approaches based solely on analysis of group averages obfuscate important individual response characteristics.

3.1 INTRODUCTION

Physiological responses, such as heart rate (HR), have often been used to index therapeutic progress (Cacioppo et al., 1991). For example, a debate that continues is whether initial HR reactivity among patients with panic disorder with agoraphobia (PD/A) undertaking situational exposure predicts the magnitude of therapeutic change (Craske et al., 2008; Meuret et al., 2012). Establishing whether specific response types exist and how these map onto other constructs could help move this debate forward. We

therefore pursue an analysis based on individual HR responses to a novel exposure situation, sorted by similarity.

There are good reasons to believe that the HR responses of PD/A patients undertaking situational exposure exhibit substantial intra- and inter-individual variability. First, the heterogeneous nature of PD/A can contribute to response variability. PD is a disorder comprised of somatic, physiological and cognitive symptoms, and some research has suggested that disorder subtypes exist. Using self-report data, respiratory and non-respiratory panic subtypes have been identified (Andor et al., 2008; Roberson-Nay et al., 2012), and other researchers have differentiated between patients on the basis of respiratory and general somatic complaints (Roberson-Nay & Kendler, 2011). Heart rate parameters such as HR variability and QT (the interval between the Q- and T-waves) variability have also been used to distinguish between panic subtypes (Sullivan et al., 2004). During a pre-treatment hyperventilation challenge, PD patients exhibited decreased HR variability and increased QT variability, which suggested that elevated, yet flattened HR responses could be expected, particularly among those with respiratory-type PD. Pairing individual agoraphobic response profiles (behavioural, physiological, and cognitive) with consonant treatments has also proved advantageous and been used to demonstrate the distinctness of individual response types (Michelson, 1986). In contrast, authors of a large review of the panic subtype literature concluded that previously identified PD subtypes were not adequately validated and thus lacked predictive validity (Kircanski et al., 2009).

A second reason to expect large response variance is that the environments in which situational exposure is conducted are complex and not all features can be standardised (cf. controlled laboratory studies). The extent to which environmental features covary with psychophysiological activity is regarded as an important dimension which affects individual responses (Cacioppo et al., 1991). At one extreme, psychophysiological activity can vary as a function of a discrete environmental feature (e.g., when a dog is presented to a person with a dog phobia). At the other extreme, activity can vary as a function of multiple environmental features (contextual features), as can be expected under complex field conditions (Fahrenberg, 1996). As a diverse range of threatening environmental features might be encountered by PD/A patients, even when situational exposure tasks are standardised, it remains likely that accompanying psychophysiological responses vary largely as a function of contextual events.

In sum, the heterogeneous nature of PD/A, as well as the influence of contextual factors inherent in field studies, produce highly variable responses. This calls for an analytic approach that puts the focus on individual responses and would help clarify

the extent to which characteristic physiological profiles are shaped by intrinsic factors. We argue that this constitutes a method of identifying disorder subtypes, and has implications for how a priori experimental conditions, or extrinsic factors are used to address specific research questions.

The Advantages of Focusing on Individual Responses

There are two practical reasons why making sense of individual response characteristics should be encouraged. First, field studies that capture multidimensional data (e.g., physiological, behavioural and psychological) in a variety of situations produce responses with high intraindividual variability. Averaging responses is often justified under laboratory conditions, where tight experimental control helps to minimise response variation attributable to uncontrolled contextual variables (Stemmler, 1996). In field studies, however, greater variability is expected, and it might be too restrictive to assume that the individual response is only driven by intrinsic factors and experimental conditions. Since form invariance and temporal stability of responses are core assumptions of averaging procedures, we maintain that responses should initially be sorted.

Second, identifying distinct response patterns provides an opportunity to determine whether subtypes are present within a diagnostic category (e.g., Furmark et al., 2000). Having discovered disorder subtypes, more could be learned about their aetiology, course and response to treatment (Hayes et al., 1996). In sum, a focus on individual responses represents an improvement on simple averaging of aggregated responses, accords with highly variable responses collected in the field, and can help to identify disorder subtypes.

The Use of Cluster Analysis to Sort Responses

Cluster analysis is a structure-revealing procedure used to identify groups of observations that are cohesive and distinct (Fraley & Raftery, 2002). There are several variants of cluster analysis, each based on a similarity criteria used to determine the probability that a data point belongs to a particular cluster (Kaufman & Rousseeuw, 1990). In contrast to a priori groupings using experimental conditions, model-based clustering makes no assumptions about the structure of data. Rather, classifications result from a combination of hierarchical agglomeration, Expectation Maximisation (EM) and Bayes Factors (Fraley & Raftery, 2002). Model-based clustering procedures are based on the idea that objects belong to one of several latent, unobserved subgroups. The observed scores of objects belonging to a subgroup are assumed to have the same underlying probability distribution (Vermunt & Magidson, 2002).

Following classification and the assessment of cluster stability, it is important to explore what meaning can be ascribed to clusters. This is referred to as cluster validation and is often neglected, which possibly explains why panic disorder subtypes lack predictive and external validity (Kircanski et al., 2009). To address this problem and demonstrate their predictive validity, it has been suggested that clusters be compared using variables external to the sorting procedure (Clatworthy et al., 2005), and ideally using measures from a range of response domains (e.g., psychophysiological, verbal, behavioural).

Introducing Our Method

To determine whether multiple HR response types existed under complex environmental situations, we examined data from PD/A patients who undertook bus exposure on repeated occasions. Patients were assigned to one of two treatments (standard exposure or augmented exposure), and took part in a fixed sequence of exposure that included therapist-accompanied and unaccompanied bus rides. We analysed an assumed salient event at the beginning of each exposure, which was marked by bus boarding. We argued that it (i) was preferable to analyse intraindividual variability during a fixed interval compared to across an entire time series, and (ii) made sense to examine HR during this circumscribed event as this allowed us to examine patterns of initial fear activation (cf. Alpers & Sell, 2008; Baker et al., 2010; Meuret et al., 2012), which plays a central, though hotly debated, role in Foa and Kozak's (1986) emotional processing theory. Further, the use of heart rate as an index of treatment progress is grounded in two descriptive facts about the disorder. First, heart rate has been found to reliably index the fearfulness of individuals during panic attacks (Barlow & Craske, 1994; Roth et al., 1986). Second, compared to other symptoms, cardiovascular disturbances are central to panic (Cohen et al., 2000; Friedman & Thayer, 1998a).

We assessed the external validity of clusters in three ways. First, momentary ratings of anxiety were collected to assess real-time changes in self-reported anxiety during exposure sessions. The benefits of experience sampling methods are well grounded in theory (for a review, see Santangelo et al., 2013). Second, we examined concordance between HR and self-reported anxiety, which has continued to receive considerable research attention (e.g., Alpers & Sell, 2008; Lewis & Drewett, 2006). Concordance refers to the strength of correlation between two response systems, and it was initially proposed that high concordance indexed successful treatment (Hodgson & Rachman, 1974). Although we did not assess the relationship between concordance and treatment outcome, the use of repeated exposure sessions permitted examination of the stability of concord-

ance across sessions. Third, we examined distress tolerance, a central goal of many psychological interventions. Distress tolerance refers to "the perceived capacity to withstand negative emotional and/or other aversive states (e.g., physical discomfort)" and also encompasses "the behavioural act of withstanding distressing internal states elicited by some type of stressor" (Leyro et al., 2011, p.4).

In summary, after individually sorting HR responses into clusters, we were specifically interested whether the mere existence of different HR patterns could be meaningfully related to experimental conditions chosen a priori, or to variables external to clustering. We therefore chose to exclude the influence of gender, symptom severity, and other patient-specific factors. We first examined the phenomenological characteristics of clustered responses to assess their distinctiveness and significance. Of particular interest was the extent to which clusters varied according to form and level criteria, and we expected bus boarding would produce a phasic HR response increase. We explored the cluster breakdown by treatment group and session number, which we predicted to systematically affect cluster assignment. Finally, we examined the capacity of several variables to explain an individual's transition between clusters.

Second, we presented results from non-sorted HR responses according to the factorial design. Of particular interest was whether level-differences or temporal trends distinguished between treatment groups. Specifically, we expected a higher average HR among those in the augmented exposure condition, who provoked bodily symptoms with interoceptive exercises during situational exposure. As the benefits of exposure are usually quickly achieved among those with PD/A, we expected to see a decrease in absolute response level across sessions.

Finally, external validation of clusters was conducted using variables external to the sorting procedure. This allowed us to establish whether examining clustered HR responses facilitated interpretation of self-reported anxiety, concordance between HR and self-reported anxiety, and tolerance of bodily symptoms. In particular, we predicted that concordance between HR and self-reported anxiety would increase across sessions, and would vary across clusters. Bodily symptom tolerance was also expected to improve across sessions, and be greatest among trials assigned to clusters with low absolute HR level. In summary, we aimed to demonstrate the incremental value of sorting psychophysiological responses to achieve a meaningful subdivision within a single diagnostic entity.

3.2 METHOD

Participants

As part of a multicentre clinical study conducted in Germany across 5 outpatient psychological treatment centres[1], participants were randomised to one of two treatment groups after they had been recruited through physician referral and via additional advertisements in various media outlets. Inclusion criteria consisted of: (a) age 18-65 years (b) a current primary diagnosis of panic disorder with agoraphobia according to DSM-IV-TR criteria; (c) clinical global impressions scale (CGI) score $\geq$ 4; (d) ability and availability to regularly attend therapy sessions. Exclusion criteria were: (a) current suicide intent; (b) comorbid psychotic or bipolar I disorder; (c) current dependence on alcohol, benzodiazepine or other psychoactive substance; (d) current psychotherapeutic or psychopharmacological treatment for another Axis I disorder; (e) serious medical illness that excluded exposure-based CBT (e.g., renal, cardiovascular or neurological disease).

Psychotherapists at each of the cooperating treatment centres coordinated recruitment, delivered treatment, and collected data. At the time of data analysis, data from 93 patients were available. After removal of cases with missing or poor quality HR data, the final sample consisted of 86 patients (age: M = 34.03; SD = 10.53; 51 females). The local ethics committees approved all data assessment procedures (for more details, see Gloster et al., 2014).

Materials

Physiological activation (heart rate, HR) and location (global positioning system, GPS coordinates) were collected during exposure using a commercial sports monitor (Garmin Forerunner 310XT). The device relied on a data compression algorithm ("smart recording"), where data points were recorded only when parameters (speed, direction or HR) changed. We found that this algorithm produced a compression ratio of about 3:1 [uncompressed : compressed, 300:300 - (.69 × 300)]. During the study, a software update released by Garmin enabled equidistant sampling of parameters (1 Hz). Ecological momentary assessment (EMA) of self-reported anxiety was conducted with a hand-held computer (Apple iPod Touch) and customised software (iDialogPad, Mutz, Cologne). Responses to self-report EMA items were provided on an 11-point Likert-type scale (Not at all – Extremely anxious); "How anxious are you now when you think about confronting this situation?" (prior to bus boarding), and "How much anxiety are you experiencing now?" (post-boarding),

[1] Bremen, Greifswald, Marburg, Münster, and Würzburg

which was thereafter automatically prompted every 3 minutes while patients were on the bus.

Procedure

The entire psychological treatment comprised 12 sessions and two follow-up booster sessions (two and four months following the last session). Therapy was delivered by advanced-level clinical psychology graduates and post-doctoral students who had received extensive training in the treatment protocol and who were experienced in CBT. Following screening, informed consent, and initial assessment, patients were randomised to one of two treatments: Standard in vivo exposure (standard exposure) or fear augmented exposure (augmented exposure). Therapy sessions were 100 minutes in duration and topics covered in the initial six sessions included psychoeducation; rationale for exposure therapy; behavioural analysis; role of avoidance behaviour; interoceptive exposure; and relapse prevention. A bus ride was chosen as the standardised exposure task as public transport is frequently avoided by PD/A patients. Patients were instructed to remain on buses until their fear reduced, and to assume a seated position in order to minimise artefacts due to exercise activation. In addition to in vivo confrontation in the standard exposure condition, those in the augmented exposure condition were instructed to additionally focus their attention towards cues that induced fear, such as bodily symptoms or specific situative aspects, and to perform interoceptive exposure exercises if fear did not occur spontaneously.

Session number, an ordered, within-subjects factor, represented four possible exposure timings: A therapist-accompanied exposure in Session 7 of therapy (Exposure 1), an unaccompanied homework exposure following Session 7 (Exposure 2), an accompanied exposure in Session 11 (Exposure 3), and an unaccompanied homework exposure following Session 11 (Exposure 4). Although this conflated session type (accompanied/unaccompanied) with exposure timing (Session 7/11), we opted to focus on change across session number as the advantages of therapist-accompanied exposure were recently demonstrated (Gloster et al., 2011). We included data from all patients, irrespective of whether they had completed all or only some of these sessions. Further, some patients repeated homework exposure, and had data from more than 4 sessions (number of completed exposure sessions per patient: min. = 1, max. = 8, M = 2.72, SD = 1.46).

The exposure task was conducted according to a standardised protocol which outlined: The exposure task instructions; device setup and recording instructions, e.g., location where recording was to commence; method of connecting watch to GPS satellites; connecting heart rate monitor; when to create an event marker (start and end of

exposure) using the watch's "lap" button; and when to use the EMA device—before, during (every three minutes) and after exposure. Prior to and during the bus ride, patients were prompted by the EMA device to rate their current anxiety level on an 11-point Likert-type scale ranging from 0 (none) to 11 (very anxious). Patients were instructed to remain on the bus until their fear reduced by itself. Following bus exposure, patients rated how well they were able to tolerate their bodily symptoms on the EMA device.

Therapists instructed patients on the use of devices and recording requirements in preparation for homework exposure tasks. Homework exposure consisted of the same exposure task carried out without the accompaniment of therapists, and patients were instructed to undertake the exercise twice. Patients were provided with a worksheet outlining the required monitoring steps. Neither patients nor therapists were blind to group assignment, however patients were not informed about the research hypotheses.

Preprocessing of the HR Data

From a total of 282 trials, we excluded 48 trials—14 contained no HR data; 30 had insufficient pre-boarding HR data; four, on inspection of GPS paths, did not appear to be bus journeys. The final sample consisted of 234 trials, of which 185 trials had to be decompressed to the common 1Hz grid. The overall missing data rate was judged from the uncompressed recordings and was 8% (SD = 17%).

We used an alternate baseline concept which capitalised on the standardised start to bus exposure (i.e., waiting for the bus). We defined baseline as the tonic HR level, which was best represented by the pre-boarding epoch. Following bus boarding, we expected a phasic HR response. Accordingly, pre-boarding epoch was the reference level in contrasts.

Selection of a Salient Event, Data Alignment and Segmentation

Following data extraction from both devices, HR and EMA data were segmented into three epochs (before, during and after situational exposure exercise). We focused on the interval 2.5 minutes before to 2.5 minutes after bus boarding during a standardised exposure task. The ambulatory assessment data were first visually inspected to control for gross artefacts in heart frequency recordings. We limited our analysis to the first boarding occasion per exposure, as patients sometimes had to change buses to reach their desired destination. The following steps of analysis were completed to identify different classes of HR responses.

Alignment of HR Responses

Although the approximate commencement of exposure was indicated by event markers from the EMA device and sports monitor, a perfect alignment of individual responses patterns with that specific time could not be expected for several reasons. First, individual responses were probably characterised by varying response latencies, and an exact coincidence between pressing a marker key and the actual exposure was unlikely. Second, although patients were instructed to press a marker button on their watch when boarding the bus, this was not always done. For these reasons, we used a combination of GPS-derived position changes, speed, and boarding markers to identify boarding time. This involved manually checking when speed had increased to above 10 km/h, when markers had been pressed on the watch and/or EMA device, and that the position changes could be construed as bus travel (i.e., on a road).

Design and Statistical Analyses

HR responses were analysed according to a repeated measures design. To assess time effects, the HR time series were first collapsed into six 50 s epochs (3 × 50 s before and 3 × 50 s after bus entry). Two strategies were used to analyse responses. First, using a conventional approach, we analysed responses according to experimental (a priori) conditions. Second, the effect of assigning each individual HR response to a specific cluster was examined. In the conventional analysis, fixed effects were based on the a priori design structure, in which treatment (standard exposure vs. augmented exposure) was a between subjects factor, and session number (Exposure 1-4) and time were within-subjects repeated measures. Main and interaction effects of conditions and epochs on HR were computed as a linear mixed effects model using the R-package "nlme" (Pinheiro et al., 2013). Linear mixed models (LMM) are well-suited to deal with violations of multisample sphericity, a common problem in unbalanced repeated-measures designs. LMMs include fixed and random model parts and are better able to account for response variance across time, and handle missing data. Two random intercept model specifications were adopted to facilitate model comparison and to account for non-systematic variance due to individual differences of subjects. In the first model, trials were nested within the participant term to account for non-systematic variance due to individual differences of subjects (at the trial level)[2]. In the second model, the participant term was replaced with a cluster assignment term[3], which represented the result of the clustering solution described below. Since the fixed-

[2] random = ~ 1 | participant / trial
[3] random = ~ 1 | cluster assignment / trial

effects structure was the same in both models, the Akaike Information Criterion (AIC, Sakamoto et al., 1986) and likelihood ratio (LR) tests were used as model comparison criteria—models with lower AIC statistics and a statistically significant LR test were preferred. HR changes over epochs were tested for linear, quadratic and cubic trends across epochs as these encompassed the most plausible excursions of the HR time course. Results of repeated measures analyses were reported by giving the exact probabilities of F-tests for main effects and interactions. We reported sequential F-tests based on Type I sums of squares. Contrasts between factor levels were assessed using t-tests, with corrections for multiple comparisons, where necessary, using the False discovery rate (FDR, Benjamini & Hochberg, 1995). The significance level was set at 5% for all analyses and where applicable, 95% confidence intervals (CI) are presented.

Clustering Procedure

Identifying response typologies was achieved through application of a latent class cluster analysis on the individual raw HR time series data (see Figure 15), using the R package "mclust" (Fraley et al., 2012). This provided a data driven sorting of the individual (300 × 1 s samples) HR time series into a set of a posteriori groups, minimising within group variance and concurrently maximising between group variance. As a consequence, similar HR time courses were grouped together in clusters assumed to represent a set of latent class distributions underlying the empirical data. Averaging the HR response within a given cluster yielded a HR time course, which we considered prototypical for that group (see Figure 4). To establish the degree to which the cluster solution accounted for non-systematic variance, a repeated measures analysis, similar to that previously outlined, was conducted in which cluster assignment was included as an additional categorical model term.

To examine changes in cluster membership (transitions) across sessions, we identified frequent and discriminant sub-sequences on the basis of treatment group and bodily symptom tolerance using the TraMineR package in R (Gabadinho et al., 2011). In addition, we validated clustered HR responses using several external measures. First, EMA was used to assess self-reported anxiety throughout exposure and these data were examined at pre-, peri- and post-boarding epochs. Second, we examined concordance between self-reported anxiety and heart rate. To test concordance between physiology and self-report, we calculated inter-individual Pearson correlations between HR, averaged across pre-, peri- and post-boarding epochs, and momentary anxiety ratings (for the same epochs) across treatment group, session number and cluster. Third, to examine variations in distress tolerance, we calculated

Pearson correlations between self-rated tolerance of bodily symptoms (reported following each exposure session), and treatment, session number and cluster.

Cluster stability measures were derived by comparing cluster solutions based on the complete data set to those based on data sets reduced consecutively by single HR series. We compared our model-based (Latent class) method with k-means and hierarchical (using Ward's minimum variance criterion) clustering algorithms using the "clvalid" package in R (Brock et al., 2011). Several metrics were inspected: Average Proportion of Non-overlap (APN), the average proportion of observations not placed in the same cluster; Average distance (AD), the average distance between observations; and Average Distance between Means (ADM), the average distance between cluster centres.

3.3 RESULTS

Data-driven Sorting of Individual HR Responses

Plots of the six clusters included in the analysis elucidated the results of the sorting procedure (see Figure 4). The average HR time course of trials within clusters showed that individual trials were grouped according to form and level characteristics and showed reduced variability within response clusters. All clusters showed a discontinuity at the estimated time of bus boarding, at time point zero, thus demonstrating that the event was both salient and correctly localised using the various marker sources.

Characteristics of level and form of average time courses within cluster

Two HR level groups could be distinguished through clustering—three low level clusters (3, 4, 5) had a mean HR less than 90 bpm, and three high level clusters (1, 2, 6) were greater than 100 bpm. Although sample size interfered with smoothness of average responses, both Clusters 5 and 6 showed increased variability, which atypically increased after the event. This was in contradiction to all other clusters, which showed the reverse pattern. Cluster 6 showed a variable response that started at extremely high levels (on average, 166 bpm) and decreased to a still high 108 bpm.

Independent of HR level, Clusters 5 and 6 were identified as two high amplitude response clusters. Concerning the association of response amplitude and overall HR level per cluster, the law of initial values would only partly explain the distribution of amplitudes. Generally, higher HR levels were associated with larger response amplitudes. The highest response change was approximately 60bpm and occurred in Cluster 6, which had the highest overall HR level. In Cluster 5, a change of about 10 bpm was observed, with average HR

being around 95 bpm. In Clusters 1, 2, 3, and 4, the pre-event level was already elevated. Following a small rise before boarding, the HR decayed, which comprised an approximate change of 5 bpm.

Most cluster prototypes followed an arousal pattern—this typically involved a small HR increase before the salient event, followed by a decrease. Individual wave forms within clusters showed high variability before the event, and reduced variability thereafter. This showed that responses in anticipation of the event were versatile, and highlighted individual differences in responding.

The effectiveness of the sorting procedure was assessed by comparing unsorted to clustered HR responses. Results of this model comparison showed that inclusion of cluster assignment as a random effects term markedly improved model fit, as indicated by a statistically significant drop in the AIC ($AIC_{unclustered}$ = 11365.56; $AIC_{clustered}$ = 10978.94), and a significant likelihood ratio test, $\chi 2$ (1:2) = 386.61, $p < .0001$.

Cluster stability

The clustering procedure yielded a stable cluster solution, which was reproduced nearly perfectly by alternative methods (k-means, hierarchical). All methods supported the assignment of individual trials to an 8-cluster solution as evidenced by a similar pattern of AD scores. APN values using all methods were close to zero (range: 0 – .11), indicating highly consistent clustering results. There were no systematic confounds between a priori clusters of trials, given by conditions, and a posteriori clusters. Following stability validation, we decided to disregard the two smallest clusters, which contained seven and two trials.

As an additional test of cluster stability, we compared Exposure 1 sessions derived from 2 clustering strategies: (i) based on all individual HR time series (described above), and (ii) based on a data set reduced to initial accompanied (Exposure 1) trials. The "fpc" R package (Hennig, 2014) was used to calculate the corrected Rand index, an index of the agreement between two cluster solutions. Higher scores indicated greater inter-cluster agreement, with a score of 1 denoting identical clusters. Cohen's kappa was also calculated to indicate agreement between the clusters. In our sample, there were 61 unique accompanied (Exposure 1) sessions. There was good agreement between both cluster strategies; corrected Rand index was 0.79, and unweighted Cohen's kappa was .63, *95% CI* = .49 – .77.

Classification of individual responses with respect to a priori groups

We examined the extent to which individual trials were systematically assigned to specific clusters (see Table 2). Fisher's exact tests indicated that there was no association between cluster assignment and session number for both treatment groups (Standard exposure:

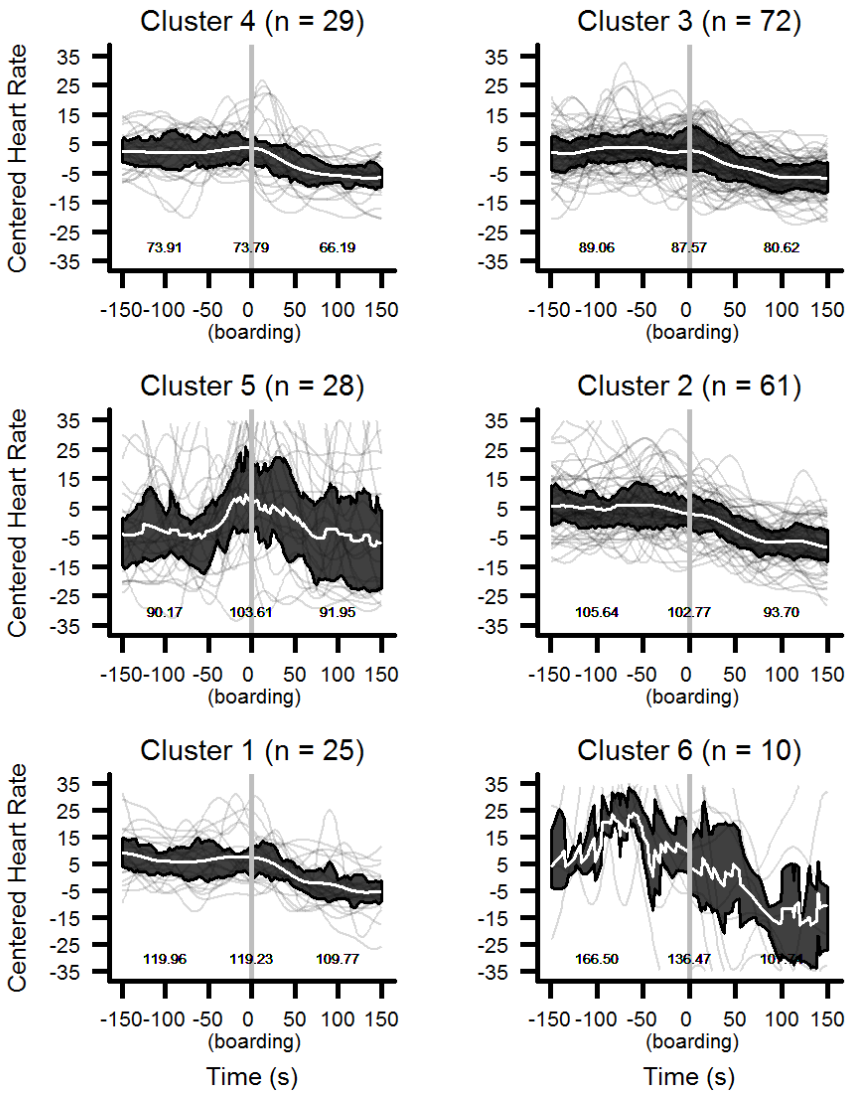

Figure 4. Mean HR responses across clusters sorted by increasing absolute level.

Note. Individual trials in each cluster were first centred by removing the mean HR level. The corresponding level information for three consecutive blocks of 100 s time spans is given in the insets as mean HR level (bpm). The variability and the average time course within each cluster are given by the interquartile range (Q1-Q3) which is the area shaded in grey, and the mean, which is indicated by a white line. Single HR responses within each cluster are super positioned in the background (light grey lines).

p = .87, Augmented exposure: p = .27). Results suggested, however, a statistically significant relationship between both gender (p < .001) and age (p < .02, grouped into tertile bins, < 27, 28-38, > 39 years), on cluster assignment. Specifically, compared to males, trials from female participants were more commonly assigned to Clusters 5 and 6 (7 vs. 21, 1 vs. 9, respectively). Of the three age groups, trials from individuals aged between 28–38 years appeared to have a distinct cluster assignment pattern; comparatively more trials were assigned to Cluster 6, and fewer trials were assigned to Cluster 5.

Intra-individual cluster transitions were also explored by identifying frequent and discriminant subsequences. Here, we found differences across treatment groups and a post-exposure measure of bodily symptom tolerance. Compared to those in the augmented exposure condition, responses from patients in the standard exposure condition were 20% more likely to transition from Cluster 3 to 2—a difference which approached statistical significance, χ^2 = 2.77, p = .10. A transition from Cluster 3 to 2 represented a change from a lower to a higher absolute HR level. In addition, after performing a median split of bodily symptom tolerance (BST), we found that compared to those with low BST, those with high BST had more Cluster 4 responses that remained in Cluster 4, χ^2 = 4.24, p = .04. Since trials in Cluster 4 had low HR level, this suggested that HR responses were likely to remain low when physiological symptoms were well tolerated. Further, it provided general support for BST promoting within-subject cluster stability.

Conventional Repeated-measures Analysis of Heart Rate Responses (a priori Grouping)

We assessed whether HR varied by treatment and session number (see Figure 5). Across treatment groups, HR responses related with bus boarding were almost identical, and were characterised by a decaying HR response that sharply declined between Epochs 4 and 5, which corresponded to 50-100 s post-boarding. Across session number, average response levels differed; responses during Exposure 3 were markedly lower than responses from Exposure 1 sessions, b = -8.38, $t(2034)$ = -3.09, p = .002, which, in turn, were statistically indistinguishable from other sessions. Although responses in Sessions 2 and 4 appeared to decay more quickly than those in Sessions 1 and 3, this did not reach statistical significance—unsurprising given the large within-group response variation reflected by wide confidence intervals. This pattern of results was confirmed by a repeated-measures analysis (see Table 3).

Table 2. Mean HR of clusters (above) and trial breakdown by session number and treatment (below).

		C1	C2	C3	C4	C5	C6
	N	25	61	72	29	28	10
	M	116.30	100.70	85.75	71.30	95.25	136.9
	(SD)	(9.32)	(10.2)	(8.25)	(8.03)	(22.88)	(35.59)
Standard Exposure	Exp. 1	2	11	8	5	2	2
	Exp. 2	5	8	10	2	4	1
	Exp. 3	2	10	11	5	2	0
	Exp. 4	1	11	9	2	2	1
	N	10	40	38	14	10	4
Augmented Exposure	Exp. 1	4	6	8	2	4	3
	Exp. 2	4	4	9	1	8	1
	Exp. 3	3	4	13	8	2	1
	Exp. 4	4	7	4	4	4	1
	N	15	21	34	15	18	6

All fixed effects parameters were first included in an overall model and revealed two significant model terms, epoch and session number (Table 1, top panel). A reduced model that included only these terms was constructed and fixed effects estimates were calculated (Table 2, lower panel). Results suggested a level difference in HR across session number such that average HR during Exposure 3 (M = 90.04, SD = 18.61, *95% CI* = 1.90) was markedly lower than that during the Exposure 1 sessions (M = 99.78, SD = 26.71, *95% CI* = 2.75). The HR difference was approximately 10 bpm, with response variation being considerably greater at Exposure 1. A statistically significant main effect of epoch indicated that HR responses were not stable during the peri-boarding interval. Polynomial trend contrasts for epoch yielded both strong negative linear and quadratic trends. Taken together, this indicated following an initial rise, HR responses generally decayed across time.

External Validation of Cluster Solution

Relationship of cluster assignment to self-reported anxiety

We first examined self-reported anxiety before, immediately after and 3 minutes after bus boarding to determine how it varied by epoch, treatment and session (see Figure 6) . Self-reported anxiety varied by treatment group and compared to the augmented condition (M = 3.48), was significantly lower in the standard exposure condition (M = 2.62), b = -0.84, $t(231)$ = -2.15, $p < .02$. A significant main effect

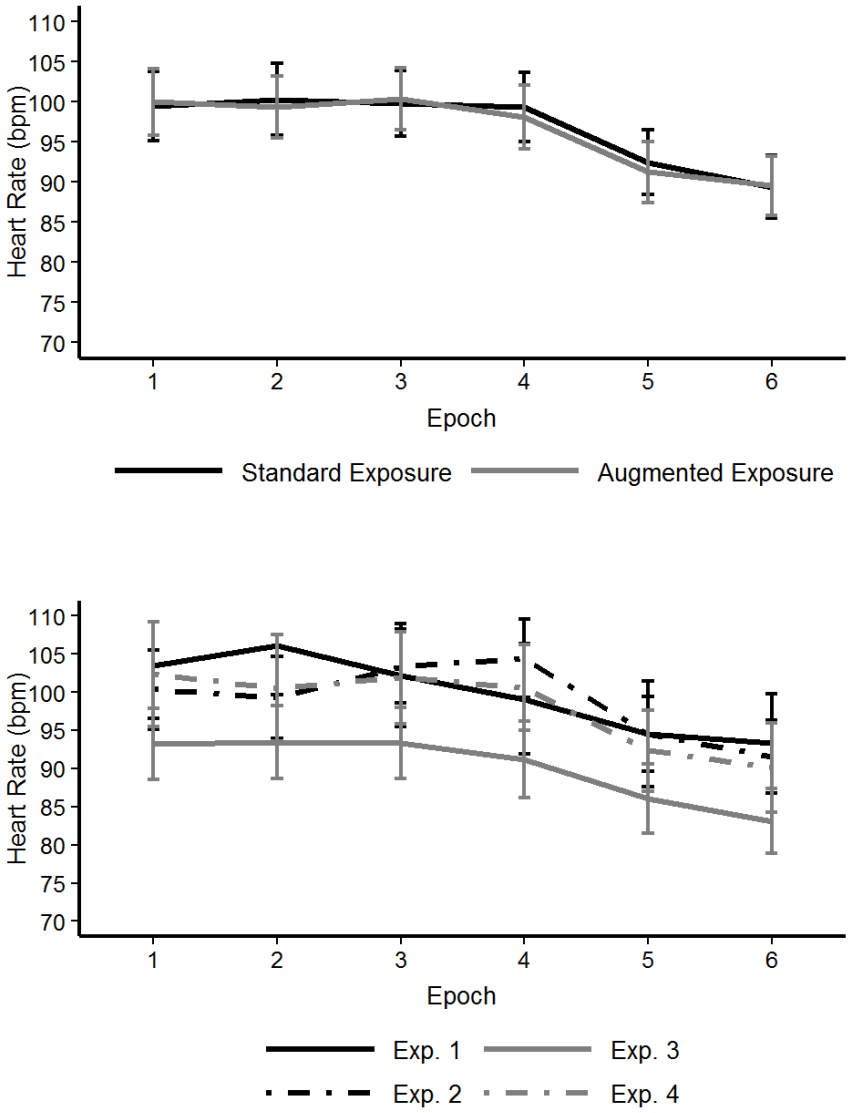

Figure 5. Mean HR by epochs across session number (above) and treatment group (below).

Note. Trial breakdown: Standard exposure (n = 113); augmented exposure (n = 121); Exposure 1, n = 61; Exposure 2, n = 59; Exposure 3, n = 62; Exposure 4, n = 52. Error bars represent confidence intervals corrected for within-subject design

of epoch was also found, with significant reductions in self-reported anxiety occurring between pre-boarding (M = 3.52) and boarding (M = 2.88), b = -0.56, $t(420)$ = -5.67, p < .001, after which no further reductions were detected.

Self-reported anxiety was most strongly determined by the additive effects of session number and epoch, both of which were within-subjects variables. Compared to Exposure 1 sessions (M =

Table 3. Results of the repeated measures analysis of HR for a priori groupings.

Grouping	*df (num)*	*df (den)*	*F*	*p*	
Between-Ss Effects					
Treatment	1	84	0.02	.893	
Within-Ss Effects					
Epoch	5	1130	37.76	<.001	
Session Number	3	142	4.52	.005	
Interactions					
Epoch × Session Number	15	1130	1.42	.132	
Epoch × Treatment	5	1130	0.32	.904	
Session Number × Treatment	3	142	0.15	.927	
Epoch × Session Number × Treatment	15	1130	0.69	.801	
(Intercept)	1	1130	3109.56	<.001	
	b	*SE*	*t*[a]	*p*	*q-val*
Contrasts					
Epochs 2 vs. 1	0.08	1.08	0.08	.940	.985
Epochs 3 vs. 1	0.35	1.08	0.33	.743	.985
Epochs 4 vs. 1	-1.06	1.08	-0.98	.329	.592
Epochs 5 vs. 1	-7.92	1.08	-7.33	<.001	<.001
Epochs 6 vs. 1	-10.26	1.08	-9.50	<.001	<.001
Exposure 2 vs. 1	-0.43	2.77	-0.16	.877	.985
Exposure 3 vs. 1	-8.38	2.71	-3.09	.002	.005
Exposure 4 vs. 1	-0.06	2.90	-0.03	.985	.985
Polynomial Contrasts					
linear [Epoch]	-9.173	0.764	-12.004	<.001	<.001
quadratic [Epoch]	-4.437	0.764	-5.807	<.001	<.001
cubic [Epoch]	0.772	0.764	1.01	.313	.402
(Intercept)	101.67	2.44	41.73	<.001	<.001

Note. The q-val is the adjusted p-values based on False Discovery Rate approach of Benjamini & Hochberg (1995).
[a] df = 2034

3.84), self-reported anxiety was lower at both Exposure 3 (M = 2.44, b = -1.41, $t(229)$ = -3.09, p = .002) and Exposure 4, (M = 1.85, b = -1.99, $t(229)$ = -4.12, p < .001) sessions. No other differences were detected across session number. Across epochs, self-reported anxiety reduced between pre-boarding (M = 3.52) and boarding (M = 2.88), b = -0.57, $t(420)$ = -5.70, p < .001, epochs, and three minutes post-boarding, the level had remained unchanged.

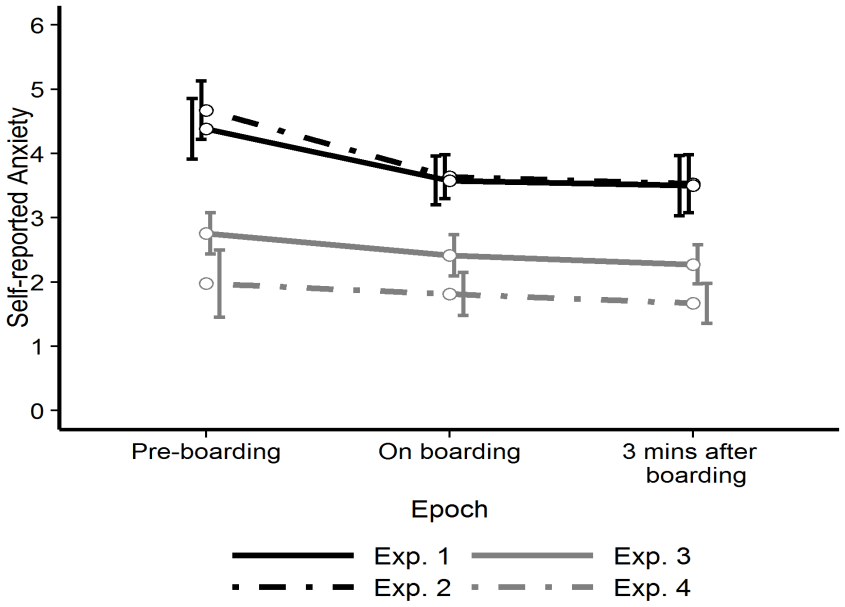

Figure 6. Self-reported anxiety across epoch by exposure session (Exposure 1-4).

Note. Error bars represent confidence intervals corrected for within-subect design and dodged to aid visibility

We examined the correspondence between cluster assignment and self-reported anxiety (see Figure 7). The average self-reported anxiety trajectory decreased linearly across epochs, b = -0.91, $t(394)$ = -4.61, p < .001. There was a significant main effect for cluster and epoch; trials assigned to Cluster 4 (M = 1.12) had the lowest average self-reported anxiety, b = -2.29, $t(218)$ = -3.18, p = .002). The interaction between epoch and cluster approached significance (p = .058), and indicated that trials assigned to Clusters 3, 4, and 5 were described by flatter, less pronounced reductions in self-reported anxiety compared to Cluster 1 trials, b = .773, $t(394)$ = 3.352, p < .001. In sum, the most robust findings based on cluster assignment suggested that HR responses characterised by low absolute level and low variability were associated with lower levels of self-reported anxiety during the initial stages of exposure.

Concordance between HR and self-reported anxiety

Concordance between physiology and self-report was assessed across treatment group, session number and cluster (see Table 4). In both treatment groups, the relationship between HR and self-reported anxiety was weak but significant, standard exposure: $r(338)$= .15, p = .005; augmented exposure: $r(313)$= .17, p = .003. Across session number, Exposure 3 trials were found to be weakly concordant, $r(176)$ = .27, p < .001, providing some evidence

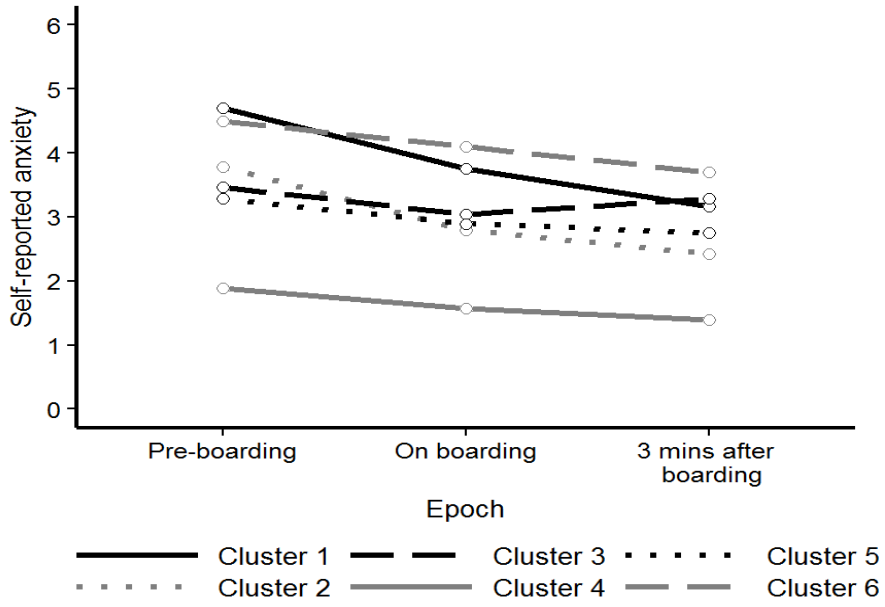

Figure 7. Average self-reported anxiety across epoch by HR clusters.

Note. Each HR represents a typical HR time course.

that physiology and self-reported responses aligned as therapy progressed. Examining this association across clusters helped clarify these findings; trials assigned to Clusters 1 and 4, both characterised by minimal variability, yet varying in absolute HR level (low and high, respectively), were most concordant, $r(69) = .22$, $p = .06$ and $r(81) = .28$, $p = .01$, respectively). It is somewhat surprising that trials in Clusters 2 and 3—also characterised by low variability—appeared more discordant. It is possible that repeated crossing of the tachycardic threshold (100 bpm), as evident in Clusters 2 and 3, may have resulted in greater dysregulation or decoupling of response systems. In sum, these findings should be conservatively interpreted; despite the large N, the wide 95% confidence intervals suggested that the associations were prone to considerable sampling variability.

Self-rated tolerance of bodily symptoms

Tolerance of bodily symptoms, rated by participants after each exposure task was first examined across treatment group and session number. Average bodily symptom tolerance was not found to differ across treatment group (standard exposure: $M = 8.02$, $SD = 2.23$; augmented exposure: $M = 7.87$, $SD = 2.42$). An analysis of variance (ANOVA) on these scores, however, yielded significant variation among session number, $F(3, 212) = 8.68$, $p < .001$. Post hoc Tukey tests were unable to detect a difference between Exposure 1, $M = 7.33$,

Table 4. Concordance: Pearson zero-order correlation between HR and self-reported anxiety across treatment, session, and cluster.

	r	*95% CI*	*df*	*p*
Treatment				
Standard Exposure	.15	[.05 – .26]	338	.005
Augmented Exposure	.17	[.06 – .27]	313	.003
Session Number				
Exposure 1	.13	[-.02 – .28]	172	.08
Exposure 2	.09	[-.06 – .25]	158	.23
Exposure 3	.27	[.13 – .40]	176	<.001
Exposure 4	.05	[-.11 – .22]	141	.53
Cluster				
1	.22	[-.01–.44]	69	.06
2	.06	[-.09 – .21]	163	.42
3	-.001	[-.14 – .14]	204	.99
4	.28	[.07 – 46]	81	.01
5	.10	[-.12 – .32]	75	.37
6	.03	[-.35 – .40]	26	.87

Note. Degrees of freedom vary across clusters as a function of both cluster size and the number of pairwise deleted data.
CI = Confidence Interval

SD = 2.46, and Exposure 2, *M* = 7.05, SD = 2.42, however, tolerance of bodily symptoms increased from Exposure 2 to Exposure 3, *M* = 8.73, *SD* = 1.92, $p < .001$, and remained unchanged in the subsequent Exposure 4 session, *M* = 8.70, *SD* = 2.02.

We were also interested in which clusters were associated with greater or lesser tolerance of bodily symptoms. Trials assigned to Cluster 5 had the lowest tolerance (*M* = 7.04, *SD* = 2.85) and those in Cluster 4 had the highest tolerance (*M* = 9.28, *SD* = 1.36). Compared to the grand mean ($M_{Overall}$ = 8.08, $SD_{Overall}$ = 2.32), Clusters 4 and 5 were statistically different, *b* = 1.36, $t(210) = 3.46$, $p < .001$, *95% CI* = [0.59, 2.14], and *b* = -0.88, $t(210) = -2.13$, $p < .03$, e= [-1.69, -0.07], respectively. In sum, these results indicated that although most participants were able to tolerate their bodily symptoms during exposure quite well, they were better able to do so as therapy progressed. Further, grouping by individual response typologies demonstrated that low and relatively unvarying HR responses (Cluster 4) were associated with better tolerance of bodily symptoms. In contrast, HR responses characterised by attenuated anticipatory activation and a pronounced rise at the time of boarding were associated with diminished tolerance levels.

3.4 DISCUSSION

We clustered HR responses collected during situational exposure to explore the extent to which versatile individual heart rate responses could be meaningfully sorted according to their phenomenology. Responses were successfully grouped by form and level characteristics, and showed minimal within-cluster variability. Clusters 1 to 4 were characterised by low variability and decayed gradually over time. Cluster 5 contained trials with a phasic HR increase that commenced shortly before (ca. 25 s) bus boarding and returned to pre-boarding levels approximately 75 s post-boarding. Trials in Cluster 6 were characterised by a very high HR level that gradually decayed over time, which suggested a slow course of relaxation following initial physiological activation.

In terms of the general cluster phenomenology, cluster types were mainly determined by the anticipatory anxiety/pre-boarding process, which is in line with the assumption that anticipatory anxiety shapes subsequent responses. With the exception of Clusters 5 and 6, which showed a pronounced HR increase at the time of bus boarding, absolute HR level across the entire 5-minute window appeared to be largely determined by pre-boarding arousal. The peri-boarding HR increases in Clusters 5 and 6 may reflect emotional activation due to bus exposure context, however alternate explanations are possible. For example, some patients may have been unable to find a seat before the bus started moving; the traffic conditions may have allowed the bus to travel at a higher speed; the bus may have been more crowded (or empty) than on previous exposure. Any of these factors may have placed an additional load on the cardiovascular system. It was also noteworthy that gender appeared to influence cluster assignment—considerably more trials from female patients were assigned to Clusters 5 and 6, which showed the greatest response variability. However, it is interesting that self-reported anxiety for trials in Clusters 5 and 6 were similar to those of other clusters, and that lowest tolerance of bodily symptoms was found for Cluster 5. One explanation for this may relate to findings that females have a diminished ability to accurately discriminate heartbeats (Katkin, 1985). This poorer autonomic self-perception may have served to attenuate self-reported anxiety in the face of large fluctuations in HR over this relatively short interval. But the discrepancy between anxiety reported during exposure and ratings of bodily symptom tolerance after exposure is puzzling. Specifically, it raises questions about whether the increased propensity of females to exhibit highly variable HR responses (as seen Clusters 5 and 6) is related to the elevated prevalence of anxiety disorders among women (for a review, see Craske, 2003).

Conventional Analysis

To assess the effect of the treatment group, we compared the responses of participants in the standard exposure to those in the augmented exposure condition. We expected those in the augmented treatment group, who undertook interoceptive exercises in addition to standard exposure, to have higher HR compared to those undertaking standard exposure. This was not the case and we were unable to detect differences in HR across treatment groups. It was also of interest whether HR responses varied across the four ordered sessions. We found partial support for our prediction that the magnitude of average response level declined across sessions; a marked decrease was found between Exposure 1 and 3, however, Exposure 2 and 4 were not significantly different from Exposure 1. Although exposure type and number were conflated, this finding provides tentative support for the idea that repeated therapist-accompanied sessions drive reductions in physiological arousal. This appears in line with an earlier finding that therapist-guided exposure plays a pivotal role in the treatment of PD/A (Gloster et al., 2011).

In general, HR responses decayed across the 5-minute interval with the most pronounced drop occurring 50 to 100 s following bus boarding. We also learned that across a priori groups, HR responses were highly variable, as indicated by wide confidence intervals. Standard analysis would try to pry apart these findings further using a variety of statistical procedures such as quantile regression or ANCOVA, however, large inter-group variance renders such approaches insufficient. The finding of no systematic cluster assignment across session number or treatment group indicated that HR responses during the initial stages of exposure probably result from complex multi-factorial interactions, not envisaged in the current study design.

External Validity of Clusters

The external validity of our HR clusters was established by comparing the results obtained using standard analyses, based on a priori groupings, to sorted responses. We found that self-reported anxiety declined across sessions, which indicated that as therapy progressed, patients were less fearful before, and during the initial stages of, bus exposure. We also confirmed that self-reported anxiety was lowest when HR was slow and unvarying (Cluster 4). Closer inspection of cluster breakdown statistics confirmed that this coincided with a general increase in the number of trials assigned to Cluster 4 between Exposure 2 to 3, especially in the augmented exposure condition.

Assessment of concordance between HR and self-reported anxiety across a priori groups revealed that these response domains were most strongly associated during Exposure 3. Although the reduction in self-reported anxiety across sessions is only a single indicator of symptom improvement, it is interesting that we could provide tentative support for Hodgson and Rachman's (1974) initial idea that symptom amelioration is associated with greater concordance between verbal and physiological responses channels. Incorporating cluster-derived information helped us learn that greater concordance occurred among individuals who had slow and unvarying (Cluster 4) HR responses. Trials in Cluster 1 also tended to exhibit greater concordance, which suggests that initial elevation of HR followed by a decaying response following confrontation of exposure context may be necessary conditions for greater alignment of response domains. Further, compared to findings from other studies (e.g., Alpers & Sell, 2008; Lewis & Drewett, 2006) which reported intra-individual correlations between self-reported anxiety and HR as high as .80, our concordance statistics were far more moderate (r = .22 and r = .28, for the two most concordant responses). Given that no systematic pattern between cluster assignment and a priori groups could be detected, this finding suggests that concordance is likely subject to several intra-individual factors. Our results contribute to the literature on concordance between response domains by suggesting that specific HR response profiles may be associated with a greater alignment of response channels.

Finally, self-rated tolerance of bodily symptoms, measured after the completion of exposure, was examined. We found a general increase in tolerance of bodily symptoms across sessions. This is in line with models that suggest that approach behaviour may result in an enhanced capacity to tolerate physical stress (Schmidt et al., 2007). We learned that greatest tolerance of bodily symptoms occurred when HR was slow and unvarying (Cluster 4) and greatest when HR was moderately high and showed a phasic increase at the time of boarding (Cluster 5). It is interesting that the most pronounced difference between Clusters 4 and 5 was the strong phasic increase at the time of boarding. This suggests that the initial reactions on entering the exposure context may weigh heavily on a patient's evaluation of their capacity to withstand uncomfortable physical sensations. Together, these findings underscore the importance of considering individual response typologies when interpreting findings.

How Clustering Supplemented Results from Conventional Analysis

Our findings underline the utility of supplementing results based on a priori variable groupings with sorted individual responses. A

central premise has been that sole reliance on independent variables selected a priori is insufficient to explain highly variable responses that can be expected in field studies. Rather, we suggest that assessing a heterogeneous disorder under complex environmental conditions should prompt researchers to invoke a phenomenological approach that heeds an individual's experience of a specific situation.

Response typology

On the whole, the results of the sorting procedure showed that a specific person could exhibit a variety of HR responses across sessions—there was no stereotypical individual HR response pattern. Thus, our findings provided support for a response-related typology, which means that HR responses varied in terms of level and form, and were not solely determined by individual factors.

To understand the meaning of this result, it is worth considering several possible scenarios. At one extreme, responses could be fully determined by intrinsic factors. Were this the case, cluster analysis would have sorted the separate trials from an individual together. Second, responses could be determined by a priori experimental conditions. Our results indicated that these conditions played a partial role. Finally, at the other extreme, responses could be determined by complex environmental conditions, commonly found in ambulatory settings. Our results partially accord with this scenario as individual responses were dispersed over clusters, suggesting that extrinsic factors dominated the intrinsic response characteristics. Stated otherwise, it appeared that aspects of the environment (e.g., random situative influences, previous physical activation, crowdedness of bus) shaped individual responses. Although extrinsic factors appeared to most strongly determine HR responses, several important findings attributable to intrinsic factors were found. For instance, clusters characterised by low HR level were associated with lower self-reported anxiety; clusters with larger amplitudes, and specifically a pronounced boarding-related increase, were associated with worse tolerance of bodily symptoms. In summary, a response typology makes sense under complex field study situations when large within-subject variability is present.

Limitations

The current investigation had several limitations. First, it was not easy to determine the time of boarding using event markers. However, we capitalised on the redundancy of information across multiple signals to identify boarding (cf. Wilhelm & Grossman, 2010). Ideally, an automatic, fail-safe method of indicating when bus boarding occurred would be used to align responses. In most cases, we were able to use a complete set of markers to determine when boarding

took place, however there were instances, when some markers were missing, which made segment definition more difficult. Second, due to a firmware upgrade by Garmin, part of our sample was based on signals with a lower sampling frequency. We tested the recording properties of the old compression algorithm with a standardised ECG and found two factors limiting acuity: HR changes of 2 bpm were necessary to trigger the sampling of a new data point; a change at the input appeared at the output only with some latency. We acknowledge that decompression may have limited the signal resolution, however believe that this was mitigated by the large HR changes we observed. Third, we discarded clusters with low sample sizes (n = 9 trials) from the final solution. Although we do not offer a suggestion about what can be done with these responses, this does not affect the interpretation of the remaining data. Fourth, as the order of therapist-accompanied and homework sessions was fixed, we were unable to neatly distinguish between repetition and companion effects. Finally, there are biases associated with relying on just one dependent variable (Fahrenberg, 1996) and it remains to be seen whether clustering based on responses combined across modalities would yield more nuanced results.

Future Research

In future research, we aim to expand the current analysis by considering other segments of interest (e.g., the complete bus journey) and assessing the ability of cluster assignment to predict various treatment outcomes. In addition, it will be important to determine the effects of gender, medication and symptom profiles. Further, it would be particularly interesting to assess whether individuals who experienced poorer therapeutic outcomes had a specific cluster transition pattern across exposure sessions.

In summary, specifying the physiological mechanisms underpinning situational exposure is important. This is particularly timely given that agoraphobia has recently been reclassified as a standalone disorder in DSM-5 (American Psychiatric Association, 2013; Wittchen et al., 2010). We believe that ambulatory assessment can help unravel these processes. Further, we maintain that a key benefit of cluster analysis is that it offers a way of sorting highly variable individual responses that result when an unknown mixture of individual and situative influences is present. As they stand, our results show that a relatively simple clustering procedure can be applied to continuous data and that external validation yields interesting results that supplement the findings from typical analyses based solely on a priori groups.

4

DOES HABITUATION MATTER? AN ASSESSMENT OF ADHERENCE TO END-OF-EXPOSURE INSTRUCTIONS AND EMOTIONAL PROCESSING THEORY PREDICTIONS

The importance of fear habituation during exposure-based therapy remains unclear. To address this issue two strategies were applied to determine whether fear reduction during situational exposure indexes anxiety reduction. First, adherence to the end-of-exposure instructions—a convenient, non-manipulated indicator of fear habituation—was used to explore the requirement that reductions in fear during exposure accompany reductions in expected anxiety across treatment. Patients were classified as always, never or sometimes adherent to instructions based on how consistently they remained in exposure until anxiety reductions were experienced. Second, we directly examined whether several indices of fear habituation (initial fear activation, within- and between-session habituation), derived from the emotional processing theory (EPT), were related to changes in expected anxiety across exposure. As part of a multi-centre treatment study, 85 patients with panic disorder with agoraphobia were recruited, and undertook repeated in vivo bus exposure in the second half of a 12-session treatment protocol. Patients who were consistently adherent to end-of-exposure instructions, on average, had higher and more persistent expected anxiety across exposures, compared to patients who never stayed on buses until their fear reduced. An interesting trend emerged for patients classified as sometimes adherent; they appeared to experience greater reductions in expected anxiety across exposure relative to other groups. EPT predictors were generally not associated with reductions in expected anxiety, however, less between-session habituation of heart rate was marginally associated with higher overall expected anxiety. The implications for habituation-based models of exposure and possible benefits of actively varying how long patients remain in exposure are discussed.

4.1 INTRODUCTION

In vivo exposure has a central role in the treatment of panic disorder with agoraphobia (PD/A) (Deacon & Abramowitz, 2004; Hofmann & Smits, 2008; Norton & Price, 2007). Debate continues, however, over the factors which maximise treatment gains for patients who undertake exposure (Craske et al., 2014). This is an important

issue, as exposure is not equally effective for all patients, and it is rare that agoraphobic symptoms completely abate (Peter et al., 2008). Determining which variants of exposure work best could therefore help to strengthen the effectiveness of the intervention.

The emotional processing theory (EPT) of Foa and Kozak (1986) is a prominent learning-based model of exposure therapy which emphasises the role of fear reduction throughout exposure. The concepts of habituation are combined with corrective learning in EPT. Exposure therapy is posited to be effective when confrontation of feared situations activates a "fear structure" (initial fear activation) and allows for the integration of incompatible information (Foa & Kozak, 1986). The fear structure is conceptualised as a set of propositions (true/false statements) about a stimulus (e.g., open spaces), response (e.g., racing heart) and their meaning ("I am going crazy") that are stored in memory (Lang, 1971). Integration of new information into the fear structure is thought to result in the development of a non-fear structure that replaces or competes with the original one. The non-fear structure is thought to be evidenced by within- and between-session habituation of physiological responses.

Support for EPT has come from a number of therapy studies, which have found a positive relationship between model components and treatment outcomes. In one, initial fear activation during, and between-session habituation (indexed by heart rate) across, imaginal and in vivo exposure, was negatively related to the severity of post-treatment fear and avoidance among a small sample of patients with obsessive compulsive disorder; within-session habituation of HR and skin conductance did not predict treatment outcome (Kozak et al., 1988). In another, initial fear activation (taken from the first imaginal exposure session) and between-session habituation, as indexed by heart rate (HR) of snake phobics who undertook 11 systematic desensitisation sessions, predicted successful treatment outcomes (Lang et al., 1970). Foa et al. (1995) also found that greater self-reported distress during initial imaginal exposure was associated with greater reductions in trauma-related pathology in a sample of assault-related trauma patients. Finally, in a study of claustrophobic individuals exposed to a confined space (also a typical agoraphobic situation), initial fear activation, indexed by elevated HR in the first exposure session, predicted superior treatment outcome (Alpers & Sell, 2008).

However, empirical evidence for the EPT is not consistent. Meuret et al. (2012) recorded HR, respiratory rate, and carbon dioxide partial pressure from 34 PD/A patients during situational exposure. Results indicated that initial (physiological) fear activation, within- and between-session habituation were not related to the slope of change in treatment outcomes; but self-report measures were inversely correlated with treatment gains. One uncontrolled, and

possibly biasing aspect of this study was the occasional presence of therapists, who may have represented safety signals for some patients. Patients who acknowledged that therapist's presence might attenuate their anxiety were required to travel unaccompanied. However, there are reasons to doubt the validity of these self-assessments. Given that therapist-accompanied exposure has been linked to superior outcomes relative to unaccompanied exposure (Gloster et al., 2011), this appears an important factor to control. In sum, despite some support for EPT predictions, the small sample sizes and inconsistent findings across different measures of fear may limit the strength of inferences that can be drawn from these results. Findings from more recent, better controlled studies have also yielded mixed findings (Meuret et al., 2012; Baker et al., 2010). Nevertheless, EPT remains an influential theory as exposure guidelines outlined in treatment manuals often recommend encouraging patients to remain in feared contexts until their fear subsides (e.g., Abramowitz et al., 2011; O'Donohue & Fisher, 2012; Puri & Treasaden, 2011; Clark & Beck, 2010; Lang et al., 2012). We therefore maintain that this warrants a re-examination of EPT predictions using a large sample of PD/A patients who undertook multiple standardised bus exposures.

Several reasons may explain why the EPT has been inconsistently supported. First, initial activation and habituation may apply only to a subset of panic disorder patients. This idea would be congruent with findings that have identified a cardio-respiratory panic subtype characterised by palpitations, shortness of breath, choking, chest pain, and numbness (Meuret et al., 2006; Roberson-Nay et al., 2012; Buller et al., 1986; Andor et al., 2008). Several patterns of improvement were found in a sample of post-traumatic stress disorder patients whose treatment involved repeated imaginal reliving of their trauma; both high initial fear activation and between-session habituation experienced greatest treatment gains (Jaycox et al., 1998). Second, large response variability under field conditions (i.e., while undertaking situational exposure) may mask the effects of specific model components. A solution we pursue here is to subset patients on the basis of a clinically-relevant responses to gain access to the effects of habituation.

Heart Rate: An Index of Phobic Arousal

Cardiovascular symptoms, and how they are interpreted by patients, are central to PD/A (Cox et al., 1994; Friedman & Thayer, 1998a,b). In several ambulatory studies of autonomic functioning, higher average heart rates were found among panic disorder patients relative to age- and sex- matched controls (Aikins & Craske, 2008; Bystritsky et al., 1995). It has been proposed that this elevated baseline arousal reflects a pervasive anticipation of threat, produced

by ongoing stimulation via external or internal cues (Craske, 2003). Evidence from an ambulatory study also confirmed that relative to healthy controls, panic disorder patients experience an upward spiral of HR, cardiac perceptions and anxiety (Pauli et al., 1991). Monitoring of HR to index arousal is therefore recommended in the case of PD/A and is well suited to ambulatory assessment since panic symptomatology typically presents outside the laboratory in circumscribed situations (Alpers, 2009). In sum, these descriptive findings support the use of HR to index patient's anxious arousal, particularly for evaluation of EPT components.

Adherence to End-of-exposure Instructions as a Window to Habituation

In the present study, patients were instructed to remain on buses until fear reduced. Following therapist instructions would, in many cases, have created a state of dissonance or psychological discomfort (Festinger, 1957)—engaging in a desired behaviour, (e.g., waiting until fear subsides) while feeling discomfort about confronting a feared situation. Constraint satisfaction theory (Shultz & Lepper, 1996) has recently been applied to explain the mechanisms behind exposure therapy. To this end, therapeutic change is argued to result through a process of constraint satisfaction that promotes consonance between external environmental input and internal states that control behaviour (Tryon, 2005). In this light, remaining in an exposure context until fear reduces represents a constraint since it entails restraining or taking control of the maladaptive cognitions, emotions and behaviours that maintain anxious avoidance. Patients who managed to adhere to end-of-exposure instructions given this constraint, were therefore reasoned to have habituated to the environment and the sensations it evoked.

This therefore warranted an examination of the association between exposure duration, patient's apparent adherence to end-of-exposure instructions, and treatment progress. Only a few studies have examined the effects of exposure duration. In a group of height phobics, longer exposure resulted in greater reduction of self-reported fear than shorter exposure durations (Marshall, 1985). In contrast, among female college students who confronted a harmless snake in a 30-step behavioural avoidance test, exposure duration (10 vs. 30 seconds) had no effect on the performance of participants (Trudel, 1979). These inconclusive results helped motivate the current analysis of exposure duration.

In addition, apparent adherence to end-of-exposure instructions was examined by identifying patients who experienced reductions in self-reported anxiety and physiological arousal during exposure. Rather than placing a focus on the degree to which responses habituated, this strategy allowed examination of intra-individual

patterns of apparent adherence across exposure. Adherence in the current context, overlaps with therapeutic compliance—highly anxious patients who remained in exposure contexts until fear reduced were regarded as compliant. Homework assignments are an integral part of manual-based treatments for PD/A (Barlow & Craske, 1994), and compliance has been strongly linked to therapy outcomes (Cammin-Nowak et al., 2013; Kazantzis et al., 2000). Given the specific instructions provided to patients in the current study, we therefore determined whether consistent adherence to end-of-exposure instructions was associated with enhanced treatment response.

The Role of Expectancies/Anticipatory Anxiety

One reason for the inconsistent findings concerning EPT may result from a focus on broad definitions of treatment outcomes using clinical scales that assess general symptom severity. Broad treatment outcomes are clearly important, however in the face of such inconsistent results, we argue that it may be beneficial to examine processes more proximal, and thus more sensitive, to exposure-related change. The maladaptive cognitions of patients with PD/A are often future-oriented perceptions of threat (Alpers, 2010). Further, there is converging evidence that a change in expectancies precedes reductions in agoraphobic avoidance (Westra et al., 2007; Whittal & Goetsch, 1997) and also plays a role in the maintenance of behavioural avoidance seen in PD/A patients (Craske & Barlow, 1988; Helbig-Lang et al., 2012; Rachman & Lopatka, 1986a). In one study, panic disorder patients with moderate to severe agoraphobia rated the probability that they would experience a panic attack prior to attempting an individually tailored behavioural test chosen from their fear and avoidance hierarchy. Expectancies were found to be closely associated with subsequent behavioural avoidance (Craske et al., 1988). Exposure appears to offer patients an opportunity to violate their expectancies which also appears to influence subsequent anxiety (Craske et al., 2014). For example, snake phobic participants who experienced unexpectedly high anxiety during exposure subsequently showed higher expected anxiety and avoidance on re-exposure (Rachman & Lopatka, 1986a,b). Expectancies therefore appear to play an important role in determining the extent of phobic avoidance. In the current study, we therefore assessed self-reported expected anxiety and fear of losing control prior to each exposure. A closer examination of factors that predict reductions in expected anxiety appears warranted as this may distinguish between who responds and who does not respond to treatment.

Study Aims

The first aim of the current study was to assess the effect of adherence to end-of-exposure instructions on treatment response. To this end, several criteria of adherence (i.e., experiencing anxiety reductions during bus exposure) were examined on the basis of exposure duration and adherence to instructions (i.e., reductions in self-reported anxiety and heart rate). Given that patients were asked to remain in exposure until anxiety reduced, and because we expected anxiety to reduce across sessions, we expected exposure duration to decrease across sessions (Criterion 1). In line with this, longer average exposure duration was therefore expected to be associated with poorer treatment response, as indicated by reductions in two measures of expected anxiety (Criterion 2). We were also interested in the consistency of each patient's adherence to end-of-exposure instructions across repeated exposure. We predicted that consistency of adherence would be associated with specific patterns of individual exposure durations (Criterion 3); and reductions in measures of expected anxiety across sessions (Criterion 4).

A second aim of the study was to determine whether EPT components (Foa & Kozak, 1986) could adequately account for changes in expected anxiety—our measure of treatment progress. Support for the theory was expected to be present when greater initial fear activation, within-session habituation, and between-session habituation were inversely associated with changes in expected anxiety across successive sessions.

4.2 METHOD

Participants

Data were collected from 93 patients who had a diagnosis of panic disorder with agoraphobia (PD/A) as part of a clinical study conducted across five treatment centres in Germany (PanikNetz). Participants were randomised to one of two treatment conditions after they had been recruited through physician referral and via additional advertisements in various media outlets. Inclusion criteria consisted of: (a) age 18-65 years (b) a current primary diagnosis of panic disorder with agoraphobia according to DSM-IV-TR criteria; (c) clinical global impressions scale (CGI) score $\geq$ 4; (d) ability and availability to regularly attend therapy sessions. Exclusion criteria were: (a) current suicide intent; (b) comorbid psychotic or bipolar I disorder; (c) current dependence on alcohol, benzodiazepine or other psychoactive substance; (d) current psychotherapeutic or psychopharmacological treatment for another Axis I disorder; (e)

serious medical illness that excluded exposure-based CBT (e.g., renal, cardiovascular or neurological disease).

Psychotherapists at each of the cooperating treatment centres coordinated recruitment, delivered treatment, and collected data. After removal of cases with missing or poor quality HR data and and whose HR was not assigned to one of six previously-identified HR clusters, the final sample consisted of 85 patients with PD/A (age: M = 34.21; SD = 10.47; 50 females) who had completed at least one bus exposure. This sample completed a total of 227 bus exposures. The local ethics committees approved all data assessment procedures (for more details, see Gloster et al., 2014).

Materials

Physiological activation (heart rate, HR) and location (global positioning system, GPS coordinates) were collected during exposure using a commercial sports monitor (Garmin Forerunner 310XT). The device relied on a data compression algorithm ('smart recording'), where data points were recorded only when parameters (speed, direction or HR) changed. This algorithm produced a compression ratio of about 3:1 [uncompressed : compressed, 300:300 - (.69 × 300)]. During the study, a software update released by Garmin enabled equidistant sampling of parameters (1 Hz) in 48 bus exposure trials. The remaining 179 trials were decompressed to a common 1 Hz grid.

Ecological momentary assessment (EMA) of self-reported anxiety was conducted with a handheld computer (Apple iPod Touch) and customised software (iDialogPad, Mutz, Cologne). Responses to self-report EMA items were provided on an 11-point Likert-type scale (Not at all anxious – Extremely anxious); "How anxious are you now when you think about undertaking the exposure task?" (prior to bus boarding), and "How much anxiety are you experiencing now?" (post-boarding). The device automatically prompted patients to answer this final question every 3 minutes during the bus ride. On exiting the bus, participants were again asked to rate their current anxiety level.

Measures

Overall treatment efficacy and outcome

The Mobility Inventory (alone subscale) (MI, Chambless et al., 1985) and Agoraphobic Cognitions Questionnaire (ACQ, Chambless et al., 1984) were used to determine treatment efficacy and treatment outcome. The MI assesses agoraphobic avoidance of various situations (in our case, while alone) and is comprised of 27 items, which are rated on a 5 point Likert-type scale from 1 = *never*

avoid to 5 = *always avoid*. The measure has demonstrated good internal consistency and discriminant validity (Chambless et al., 2011). In addition, the ACQ was used to measure the severity of maladaptive thoughts when anxiety is experienced. It has 15 items, which are scored on a 5-point Likert-type scale ranging from 1 = *thought never occurs* to 5 = *thought always occurs*. To examine treatment efficacy, we assessed the linear slope of change of MI- and ACQ-scores over assessments.

Treatment progress

Treatment progress was assessed using two measures of self-reported anxiety collected shortly prior to bus exposure—expected maximum anxiety and fear of losing control. Principal components analysis without rotation was carried out to provide an empirical summary of the self-reported items assessed prior to bus boarding (see Appendix section C.2). The two components accounted for 78.4% of the total variance in pre-boarding responses, and were characterised by "Expecting the worst" and "Losing control". Based on these results and an inspection of the scree plot, we decided to base treatment progress on expected maximum anxiety and fear of losing control. Because these were proximal in time to bus exposure (unlike treatment outcome measures), they had the benefit of being able to document change between repeated exposure. Ratings were obtained for each session (Exposure 1-4), and mean values were calculated when multiple unaccompanied exposure sessions were completed.

Adherence to end-of-exposure instructions

Patients were asked to remain in the exposure context until their fear reduced. We operationalised adherence to instructions using two methods: (i) reductions in self-reported anxiety and (ii) reductions in heart rate, from the start (bus boarding) until the end (exiting bus) of bus exposure. Sessions in which self-reported anxiety reduced by at least 1 point on the 11-point Likert scale were classified as adherent to end-of-exposure instructions. Robust linear regression models were fit to HR data; Sessions in which slope coefficients were less than zero were classified as responsive.

Consistency of adherence was also examined. Using classifications of each patient's individual exposure sessions, we categorised individuals as *always, never*, or *sometimes* adherent. Before assessing its effect on expected anxiety trajectories, we assessed whether consistent adherence was a function of initial symptom severity, which was determined from Mobility Inventory and Agoraphobic Cognitions Questionnaire scale scores at intermediate assessment (just prior to exposure).

Emotional processing theory variables as predictors of improvement

The combined effect of EPT components on treatment progress was examined. Initial fear activation (IFA) was operationalised as the mean peri-boarding HR (±2.5 minutes)—the same interval on which clustering was based. We argued that this segment was particularly salient for patients as it marked the beginning of exposure; was a circumscribed, standardised segment; and contained both anticipatory and initial confrontation with the exposure context.

As intra-individual HR responses were highly variable and did not follow a simple pattern (for a full description, see chapter 3), this suggested that time-varying predictors should be used to assess relationships with treatment progress. This approach is not uniformly adopted. For example, Meuret et al. (2012) calculated within-session activation and habituation by averaging HR across repeated exposures; inter-session changes were therefore discarded.

We therefore derived within-session habituation from the robust linear regression slope coefficient, which represented an individual's average rate of change in HR across a single bus exposure. Finally, between-session habituation was also based on individual robust linear regression slopes, which provided an index of the change in maximum HR values across repeated exposure sessions (for a depiction of this operationalisation, see Appendix section B.3) . We restricted the segments from which between-session habituation was calculated to the time in which patients were physically situated on the bus (for a summary of grand averages, see Table 5). Since each EPT parameter depended on the availability of different HR segments, there was a variable number of trials associated with each of these parameters. For example, initial fear activation required an additional 2.5 min of pre-boarding HR, however, within-session habituation depended only on HR while patients were on the bus.

EPT parameters revealed that on average, participants had moderate levels of fear activation, with HR generally in a sub-tachycardic range. On average, HR decreased by 0.30 bpm throughout exposure. Between sessions, maximum HR decreased on average by 0.70 bpm per session, however, large between-patient variation was apparent (SD = 6.13).

Procedure

In-vivo exposure therapy

The entire psychological treatment comprised 12 sessions and two follow-up booster sessions (two and four months following the last session). Therapy was delivered by advanced-level clinical psychology graduates and post-doctoral students who had received extensive training in the treatment protocol and who were experienced

Table 5. Descriptive statistics for emotional processing theory predictors.

Measure	*M* (*SD*)	*n* (trials)	*n* (person)
Initial Fear Activation [bpm]	95.46 (16.10)	207	85
Within-session habituation [Δ bpm/min]	-0.30 (0.49)	221	85
Between-session habituation [Δ bpm/session]	-0.70 (6.13)	207	65

Note. Initial Fear activation = Mean HR at bus boarding ±2.5 minutes ; Within-session habituation = robust regression trend [linear] based on HR within ; Between-session habituation = robust regression trend [linear] based on maximum HR in each exposure

in CBT. Following screening, informed consent, and initial assessment, patients were randomised to one of two treatment conditions: Standard in vivo exposure (bus exposure) or fear augmented exposure, which, in addition to standard in vivo exposure, involved focussing attention towards fear inducing aspects, (e.g., bodily symptoms) or specific situational fear cues, and sometimes performing interoceptive exposure exercises if fear did not occur spontaneously. Treatment condition effects were not examined in the current study. Therapy sessions were 100 minutes in duration and topics covered in the initial six sessions included psychoeducation; rationale for exposure therapy; behavioural analysis; role of avoidance behaviour; interoceptive exposure; and relapse prevention. Clinical assessments were conducted as part of the PanikNetz study protocol at intake, baseline, mid-way through treatment (intermediate; after session 6), following the final (12th) session (post-assessment) and at follow-up (6 months following the final session).

After completing the initial six therapy sessions, patients undertook several bus exposure sessions. A bus ride was chosen as the standardised exposure task as public transport is frequently avoided by PD/A patients. Four session types were possible: In Session 7, therapist-accompanied (Exposure 1), was followed by an unaccompanied session (Exposure 2), and in Session 11, therapist-accompanied (Exposure 3) was followed by an unaccompanied session (Exposure 4). Repetitions of unaccompanied exposure were possible (maximum number of repetitions: Exposure 2 = 4; Exposure 3 = 3).

Patients were instructed to remain in the exposure situation until the fear they experienced in the situation reduced by itself. Exposure sessions therefore had variable duration (M = 51.91 min, SD = 35.15 min, range = 2.67 – 253.25 min).

Preprocessing

Heart rate responses and ecological momentary assessment data were aligned using boarding markers from both devices and verified by examining plots of GPS-position and speed changes. In the current study, we were also interested in the duration of bus exposure, so HR, GPS-position and self-reported data between boarding and disembarking the bus were extracted. Single exposures sometimes comprised multiple bus excursions (e.g., when a patient needed to change bus), so time on both buses was combined.

Raw data were collapsed across time and a log transformation was applied to bus ride duration to help correct deviations from normality. Outliers (± 3 *SD*) were identified among treatment outcome measures and predictors, and were replaced with the nearest non-outlier value according to the Winsor method (Guttman, 1973). No more than 10% of values were Winsorised (initial fear activation= 5%; within-session habituation= 10%, between-session habituation= 10%) to correct outliers (for distributional plots of EPT components, see Appendix section B.4).

Analytic Strategy

Robust linear regression, based on Huber's M-estimator (Huber, 1981), was used to calculate HR habituation (within- and between-sessions). This procedure has the advantage of being resistant to outliers, which, despite preprocessing of HR signal, still may have biased typical ordinary least-squares methods (Tabachnick & Fidell, 2007).

Before fitting statistical models, treatment efficacy was first evaluated by estimating the linear trend across measures of treatment progress and treatment outcome. Doing so allowed us to determine whether patients generally improved in response as treatment progressed. Empirical growth plots for expected anxiety and fear of losing control trajectories were plotted for each patient to examine the relationship between outcome measure and time. For expected anxiety, it was evident that the individual growth process did not always follow a simple linear trajectory, but seemed to include a quadratic trend (see Appendix section C.1). However, a formal comparison of first-order (time) and second-order (time + time^2) polynomial models revealed that inclusion of a quadratic growth component was not justified, Likelihood Ratio (LR, $df = 4$) = .18, p = .99. An individual growth model that only included linear change trajectories was thus adopted for subsequent analyses. In contrast, including linear and quadratic growth processes for fear of losing control appeared justified, LR ($df = 4$) = 10.88, p = .03 (see Appendix section C.1).

We first explored variation in EPT components across binary treatment response, which we calculated from expected anxiety trajectories. OLS slope parameters were extracted and a cut-off of -1, corresponding to an average reduction of 1-point across repeated exposures, was used to divide the sample into responders and non-responders to exposure treatment[1]. Box plots revealed that EPT components did not vary across responder/non-responder groups. Logistic regression was used to test whether EPT predictors were associated with binary outcome (responder/non-responder) at each exposure time (for plots and analysis, see Appendix section B.5).

A finer-grained model of change was subsequently used to describe expected anxiety trajectories. Measures of treatment progress were assessed across the four sessions—Exposure 1, 2, 3, and 4[2]. Since the order of exposure sessions was fixed, session type (unaccompanied vs. accompanied) and session number (1-4) were conflated. This meant that it was not possible to disentangle the effect of accompanied and unaccompanied exposure from the effects of exposure repetition.

Because expected anxiety outcome variables were collected before exposure, and physiological and behavioural predictors were collected within exposure, we included the outcomes from Exposure 1 as a covariate and modelled change from Exposure 2 to 4 using lagged predictors (e.g., within-session habituation from Exposure 1 was used to predict treatment progress at Exposure 2). Including outcome at Exposure 1 as a covariate helped to more accurately account for baseline differences among patients and thus helped minimise variance in outcomes generally (Tabachnick & Fidell, 2007). In addition, because only t - 1 random effects can be estimated when t time points are present (McCoach & Kaniskan, 2010), only the random effects for intercept and linear trajectory were included when analysing fear of losing control.

Changes in outcome variables were analysed using multilevel growth modelling. This procedure is well-suited to dealing with missing data through use of maximum likelihood estimation, which unlike ordinary least squares procedures, avoids the usual case omission. Further, multilevel growth modelling accommodates repeated-measures data with varying number of measurements per person, and time-varying covariates, and are the recommended analytic approach for modelling change (Singer & Willett, 2003; McCoach & Kaniskan, 2010; Kristjansson et al., 2007). The Level 1 (individual growth model) component of the model described how each patient changed over time. The Level 2 model component

[1] A cut-off value of -1 represented a rather modest 1-point reduction across each session

[2] Because only six participants had completed three or more unaccompanied exposures, data were aggregated across repeated sessions and models were fit across four time points.

represented how these changes differed across patients. The effect of various predictors on individual growth profiles (Level 1) were assessed by inspecting fixed effects parameters of the Level 2 submodel (Singer & Willett, 2003).

Modelling was employed under the assumption that missing data in our data set were missing at random or missing completely at random (Bell et al., 2013; West et al., 2007). Multilevel models were calculated using the statistics program, R (R Development Core Team, 2013), using the "nlme" package (Pinheiro et al., 2013). Predictors (excluding time) were centred on their sample mean to facilitate interpretation of intercepts. Welch's t-test (which assumes samples have unequal variances) was used to compare the means of two groups. The level of significance was set at $p < .05$ for all comparisons. Where appropriate, 95% confidence intervals are presented in figures to aid interpretation. When figures illustrate intra-individual change, error bars represent 95% confidence intervals (CIs) corrected for within-subjects design as per Morey's (2008) method.

Missing Data Analysis

Clinical scales scores from patients with data from all exposure sessions (1-4) were compared with those who had incomplete data. Omnibus tests failed to reveal any statistically significant group differences in both MI and ACQ scores between baseline and follow-up assessment, F(3,303) = .10, p = .96, and F(3,303) = .70, p = .56, respectively.

4.3 RESULTS

Overall Treatment Efficacy

Treatment efficacy was determined by fitting separate unconditional growth models for each measure of treatment progress and outcome. Fixed effects parameters for average slopes of change across the four assessment periods indicated that that patients generally avoided fewer situations (MI) and experienced less severe maladaptive thoughts when they became anxious (ACQ), b = -0.42, $t(225)$ = -15.99, $p < .001$ and b = -0.24, $t(225)$ = -12.75, $p < .001$, respectively.

The influence of several demographic variables on treatment outcome was assessed. The magnitude of changes in MI scores between intermediate and follow-up assessments was not associated with patient age, $r(76) = .11$, p = .33, CI = [-.11 – .33], or sex (M_{Male}= 0.78, M_{Female}= 0.98), $t(50.51) = 1.35$, p = .18. Patient's self-reported avoidance of buses at baseline, did not systematically vary with age, $r(116)$ = -.01, p = .88, CI = [-.19 – .17], or sex (M_{Male}= 2.99, M_{Female}= 3.26), $t(79.05) = 0.97$, p = .33.

Adherence to End-of-exposure Instructions

Criterion 1: Decrease in exposure duration across sessions

We predicted that across sessions, participants would spend less time undertaking bus exposure (see Figure 8). A repeated-measures analysis with forward difference coding was used to examine duration changes across adjacent sessions. An overall test confirmed that exposure duration changed across sessions, F (3,138) = 42.93, $p < .001$. Follow-up tests revealed a significant decrease from session Exposure 1 to 2, b = 0.61, t(118) = 8.71, $p < .001$) an increase from Exposure 2 to 3, b = -0.26, t(118) = -3.06, p = .003, and a decrease from Exposure 3 to 4 (b = 0.32, t(118) = 3.36, p = .001).

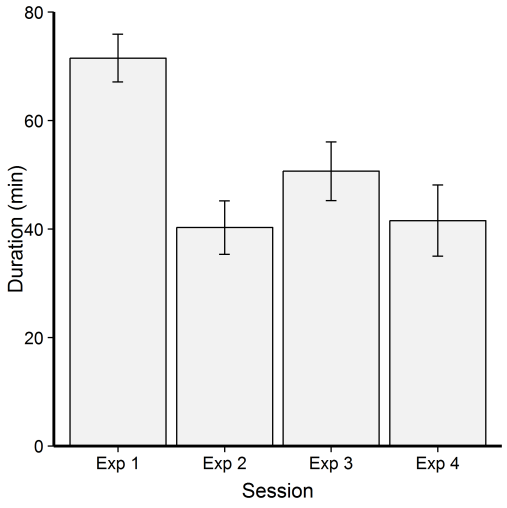

Figure 8. Mean exposure duration (in min) across each exposure session (Exp. 1-4) aggregated across all patients.

Note. Error bars represent 95% within-subjects confidence intervals

Criterion 2: Effect of average exposure duration on treatment progress and outcome

The effect of a patient's average exposure duration on treatment progress and treatment outcome was examined. Expected anxiety was generally higher among participants with longer exposure durations, b = 3.02, t(83) = 3.36, p = .001. Analyses did not reveal, however, that exposure durations had an influence on the rate of change of expected anxiety across treatment, b = 0.35, t(126) = 1.00, p = .32. In contrast, there was a marginal effect of average duration on fear of losing control trajectories, b = -0.56, t(126) = -1.91, p = .06. This indicated that by the end of exposure, patients with longer average

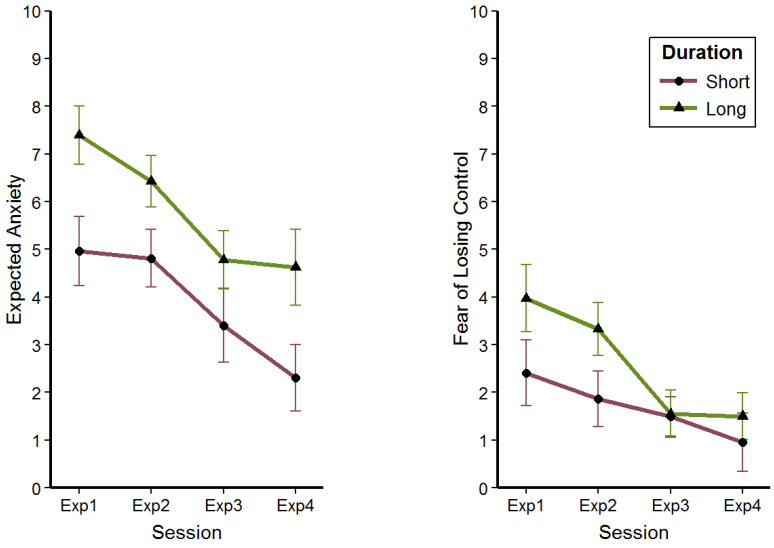

Figure 9. Changes in expected anxiety and fear of losing control as a function of mean exposure duration (based on a median split).

Note. Error bars represent CIs corrected for within-subjects design.

exposure durations experienced reductions that brought their fear of losing control down to the level of patients with shorter exposure durations. Floor effects are likely to have influenced these results however. (see Figure 9).

We reasoned that this general reduction in exposure duration constituted a necessary, but not sufficient condition of adherence to instructions. Other factors, such as the intra-individual stability of adherence across repeated exposure and baseline symptom severity, were reasoned to account for this pattern. In addition, there was no evidence that a patient's average exposure duration was associated with reductions in agoraphobic avoidance or the severity of negative anxious cognitions between intermediate and post-assessment periods, Mobility Inventory: b = -0.05, $t(75)$ = -0.30, p = .77, Agoraphobic Cognitions Questionnaire: b = -0.04, $t(75)$ = -0.36, p = .72.

Descriptive: Patterns of adherence to instructions across sessions

Patterns of adherence to end-of-exposure instructions were first assessed by calculating the number of exposure sessions in which self-reported anxiety and physiological arousal decreased (see Table 6). Separate χ^2 goodness-of-fit tests were performed to assess whether classifications of adherence were equally distributed among the two categories (decreasing and increasing) for each exposure session (Exp 1-4). Results revealed that self-reported anxiety was not equally

Table 6. Adherence to end-of-exposure instructions throughout exposure: Counts of exposure sessions in which self-reported and HR anxiety reduced.

	Exp 1	Exp 2	Exp 3	Exp 4
Decrease in self-reported anxiety				
No	18	20	26	29
Yes	38	34	32	16
Decrease in heart rate				
No	18	14	22	15
Yes	40	44	40	34

Note. N = 213 exposure trials.

Changes in self-reported anxiety were calculated between boarding and on exiting bus. Here, decreases referred to reductions of at least 1 point on the 11-point Likert-type scale between the start and end of exposure; Changes in heart rate were derived from robust linear regression coefficients for the average rate of change in HR across an entire bus exposure. Decreasing arousal referred to coefficients < 0; increasing arousal referred to coefficients greater than or equal to 0; N = 227 exposure trials.

distributed in Exposure 1, 2 and 4, $\chi^2(1) = 7.14$, $p = .007$, $\chi^2(1) = 3.63$, $p = .007$, and $\chi^2(1) = 3.75$, $p = .05$, respectively. Interestingly, compared to other exposure sessions, there was a greater proportion of patients whose self-reported anxiety increased in Exposure 4.

Concerning within-session reductions in heart rate, the null hypothesis of equal distribution between decreasing and increasing HR was rejected—all p-values were $\leq .02$—the majority of patients experienced decreasing physiological arousal during exposure at each measurement occasion. These findings revealed a more unstable pattern of assignment for self-reported anxiety compared to physiological arousal.

Criterion 3: Relationship between consistency of adherence to instructions and exposure duration

Patients who consistently adhered to end-of-exposure instructions were expected to remain in exposure longer. There was some support for this notion. Patients who never remained in buses until their self-reported anxiety reduced had significantly shorter exposure durations compared to those who always waited for self-reported anxiety to drop before exiting, $b = -0.55$, $t(78) = -3.67$, $p < .001$. A marginal effect of change in duration across sessions emerged for patients who only sometimes waited for their self-reported anxiety to drop before exiting buses, $b = -0.12$, $t(135) = -1.88$, $p = .06$. This trend suggested that compared to the other two groups, patients who

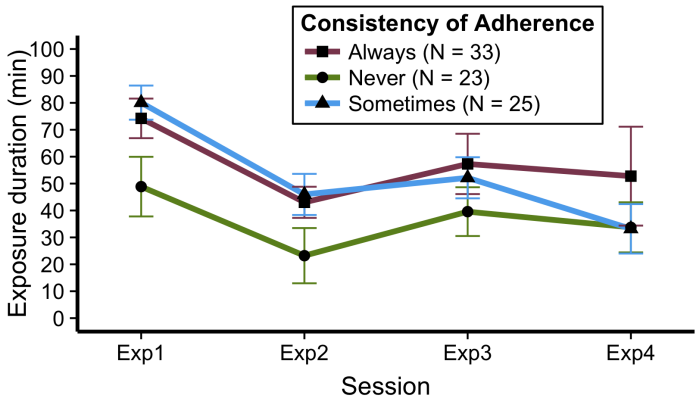

Figure 10. The relationship between exposure duration and consistency of adherence to end-of-exposure instructions (i.e., patients whose self-reported anxiety always, never or sometimes decreased during exposure).

Note. Error bars represent CIs corrected for within-subjects design. Four patients with missing self-reported anxiety ratings could be not included in this analysis.

only sometimes adhered instructions appeared to have the greatest reductions in exposure duration across sessions.

In contrast, no differences in exposure duration were detected for patients who always remained in buses until their HR dropped, compared to those who were never or only sometimes adherent, b = -0.08, $t(82)$ = -0.30, p = .77 and b = -0.03, $t(82)$ = -0.28, p = .78, respectively. This suggested that a patient's self-reported anxiety appeared to guide their decision to remain on buses (see Figure 10).

Descriptive: Relationship between initial symptom severity and consistency of adherence to instructions

We examined whether patients who always adhered to instructions (i.e., whose self-reported anxiety always decreased during exposure) were those with greater initial symptom severity measured prior to the start of Exposure 1. Results confirmed this was the case for agoraphobic avoidance (MI: $M_{Always\ adhered}$= 2.75, $M_{Never\ adhered}$= 1.83, $M_{Sometimes\ adhered}$= 2.44), and severity of negative anxious cognitions (ACQ: $M_{Always\ responsive}$= 2.15, $M_{Never\ adhered}$= 1.80, $M_{Sometimes\ adhered}$= 2.17), $F(2, 77)$ = 11.81, $p < .001$, and $F(2,77)$ = 4.04, p = .02, respectively. Initial symptoms were therefore significantly higher for patients who always adhered to instructions, compared to those who never remained on buses until their self-reported anxiety decreased. In line with the law of initial values (Lacey, 1956; Lacey & Lacey, 1962), it appeared likely that initial differences in baseline symptom severity moderated

the subsequent relationship between expected anxiety and adherence to instructions.

Criterion 4: Relationship between consistency of adherence to end-of-exposure instructions and expected anxiety

Controlling for initial agoraphobic avoidance (assessed with the MI at intermediate assessment), we examined the relationship between consistency of adherence to end-of-exposure instructions and expected anxiety across repeated exposure. Inspection of the overall effect of these factors on the growth model supported the inclusion of both additive and interactive effects of adherence and session repetition (time), although the interactive term appeared less strongly supported, F (2,130) = 3.06, p = .05.

Inspection of parameter estimates revealed that patients who always remained on buses until their self-reported anxiety reduced, had, on average, higher expected anxiety across exposure compared to patients who never remained on buses until their self-reported anxiety reduced, b = -4.76, t (76) = -5.59, p < .001 (see Figure 11). Greater reductions in expected anxiety across sessions appeared to be experienced by patients who sometimes remained in exposure until their self-reported anxiety reduced, relative to those who never stayed until self-reported anxiety dropped, b = -0.89, t (130) = -2.43, p < .02. Expected anxiety trajectories were not found to differ among patients who occasionally and consistently adhered, b = -0.51, t (130) = -1.43, p = .15. There were, however, very few patients who always adhered and who completed Exposure 4, which may help explain this null result and the non-overlapping average trajectories seen in Figure 11 (Always adhered, Exp. 1-4 n: 23, 21, 21, 8 vs. sometimes adhered, Exp. 1-4 n: 21, 25, 23, 23). It appeared that patients who always adhered were less likely to complete a fourth unaccompanied exposure.

In summary, this suggested that waiting for self-reported anxiety to reduce before leaving buses was not associated with reductions in expected anxiety. Further, there was a trend that occasionally waiting for self-reported anxiety to reduce was associated with greater reductions in expected anxiety.

Evaluation of the Emotional Processing Theory

Treatment progress

Overall, measures of treatment progress decreased across exposure sessions—as treatment advanced, expected maximum anxiety, and the belief that a loss of control would be experienced, declined significantly between Exposure 1 to 4 (see Table 7).

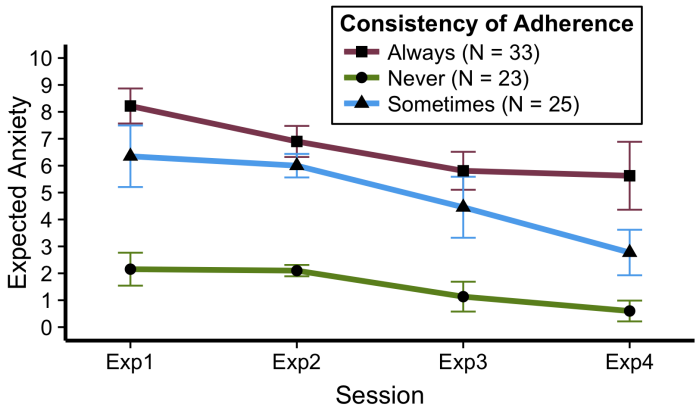

Figure 11. The relationship between consistency of adherence to end-of-exposure instructions (indexed by reductions in self-reported anxiety during exposure) and expected anxiety.

Note. Each line represents a subgroup is comprised *N* individuals, however, not all individuals completed exposure in each session (Exp 1-4: Always adhered, *n* = 23, 21, 21, 8; Never adhered, *n* = 13, 11, 15, 15; Sometimes adhered, *n* = 21, 25, 23, 23). Error bars represent CIs corrected for within-subjects design.

Unconditional means and growth models of expected anxiety

For expected anxiety trajectories, the unconditional means model (UMM, across people disregarding time) and the unconditional growth model (UGM, across people and time) were examined across all four time points (Exposure 1-4). Fitting the UMM allowed us to identify sources of response variability. The estimated within-person variance at Level 1 was 4.52, and the estimated between-person variance was 7.51, indicating that expected anxiety varied considerably between-individuals. This was confirmed by the intraclass correlation coefficient (ICC) of .38 which indicated that 38% of the total outcome variation resided between patients[3].

For fear of losing control, the within- and between-person variance was 3.53 and 3.78, respectively, suggesting that fear of control trajectories were generally less dominated by between-person variance relative to expected anxiety. The ICC was .48. Together, these relatively high ICC values helped justify a two-level analysis. These diagnostic models also showed that systematic variation existed in the outcome and that the predominant source of that variation existed within individuals.

Unconditional growth model (UGM) results indicated that the average change trajectory for expected anxiety had an intercept of

[3] 4.52/(4.52 + 7.51). This statistic also helps justify the use of a clustering procedure in study 2 to sort this variance, albeit on the basis of heart rate.

Table 7. Average (observed) treatment progress measures and estimated linear rates of change across exposure session (Exposure 1-4) derived from Level 2 fixed effects parameters.

Measure	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Rate of change (slope)
Expected Anxiety	6.25 (3.31)	5.56 (3.03)	4.08 (3.13)	3.26 (3.23)	-.94***
Loss of control	3.23 (3.13)	2.52 (2.42)	1.52 (2.21)	1.17 (1.96)	-0.76***

Note. N = 85; df = 127 for expected maximum anxiety and loss of control; df = 225 for MI-Alone and ACQ

$^{\sim}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

6.10, a linear slope of -0.94, both of which significantly differed from zero. The Level 1 residual variance was 2.27—a 50% decline from the UMM. The UGM for fear of losing control that included both linear and quadratic trends revealed that within-person variance at Level 1 was 1.46, which was a 59% decline from the Level 1 residual variance in the UMM. Together, these large reductions in Level 1 residual variance suggested that at least half of the within-person variance in treatment progress was associated with time. As within-person variance was still quite high, this justified the inclusion of some Level 1 time-varying predictors.

EPT components

Individual (lagged) EPT components were included in separate models of expected anxiety trajectories between Exposure 2 and 4, and that included expected anxiety ratings at Exposure 1 as a covariate (for complete model summaries, see Appendix section B.6). Results suggested that initial fear activation (IFA) was not strongly associated with the initial status of, or rate of change in, expected anxiety, b_{IFA}= -0.01, $t(40)$ = -0.55, p = .59, and $b_{\text{IFA x Session}}$= -0.003, $t(40)$ = -0.17, p = .87, respectively. This suggested that expected anxiety trajectories were not systematically associated with variations in initial fear activation. More specifically, initial fear activation did not appear to exert a lagged effect on subsequent expected anxiety ratings.

The effect of within-session habituation (WSH) did not appear to influence the initial status of, or the rate at which expected anxiety changed across time, b_{WSH}= 0.44, $t(46)$ = 0.65, p = .52, and $b_{\text{WSH x Session}}$= -0.42, $t(46)$ = -0.70, p = .49, respectively.

Finally, the effect of between-session habituation on expected anxiety was examined. We did not detect a statistically significant effect on the slope of change across sessions, $b_{\text{BSH x Session}}$= -0.02, $t(52)$

= -0.35, p = .72. However, it appeared that BSH exerted a marginal effect on the initial status of expected anxiety, b_{BSH}= 0.10, $t(47)$ = 1.79, p = .08, suggesting that less BSH was somewhat associated with higher overall expected anxiety.

A similar pattern of results was found for separate models of fear of losing control that included EPT components (see Table 10). No individual component sufficiently accounted for either initial status or slope of change in fear of losing control trajectories over Exposures 1 to 4.

4.4 DISCUSSION

We examined whether initial fear activation and habituation of fear during exposure are required to promote progressive reductions in expected anxiety. Overall, individual emotional processing theory (EPT) components did not predict reductions in expected anxiety. Using adherence to end-of-exposure instructions to explicitly target habituation, however, we found some evidence that patients who always remained on buses until their self-reported anxiety reduced had persistently high levels of expected anxiety across exposure therapy. In contrast, patients who occasionally waited until their anxiety reduced tended to experience greater reductions in expected anxiety. These results suggest that inconsistent adherence may act as occasional reinforcement across sessions, which has been shown to promote more robust extinction learning in laboratory studies (Bouton, 2004).

The exposure instructions provided by therapists were expected to influence how long patients would stay on buses—patients were instructed to remain on buses until their fear reduced. This prediction was based on several observations. First, patients had, on average, successively lower levels of expected anxiety across sessions, and experienced positive overall treatment outcomes. Second, good compliance with therapist instructions was generally found in an earlier phase of the same project (Cammin-Nowak et al., 2013) and in study with a similar procedure (White et al., 2014). We further believed that this provided a window into habituation; some patients would have allowed enough time for fear to habituate, and others would have exited the bus before this occurred. Several criteria were addressed to examine the idea that exposure duration and adherence to end-of-exposure instructions indexed habituation.

Exposure duration effects

Exposure duration generally reduced across exposure sessions; patients remained in the exposure context for increasingly shorter periods. A noteworthy exception was the increase in average duration from Exposure 2 to 3. This increase may have resulted from the differ-

ent conditions under which exposure was undertaken—Exposure 2 was unaccompanied, but Exposure 3 was accompanied. It is likely that the presence of therapists helped encourage longer exposure. This may relate to previous findings that therapist-accompanied exposure is more effective than unaccompanied exposure (Gloster et al., 2011). However, this idea is predicated on the notion that longer exposure promotes greater symptom reduction.

It was also found that patients who remained on buses longer had, on average, higher expected anxiety and fear of losing control. In Exposure 3, fear of losing control decreased sharply among those with longer exposure duration. However, floor effects may have prevented further reduction in negative anxiety cognitions among patients with shorter exposure durations. These findings should therefore be cautiously interpreted. Studies in which exposure duration has been actively manipulated have yielded inconsistent results (Marshall, 1985; Trudel, 1979; Rabavilas et al., 1976; Chaplin & Levine, 1981). Unlike these studies, where exposure duration was a manipulated variable, exposure duration in the present study was largely under patient control. For this reason, and as we used a clinical sample, it was likely that exposure duration effects were confounded with symptom severity. The present findings supported this notion; patients with more severe symptoms were more likely to remain longer on buses. Closer examination of adherence to instructions was therefore warranted.

Adherence to end-of-exposure instructions and treatment progress

Patient's apparent adherence to end-of-exposure instructions was also examined. This was indicated by changes in self-reported anxiety and physiological arousal within individual exposures sessions. Substantial intra- and inter-individual variation in adherence was expected, in line with the idea that escape from fear-provoking situations is quite common among panic disorder patients (Richter et al., 2012). In addition, it was posited that the discomfort of remaining in buses despite experimenter instructions would, in some cases, be very challenging and thus result in variable degrees of constraint satisfaction across patients (Tryon, 2005; Shultz & Lepper, 1996). We initially expected consistent adherence, and thus greater constraint satisfaction, to be central to therapeutic change.

Data indicated that the majority of patients remained on buses until either their HR or self-reported anxiety decreased. HR and self-reported anxiety appeared concordant in the first three exposure sessions, however notable discordance between self-reported and physiological indices was apparent in Exposure 4; the majority of patients did not experience decreases in self-reported anxiety, although most experienced HR reductions. A lack of synchrony between different measures of fear has been noted by researchers

(Lang, 1979, 1993; Rachman & Hodgson, 1974), with some positing that concordance between response systems is greatest during strong emotional arousal (Hodgson & Rachman, 1974). By the final exposure, buses may have only evoked mild fear causing response systems to be weakly aligned, possibly due to unspecific activity in the autonomic domain (Lang, 1971). However, as these results did not account for within-person patterns of adherence to instructions, an examination of consistently adherent patients was required.

Approximately 40% of patients consistently adhered to end-of-exposure instructions in all sessions. Patients who always adhered tended to remain in exposure longer than those who never adhered. Patients who sometimes adhered, however, tended to have long exposures to begin with, but by the final exposure session, remained in buses only as long as those who only sometimes adhered. Since exposure duration and consistency of adherence were found to depend on initial symptom severity (in line with the law of initial values, Lacey, 1956; Lacey & Lacey, 1962), we controlled for baseline symptom severity when analysing the relationship between consistency of adherence and expected anxiety.

Compared to those who only sometimes or never adhered with end-of-exposure instructions, patients who always adhered stayed on buses longer, and had expected anxiety that was high at the beginning of exposure and that was resistant to reduction. There was also some evidence that patients who only sometimes adhered with end-of-exposure instructions experienced greatest reductions in expected anxiety across sessions. This indicates that patients with more severe initial symptoms doggedly remained in exposure until fear reduced (in accordance with their therapist's instructions), even when this proved to be ineffective at reducing their expected anxiety. Assuming that decisions to remain on buses were always made consciously, the assertiveness of patients might be involved in consistent adherence, with some research demonstrating an inverse relationship between agoraphobia and assertion (Chambless, 1985; Rapee & Murrell, 1988). Alternatively, expectancies may have driven decisions to remain in exposure. Specifically, only occasionally waiting for anxiety to reduce before leaving the bus may have resembled occasional reinforced extinction (Bouton et al., 2004; Woods & Bouton, 2007). This makes sense when one considers that reductions in anxiety during exposure, were likely experienced as a form of reinforcement by patients, and possibly as a form of constraint satisfaction (Tryon, 2005). Patients who were only sometimes adhered to instructions would therefore have experienced variable reinforcement across sessions—experiencing a mixture of reinforced (bus ride + anxiety reduced) and non-reinforced (bus ride + anxiety unaltered/increased). This may have served to highlight a mismatch between their expectancies and outcome and thus promoted new learning (Rescorla & Wagner, 1972;

Craske et al., 2014). In summary, these findings support the the idea that inconsistent habituation may promote reductions in expected anxiety.

Emotional processing theory

Another central aim of this study was to examine the predictions of the emotional processing theory. Given that several studies have found that EPT predictors do not predict overall treatment outcome (e.g., Meuret et al., 2012; Baker et al., 2010), we examined changes in expected anxiety across repeated exposure (our measure of treatment progress). However, EPT components did not account for the sizeable within- and between-person variation in treatment progress. The strongest EPT predictor of expected anxiety was between-session habituation, which was inversely related with higher overall expected anxiety. This suggested that high expected anxiety was related to persistently high maximum HR levels across sessions. This effect was found for ratings of expected maximum anxiety but not fear of losing control, suggesting that patient predictions about general anxiety are more tightly coupled with HR.

Our operationalisation of components may have contributed to these results. We chose to index initial fear activation with mean peri-boarding HR. Given that bus boarding was regarded as a salient cue for patients and that bus routes were reasoned to differ in terms of the number of phobic environmental cues (e.g., bridges) that were passed, we opted to calculate IFA from this restricted segment. However, the disadvantage of doing so was that fear may have been activated after 2.5 minutes had elapsed. Our operationalisations of within- and between-session habituation were similar to those of other researchers (Meuret et al., 2012; Baker et al., 2010). But unlike Meuret and colleagues (2012), who encouraged patients to choose the most fear-provoking situation, the exposure situation was fixed in the current study—buses would not have been at the top of all patient's fear hierarchies. We expected that this variability would provide useful variance in both criterion and predictor variables. A variety of exposure tasks, however, ranked as the most fear-provoking situation for each participant, may actually have helped to partition some of the variance. Finally, within the current sample, it may be that the act of walking to the bus station may have interfered with the clarity of the initial fear responses, which could have affected how much habituation was possible. This may explain why support for EPT components has predominantly been found in samples undertaking imaginal exposure, or confrontation of phobic stimuli such confined spaces (e.g., Kozak et al., 1988; Alpers & Sell, 2008).

Limitations

A consequence of our operationalisation of EPT predictors was that it resulted in several, non-identical datasets for each separate test. Formal model comparisons using log likelihood, AIC or BIC statistics were therefore not possible (Singer & Willett, 2003). An even larger sample or an alternate way of parametrising initial fear activation would be required to overcome this issue.

A practical constraint associated with the exposure instructions should also be noted. Asking patients to remain in buses until fear reduced might have conflicted with the desire of some patients not to exit the bus at undesirable locations. For example, a patient who experienced anxiety reduction and felt that enough time had elapsed on the bus, might have been ready to disembark, but might not have wished to do so in their current location. This may have been more likely in remote rural settings, where buses were irregularly scheduled. This would primarily have affected interpretation of exposure duration, but may also have affected the extent to which HR habituated within a single exposure. Had a patient remained on a bus longer than desired, this may have served to flatten the robust linear regression slope on which within-session habituation was based.

Clinical Implications and Future Directions

Several implications can be drawn from these results. Our findings cast doubt on effectiveness of remaining in fear-provoking situations until fear subsides in order to promote symptom reductions. Rather, it appears preferable to encourage patients to occasionally leave exposure contexts even when anxiety levels have not attenuated. In making this decision, therapists should consider the number of previous attempts to adhere to end-of-exposure instructions and the average duration of exposure. By individually tailoring exposure to individuals, therapists could encourage patients who consistently undertake longer exposures without experiencing reductions in expected anxiety to attempt shorter exposures. This approach appears congruent with the idea that the homework quality, rather than quantity, is a better predictor of treatment outcome (Cammin-Nowak et al., 2013).

Our findings also implicate adherence to end-of-exposure instructions as a useful target for examining exposure processes. Doing so helps focus attention on the capacity of patients to overcome their psychological discomfort and satisfy a particular constraint (Tryon, 2005). Homing in on the process of overcoming this dissonance is important, since patients with phobic avoidance often recognise the threats and rewards associated with avoidance (Pittig et al., 2014; Kashdan et al., 2006, 2013). To this end, we argue that assessing patients adherence to

specific exposure instructions using objective ambulatory assessment devices can provide useful insights into how patients learn during treatment.

5

MIX IT UP: THE ADVANTAGES OF VARYING SITUATIONAL EXPOSURE CONTEXTS

Although generally an effective invervention, it is quite common for people to experience relapse following situational exposure therapy. Results from fear extinction studies have identified several factors that can strengthen extinction learning and prevent return of fear, however whether these findings transfer to in vivo exposure remains unclear. The current study examines the effect of multiple exposure contexts and compound stimuli on relapse prevention in a group of panic disorder with agoraphobia (PD/A) patients undertaking exposure. The current sample consisted of 85 patients with PD/A who undertook repeated bus exposure as part of a 12-session treatment program. Multiple contexts (number of unique unaccompanied exposure paths; rural/urban settings) and compound stimuli (standard situational exposure vs. fear augmented situational exposure) were assessed. Results revealed that variable exposure contexts, as measured by the proportion of unique unaccompanied exposure paths, was predictive of greater maintenance of gains six months after the end of treatment. These findings suggest that superior long-term outcomes can be achieved by encouraging patients to vary exposure contexts.

5.1 INTRODUCTION

In vivo exposure has a central role in the treatment of panic disorder with agoraphobia (PD/A) (Chambless et al., 1998; Norton & Price, 2007; Deacon & Abramowitz, 2004). Despite being generally effective, exposure is not equally effective for all patients, and it is quite common for fear to return following treatment (Peter et al., 2008). Three to nine years following exposure therapy, 36% of treated patients still had mild-to-moderate agoraphobic symptoms, and 14% of patients experienced a worsening of symptoms (Peter et al., 2008). Identifying factors which promote treatment gains and help reduce relapse is therefore of great importance.

Theories of exposure have been heavily influenced by laboratory-based fear conditioning. Fear conditioning and extinction serves as the laboratory analogue for exposure therapy, and Pavlovian conditioning procedures are particularly well suited for examining how fear is acquired and extinguished. Fear acquisition involves the repeated pairing of a neutral stimulus with an unpleasant unconditioned stimulus (US, e.g., electric shock) until the US becomes

"conditioned" as a predictor of the aversive stimulus. Thereafter, the newly conditioned stimulus (CS) is capable of eliciting cognitions, physiological and behavioural responses in anticipation of the aversive stimulus (Vervliet & Raes, 2013). Fear extinction is seen as the gradual reduction of anticipatory fear reactions when CSs are repeatedly presented in the absence of an aversive stimulus (US). But return of fear (renewal, reinstatement, and spontaneous recovery) following fear extinction (Bouton et al., 2001) suggests that the original CS-US association is not erased. For example, fear conditioned (through CS-US pairings) in one context, followed by extinction (CS presented alone) in a separate context, has been shown to return if the original CS is presented in the original context (termed "ABA renewal", Neumann, 2006; Bouton, 2004). This and other studies have provided evidence that the CS retains its original excitatory meaning (CS-US) and also acquires a new inhibitory meaning (CS-noUS). A focus of research has therefore been to identify factors that strengthen the inhibition of the original CS-US association (Bouton, 2002, 2004). Transferring these findings to exposure therapy holds the promise of reducing the risk of relapse following exposure.

Inhibitory Learning Variables: Multiple contexts and Compound Stimuli

Fear can be renewed by presenting the CS in a context that is different to that in which extinction was conducted (Bouton, 2004). To account for this phenomena, Bouton's theory of extinction (Bouton, 1994) states that context modulates the inhibitory CS-US association. Specifically, the inhibitory CS-US association is modulated by both the CS and the context in which extinction occurs (for a review, see Vervliet et al., 2012). Building on this finding, several fear extinction studies have shown that conducting extinction in multiple contexts can help to reduce relapse (Balooch et al., 2012; Gunther et al., 1998; Neumann, 2006), although some exceptions have also been found (Bouton et al., 2006; Neumann et al., 2007).

In applied research settings, the effect of multiple contexts has been examined in a few studies (for a review, see Craske et al., 2008). Participants were better able to retrieve newly learned motor skills and verbal tasks following a hiatus when they engaged in random and variable practice compared to blocked schedules of practice (Schmidt & Bjork, 1992). Extinction conducted in multiple virtual reality contexts exposure with spider-phobic (Shiban et al., 2013) and healthy participants (Dunsmoor et al., 2014) appeared to reduce return of fear (renewal) relative to extinction in a single context. Further, Lang and Craske (2000) assessed the effect of multiple exposure contexts by exposing participants to different sequences of exposure contexts that varied in difficulty (balconies in a 10-story building; with the top balcony being the most challenging).

Participants who were exposured to heights in a random order of difficulty had less anxiety following treatment, relative to a control group who undertook exposure in a fixed, increasing order of difficulty. A limiting factor in this study was the use of a non-phobic sample and the constant accompaniment of researchers during tasks. Experimenter demand characteristics and mode of instruction (personal vs. impersonal) have been found to substantially influence the extent to which participants confront feared stimuli (Bernstein & Nietzel, 1974; Speltz & Bernstein, 1976; Trudel, 1979). Further research is therefore required to determine whether these finding generalise to the exposure contexts typically avoided by agoraphobics.

Benefits associated with the use of multiple excitatory conditioned stimuli (compound stimuli) during extinction suggests that this may be another method of strengthening inhibitory learning. In a series of animal experiments, Rescorla (2006) found that in comparison to extinction trials in which separate stimuli were presented, presenting two extinguished excitatory stimuli in compound served to attenuate subsequent return of fear. Results from another animal study also supported the benefits of presenting compound stimuli to strengthen extinction learning (Janak et al., 2012). Incorporating this into exposure by combining interoceptive (i.e. systematic confrontation of feared bodily cues) with situational exposure may be promising approach (Craske et al., 2008). The generalisability of these initial findings requires research attention since they have the potential to further enhance treatment outcomes for those undertaking exposure therapy. However, a limitation of previous studies is that these effects have not always been investigated in clinical samples (e.g., Lang & Craske, 2000), and require further validation in naturalistic settings.

These inhibitory learning processes—multiple exposure contexts and compound stimuli—were examined in the present study using a sample of PD/A patients who undertook repeated bus exposure. As patients were free to choose where they undertook unaccompanied bus exposure, this allowed us to examine the effect of variable exposure contexts. Patients who completed exposure in a greater variety of contexts were expected to show maintained, if not improved, outcomes following the last exposure. Here, we were specifically concerned with the number of unique paths and whether the bus exposure was conducted in urban or a mixture of rural and urban settings. In the current study, Global Positioning System (GPS) technology was used to record patients' position during exposure and was an objective measure of where exposure was undertaken. The effects of compound stimuli were also assessed—patients were randomised to one of two active treatment conditions. One group undertook standard situational exposure, and those in an augmented exposure condition undertook exercises throughout in vivo exposure to provoke greater fear.

An additional aim of the current study was to determine whether multiple exposure contexts could be effectively operationalised using plots of GPS-trajectories. This has some precedent—documenting movement trajectories during unaccompanied driving exposure was effective in documenting where the patient travelled during unaccompanied exposure (White et al., 2014). Determining whether meaningful movement parameters provide insights into therapy progress when larger samples was therefore of considerable interest.

5.2 METHOD

Participants

Data were collected from 93 patients with PD/A as part of a clinical study conducted across five treatment centres in Germany (PanikNetz). Participants were randomised to one of two treatment conditions after they had been recruited through physician referral and via additional advertisements in various media outlets. Inclusion criteria consisted of: (a) age 18-65 years (b) a current primary diagnosis of panic disorder with agoraphobia according to DSM-IV-TR criteria; (c) clinical global impressions scale (CGI) score ≥ 4; (d) ability and availability to regularly attend therapy sessions. Exclusion criteria were: (a) current suicide intent; (b) comorbid psychotic or bipolar I disorder; (c) current dependence on alcohol, benzodiazepine or other psychoactive substance; (d) current psychotherapeutic or psychopharmacological treatment for another Axis I disorder; (e) serious medical illness that excluded exposure-based CBT (e.g., renal, cardiovascular or neurological disease).

Psychotherapists at each of the cooperating treatment centres coordinated recruitment, delivered treatment, and collected data. After removal of cases with missing or poor quality heart rate (HR) data and which was not assigned to one of six previously-identified HR clusters, the final sample consisted of 85 patients with PD/A (age: $M = 34.21$; $SD = 10.47$; 50 females) who had completed at least one bus exposure. This sample completed a total of 227 (120 therapist-accompanied, 107 unaccompanied sessions) bus exposures. The local ethics committees approved all data assessment procedures.

Materials

GPS-derived position was collected during exposure using a commercial sports monitor (Garmin Forerunner 310XT). The device relied on a data compression algorithm ("smart recording"), where data points were recorded only when parameters (speed, direction or HR) changed. This algorithm produced a compression ratio of about 3:1 [uncompressed : compressed, 300:300 - (.69 × 300)].

During the study, a software update released by Garmin enabled equidistant sampling of parameters (1 Hz) in 48 bus exposure trials. The remaining 179 trials were decompressed to a common 1 Hz grid.

Measures

Treatment outcome measures

Treatment outcome was assessed with two clinical scales administered as part of the PanikNetz study protocol at intake, baseline, mid-way through treatment (intermediate; after session 6), following the final (12th) session (post-assessment) and at follow-up (6 months following the final session). The Mobility Inventory (alone subscale) (MI, Chambless et al., 1985), was used to assess agoraphobic avoidance in various settings while alone. This subscale was chosen because unaccompanied confrontation of feared situations following treatment was of primary interest. The inventory is comprised of 27 items, which are rated on a 5 point Likert-type scale from 1 = *never avoid* to 5 = *always avoid*. The measure has demonstrated good internal consistency and discriminant validity (Chambless et al., 2011). In addition, the Agoraphobic Cognitions Questionnaire (ACQ, Chambless et al., 1984) was administered; it has 15 items, which are scored on a 5-point Likert-type scale ranging from 1 = *thought never occurs* to 5 = *thought always occurs*. The ACQ measures the severity of maladaptive thoughts when anxiety is experienced.

Multiple exposure contexts and compound stimuli

The extent to which patients undertook exposure in multiple contexts was operationalised with the aid of individual GPS trajectory plots (see Figure 14). We were predominantly interested in patient's choice about where they undertook unaccompanied exposure, as this involved them making a choice, independent of therapists. Two raters were used to inspect and classify exposure path plots. Raters counted the number of unique unaccompanied exposure paths per participant. Unaccompanied bus exposure paths were classified as unique when they were different from all other exposure paths for that individual—when they reflected different bus routes and when the majority of the exposure was in a distinct location. As such, it was possible that there was some minimal overlap between two paths that were both classified as unique, but raters were instructed that paths which overlapped by more than 50% were not to be considered unique.

The number of exposure paths varied for each patient and there was a high correlation between the total number of exposures and the number of unique unaccompanied exposure paths, $r(83) = .92$, $p < .001$, $CI = [.89\text{-}.95]$. We therefore calculated the proportion of unique unaccompanied exposure paths for each patient relative to

the total number of all exposure sessions per patient. When patients had a missing therapist-accompanied session (e.g., due to technical failure), adjustments were made to ensure that proportions accurately reflected the actual number of exposures undertaken[1]. This correction served to make this parametrisation more conservative. For path uniqueness classifications, Cohen's Kappa indicated very good agreement between raters, κ = .92, 95% *CI* = [.88, .95].

In addition, each plotted exposure path was inspected and classified as encompassing rural, urban or a mixture of rural and urban settings. Data were then aggregated and each patient's set of exposure paths were classified as either urban or a mixture of rural and urban[2]. Here, inter-rater agreement was also very high, κ = .96, 95% *CI* = [.93, 1].

Procedure

In-vivo exposure therapy

Psychological treatment comprised 12 sessions and two follow-up booster sessions (two and four months following the last session). Therapy was delivered by advanced-level clinical psychology graduates and post-doctoral students who had received extensive training in the treatment protocol and who were experienced in CBT. Following screening, informed consent, and initial assessment, patients were randomised to one of two treatment conditions: Standard in vivo exposure (bus exposure) or fear augmented exposure, which, in addition to standard in vivo exposure, involved focussing attention towards fear inducing aspects, (e.g., bodily symptoms) or specific situational fear cues, and sometimes performing interoceptive exposure exercises if fear did not occur spontaneously. Treatment condition effects of which were not examined in the current study. Therapy sessions were 100 minutes in duration and topics covered in the initial six sessions included psychoeducation; rationale for exposure therapy; behavioural analysis; role of avoidance behaviour; interoceptive exposure; and relapse prevention (for a full description, see Gloster et al., 2014).

After completing the initial six therapy sessions, patients undertook several bus exposure sessions. Four session types were possible: In Session 7, therapist-accompanied (Exposure 1), was followed by an unaccompanied session (Exposure 2). This was repeated again in

[1]For example, if Exposure 2, 3 and 4 were present, but Exposure 1 was missing due to technical failure, the proportion was calculated as the *number of unaccompanied paths / 4*. Since we were only interested in the number of unique accompanied paths and accompanied sessions were always a standardised route, this served to increase the accuracy of this statistic.

[2]It was not possible that all exposures were conducted in rural settings, since all therapist-accompanied exposures were undertaken in the vicinity of outpatient clinics, located in urban settings.

Session 11, a therapist-accompanied (Exposure 3) was followed by an unaccompanied session (Exposure 4). Repetitions of unaccompanied exposure were possible (maximum number of repetitions: Exposure 2 = 4; Exposure 3 = 3).

Patients were instructed to remain in the exposure situation until the fear they experienced in the situation reduced by itself. Exposure sessions therefore had variable duration (M = 51.91 min, SD = 35.15 min, range = 2.67 – 253.25 min).

Preprocessing

Data from the present study were based on the same patients as in studies two and three. As such, the preprocessing of positional data was identical. As we were interested in categorising bus exposure paths, GPS-position data between boarding and disembarking the bus were extracted. Separate GPS plots were created using the "RgoogleMaps" package (Loecher, 2014) within R to allow analysis of trajectories (see Appendix section A.1). When single exposures comprised multiple bus excursions (e.g., when a patient needed to change bus), paths from both bus rides were combined. These separate segments did not contribute towards counts of unique exposure contexts.

Analytic Strategy

Changes in outcome variables were analysed using linear mixed models. These models are well suited to deal with missing data, handle repeated-measures data with varying number of measurements per person, and are the recommended analytic approach for modelling change (Singer & Willett, 2003; McCoach & Kaniskan, 2010; Kristjansson et al., 2007). Models consisted of two levels: Level 1 (individual growth model) described how each patient changed over time; Level 2 model represented how these changes differed across patients. The effect of inhibitory learning predictors on individual growth profiles (Level 1) were assessed by inspecting fixed effects parameters of the Level 2 submodel (Singer & Willett, 2003). The variance components in multilevel models were estimated by maximum likelihood. When reporting on the change across more than two time points, we report the effect of predictors on the initial status (intercept) and the rate of change.

The effect of independent variables on treatment outcome was assessed during (between intermediate- and post-assessment periods) and following exposure (between post- and follow-up assessment periods) in separate tests. When multiple comparison were performed, the false discovery rate (the expected proportion of

false discoveries amongst the rejected hypotheses) was controlled as per Benjamini and Hochberg's (1995) recommendations.

Multilevel models were calculated using the statistics program, R (R Development Core Team, 2013), using the "nlme" package (Pinheiro et al., 2013). The level of significance was set at $p < .05$ for all comparisons. Where appropriate, 95% confidence intervals are presented in figures to aid interpretation. When figures illustrate intra-individual change, error bars represent 95% confidence intervals (CIs) corrected for within-subjects design as per Morey's (2008) method.

5.3 RESULTS

The Influence of Demographic Variables on Outcome Variables

Overall tests were used to examine the influence of age and sex on changes in Mobility Inventory (alone subscale) and Agoraphobic Cognitions Questionnaire (ACQ) scores between intermediate and follow-up assessments. Age and sex did not appear be associated with differential changes in ACQ scores, $b_{\text{Age x Assessment Time}}$= -0.001, $t(140)$ = -0.36, p = .71 and $b_{\text{Sex x Assessment Time}}$= 0.01, $t(140)$ = 0.19, p = .85, respectively. For MI, patient age did not appear related to scale scores , $b_{\text{Age x Assessment Time}}$= -0.001, $t(140)$ = -0.29, p = .76. However, patient sex did appear influence average level of MI across these assessments, b_{Sex}= -0.41, $t(82)$ = -1.96, p = .05, and there was a marginal effect of sex on the rate of change in MI scores, $b_{\text{Sex x Assessment Time}}$= 0.13, $t(140)$ = 1.72, p = .09. Follow-up tests revealed that males had lower MI scores at intermediate assessment, M_{Male}= 2.12, M_{Female}= 2.64, $t(64)$ = 2.86, p = .02, but at post-assessment, no differences were apparent, M_{Male}= 1.60, M_{Female}= 1.74, $t(62)$ = 0.91, p = .55. At follow-up, there were also no differences between males and females, $t(51)$ = 0.47, p = .64. Thus, although males were found to have lower agoraphobic avoidance at intermediate assessment, no difference was detected at later assessment points. As subsequent results predominantly focus on the period between post- and follow-up assessment, we could rule out the influence of sex and age on scale score changes during this period.

Treatment Efficacy

The efficacy of treatment was determined by fitting separate unconditional growth models for each measure of treatment outcome. Fixed effects parameters for average slopes of change across all four bus exposures were extracted (see Table 8). MI and ACQ results revealed that following exposure patients generally avoided fewer

Table 8. Average (observed) treatment outcome and estimated rates of change throughout treatment derived from fixed effects parameters.

	Baseline	Intermediate	Post	Follow-up	Rate of change (slope)
Measure	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*b*
Mobility Inventory (alone subscale)	2.66 (0.79)	2.43 (0.75)	1.69 (0.61)	1.53 (0.63)	-.42***
Agoraphobic Cognitions Quesionnaire	2.24 (0.52)	2.04 (0.50)	1.75 (0.47)	1.54 (0.43)	-0.25***

Note. $N = 85$; $df = 225$ for MI-Alone and ACQ; Slopes represent linear trends based on raw data.

***$p < .001$

situations and experienced less severe negative anxious cognitions. In sum, results supported the efficacy of treatment.

Effect of Inhibitory Learning Variables on Treatment Progress

A similar analytic strategy was employed to explore whether factors associated with inhibitory learning accounted for reductions in expected anxiety. Compound stimuli, exposure setting (urban/mixture of urban and rural), and the number of unique exposure paths were included as model parameters.

Compound stimuli

We assessed the effect of compound stimuli by examining differences in treatment outcome trajectories (between intermediate and follow-up) across the two treatment conditions—standard and augmented exposure. For MI, we detected no differences in initial Mobility Inventory (MI) scores, or rate of change across sessions varied between treatment conditions, $b_{\text{Condition}}$= -0.14, $t(82)$ = -0.65, p = .52 and $b_{\text{Condition x Time}}$= -0.02, $t(140)$ = 0.25, p = .80, respectively[3]. Similarly, condition did not influence Agoraphobic Cognitions Questionnaire (ACQ) trajectories, $b_{\text{Condition}}$ = 0.06, $t(82)$ = 0.42, p = .67 and $b_{\text{Condition x Time}}$= -0.02, $t(140)$ = -0.39, p = .70. Combining interoceptive exercises with situational exposure did not

[3]In all these models, standard exposure was set as the reference group

appear to foster reductions in agoraphobic avoidance or in the severity of negative thoughts when anxiety is experienced.

Variable exposure: Rural/urban exposure context

On a descriptive level, 51 patients completed exposure in a mixture of rural and urban settings, and 34 patients completed exposure exclusively in urban settings. Patients who undertook a mixture of rural and urban exposures were posited to have greater anxiety during exposure, but less anxiety following treatment, compared to those who only undertook urban exposure. However, exposure context (rural/mixture of rural and urban) was not found to be associated with different patterns of clinical scale scores between intermediate and follow-up assessments—MI: b_{Context}= 0.19, $t(64)$ = 0.76, p = .45, and $b_{\text{Context x Assessment Time}}$= -0.08, $t(115)$ = -0.86, p = .39, ACQ: b_{Context}= -0.04, $t(64)$ = -0.25, p = .81, and $b_{\text{Context x Assessment Time}}$= 0.05, $t(115)$ = 0.80, p = .43. In sum, variation in the geographic location of exposure did not result in greater reductions in agoraphobic avoidance or the severity of negative anxious cognitions, both during and following treatment.

Variable exposure: Proportion of distinct exposure paths

The influence of total number of exposures on symptom changes was first examined in overall tests (i.e., across all three assessment times). These suggested that a greater number of exposure sessions was not associated with larger reductions in MI or ACQ scores between intermediate and follow-up assessment, b= -0.04, $t(140)$ = -1.32, p = .18 and b= -0.03, $t(140)$ = -1.61, p = .11, respectively. However, more exposures sessions appeared weakly associated with higher average MI and ACQ scores across each assessment point, b= .13, $t(82)$ = 1.77, p = .08 and b= 0.11, $t(82)$ = 2.04, p = .04, respectively (see Figure 12).

Overall tests were also used to assess the relationships between variable exposure context and change in clinical scale scores. These revealed that across these three assessments, patients with larger proportions of unique unaccompanied exposure paths had higher overall MI and ACQ scores, b= 1.37, $t(82)$ = 2.91, p = .005 and b= 0.81, $t(82)$ = 0.02, p = .02, respectively. Of particular note, however, the reductions in these scores across time appeared greatest among patients with larger proportions of unique unaccompanied paths; MI: b= -0.40, $t(140)$ = -2.32, p = .02, and ACQ: b= -0.30, $t(140)$ = -2.30, p = .02.

Additional tests were conducted to establish whether this latter relationship manifested both during and after exposure therapy. In line with our predictions, the proportion of unique unaccompanied paths did not appear to exert an influence on changes in MI or ACQ scores between intermediate and post-assessment, b= -0.07, $t(75)$ =

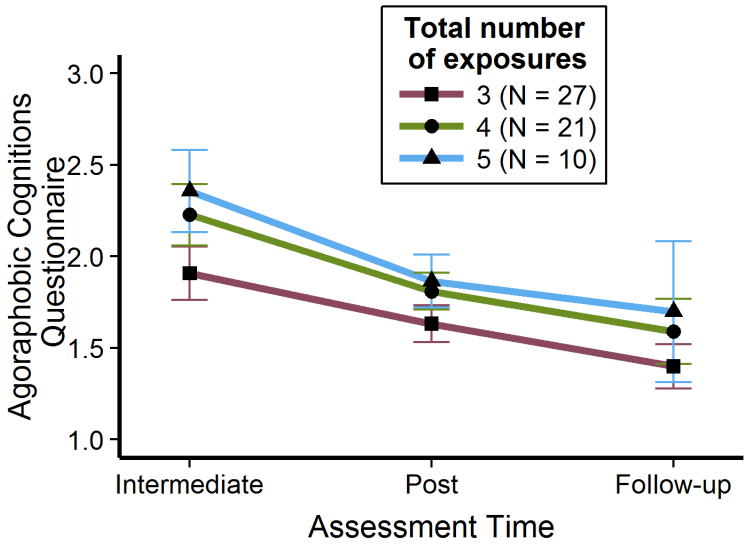

Figure 12. The relationship between total number of exposures undertaken and changes in agoraphobic cognitions questionnaire scores (for a subset of participants with 3, 4, and 5 exposures).

Note. Error bars represent 95% within-subjects confidence intervals

-0.22, p = .83 and b= -0.17, $t(75)$ = -0.78, p = .44, respectively. In contrast, a statistically significant negative association was detected between variable exposure context and reductions in agoraphobic avoidance (MI) between post- and follow-up assessment, b= -0.63, $t(63)$ = -2.56, p = .03. However, the relationship between variable exposure context and negative cognitions (ACQ) between post- and follow-up assessment was not statistically significant, b= -0.33, $t(63)$ = -1.75, p = .17.

Since it appeared that patients with a greater number of exposures had higher overall avoidance, we decided to subset the dataset to include only patients with between four and six exposures. This also served to remove patients who had completed no unaccompanied exposure. Findings for this subset again revealed a statistically significant negative relationship between proportion of unique unaccompanied exposures and change in MI scores; the greater the proportion of unique paths, the greater the reduction in agoraphobic avoidance between post- and follow-up assessment (see Figure 13).

In summary, these results revealed two interesting patterns of findings. First, it appeared that patients who undertook a greater number of exposures sessions tended to have higher average agoraphobic avoidance and more severe negative anxiety cognitions across sessions; this may point to patients with more

severe symptoms being motivated to repeat exposure tasks. Further, maintenance, or further reduction, of agoraphobic avoidance was most apparent among those with more variable exposure paths. This relationship persisted when the analysis was restricted to patients who completed between four and six exposures.

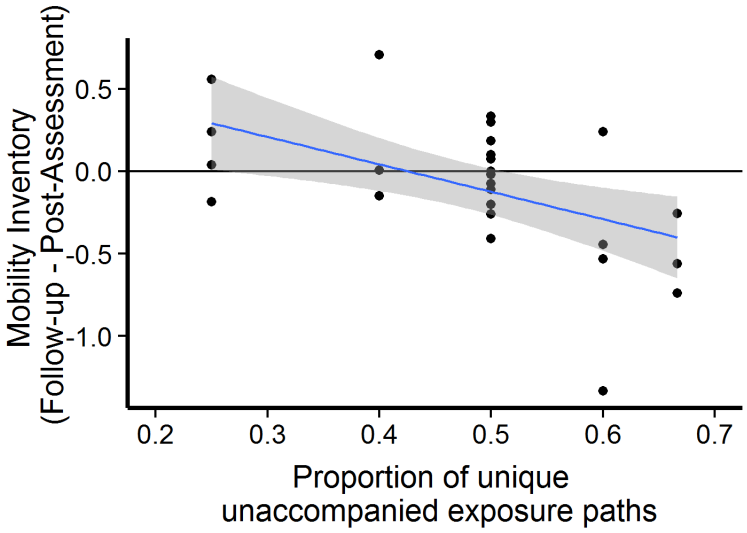

Figure 13. The association between multiple exposure contexts (proportion of unique unaccompanied exposure paths) and maintenance of treatment gains for patients who completed between four and six exposure sessions.

Note. $N = 36$. The linear regression line (blue line) and its accompanying 95% confidence region (shaded region) reflect a moderate negative relationship between relapse and variable exposure paths. Points above the horizontal black line represent patients whose avoidance increased after treatment ended.

5.4 DISCUSSION

We assessed whether variables derived from inhibitory learning models of fear extinction (Bouton, 1994; Rescorla, 2006), accounted for reductions in treatment response or maintenance of treatment gains. The most remarkable result in the current study was that patients who undertook unaccompanied exposure in a greater variety of exposure contexts tended to experience less agoraphobic avoidance following, but not during, treatment. Specifically, compared to patients with less variable exposure, those who varied bus routes experienced better maintenance of, and in some cases further reductions in, self-reported avoidance. These findings therefore

support the notion that exposure conducted in multiple contexts can help to guard against return of fear (Craske et al., 2014; Lang & Craske, 2000). This is an important finding because relapse following exposure therapy is not uncommon (Boschen et al., 2009; Peter et al., 2008).

In contrast, combining interoceptive with situational exposure did not appear to affect the rate of improvement during and following treatment. Rescorla (2006) noted that presenting two extinguished excitatory stimuli in compound may only be beneficial when one conditioned stimulus does not overshadow the other. In our study, it is possible that the situation was a more salient cue than the deliberately provoked bodily symptoms. As increased interoceptive sensitivity is common among panic patients (Domschke et al., 2010), the inverse is also possible. Another possibility is that patients undertook interoceptive tasks with varying degrees of success. The interoceptive task chosen was individually tailored, so it was possible that the ease of integrating tasks with situational exposure varied across patients. Closer examination of the therapists' records may help clarify how well these tasks were integrated. Anecdotal reports from several therapists indicated that interoceptive exercises were not always undertaken by many patients during unaccompanied exposure. This may have contributed to the current pattern of results. The present results therefore suggest that augmenting fear during situational exposure does not promote greater reductions in agoraphobic avoidance, nor does it help reduce the severity of anxious cognitions.

Exposure location (urban vs. a mixture of urban and rural) was not found to exert a systematic influence on treatment outcome. In line with findings that multiple exposure contexts help prevent return of fear (Balooch et al., 2012; Gunther et al., 1998; Neumann, 2006), completing exposure in a mixture of rural and urban contexts was expected to promote better maintenance of treatment gains. It is likely, however, that our operationalisation was confounded with other variables. For example, it is possible that patient's residential location interacted with symptom severity, in line with findings that people living in remote areas take longer to seek initial treatment for anxiety disorders compared to their urban counterparts (Green et al., 2012). Further, patients living in rural areas would have automatically been exposed to a greater variety of contexts since they were required to travel to urban centres to receive treatment. This would not, however, have been an active choice on their part. Controlling for residential location and randomising patients to urban or urban/rural conditions may help to overcome the drawbacks of correlation-based analysis of exposure location. This appears warranted given that urban upbringing and habitation have

been associated with specific patterns of social stress processing (Lederbogen et al., 2011).

Clinical Implications Relating to Variable Exposure

The clearest, and perhaps most clinically relevant findings that can be drawn from the current study relate to the use of variable exposure contexts. Patients who undertook exposure in a variety of contexts were better at maintaining treatment gains relative to those with a less variable set of exposure contexts. These findings lend support to the notion that therapists can promote superior long-term outcomes by encouraging patients to vary exposure contexts (Gunther et al., 1998; Rowe & Craske, 1998).

These results compliment and build on those obtained by Lang and Craske (2000), who found that the use of random and variable exposure contexts was more beneficial than blocked and constant exposure contexts for height phobias. Our study is the first to demonstrate these effects in a large sample of anxious patients undertaking unaccompanied situational exposure. Given that patients do not all improve at the same rate, one possibility is to individually tailor the sequence of exposure sessions in order to promote symptom reductions. This seems justified in light of the current finding that variable exposure contexts mainly helped to reduce agoraphobic avoidance but not the severity of anxious cognitions following therapy. By monitoring a patient's agoraphobic avoidance throughout treatment, therapy could be adapted to patients by identifying times when they stand to profit most from a greater variety of exposure contexts. This appears compatible with approaches that emphasise the importance of accurately characterising an individual's panic in terms of the degree and specificity of physiological change, catastrophic cognitions and somatic sensations (Whittal et al., 1996).

The current findings also have implications for how change is monitored between post-assessment and follow-up. Although considerably shorter than the intervening period adopted by Peter et al. (2008) in their study of treatment relapse, our 6-month follow-up nonetheless identified differential maintenance of treatment gains in our sample. As maintenance of gains is a key criterion of treatment success, identifying the greatest periods of risk following treatment is worthy of future research attention. To this end, it would be beneficial to monitor post-treatment change at shorter intervals to identify when patients are at greatest risk of relapse. To this end, mobile-phone based applications that use GPS-derived position could be used to objectively assess a patient's movement patterns following treatment (White et al., 2013). It is important to note that the number of unique exposure paths was a non-

manipulated variable in the current study, which limits the extent to which causal descriptions can be made (Shadish et al., 2002). For this reason, future research should extend the current findings by determining whether maintenance of treatment gains varies when number of unique exposure contexts is systematically varied. In summary, conducting exposure in multiple contexts appears to improve the likelihood that the extent of agoraphobic avoidance following treatment is maintained, and possibly improved. Increasing the number of post-treatment assessments, systematically varying the number of exposure contexts, and obtaining more objective measures of avoidance appears warranted.

Future Directions and Conclusion

Our findings regarding context variability were based on GPS-derived position. In line with White et al. (2014), it therefore appears justified to conclude that incorporating simple, readily available position sensors can help to document important dimensions of exposure (for a review see, Boschen, 2009a,b). Given a different form of exposure, it would make sense to broaden the range of GPS-derived parameters included as predictors (for a review, see Kerr et al., 2011). For example, if patients entered exposure contexts on foot, then it would make sense to examine additional parameters such as walking speed and path complexity, as feared locations were confronted. For this reason, it remains an open question whether the current findings generalise to other forms of situational exposure.

Although promising, some caution is needed when drawing conclusions from the current findings. Despite very good agreement between raters, our method of determining the variability of exposure from plots of GPS trajectories was based on subjective assessments. In future research, this could be improved by developing automated algorithms to quantify path uniqueness, or objective measurement tools, such as Geographic Information Systems to precisely determining boundaries between urban and rural area (Badland et al., 2010). In addition, we assumed that exposure context was most strongly influenced by properties of the environment outside the bus. This may not have been the case for all patients. Some may have attended predominantly to events within the bus, such as interactions with neighbouring passengers. Determining what constitutes context is of central importance when making claims about the effect of multiple contexts and is a challenge that could benefit from interdisciplinary approaches. To this end, exchanges between researchers who employ ambulatory assessment and experimental methods, especially those that use virtual reality to manipulate context (e.g., Shiban et al., 2013), should be encouraged.

As they stand, our findings support the use of multiple exposure contexts to prevent a return of agoraphobic avoidance following treatment. More generally, the current approach of examining whether a theory generated in the laboratory withstands scrutiny under natural conditions should encourage other researchers to study psychological interventions in everyday settings.

Part III

GENERAL CONCLUSION

6 GENERAL CONCLUSION

The overarching aim of this thesis was to clarify how ambulatory assessment of anxiety symptoms can help to develop a more comprehensive understanding of who responds to exposure therapy. To this end, I have outlined several approaches that revealed how panic disorder with agoraphobia (PD/A) patients undertake situational exposure. Each approach revealed psychophysiological phenomena characterised by a complex interplay between self-report, physiological, and behavioural systems (Cacioppo et al., 2007). In contrast to studies based on actively manipulated treatment variables, each of the herein presented studies predominantly involved non-manipulated treatments-related aspects (Shadish et al., 2002). This was a deliberate choice. Aside from the treatment condition within the PanikNetz study, there was no randomisation of exposure duration, adherence to end-of-exposure instructions, exposure context or exposure timing. Rather, I chose to explore the relationship between naturally varying variables and treatment response. Given the large quantity of repeated measures data collected as part of the PanikNetz project, this allowed analysis of the temporal course of variables such as heart rate and self-reported anxiety. To this end, theories of exposure were used to identify meaningful predictors and outcomes (e.g., Bouton, 1994; Foa & Kozak, 1986). By capitalising on the strengths of specific methodological approaches (e.g., HR monitoring, GPS-position) it was possible to explore how specific variables impinged on treatment responses during situational exposure. In sum, the current findings demonstrate that assessing patients under complex field conditions can be used to home in on factors that shape individual responses to situational exposure. The current series of studies thus reflected a discovery-oriented approach to addressing these aims.

6.1 MAIN FINDINGS

Study 1: A Novel Application of Technology

Study 1 provided evidence that GPS-derived position can be integrated into situational exposure and can reveal clinically-relevant information about treatment progress. Although GPS has been applied in number of health settings (Gustafson et al., 2014; Kirchner et al., 2013; Epstein et al., 2014), our idea to assess movement with GPS during situational exposure was novel. This therefore required

that we assess the feasibility and benefits of documenting movement and accompanying physiology. We used a single-case design, which is ideal for elucidating in-depth description of cases, describing how interventions are implemented, and identifying possible threats to the validity of findings (Shadish et al., 2002). Study 1 highlighted several practical implications of this new approach that were relevant for subsequent studies. First, the importance of exposure instructions was underscored—in our case, providing specific task instructions (using a drawn city map) allowed the patient and therapist to set clear goals against which compliance and progress could be gauged. We learned, however, that care should be taken to ensure that instructions are not overly restrictive or too difficult. Providing objective feedback to patients and therapists is one way of guarding against this. It is also in line with recommendations that feedback is most effective when it is objective, specific, and linked to personal goals (Archer, 2010).

In addition, analysis of movement and accompanying HR revealed that fear reactivity varied across exposure contexts. This suggested that the patient's fears were linked to specific environments—the two driving exposures had a similar structure (i.e., driving exposure, then shopping centre exposure, followed by driving exposure), yet the second exposure was undertaken in a less familiar environment and regarded as more challenging. This provided firm evidence that the familiarity of exposure location can influence the levels of physiological arousal. Compared to clinical scales measuring agoraphobia (e.g., Chambless et al., 1984, 1985), this suggests that objectively documenting confrontation of specific environments might, with further refinement, offer a more accurate picture of phobic avoidance. The findings also revealed that the symptom profile of even single patient is complex and influenced by specific background factors. For example, the patient's increased reliance on her spouse to drive her to shopping centres, and her worries about not improving given a previous relapse, served to increase the demands that she placed on herself. This study also helped us develop a detailed knowledge of the recording device. Specifically, we learned about its respective strengths (e.g., its compactness, integration of GPS and HR, ease of use) and weaknesses (e.g., 1 Hz sampling frequency limit, occasional signal dropout). In summary, this proof-of-concept demonstrated the feasibility of using GPS-derived movement to assess and refine exposure tasks, and represents a novel application of technology to the study of psychopathology (for other examples, see Eonta et al., 2011).

Study 2: Quantifying and Describing the Response Variability

As part of the PanikNetz study, on which Studies 2, 3, and 4 were based, a large quantity of psychophysiological data were

collected during exposure. Given the high between-session response variability in our single-case study, it was expected that the larger multicentre sample would produce highly variable individual responses. Inspection of HR data confirmed this. Complex field conditions are known to result in high intraindividual variability (Fahrenberg, 2006; Deboeck, 2011). In our case the conditions were particularly complex—exposure was conducted at various treatment centres, with different bus routes, under the guidance of different therapists, and with patients whose fear of buses was variable.

The main result of the clustering procedure was that individual HR responses were not systematically assigned to individuals (which we termed a "response typology"). This indicated that an individual patient could exhibit a variety of distinct HR responses across sessions—there was no stereotypical individual HR response pattern. This lent support to the idea that environmental factors (e.g., random situative influences, previous physical activation, crowdedness of bus) outweighed the influence of individual factors in determining the level and form of HR responses. As such, these data did not support the existence of panic subtypes (Kircanski et al., 2009). Further, clusters with low absolute level and low variability were associated with lower levels of self-reported anxiety during the initial stages of exposure, which suggested moderate concordance between self-report and physiological systems. These findings also had implications for the subsequent studies. First, as individual responses were so versatile, this reinforced the idea that meaningful groups of patients would need to be formed—ideally based on current theories of panic disorder and exposure therapy (Bouton, 1994; Foa & Kozak, 1986). Second, results indicated that cluster types were mainly determined by the anticipatory anxiety/pre-boarding process. This encouraged a focus on expected anxiety in later studies which was supported by findings linking anticipatory anxiety and avoidance behaviour (Craske et al., 1988). Third, conventional analysis of results revealed that repeated therapist-accompanied sessions drove reductions in physiological arousal, which supports the idea that therapist-guided exposure is centrally involved in promoting treatment response (Gloster et al., 2011). In summary, bus boarding produced an unstable pattern of intraindividual HR responses across repeated exposure. This was a first sign that it would be unlikely that HR responses would fully account for patient's treatment response. This, in turn, helped support the grouping patients on the basis of remaining responses (GPS-derived behaviours and self-reported anxiety).

Study 3: The Role of Habituation

The aim of Study 3 was to examine the degree to which treatment response was influenced by fear habituation during exposure, which is a central feature of several exposure theories (Foa & Kozak, 1986; Clark, 1986). Results revealed that EPT components were weak predictors of treatment progress and outcome. Specifically, initial fear activation and subsequent within- and between-session habituation of HR were not related with progressive reductions in pre-boarding self-reports. Less between-session habituation of heart rate was marginally associated with higher overall expected anxiety. This suggested that patients who persistently experienced high maximum HR levels across sessions tended to have high expected anxiety across sessions. This does not suggest that habituation plays no role in promoting treatment gains (i.e., reducing expectancies or reducing avoidance), but that decreased HR responses during exposure weakly predict response. These findings therefore support results from other exposure therapy studies which did not identify HR habituation as a predictor of therapeutic change (Meuret et al., 2012; Baker et al., 2010).

Adherence to the end-of-exposure instructions—a convenient, non-manipulated indicator of fear habituation—was also used to explore whether fear habituation accompanies reductions in expected anxiety across treatment. Patients who consistently adhered to end-of-exposure instructions, had higher and more persistent expected anxiety throughout treatment and remained on buses longer, compared to patients whose fear never reduced during exposure. Further, patients classified as sometimes adherent appeared to experience greater reductions in expected anxiety across exposure relative to the other groups. This raised the possibility that inconsistent adherence was akin to occasional reinforced extinction (Bouton et al., 2004; Woods & Bouton, 2007)—that reductions in anxiety during exposure were likely experienced as a form of reinforcement by patients, and possibly as a form of constraint satisfaction (Tryon, 2005). This variable reinforcement may have served to highlight a mismatch between their expectancies and outcome, thus promoting new learning (Rescorla & Wagner, 1972; Craske et al., 2014). These findings provided initial evidence that reductions in expected anxiety can be achieved by encouraging patients to occasionally leave exposure contexts even if their fear has not habituated. Further, adherence to end-of-exposure instructions, objectively assessed with ambulatory monitoring devices, emerged as a useful target for examining exposure processes. In addition, a broader implication of these results is that ambulatory assessment of position can increase the precision with which compliance can be evaluated. These more detailed accounts of *how* patients undertake

exposure should help to distil the treatment components associated with enhanced treatment gains.

Study 4: Factors That Prevent Relapse

In Study 4, we examined several aspects of exposure therapy thought relevant to relapse prevention. Here, we derived factors from inhibitory learning models that underpinned fear extinction (Craske et al., 2014; Bouton, 1993). Specifically, the effects of multiple exposure contexts and compound stimuli were studied. Results revealed that variable exposure contexts, as measured by the proportion of unique unaccompanied (homework) exposure paths, was predictive of greater maintenance of gains following post-assessment. Results concerning compound stimuli, however, did not support the idea that combining interoceptive and situational exposure was more beneficial than situational exposure alone. In summary, it appears possible to guard against relapse by encouraging patients to vary exposure contexts. This raises the possibility that there are certain times during therapy when it is important to encourage patients to vary exposure contexts. For this reason, we advocate an adaptive therapeutic approach whereby therapists monitor agoraphobic avoidance with a view to identifying when the variability of exposure contexts should be increased. Similar to Study 3, these findings also highlight the incremental validity of using GPS to track a patient's position during exposure. This enabled the study of path uniqueness, which holds promise as a method of quantifying multiple exposure contexts.

6.2 DISCUSSION AND IMPLICATIONS

The Influence of Study Design on Inferences

The research designs employed in the current dissertation predominantly involved non-experimental manipulations, which influenced the type of research questions that could be addressed (Shadish et al., 2002). Another notable feature of the studies was that they involved continuous sampling of multiple responses across repeated occasions. These two features resulted in our studies being particularly well-suited to detecting patterns of association between non-manipulated variables. This supports the observation that ambulatory studies afford a discovery-oriented research approach—it may be more difficult to exert experimental control, but the natural covariation among variables is largely preserved (Reis, 2012). For example, quantifying the natural occurrence of unique exposure paths might allow researchers to capture the patient characteristics that facilitate decisions to confront new environments. Strategies

of developing these characteristics could then be examined in a controlled therapeutic study (e.g., Salkovskis et al., 2007)

Study one was a single-subject case study, which limited the extent to which causal inferences could be drawn about the impact of our novel application of GPS (Shadish et al., 2002). This design, however, permitted a detailed description of how the patient profited from the new approach. An alternative approach that allows greater control, is the "changing criterion design", which involves assessing the baseline level of target behaviour followed by repeated implementation of a treatment. For each treatment phase, the performance criterion for the target behaviour is set until subgoals are attained (Hartmann & Hall, 1976; Meltzoff, 1998). As the study progresses, the criterion is changed in a step-wise fashion, and each treatment phase provides a baseline for the next phase (Hartmann & Hall, 1976). Changing criterion design has been found effective in reducing phobic (Wolff & Symons, 2013) and excessive checking behaviours (Allen & Evans, 2001). This design partially overlaps with an aim of the Study 1—to use depictions of previous exposure to choose an increasingly challenging exposure situation. In this case, the target behaviour is approach behaviour. By including more repetitions, it would be possible to strengthen our claim that depictions of GPS-trajectories and accompanying physiology can facilitate the planning of subsequent exposures. For example, the criterion (e.g., length of time spent in a particular area, or number of exposure occasions) could be updated after each exposure by modifying exposure to include areas associated with increasingly higher levels of physiological arousal. Alternatively, were a patient's safety behaviours (central to conceptions of agoraphobia, Rachman, 1984a) trackable with GPS (e.g., slowed walking pace to avoid increasing anxiety), a changing criterion design could be applied to demonstrate the utility of objectively assessing the target behaviour. In sum, the design of Study 1 was appropriate for demonstrating the feasibility and utility of a novel method, however depending on the target behaviour, alternate designs could be employed to establish the benefits of novel applications of technology.

Studies 2 to 4 were based on data collected within a clinical treatment study with a small number of a priori experimental conditions—treatment condition, session type (accompanied vs unaccompanied) and session number (1-4). The remaining variables which we examined were not randomised or controlled (i.e., exposure duration, adherence to instructions, number of unaccompanied sessions completed). As it is difficult to determine the effects of non-manipulated causes (Shadish et al., 2002) it is worth considering the benefits and limitations of this design feature. In Study 2, for example, descriptive modelling was adopted to summarise the data in a compact manner and as such, was not designed to test a causal theory

(Shmueli, 2010). After performing the cluster analysis, associations between cluster assignment and variables external to clustering were used to attribute meaning to the HR groups. Had clustering sorted the separate trials from an individual together, it would have been possible to determine whether specific group characteristics were responsible for differential treatment response. This, in turn, would support the existence of discrete disorder subtypes (Andor et al., 2008; Roberson-Nay et al., 2012; Sullivan et al., 2004). The results of Study 2, however, suggested that it would be somewhat far-fetched to expect that physiological responses would strongly predict treatment outcome. This was borne out in Studies 3 and 4, where the majority of supported hypotheses were derived from a combination of GPS or self-report measures. In these studies, the natural variation in participant's responses and current theories of extinction were used to guide variable selection and construct operationalisation. Association-based statistical models were then applied to examine relationships between these newly created factors and aspects of treatment response. The merit of this approach is that it is hypothesis-generating. Future research can thus build on these findings by actively manipulating the variables that we identified as being associated with desirable outcomes. For example, in a follow-up field study, an ambulatory device could be used to prompt patients to exit the exposure context on the basis of HR habituation (e.g., habituated vs not habituated). Here, different modes of instruction (e.g., automatic prompt vs. a phone call from therapist) could be compared to examine the mode of instruction effects (Bernstein & Nietzel, 1973; Speltz & Bernstein, 1976). This would be particularly helpful as there is a paucity of well-controlled research on the effects of demand characteristics (McCambridge et al., 2012). As they stand, our studies played to the strengths of their underlying research design. Future studies with greater experimental control can use the targets that we identified to generate causal descriptions that explain which factors drive superior treatment outcomes.

Crossovers Between Laboratory and Ambulatory Studies

A central aim of this dissertation was to demonstrate that ambulatory studies can build on, and promote an exchange with, research conducted in controlled settings. Wittmann and Klumb's (2006) description of experimental and non-experimental treatments provides a framework for understanding how the current studies compliment laboratory-based research. In their five-data-box conceptualisation, they distinguish between two broad approaches to examining relationships between predictors, causes and effects: (i) predictors can be actively manipulated, randomised to treatment conditions, or (ii) predictors can consist of non-

manipulated treatment aspects (Wittmann & Klumb, 2006). Further, poor correspondence between predictors and criterion constructs (violations of symmetry) should be examined and can identify where research methods can be refined (Wittmann, 2012). Sources of asymmetry can arise from differences in the psychometric properties of measures; how data have been aggregated; the hierarchy of predictors and criteria, and the level of generality at which constructs are studied (Wittmann & Klumb, 2006; Wittmann, 2012). Ultimately, understanding the sources of this asymmetry can help to identify discrepancies between laboratory and field research and sharpen the methods used to address research questions.

This framework helps to clarify sources of asymmetry in our studies. A key finding in Study 4 was that multiple exposure contexts may be associated with a reduced risk of relapse, which relates to fear extinction research that has examined context effects (Bouton, 1993, 1994, 2002). Under laboratory conditions, visual context has been varied using photographs of different scenes (Balooch et al., 2012), with different coloured room lights (Neumann et al., 2007), and by altering the room colour in virtual reality environment (Shiban et al., 2013). These manipulations have provided high levels of internal validity, but raise the question whether these findings extend to natural environments. Specifically, it is questionable whether context, as it is manipulated under laboratory manipulations, operates at the same level of generality as variations in natural settings. Data from Study 3 indicated that a greater proportion of unique, unaccompanied exposure paths was associated with less relapse. This compliments findings on the benefits of multiple exposure contexts and also serves to generate further hypotheses. For example, it remains unclear what aspect of the bus exposure context was actually important. Perhaps a combination of passengers effects, differences in built environment, and even road features (e.g., straight vs. with many bends) exert an influence on this association. Further, it has been noted that salient interoceptive events (e.g., expectation of events, Bouton, 1993) provide an internal context that may be relevant to the treatment of anxiety (Bouton, 2002). The way in which interoceptive contexts interact with external contextual stimuli to create an overall context is no doubt complex and requires further research to disentangle these factors. Using the example of exposure context, I maintain that both laboratory and ambulatory assessment researchers should strive to find parsimony when operationalising complex, multi-faceted psychological constructs, and agree on comparable methods of aggregating data before analysis. Drawing on Brunswik's notion of representative design (1955; 1956), I propose that this can be achieved by actively testing ideas derived from each of these fields in a variety of research settings.

Procedural Implications

A study with a suitably large sample, modern measurement devices and a clear research question will still fail if procedural matters are not adequately refined. A number of important lessons can be drawn from the current studies.

In laboratory settings, event markers are typically programmed by researchers before the experiment to allow subsequent extraction of precise segments of interest (Coan & Allen, 2007). In unsupervised ambulatory studies, participants can be instructed to manually set event markers, or they can be programmed to be automatically set. In the PanikNetz project, patients were required to manually set markers on both the ecological momentary assessment (EMA) device and the sports watch. This presented a challenge when patients forgot to set markers on entering and exiting buses. When this occurred, the redundancy from additional channels (GPS-derived speed, position, and EMA input) was used to verify time of boarding and exiting (cf. Wilhelm et al., 2012a). This, however, raised some questions about the burden of setting markers for patients while undertaking exposure. Given the relationship between the burden of recording and instrument reactivity (Santangelo et al., 2013), and that devices can represent a safety signal for some patients (Flynn et al., 1992), a strong case should be made for the use of automated, failsafe methods of setting markers. In future studies, this could be overcome using mobile devices that use algorithms based on multi-modal sensors (e.g., audio signals, accelerometer, GPS position changes) to detect the occurrence of specific events. Although this may sound ambitious, mobile-based applications that monitor physical activity, such as Moves (`www.moves-app.com`), have already demonstrated that it is possible to detect a range of activities, such as bike riding, travelling on public transport, and walking. Applying similar algorithms would also enable delivery of context-dependent prompts which could allow acquisition of event-triggered responses.

The current studies also have implications for the interpretation of adherence to exposure instructions. In Study 3, adherence to end-of-exposure instructions were argued to reveal the impact of fear habituation, but may also have been influenced by demand characteristics. The patients that comprised samples in all studies were provided with a set of instructions as to how to complete exposure. In Study 1, the instructions were to drive to an agreed upon area, enter a shopping centre, and then continue driving. Here, it was argued that adherence with instructions reflected compliance. In subsequent studies, patients were instructed to remain on buses until their fear reduced; adherence was proposed to reflect compliance and as a consequence, represent the effects of fear habituation. Although adherence to instructions can be viewed

as therapeutic compliance, the benefits of which are central to CBT (Kazantzis et al., 2000; Westra et al., 2007), they may also be a product of demand characteristics, especially when therapists are accompany patients during exposure. Demand characteristics, in this context, refer to the range of cues and expectations that influence the responses of participants. As the client-therapist relationship engenders expectations that support improvement (Dew & Bickman, 2005), within the context of exposure therapy, it is worth considering which factors discriminate between compliance-driven and expectation-driven change. I define compliance as acting with volition and control to bring about a change of state. In contrast, actions driven by expectations have an external focus and are probably subject to demand characteristics. Early findings involving undergraduate samples approaching snakes revealed somewhat inconsistent findings concerning demand characteristics. Imposing strong demands was shown to encourage greater approach behaviour (Speltz & Bernstein, 1976), yet in another study, high experimental demands only served to increase HR (Odom & Nelson, 1977), relative to low demand conditions. Applying modern ambulatory assessment techniques could refine this procedure and would allow researchers to examine these effects during unaccompanied situational exposure. For instance, using context-aware ambulatory devices (e.g., Ho & Intille, 2005), would allow manipulation of both exposure task difficulty and the level of demand (e.g., by varying the source of information: sent from therapist vs. determined by a computer algorithm or previous behaviour). This could help to disentangle these complex relationships. In summary, results of the current studies suggest that modern context-aware devices can be used to reduce the burden placed on patients. Further, the experimental demands associated with exposure-based procedures could be manipulated to separate autonomous, active compliance from mere acquiescence to instructions.

Implications for Analytic Methods

Our findings support the adoption of analytic methods which accurately describe the time course of treatment effects (Haynes & Yoshioka, 2007). Multilevel growth models are well suited to modelling change since they can efficiently handle missing data and allow inclusion of time-varying predictors (Singer & Willett, 2003; West et al., 2007). However, these models require that researchers carefully consider whether predictors vary systematically or randomly across individuals. This is a departure from traditional methods of analysing change such as repeated-measures ANOVA, which do not permit such precise partitioning of variance. Another strength of multilevel approaches is that they allow researchers to

examine how assessments for each individual are related across time—it is possible to specify these predictions in a variety of covariance structures (Hamaker, 2012). For example, within-person processes can be described as unstructured, exchangeable or autoregressive. This has allowed some researchers to investigate "emotional inertia", the time-lagged effects of emotional states (Kuppens et al., 2010). In sum, the use of multilevel models, in which specific forms of within-person change is specified, can greatly help to examine how processes unfold during exposure-based therapy.

Although instructions were provided to measure baseline heart rate after each exposure, patients inconsistently obtained these measures. Among those who did record baseline, it was evident that HR responses were, in many cases, still elevated and resembled a phasic response to post-exposure context. This cast doubt on the utility of measuring change from this "basal condition" and highlighted the possible advantage of using an alternate baseline (Jennings et al., 1992). Jennings et al. (1992, p. 743), recommend a "vanilla baseline", where participants are "measured during a stable comparison period during the experimental sessions". This is a departure from resting baseline, and should encourage ambulatory assessment researchers to identify meaningful points of comparison. For example, exposure to less feared situations (i.e., with a low position on an individual's fear hierarchy) could be used as a baseline for more challenging tasks (cf. multiple baseline procedures used in changing criterion designs).

Our findings also have implications for the analysis of aggregated data. In Study 2, clustering individual responses revealed a response typology, suggesting that no individual response subtypes could be identified on the basis of HR. This provided an early sign that it would be ambitious to expect that facets of individual responses based on HR alone (e.g., our operationalisation of initial fear activation) could account for treatment response. This, in turn, helped justify the use of a flexible modelling strategy that included random intercepts and slopes of change. An extension of our method would have been to use a multilevel latent class analysis (MLCA) approach, which permits a more direct identification of latent classes based on within-person profiles. Simulation studies have revealed, however, that identifying clusters of individuals in MLCA are underpowered unless larger sample sizes, typically exceeding 500 individuals, are used (Ng & McLachlan, 2014; Vermunt, 2003; Finch & French, 2013). The upshot of this is that researchers should consider the level of individual response variability that can be expected under particular conditions. When variability is particularly high, it is worthwhile grouping similar responses before aggregation and averaging procedures are applied.

Finally, the depictions of raw data and associations among variables presented throughout this dissertation attest to the utility of data visualisation. It has been noted that "visualisation is critical to data analysis. It provides a front line of attack, revealing intricate structure in data that cannot be absorbed in any other way. We discover unimagined effects, and we challenge imagined ones" (Cleveland, 1993, p.1). Data visualisation is essential to discovery-oriented research approaches. In Study 1, plots helped the therapist and patient to understand the association between geophysical location and physiological arousal (Figure 2), and in Study 2, a novel depiction was used to inspect the complex pattern of cluster assignment (Figure 16). In the Studies 3 and 4, individual exposure paths were explored, which guided the selection of parameters that were expected to be concomitant with treatment response. In summary, these results encourage the use of data visualisation techniques to guide analysis and suggest that they can be a useful hypothesis-generation tool.

Clinical Implications

Diagnosis and treatment

Although generally effective, the current set of findings suggest that greater therapeutic benefits can be achieved by tailoring treatment to individual patients. This has implications for the way in which mental illness is diagnosed. Classifying disorders is argued to be a necessary precursor to identifying disorders with a known aetiology, course and treatment response. The diagnosis a patient receives is important as it determines the treatment that is administered. Currently, syndromal classification is used establish illness categories such as those found in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, American Psychiatric Association, 2013). Here, classification involves identifying groups of signs and symptoms (the client's complaint/s) that purportedly characterise the condition (Hayes et al., 1996). A problem with this approach is that the same biobehavioural phenomena are common to several distinct disorders (e.g., fear circuits are involved in anxiety disorders and the activation of stress responses, Shin & Liberzon, 2010), which may explain the difficulty in pinpointing the distinct mechanisms of change for successful treatments (Hayes et al., 1987). To address this shortcoming, it has been proposed that behavioural and neurobiological dimensions should form the basis of mental disorder classification (Cuthbert & Insel, 2013; Insel, 2012).

The results of Studies 3 and 4 provide support for the notion that treatment should be tailored to patients and thus be based on the responses that characterise an individual. For example, we recommend that therapists monitor the duration of time patients

spend on buses, and vary the number of distinct exposure contexts to promote better treatment outcomes. This could be taken much further by systematically examining a larger range of therapeutic targets. This is not easy as psychological treatments such as situational exposure are multifaceted, however, one possibility is to choose targets that already have some empirical support. The research domain criteria (RDoC) provides a common set of targets (e.g., negative valence domain) which are regarded as central to mental disorders and their treatment (Cuthbert & Insel, 2013). The task force which developed this framework, in line with a dimensional classification approach, identified two challenges that require addressing: (i) to "determine the full range of variation, from normal to abnormal, among the fundamental components to improve understanding of what is typical versus pathological," and (ii) to "develop reliable and valid measures of these fundamental components[1] of mental disorders for use in basic studies and in more clinical settings" (Cuthbert & Insel, 2013, p.4). Measures developed to study behaviour in natural settings (e.g., GPS measures of avoidance) require further validation, but show great potential to elucidate the full range of variation among these components.

In summary, a complex range of factors appear to determine how participants respond to exposure therapy—a few of which we identified by drawing on modern learning accounts of fear. Examining patients with a variety of anxiety disorders on the basis of components derived from studies of fear and anxiety circuits (Shin & Liberzon, 2010; LeDoux, 2014) and that accord with the research domain criteria would greatly extend the current approach. Extracting the essential components of heterogeneous diagnostic categories such as agoraphobia could then help to identify the range of mechanisms that require targeting during treatment.

Novel technological adjuncts to therapy

Novel applications of technology hold promise for promoting important psychotherapeutic aims (Eonta et al., 2011; Boschen & Casey, 2008; Clough & Casey, 2011). Findings from Study 1 indicated that technology can play a central role in delivering feedback during therapy and can help patients to develop insight into a dysfunctional process—an important therapeutic goal (Shapiro, 1995). Technological adjuncts to therapy that facilitate the objective feedback of patient's behavioural and physiological measures can help to elucidate the factors which contribute to their problems, and provide an indication of treatment response (Newman, 2004). Designing

[1]The five research domain criteria are: Negative valence domain (e.g., potential threat); positive valence systems (e.g., approach motivation); Cognitive systems (e.g., attention); systems for social processes (e.g., affiliation and attachment); arousal/modulatory systems (e.g., biological rhythms)

effective tools for patients could be achieved by drawing on reviews of feedback interventions (e.g., Archer, 2010; Kluger & Denisi, 1996). For example, feedback is most effective when it encourages reflection-in-action, supports the setting of learning-oriented goals, and when conceptualised as an iterative process (Archer, 2010). When attention is given to these issues, applying technology within clinical settings has great potential to support patients to learn more about their problems and can help convey the rationale behind therapy.

The development of technological adjuncts to psychotherapy is already quite advanced. Research at the intersection of psychology, design and computer science has yielded an increasingly sophisticated range of mobile applications and wearable sensing devices that will become increasingly relevant for clinical practice (for reviews, see Morris & Aguilera, 2012; Maheu et al., 2012). Several innovative proof-of-concepts have been described in the literature including using deviations in heart rate variability to trigger mobile-based breathing visualisation and cognitive reappraisal application (Morris & Guilak, 2009), predicting stereotypical motor movements in autistic children using real-time psychophysiological responses to enable timely intervention (Goodwin et al., 2011), and monitoring physiology and behaviour to track depression levels among clinically depressed patients throughout treatment (Sung et al., 2005). Simpler mobile interventions that involve offering encouragement to patients via text messages have also been positively received and rated as helpful by patients and therapists (Aguilera & Muñoz, 2011; Bauer et al., 2003). But this is not always case. Simply applying technology does not guarantee that it will be adopted by patients. In one study, a post-treatment relapse prevention intervention was delivered to a sample of bulimia nervosa sufferers via text-messages (Robinson et al., 2006). The limited acceptance of the intervention and high attrition, however, raises questions about whether patients received adequate training in how to use the tool; the suitability of the intervention for patients who show limited improvement during therapy; and whether interventions that promote greater interaction and offer useful feedback might be better accepted. When barriers to engagement are overcome, I maintain that adjunct technologies can serve to enhance clinical interventions. For example, preventing relapse among patients who have completed exposure-based therapies is one area that is amenable to technological support. For example, we recently developed a prototype of a post-exposure relapse prevention application which interactively creates fear maps based on GPS-movement and thus encourages continued approach of feared locations (White et al., 2013). Identifying suitable times to integrate technology (e.g., during or after treatment), ensuring that applications are based on sound theoretical concepts, and verifying that features are well-implemented will help to leverage the benefits of novel interventions. A further examination

of factors that enable adoption as well as the ethical challenges that present when large quantities of patient data are collected, requires further research attention (Clough & Casey, 2011; Eonta et al., 2011; Maheu et al., 2012).

6.3 CONCLUSION

An important goal for ambulatory assessment researchers will be to better understand the associations between psychophysiological responses as they unfold in specific environments. When disorders are strongly influenced by environmental features, then considering *where*, as well as *how* people respond will become increasingly relevant.

The development of multi-sensor devices that allow real-time monitoring and communication with participants will give researchers the opportunity to examine person-environment relationships with greater precision (Intille, 2012). Further, as increasingly innovative technology is developed, it will soon be a reality for clinicians to use technology to deliver interventions based on changes in psychophysiological events within specific environments. This will greatly help to expand the range of novel applications of technology psychotherapy (Morris & Aguilera, 2012), will allow interventions to be tailored to individuals, and will assist efforts to identify the mechanisms by which psychological treatments exert their influence.

Ultimately, the advancement of psychological treatments, and particularly exposure-based interventions, will result from a closer interplay between laboratory and clinical researchers (Holmes et al., 2014). But clinical researchers should bolster their efforts to conduct studies in the daily settings where their patients' problems manifest. In these settings, great opportunities exist to conduct well-designed, procedurally robust, and clinically-relevant research that can identify new, and test established, mechanisms thought responsible for treatment gains. Creative, discovery-oriented methodological approaches should therefore be encouraged, especially those which consider person-environment interactions. These, I argue will move us closest to understanding the mapping of thoughts, physiology, and behaviour that help define the mental illness. Hitherto, basic learning models of fear based on human and animal laboratory studies have provided useful causal descriptions of factors relevant to treatment. Yet further research is needed before patients can expect to receive interventions that are tailored to their needs and not simply applied in a trial-and-error fashion (Cuthbert & Insel, 2013). Despite initially grappling to resolve basic definitional and methodological issues (Fahrenberg, 1996, 2006), ambulatory assessment researchers, are now well positioned to productively exchange ideas with their laboratory counterparts. Juxtaposing findings gathered from laboratory and

ambulatory studies will help to generate questions that compel researchers to extend current methods in both research domains. Combined with the rapidly developing field of mobile health technologies, exposure-based therapies will surely have a future with unbounded possibilities.

Part IV

APPENDIX

A

GPS-RELATED SUPPLEMENTS

A.1 A COLLECTION OF GPS TRAJECTORIES FROM A SINGLE PATIENT

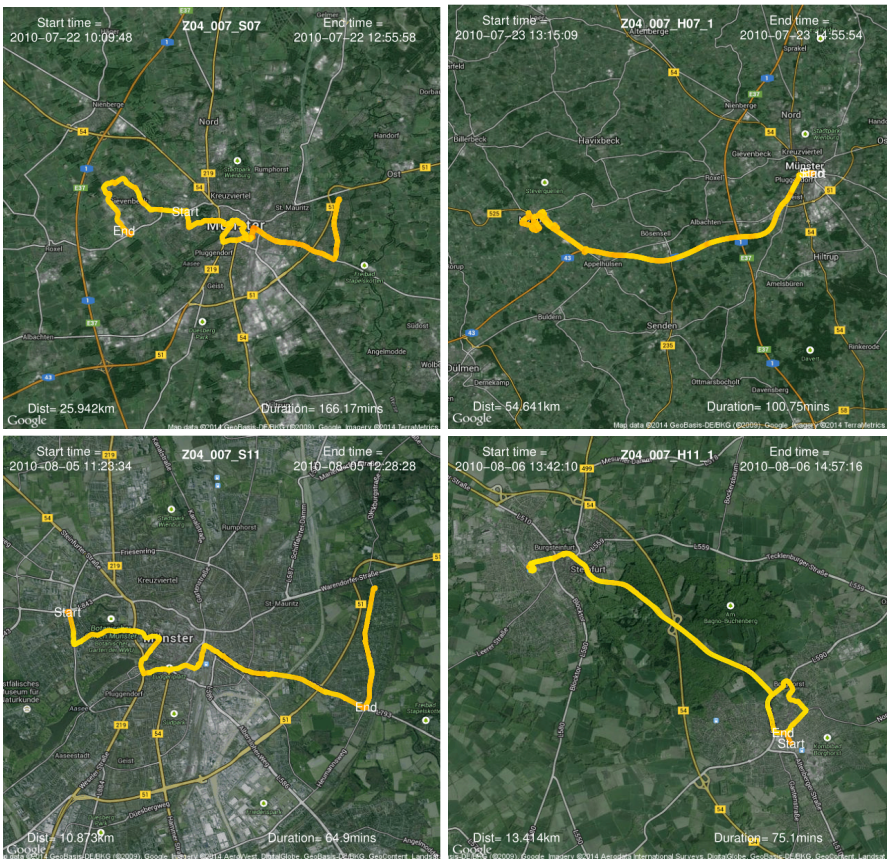

Figure 14. GPS trajectories from exposure sessions undertaken by a single patient from the PanikNetz study.

Note. Plots produced using RgoogleMaps (Loecher, 2014) using maps obtained from Google.

B

HEART RATE SUPPLEMENTS

B.1 SPAGHETTI PLOT: UNSORTED HEART RATE

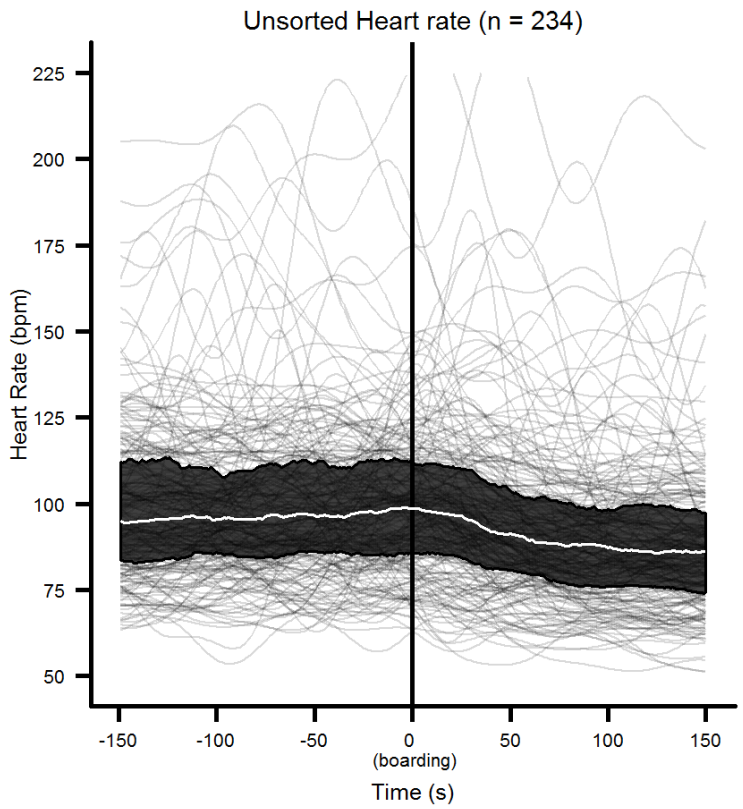

Figure 15. Spaghetti plot of unsorted, unstandardised heart rates for the circumscribed segment of interest.

B.2 DICEPLOT: CLUSTER ASSIGNMENT PATTERNS

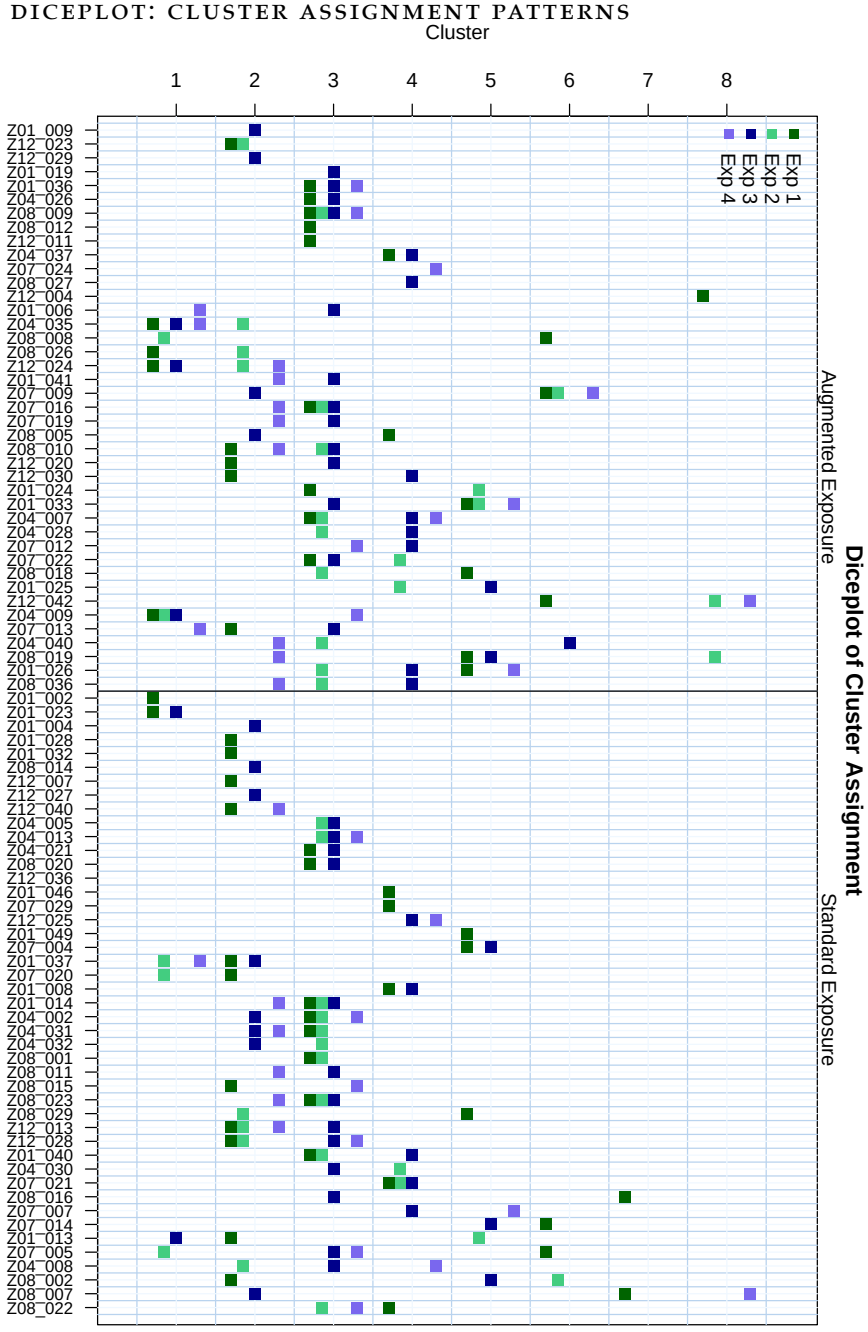

Figure 16. Diceplot of cluster assignment and transitions. Square represents a single bus exposure and a patient's complete set of exposures is depicted along each horizontal line.

B.3 OPERATIONALISATION OF EMOTIONAL PROCESSING THEORY COMPONENTS

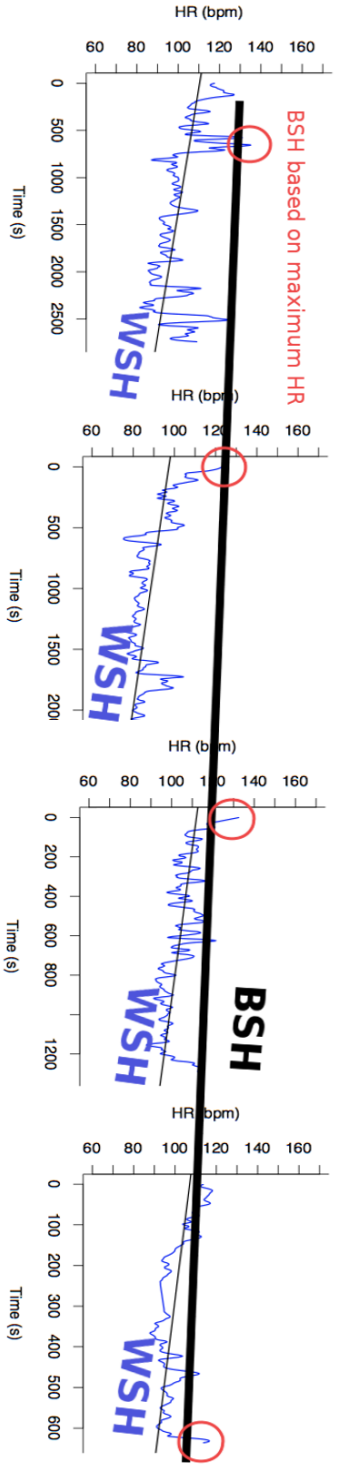

Figure 17. Operationalisation of within- and between-session habituation, derived from the emotional processing theory components.

Note. Each plot depicts HR responses from Exposure 1 (far left) to 4 (far right). BSH = between-session habituation, WSH = within-session habituation. Both BSH and WSH were derived from robust regression coefficients.

B.4 DISTRIBUTIONAL PLOTS OF EPT COMPONENTS

B.4.1 *Initial fear activation*

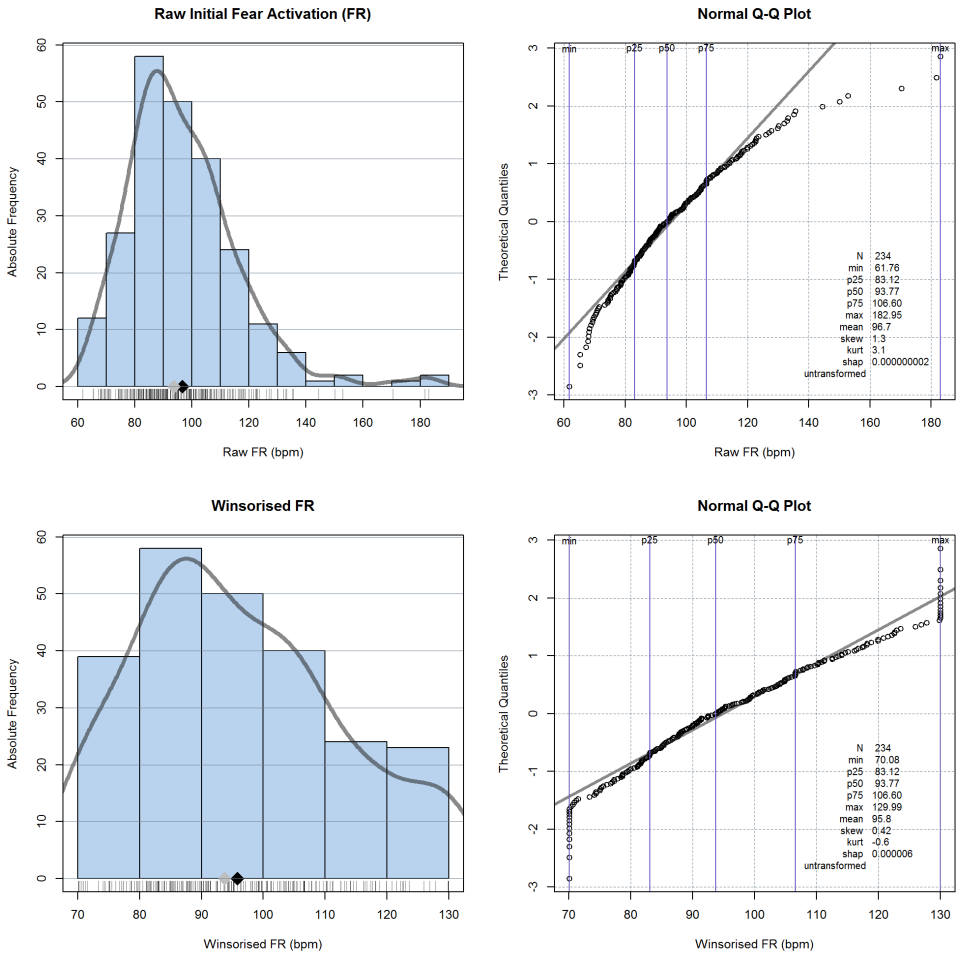

Figure 18. Distributional plots for raw (top row) and winsorised (bottom row) initial fear activation.

Note. 5% Winsoring applied to raw data.

B.4.2 *Within-session habituation*

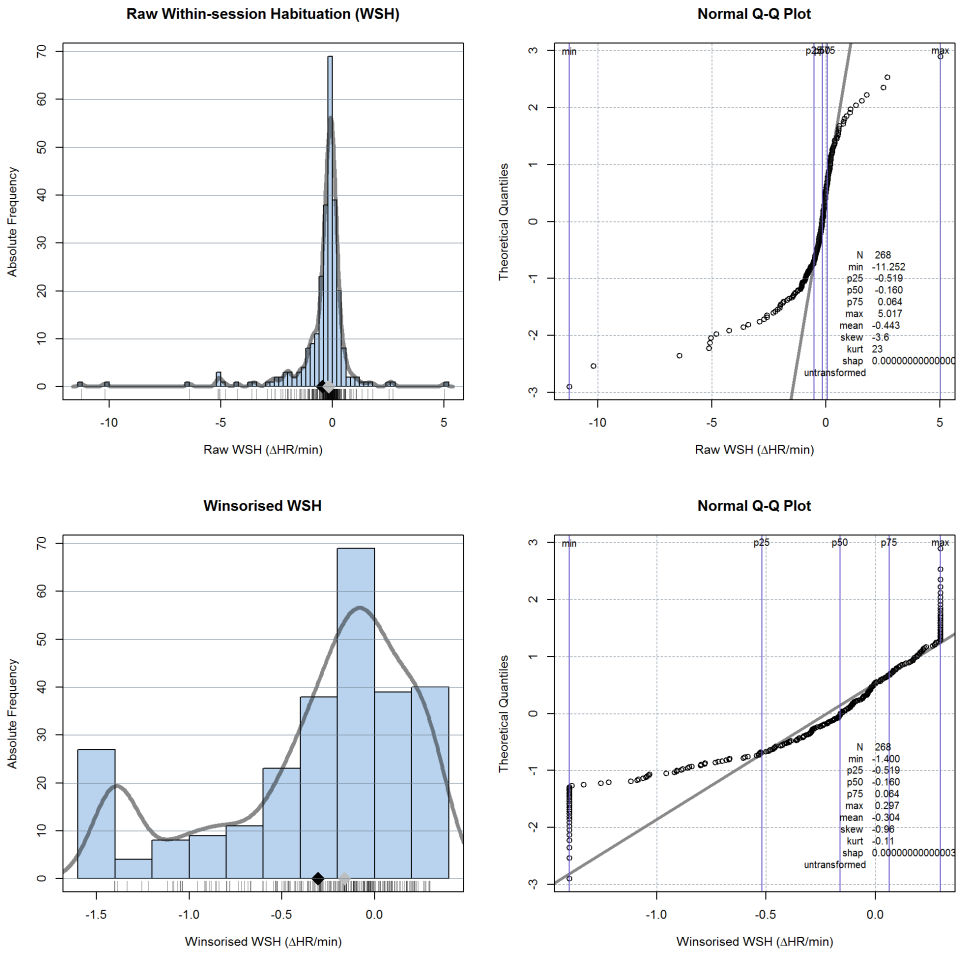

Figure 19. Distributional plots for raw (top row) and winsorised (bottom row) within-session habituation.

Note. 5% Winsoring applied to raw data.

B.4.3 *Between-session habituation*

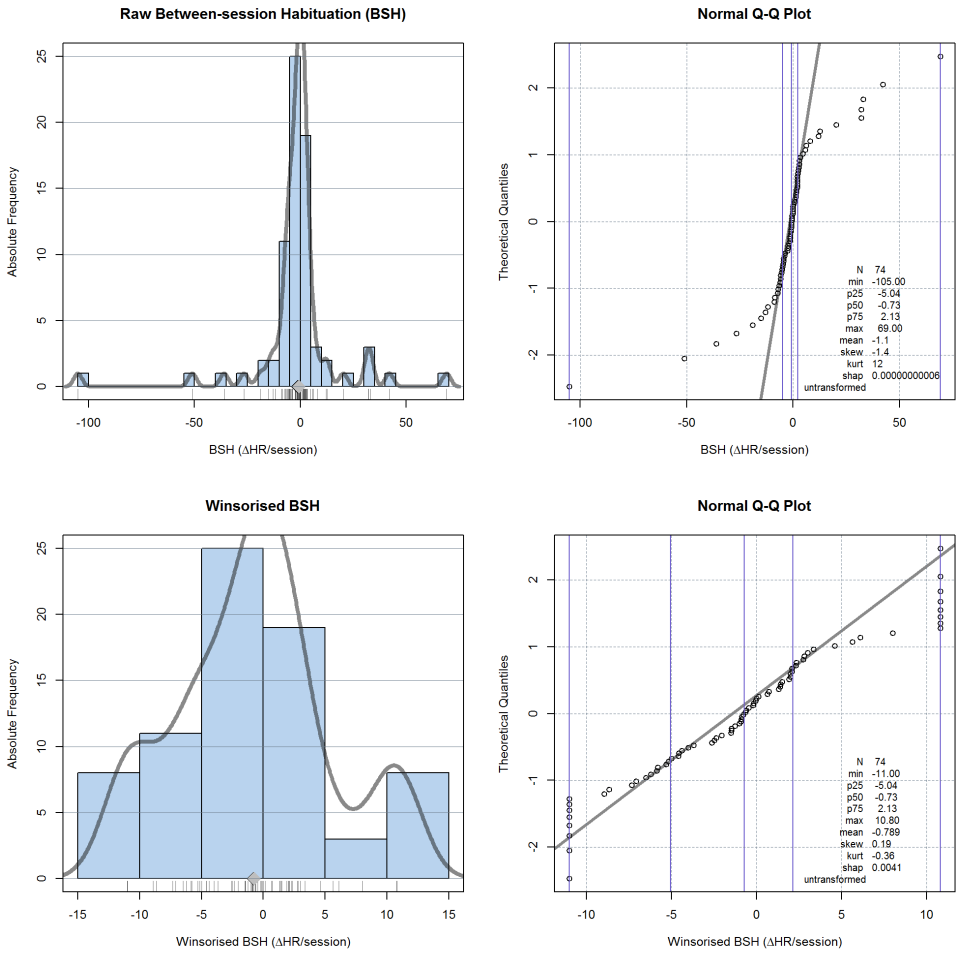

Figure 20. Distributional plots for raw (top row) and winsorised (bottom row) between-session habituation

Note. 5% Winsoring applied to raw data.

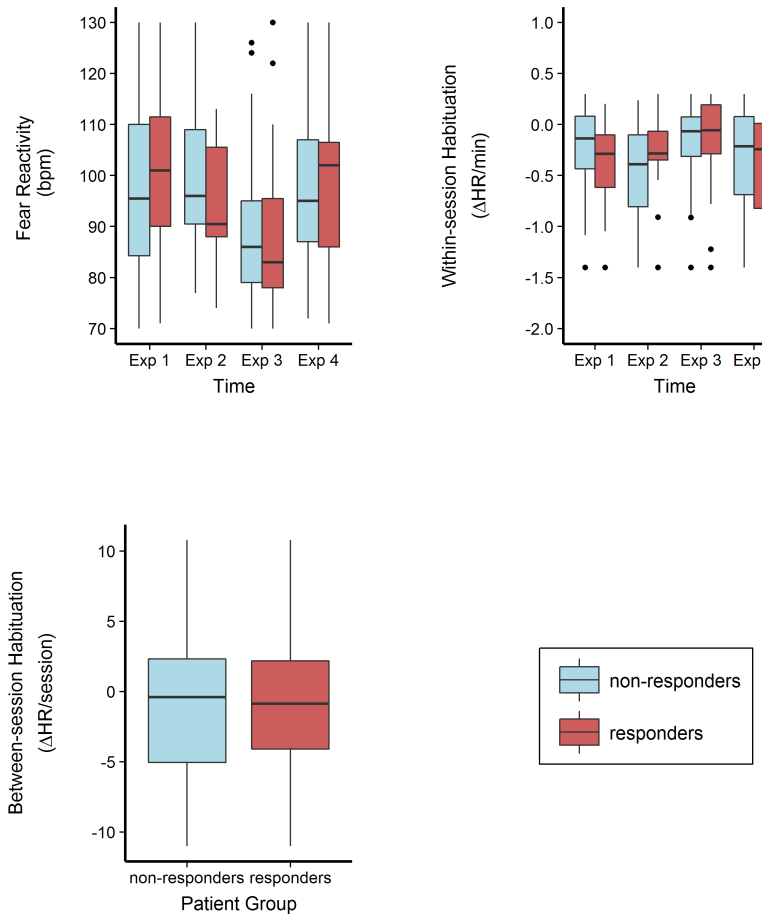

Figure 21. Emotional processing theory predictors (initial fear activation, within- and between-session habituation) across treatment non-responders and responders.

Note. Binary treatment response based on reduction in expected anxiety scores.

The ability of the time-varying parameters, IFA and WSH, to predict binary treatment response (responders vs. non-responders, based on expected anxiety) was assessed in series of logistic regression models. Results revealed that initial fear activation and within-session habituation did not predict treatment response in any specific exposure session. Deviance reductions from the null model at each time point were also non-significant (Exp 1: $D = 1.94$, $p = .38$; Exp 2: $D = 2.10$, $p = .35$; Exp 3: $D = .20$, $p = .90$; Exp 4: $D = 0.34$, $p = .84$, each with $df = 2$). The time-invariant variable, BSH, was also a weak predictor of treatment response ($D = 1.00$, $df = 1$, $p = .32$). In sum, EPT parameters, when aggregated across individuals were poor predictors of treatment response. This suggested that treatment progress trajectories were idiosyncratic and that other contextual variables might be required to better predict a person's response to exposure.

B.6 EPT COMPONENT COMPARISON

Table 9. Results of fitting several multilevel models for change to expected anxiety.

	FR Model	WSH Model	BSH Model
Fixed effects			
Intercept (Initial status)	4.43 *** (0.38)	4.29 *** (0.37)	4.42*** (0.34)
Expected Anxiety (Exp 1)	0.82*** (0.11)	0.80*** (0.11)	0.69*** (0.09)
Session (Exp 1-4) (Rate of change)	-0.99*** (0.28)	-0.89 ** (0.28)	-0.86** (0.25)
EPT Component	-0.01 (0.02)	0.44 (0.69)	0.10~ (0.06)
EPT Component x Session	-0.003 (0.02)	-0.42 (0.60)	-0.02 (0.06)
Variance Components			
Within-subject (Level 1)	1.89	2.16	0.77
Initial Status	1.65	1.28	2.11
Rate of Change (Session)	0.71	0.80	16.88
Model Information Criteria			
-2 log-likelihood	-179.33	-193.48	-224.80
AIC	376.65	404.95	467.61
BIC	398.53	427.45	491.41

Note. Sample sizes for each model: Model A, B = 41; Model C = 50
~$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$
Standard errors of fixed effect coefficients presented in brackets.

Table 10. Results of fitting several multilevel models for change to fear of losing control.

	FR Model	WSH Model	BSH Model
Fixed effects			
Intercept	2.20 ***	2.17 ***	1.99***
(Initial status)	(0.35)	(0.32)	(0.30)
Fear of Losing Control	0.27**	0.26***	0.27***
(Exp 1)	(0.08)	(0.07)	(0.06)
Session (Exp 1-4)	-0.77***	-0.74 ***	-0.70***
(Rate of change)	(0.18)	(0.18)	(0.18)
EPT Component	-0.01	0.40	0.07
	(0.02)	(0.65)	(0.06)
EPT Component x	-0.01	-0.001	-0.06
Session	(0.01)	(0.47)	(0.04)
Variance Components			
Within-subject (Level 1)	1.23	1.39	1.43
Initial Status	2.48	2.06	2.05
Rate of Change (Session)	0.13	0.18	0.20
Model Information Criteria			
-2 log-likelihood	-153.96	-164.33	-189.88
AIC	325.91	346.67	397.75
BIC	347.79	369.17	421.55

Note. Sample sizes for each model: Model A, B = 41; Model C = 50
$^{\sim}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$
Standard errors of fixed effect coefficients presented in brackets.

C

ECOLOGICAL MOMENTARY ASSESSMENT SUPPLEMENTS

C.1 EMPIRICAL CHANGE PLOTS FOR A SUBSAMPLE OF PATIENTS

C.1.1 *Expected anxiety*

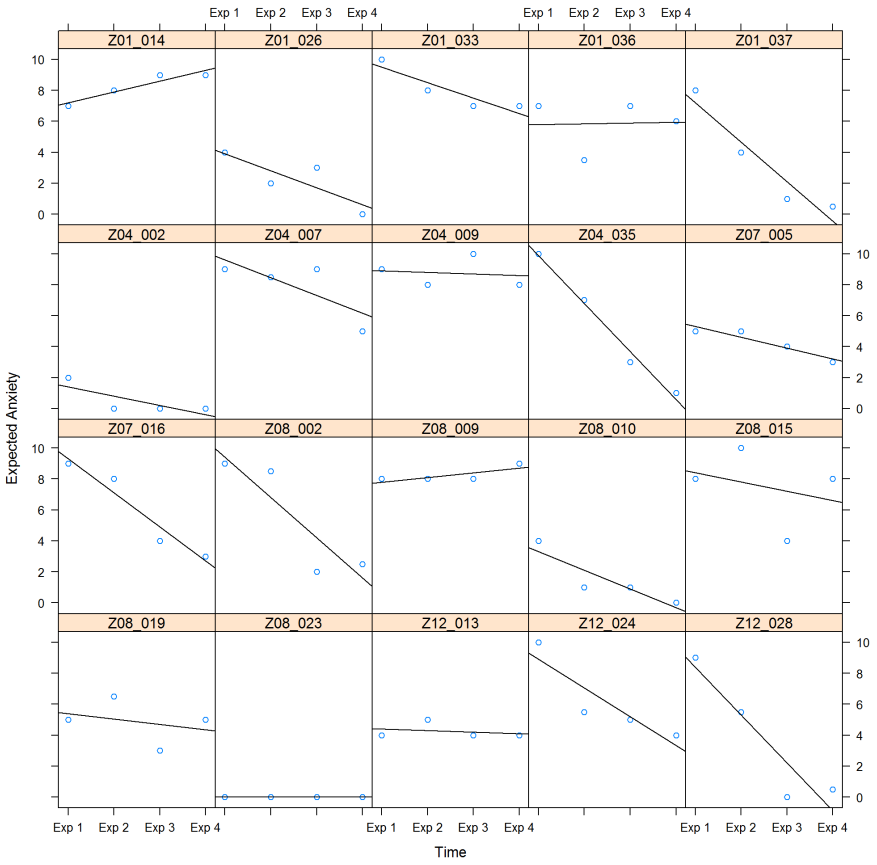

Figure 22. Empirical growth plots for expected anxiety overlaid with ordinary least squares trajectories.

C.1.2 *Fear of losing control*

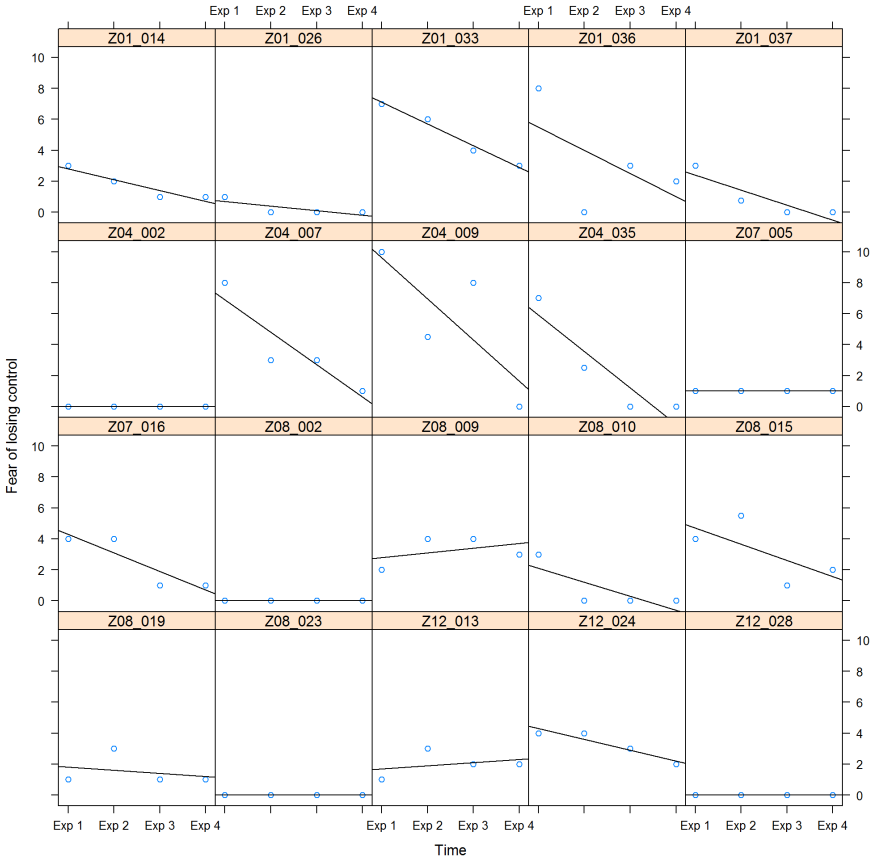

Figure 23. Empirical growth plots for fear of losing control overlaid with ordinary least squares trajectories.

C.2 SAMPLE ECOLOGICAL MOMENTARY ASSESSMENT QUESTIONS

Table 11. Sample questions presented on the ecological momentary assessment device (translated from German).

Prior to exposure commencement:	Scale
This is a therapist assisted session?	Y/N
How prepared are you to seek out this situation?	0-10 (completely prepared)
I will faint	0-10 (very sure)
I will lose control	0-10 (very sure)
Another catastrophe will occur	0-10 (very sure)
How anxious are you now, as you think about confronting this situation?	0-10 (very anxious)
What is the maximum anxiety you will experience in the situation?	0-10 (very anxious)
Start of exposure (Automatically prompted questions every 3 mins):	
How much anxiety are you experiencing now?	0-10 (very anxious)
Have you tried to augment your anxiety? (asked to those in the Fear Augmented condition)	Y/N
Are you still in the exercise?	Y/N
After exposure:	
During the exercise, what was the highest level of anxiety you experienced?	0-10 (very high)
How strong were your bodily symptoms?	0-10 (very strong)
During the exercise, did you engage in any safety behaviours? If yes, what safety behaviours?	Y/N + open-ended response
How well did you manage to tolerate the emergence of physical symptoms?	0-10(very well)
How well did you manage to tolerate the anxiety?	0-10(very well)
During the exercise, did you try to augment your anxiety?	Y/N
How strenuous was the implementation of the exercise for you?	0-10 (very strenuous)
How willing are you now to seek out the same situation?	0-10 (very willing)

Abramowitz, J. S., Deacon, B. J., Whiteside, S. P. H., Davis, M., & Powers, M. (2011). *Exposure therapy for anxiety: Principles and practice*. New York: Guilford. (Cited on pages 10 and 64.)

Aguilera, A. & Muñoz, R. F. (2011). Text messaging as an adjunct to CBT in low-income populations: A usability and feasibility pilot study. *Professional Psychology: Research and Practice*, 42(6), 472–478. (Cited on page 118.)

Aikins, D. E. & Craske, M. G. (2008). Sleep-based heart period variability in panic disorder with and without nocturnal panic attacks. *Journal of Anxiety Disorders*, 22(3), 453–463. (Cited on page 64.)

Allen, K. D. & Evans, J. H. (2001). Exposure-based treatment to control excessive blood glucose monitoring. *Journal of Applied Behavior Analysis*, 34(4), 497–500. (Cited on page 110.)

Alpers, G. W. (2009). Ambulatory assessment in panic disorder and specific phobia. *Psychological Assessment*, 21(4), 476–485. (Cited on pages 3, 6, 14, 33, and 65.)

Alpers, G. W. (2010). Avoiding treatment failures in specific phobias. In M. W. Otto & S. Hofmann (Eds.), *Avoiding treatment failures in the anxiety disorders* (pp. 209–227). New York: Springer. (Cited on page 66.)

Alpers, G. W., Abelson, J., Wilhelm, F. H., & Roth, W. T. (2003). Salivary cortisol response during exposure treatment in driving phobics. *Psychosomatic Medicine*, 65(4), 679–687. (Cited on pages 6, 25, and 33.)

Alpers, G. W. & Sell, R. (2008). And yet they correlate: Psychophysiological activation predicts self-report outcomes of exposure therapy in claustrophobia. *Journal of Anxiety Disorders*, 22, 1101–1109. (Cited on pages 6, 12, 13, 23, 40, 59, 63, and 85.)

Alpers, G. W., Wilhelm, F. H., & Roth, W. T. (2005). Psychophysiological assessment during exposure in driving phobic patients. *Journal of Abnormal Psychology*, 114(1), 126–139. (Cited on pages 6, 23, 24, 25, and 33.)

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders (DSM-IV-TR)*. Washington, DC: American Psychiatric Publishing. (Cited on pages 7 and 23.)

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. Arlington, VA: American Psychiatric Publishing. (Cited on pages 7, 8, 14, 18, 35, 61, and 116.)

Andor, F., Glöckner-Rist, A., Gerlach, A. L., & Rist, F. (2008). Symptomspezifische Subgruppen der Panikstörung. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 37(3), 161–171. (Cited on pages 17, 38, 64, and 111.)

Antony, M. & Barlow, D. H. (2010). *Handbook of assessment and treatment planning for psychological disorders (2nd ed.)*. New York: The Guilford Press. (Cited on pages 10, 15, and 24.)

Archer, J. C. (2010). State of the science in health professional education: Effective feedback. *Medical Education*, 44, 101–108. (Cited on pages 25, 106, and 118.)

Badland, H., Duncan, M. J., Oliver, M., Duncan, J. S., & Mavoa, S. (2010). Examining commute routes: Applications of GIS and GPS technology. *Environmental Health and Preventive Medicine*, 15(5), 327–330. (Cited on pages 24 and 102.)

Baker, A., Mystkowski, J. L., Culver, N., Yi, R., Mortazavi, A., & Craske, M. G. (2010). Does habituation matter? Emotional processing theory and exposure therapy for acrophobia. *Behaviour Research and Therapy*, 48, 1139–1143. (Cited on pages 7, 15, 18, 23, 24, 40, 64, 85, and 108.)

Balooch, S. B., Neumann, D. L., & Boschen, M. J. (2012). Extinction treatment in multiple contexts attenuates ABC renewal in humans. *Behaviour Research and Therapy*, 50(10), 604–609. (Cited on pages 19, 89, 100, and 112.)

Bandelow, B. (1995). Assessing the efficacy of treatments for panic disorder and agoraphobia: II. The Panic and Agoraphobia Scale. *International Clinical Psychopharmacology*, 10, 73–81. (Cited on page 13.)

Bandelow, B., Hajak, G., Holzrichter, S., & Kunert, H. (1995). Assessing the efficacy of treatments for panic disorder and agoraphobia: I. Methodological problems. *International Clinical Psychopharmacology*, 10(2), 83–93. (Cited on page 17.)

Barlow, D. H., Brown, T. A., & Craske, M. G. (1994). Definitions of panic attacks and panic disorder in the DSM-IV: Implications for research. *Journal of Abnormal Psychology*, 103, 553–564. (Cited on pages 12 and 17.)

Barlow, D. H. & Craske, M. G. (1994). *Mastery of your anxiety and panic (2nd ed.)*. San Antonio, TX: The Psychological Corporation. (Cited on pages 26, 40, and 66.)

Barlow, D. H., Mavissakalian, M., & Schofield, L. D. (1980). Patterns of desynchrony in agoraphobia: A preliminary report. *Behaviour Research and Therapy*, 18(5), 441–448. (Cited on page 13.)

Bauer, S., Percevic, R., Okon, E., Meermann, R., & Kordy, H. (2003). Paper Use of Text Messaging in the Aftercare of Patients with Bulimia Nervosa. *European Eating Disorders Review*, 11, 279–290. (Cited on page 118.)

Beck, A. & Emery, T. (1985). *Anxiety disorders and phobias*. New York: Basic Books. (Cited on page 8.)

Beck, A., Steer, R., & Brown, G. (1996). *Manual for the BDI-II*. San Antonio: The Psychological Corporation. (Cited on page 26.)

Bell, M. L., Kenward, M. G., Fairclough, D. L., & Horton, N. J. (2013). Differential dropout and bias in randomised controlled trials: When it matters and when it may not. *British Medical Journal*, 346, 1–7. (Cited on page 74.)

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300. (Cited on pages 46 and 95.)

Bernstein, D. A. & Nietzel, M. T. (1973). Procedural variation in behavioral avoidance tests. *Journal of Consulting and Clinical Psychology*, 41, 165–174. (Cited on page 111.)

Bernstein, D. A. & Nietzel, M. T. (1974). Behavioral avoidance tests: The effects of demand characteristics and repeated measures on two types of subjects. *Behavior Therapy*, 5(2), 183–192. (Cited on pages 15 and 90.)

Bibb, J. (1988). *Parental bonding, pathological development, and fear of losing control among agoraphobics and normals*. Unpublished doctoral dissertation, The American University, Washington, DC. (Cited on page 28.)

Bornas, X., Gelabert, J. M., Llabrés, J., Balle, M., & Tortella-Feliu, M. (2011). Slope of change throughout exposure treatment for flight phobia: The role of autonomic flexibility. *Journal of Clinical Psychology*, 67, 550–560. (Cited on page 23.)

Boschen, M. J. (2009a). Mobile telephones and psychotherapy—I: Capability and applicability. *The Behavior Therapist*, 32(8), 168–175. (Cited on page 102.)

Boschen, M. J. (2009b). Mobile telephones and psychotherapy—II: A review of empirical research. *The Behavior Therapist*, 32(8), 175–181. (Cited on page 102.)

Boschen, M. J. & Casey, L. M. (2008). The use of mobile telephones as adjuncts to cognitive behavioral psychotherapy. *Professional Psychology: Research and Practice*, 39, 546–552. (Cited on pages 23 and 117.)

Boschen, M. J., Neumann, D. L., & Waters, A. M. (2009). Relapse of successfully treated anxiety and fear: Theoretical issues and recommendations for clinical practice. *The Australian and New Zealand Journal of Psychiatry*, 43(2), 89–100. (Cited on pages 17 and 100.)

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, 114(1), 80–99. (Cited on pages 109 and 112.)

Bouton, M. E. (1994). Conditioning, remembering, and forgetting. *Journal of Experimental Psychology: Animal Behavior Processes*, 20, 219–231. (Cited on pages 10, 19, 89, 99, 105, 107, and 112.)

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, 52(02), 976–986. (Cited on pages 9, 89, and 112.)

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, 11(5), 485–494. (Cited on pages 10, 82, and 89.)

Bouton, M. E., García-Gutiérrez, A., Zilski, J., & Moody, E. W. (2006). Extinction in multiple contexts does not necessarily make extinction less vulnerable to relapse. *Behaviour Research and Therapy*, 44(7), 983–994. (Cited on page 89.)

Bouton, M. E., Mineka, S., & Barlow, D. H. (2001). A modern learning theory perspective on the etiology of panic disorder. *Psychological Review*, 108(1), 4–32. (Cited on pages 8, 9, 10, 11, and 89.)

Bouton, M. E. & Swartzentruber, D. (1991). Sources of relapse after extinction in Pavlovian and instrumental learning. *Clinical Psychology Review*, 11(2), 123–140. (Cited on pages 10 and 19.)

Bouton, M. E., Woods, A. M., & Pineño, O. (2004). Occasional reinforced trials during extinction can slow the rate of rapid reacquisition. *Learning and Motivation*, 35, 371–390. (Cited on pages 84 and 108.)

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, 1, 276–298. (Cited on page 12.)

Brock, G., Pihur, V., Datta, S., & Datta, S. (2011). clValid: An R package for cluster validation. Version 0.6-4. (Cited on page 47.)

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193–217. (Cited on pages 3, 4, and 112.)

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Los Angeles: Univerisity of California Press. (Cited on pages 4 and 112.)

Buller, R., Maier, W., & Benkert, O. (1986). Clinical subtypes in panic disorder: Their descriptive and prospective validity. *Journal of Affective Disorders*, 11(2), 105–114. (Cited on page 64.)

Bystritsky, A., Craske, M. G., Maidenberg, E., Vapnik, T., & Shapiro, D. (1995). Ambulatory monitoring of panic patients during regular activity: A preliminary report. *Biological Psychiatry*, 38(10), 684–689. (Cited on page 64.)

Cacioppo, J., Berntson, G., & Andersen, B. (1991). Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome, 1991: Contributions from social psychophysiology. *Psychological Assessment*, 3(3), 321–336. (Cited on pages 2, 24, 37, and 38.)

Cacioppo, J., Tassinary, L., & Berntson, G. (2007). *Handbook of psychophysiology (3rd ed.)*. New York: Cambridge University Press. (Cited on pages 2 and 105.)

Cammin-Nowak, S., Helbig-Lang, S., Lang, T., Gloster, A. T., Fehm, L., Gerlach, A. L., Ströhle, A., Deckert, J., Kircher, T., Hamm, A. O., Alpers, G. W., Arolt, V., & Wittchen, H.-U. (2013). Specificity of homework compliance effects on treatment outcome in CBT: Evidence from a controlled trial on panic disorder and agoraphobia. *Journal of Clinical Psychology*, 69(6), 616–629. (Cited on pages 25, 66, 82, and 86.)

Ceja, L. & Navarro, J. (2012). 'Suddenly I get into the zone': Examining discontinuities and nonlinear changes in flow experiences at work. *Human Relations*, 65(9), 1101–1127. (Cited on page 13.)

Chambless, D. (1985). The relationship of severity of agoraphobia to associated psychopathology. *Behaviour Research and Therapy*, 23, 305–310. (Cited on pages 9 and 84.)

Chambless, D., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., Daiuto, A., Derubeis, R., Detweiler, J., Haaga, D. A., Johnson, S. B., McCurry, S., Mueser, K. T., Pope, K. S., Sanderson, W. C., Shoham, V., Stickle, T., Williams, D. A., & Woody, S. R. (1998). Update on empirically validated therapies, II. *The Clinical Psychologist*, 51, 3–16. (Cited on pages 2, 23, and 88.)

Chambless, D., Caputo, G., Bright, P., & Gallagher, R. (1984). Assessment of fear of fear in agoraphobics: The Body Sensations Questionnaire and the Agoraphobic Cognitions Questionnaire. *Journal of Consulting and Clinical Psychology*, 52, 1090–1097. (Cited on pages 13, 26, 68, 92, and 106.)

Chambless, D., Caputo, G., Jasin, S., Gracely, E., & Williams, C. (1985). The Mobility Inventory for agoraphobia. *Behaviour Research and Therapy*, 23, 35–44. (Cited on pages 13, 14, 15, 17, 23, 26, 68, 92, and 106.)

Chambless, D., Sharpless, B. A., Rodriguez, D., McCarthy, K. S., Milrod, B. L., Khalsa, S.-R., & Barber, J. P. (2011). Psychometric properties of the Mobility Inventory for agoraphobia: Convergent, discriminant, and criterion-related validity. *Behavior Therapy*, 42(4), 689–699. (Cited on pages 69 and 92.)

Chaplin, E. W. & Levine, B. A. (1981). The effects of total exposure duration and interrupted versus continuous exposure in flooding therapy. *Behavior Therapy*, 12(3), 360–368. (Cited on page 83.)

Clark, D. M. (1986). A cognitive approach to panic. *Behaviour Research and Therapy*, 24(4), 461–470. (Cited on pages 8, 9, 10, and 108.)

Clark, D. M. & Beck, A. (2010). *Cognitive therapy of anxiety disorders: Science and practice*. New York: Guilford Publications. (Cited on pages 10 and 64.)

Clatworthy, J., Buick, D., Hankins, M., Weinman, J., & Horne, R. (2005). The use and reporting of cluster analysis in health psychology: A review. *British Journal of Health Psychology*, 10, 329–358. (Cited on page 40.)

Cleveland, W. (1993). *Visualizing data*. Summit, New Jersey: Hobart Press. (Cited on page 116.)

Clough, B. A. & Casey, L. M. (2011). Technological adjuncts to enhance current psychotherapy practices: A review. *Clinical Psychology Review*, 31, 279–292. (Cited on pages 23, 117, and 119.)

Coan, J. & Allen, J., Eds. (2007). *Handbook of emotion elicitation and assessment*. Oxford, UK: Oxford University Press. (Cited on pages 4 and 113.)

Cohen, H., Benjamin, J., Geva, A., Matar, M., Kaplan, Z., & Kotler, M. (2000). Autonomic dysregulation in panic disorder and in post-traumatic stress disorder: Application of power spectrum analysis of heart rate variability at rest and in response to recollection of trauma or panic attacks. *Psychiatry Research*, 96(1), 1–13. (Cited on pages 14 and 40.)

Cox, B., Swinson, R. P., Endler, N. N. S., & Norton, G. R. (1994). The symptom structure of panic attacks. *Comprehensive Psychiatry*, 35(5), 349–353. (Cited on pages 14, 24, and 64.)

Craske, M. G. (2003). *Origins of phobias and anxiety disorders: Why more women than men?* Oxford, UK: Elsevier. (Cited on pages 57 and 65.)

Craske, M. G. & Barlow, D. H. (1988). A review of the relationship between panic and avoidance. *Clinical Psychology Review*, 8(6), 667–685. (Cited on pages 16 and 66.)

Craske, M. G., Kircanski, K., Epstein, A., Wittchen, H.-U., Pine, D. D. S., Lewis-Fernández, R., & Hinton, D. (2010). Panic disorder: A review of DSM-IV panic disorder and proposals for DSM-V. *Depression and Anxiety*, 27, 93–112. (Cited on page 8.)

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J. L., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, 46(1), 5–27. (Cited on pages 9, 10, 12, 37, 89, and 90.)

Craske, M. G., Rapee, R. M., & Barlow, D. H. (1988). The significance of panic-expectancy for individual patterns of avoidance. *Behavior Therapy*, 19, 577–592. (Cited on pages 66 and 107.)

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, 58, 10–23. (Cited on pages 2, 10, 19, 62, 66, 85, 100, 108, and 109.)

Cuthbert, B. N. & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Medicine*, 11, 126. (Cited on pages 13, 116, 117, and 119.)

Cuthbert, B. N., Lang, P. J., Strauss, C., Drobes, D., Patrick, C. J., & Bradley, M. M. (2003). The psychophysiology of anxiety disorder: Fear memory imagery. *Psychophysiology*, 40, 407–422. (Cited on pages 12 and 13.)

de Beurs, E., Chambless, D., & Goldstein, A. J. (2002). The match-mismatch model and panic patients' accuracy in predicting naturally occurring panic attacks. *Depression and Anxiety*, 16(4), 172–181. (Cited on page 35.)

de Beurs, E., Lange, A., van Dyck, R., & Koele, P. (1995). Respiratory training prior to exposure in vivo in the treatment of panic disorder with agoraphobia: Efficacy and predictors of outcome. *The Australian and New Zealand Journal of Psychiatry*, 29(1), 104–113. (Cited on page 10.)

Deacon, B. J. & Abramowitz, J. S. (2004). Cognitive and behavioral treatments for anxiety disorders: A review of meta-analytic findings. *Journal of Clinical Psychology*, 60, 429–441. (Cited on pages 23, 62, and 88.)

Deboeck, P. R. (2011). Modeling nonlinear dynamics in intraindividual variability. In M. Mehl & T. Conner (Eds.), *Handbook of research methods for studying daily life* chapter 24, (pp. 440–458). New York: Guilford Press. (Cited on page 107.)

Derogatis, L. (1993). *Brief Symptom Inventory (BSI) administration, scoring, and procedures manual*. Minneapolis: National Computer Systems, 3rd ed. edition. (Cited on page 26.)

Dew, S. E. & Bickman, L. (2005). Client expectancies about therapy. *Mental Health Services Research*, 7(1), 21–33. (Cited on page 114.)

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin*, 130(6), 959–88. (Cited on page 4.)

Domschke, K., Stevens, S., Pfleiderer, B., & Gerlach, A. L. (2010). Interoceptive sensitivity in anxiety and anxiety disorders: An overview and integration of neurobiological findings. *Clinical Psychology Review*, 30(1), 1–11. (Cited on page 100.)

Dunsmoor, J. E., Ahs, F., Zielinski, D. J., & Labar, K. S. (2014). Extinction in multiple virtual reality contexts diminishes fear reinstatement in humans. *Neurobiology of Learning and Memory*, 113(2014), 157–164. (Cited on page 89.)

Ebner-Priemer, U. W., Kubiak, T., & Pawlik, K. (2009). Ambulatory assessment. *European Psychologist*, 14, 95–97. (Cited on page 13.)

Ebner-Priemer, U. W. & Trull, T. J. (2009). Ecological momentary assessment of mood disorders and mood dysregulation. *Psychological Assessment*, 21, 463–475. (Cited on page 13.)

Elgethun, K., Fenske, R. A., Yost, M. G., & Palcisko, G. J. (2002). Time-location analysis for exposure assessment studies of children using a novel Global Positioning System instrument. *Environmental Health Perspectives*, 111, 115–122. (Cited on page 23.)

Eonta, A. M., Christon, L. M., Hourigan, S. E., Ravindran, N., Vrana, S. R., & Southam-Gerow, M. A. (2011). Using everyday technology to enhance evidence-based treatments. *Professional Psychology: Research and Practice*, 42, 513–520. (Cited on pages 23, 106, 117, and 119.)

Epstein, D. H., Tyburski, M., Craig, I. M., Phillips, K. A., Jobes, M. L., Vahabzadeh, M., Mezghanni, M., Lin, J.-L., Furr-Holden, C. D. M., & Preston, K. L. (2014). Real-time tracking of neighborhood surroundings and mood in urban drug misusers: Application of a new method to study behavior in its geographical context. *Drug and alcohol dependence*, 134, 22–29. (Cited on page 105.)

Fahrenberg, J. (1996). Ambulatory assessment: Issues and perspectives. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies.* (pp. 3–20). Ashland, OH US: Hogrefe & Huber Publishers. (Cited on pages 38, 61, and 119.)

Fahrenberg, J. (2006). *Assessment in daily life: A review of computer-assisted methodologies and applications in psychology and psychophysiology, years 2000-2005*. Technical report, University of Freiburg, Freiburg. (Cited on pages 107 and 119.)

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press. (Cited on page 65.)

Fiedler, K. & Juslin, P. (2006). *Information sampling and adaptive cognition*. Cambridge, UK: Cambridge University Press. (Cited on page 4.)

Finch, W. H. & French, B. F. (2013). Multilevel latent class analysis: Parametric and nonparametric models. *The Journal of Experimental Education*, 82(3), 307–333. (Cited on page 115.)

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1996). *Structured clinical interview for DSM-IV axis I disorders, clinician version (SCID-CV)*. Washington, DC: American Psychiatric Press, Inc. (Cited on page 26.)

Flynn, T. M., Taylor, P., & Pollard, C. A. (1992). Use of mobile phones in the behavioral treatment of driving phobias. *Journal of Behavior Therapy and Experimental Psychiatry*, 23(4), 299–302. (Cited on page 113.)

Foa, E. B. & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, 99(1), 20–35. (Cited on pages 11, 18, 40, 63, 67, 105, 107, and 108.)

Foa, E. B., Riggs, D., Massie, E., & Yarczower, M. (1995). The impact of fear activation and anger on the efficacy of exposure treatment for posttraumatic stress disorder. *Behavior Therapy*, 26(3), 487–499. (Cited on pages 18 and 63.)

Fraley, C. & Raftery, A. E. (2002). Model-Based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association*, 97(458), 611–631. (Cited on page 39.)

Fraley, C., Raftery, A. E., Murphy, T. B., & Scrucca, L. (2012). mclust version 4 for R: Normal mixture modeling for model-based clustering, classification, and density estimation. Technical Report No. 597. (Cited on page 46.)

Friedman, B. & Thayer, J. F. (1998a). Anxiety and autonomic flexibility: A cardiovascular approach. *Biological Psychology*, 47(3), 243–263. (Cited on pages 14, 40, and 64.)

Friedman, B. & Thayer, J. F. (1998b). Autonomic balance revisited: Panic anxiety and heart rate variability. *Journal of Psychosomatic Research*, 44(1), 133–151. (Cited on pages 14 and 64.)

Furmark, T., Tillfors, M., Stattin, H., Ekselius, L., & Fredrikson, M. (2000). Social phobia subtypes in the general population revealed by cluster analysis. *Psychological Medicine*, 30(6), 1335–1344. (Cited on page 39.)

Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software*, 40(4), 1–37. (Cited on page 46.)

Gloster, A. T., Klotsche, J., Gerlach, A. L., Hamm, A., Ströhle, A., Gauggel, S., Kircher, T., Alpers, G. W., Deckert, J., & Wittchen, H.-U. (2014). Timing matters: Change depends on the stage of treatment in cognitive behavioral therapy for panic disorder with agoraphobia. *Journal of Consulting and Clinical Psychology*, 82(1), 141–153. (Cited on pages 42, 68, and 93.)

Gloster, A. T., Wittchen, H.-U., Einsle, F., Höfler, M., Lang, T., Helbig-Lang, S., Fydrich, T., Fehm, L., Hamm, A. O., Richter, J., Alpers, G. W., Gerlach, A. L., Ströhle, A., Kircher, T., Deckert, J., Zwanzger, P., & Arolt, V. (2009). Mechanism of action in CBT (MAC): Methods of a multi-center randomized controlled trial in 369 patients with panic disorder and agoraphobia. *European Archives of Psychiatry and Clinical Neuroscience*, 259 Suppl, S155–166. (Cited on page 8.)

Gloster, A. T., Wittchen, H.-U., Einsle, F., Lang, T., Helbig-Lang, S., Fydrich, T., Fehm, L., Hamm, A. O., Richter, J., Alpers, G. W., Gerlach, A. L., Ströhle, A., Kircher, T., Deckert, J., Zwanzger, P., Höfler, M., & Arolt, V. (2011). Psychological treatment for panic disorder with agoraphobia: A randomized controlled trial to examine the role of therapist-guided exposure in situ in CBT. *Journal of Consulting and Clinical Psychology*, 79, 406–420. (Cited on pages 43, 58, 64, 83, and 107.)

Goldstein, A. & Chambless, D. (1978). A reanalysis of agoraphobia. *Behavior Therapy*, 9, 47–59. (Cited on pages 9 and 10.)

Goodwin, M. S., Intille, S. S., Albinali, F., & Velicer, W. F. (2011). Automated detection of stereotypical motor movements. *Journal of Autism and Developmental Disorders*, 41(6), 770–782. (Cited on page 118.)

Goodwin, R., Fergusson, D. M., & Horwood, L. J. (2004). Panic attacks and the risk of depression among young adults in the community. *Psychotherapy and Psychosomatics*, 73(3), 158–165. (Cited on page 8.)

Green, A. C., Hunt, C., & Stain, H. J. (2012). The delay between symptom onset and seeking professional treatment for anxiety and depressive disorders in a rural Australian sample. *Social Psychiatry and Psychiatric Epidemiology*, 47(9), 1475–1487. (Cited on page 100.)

Groves, P. M. & Thompson, R. F. (1970). Habituation: A dual-process theory. *Psychological Review*, 77, 419–450. (Cited on page 10.)

Gunther, L. M., Denniston, J. C., & Miller, R. R. (1998). Conducting exposure treatment in multiple contexts can prevent relapse. *Behaviour Research and Therapy*, 36(1), 75–91. (Cited on pages 89, 100, and 101.)

Gustafson, D. H., McTavish, F. M., Chih, M.-Y., Atwood, A. K., Johnson, R. A., Boyle, M. G., Levy, M. S., Driscoll, H., Chisholm, S. M., Dillenburg, L., Isham, A., & Shah, D. (2014). A smartphone application to support recovery from alcoholism: A randomized clinical trial. *JAMA Psychiatry*, 71(5), 566–72. (Cited on page 105.)

Guttman, I. (1973). Care and handling of univariate or multivariate outliners in detecting spuriosity: A Bayesian approach. *Technometrics*, 15, 723–738. (Cited on page 72.)

Hamaker, E. (2012). Why researchers should think "within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43–61). New York: The Guilford Press. (Cited on pages 5 and 115.)

Hammond, K. R. (1996). Upon reflection. *Thinking & Reasoning*, 2(2-3), 239–248. (Cited on page 4.)

Hammond, K. R. (2001). Introduction. In K. R. Hammond & T. R. Stewart (Eds.), *The essential Brunswik: Beginnings, explications, applications* (pp. 3–10). New York: Oxford University Press. (Cited on page 4.)

Hartmann, D. & Hall, R. (1976). The changing criterion design. *Journal of Applied Behavior Analysis*, 9(4), 527–532. (Cited on page 110.)

Hayes, S., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment. A functional approach to evaluating assessment quality. *The American Psychologist*, 42(11), 963–974. (Cited on page 116.)

Hayes, S., Wilson, K., Gifford, E., Follette, V., & Strosahl, K. (1996). Experiental avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology*, 64(6), 1152–1168. (Cited on pages 39 and 116.)

Haynes, S. N. & Yoshioka, D. T. (2007). Clinical assessment applications of ambulatory biosensors. *Psychological Assessment*, 19(1), 44–57. (Cited on page 114.)

Helbig-Lang, S., Lang, T., Petermann, F., & Hoyer, J. (2012). Anticipatory anxiety as a function of panic attacks and panic-related self-efficacy: An ambulatory assessment study in panic disorder. *Behavioural and Cognitive Psychotherapy*, 40(5), 590–604. (Cited on pages 5, 16, and 66.)

Hennig, C. (2014). fpc: Flexible procedures for clustering. Version 2.1-7. (Cited on page 48.)

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, 33(2-3), 61–83. (Cited on page 4.)

Ho, J. & Intille, S. S. (2005). Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In *Proceedings of the SIGCHI conference on human factors in computing systems - CHI '2005* (pp. 909–909). New York, New York, USA: ACM Press. (Cited on page 114.)

Hodgson, R. & Rachman, S. (1974). Desynchrony in measures of fear. *Behaviour Research and Therapy*, 12, 319–326. (Cited on pages 13, 40, 59, and 84.)

Hoffart, A., Sexton, H., Hedley, L. M., & Martinsen, E. W. (2008). Mechanisms of change in cognitive therapy for panic disorder with agoraphobia. *Journal of Behavior Therapy and Experimental Psychiatry*, 39, 262–275. (Cited on page 9.)

Hofmann, S., Alpers, G. W., & Pauli, P. (2009). Phenomenology of panic and phobic disorders. In M. Antony & M. B. Stein (Eds.), *Oxford handbook of anxiety and related disorders* (pp. 34–46). New York, NY: Oxford University Press. (Cited on page 23.)

Hofmann, S. & Smits, J. A. J. (2008). Cognitive-behavioral therapy for adult anxiety disorders. *The Journal of Clinical Psychiatry*, 69(4), 621–632. (Cited on page 62.)

Holmes, E. A., Craske, M. G., & Graybiel, A. M. (2014). Psychological treatments: A call for mental-health science. *Nature*, 511(7509), 287–289. (Cited on page 119.)

Huber, P. (1981). *Robust statistics*. New York: Wiley. (Cited on page 72.)

Insel, T. R. (2012). Rethinking mental illness: The view from 2022. In *42nd National Council Mental Health and Addictions Conference*. (Cited on pages 2 and 116.)

Intille, S. S. (2012). Emerging technology for studying daily life. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 267–282). New York, NY US: Guilford Press. (Cited on page 119.)

Janak, P. H., Bowers, M. S., & Corbit, L. H. (2012). Compound stimulus presentation and the norepinephrine reuptake inhibitor atomoxetine enhance long-term extinction of cocaine-seeking behavior. *Neuropsychopharmacology*, 37(4), 975–985. (Cited on pages 19 and 90.)

Jaycox, L. H., Foa, E. B., & Morral, A. R. (1998). Influence of emotional engagement and habituation on exposure therapy for PTSD. *Journal of Consulting and Clinical Psychology*, 66(1), 185–192. (Cited on page 64.)

Jennings, J. R., Kamarck, T. W., Stewart, C., Eddy, M., & Johnson, P. (1992). Alternate Cardiovascular Baseline Assessment Techniques: Vanilla or Resting Baseline. *Psychophysiology*, 29(6), 742–750. (Cited on page 115.)

Jun, J., Guensler, R., & Ogle, J. (2006). Smoothing methods to minimize impact of global positioning system random error on travel distance, speed, and acceleration profile estimates. *Transportation Research Record*, 1972, 141–150. (Cited on page 35.)

Kahle, D. & Wickham, H. A. (2013). ggmap: A package for spatial visualization with Google Maps and OpenStreetMap. (Cited on page 27.)

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. *Psychophysiology*, 40(5), 675–683. (Cited on pages 6 and 7.)

Kashdan, T. B., Barrios, V., Forsyth, J. P., & Steger, M. F. (2006). Experiential avoidance as a generalized psychological vulnerability: Comparisons with coping and emotion regulation strategies. *Behaviour Research and Therapy*, 44(9), 1301–1320. (Cited on page 86.)

Kashdan, T. B., Farmer, A. S., Adams, L. M., Ferssizidis, P., McKnight, P. E., & Nezlek, J. B. (2013). Distinguishing healthy adults from people with social anxiety disorder: Evidence for the value of experiential avoidance and positive emotions in everyday social interactions. *Journal of Abnormal Psychology*, 122(3), 645–655. (Cited on page 86.)

Katkin, E. (1985). Blood, sweat, and tears: Individual differences in autonomic self-perception. *Psychophysiology*, 22(2), 125–137. (Cited on page 57.)

Kaufman, L. & Rousseeuw, P. (1990). *Finding groups in data: An introduction to cluster analysis*. Hoboken, New Jersey: John Wiley & Sons, Inc. (Cited on page 39.)

Kazantzis, N., Zealand, N., & Ronan, K. R. (2000). Homework assignments in cognitive and behavioral therapy : A meta-analysis. *Clinical Psychology: Science and Practice*, 7, 189–202. (Cited on pages 66 and 114.)

Kenardy, J., Oei, T. P., Weir, D., & Evans, L. (1993). Phobic anxiety in panic disorder: Cognition, heart rate, and subjective anxiety. *Journal of Anxiety Disorders*, 7, 359–371. (Cited on page 24.)

Kerr, J., Duncan, S., & Schipperjin, J. (2011). Using global positioning systems in health research: A practical approach to data collection and processing. *American Journal of Preventive Medicine*, 41, 532–540. (Cited on pages 24, 35, and 102.)

Kessler, R. C., Chiu, W. T., Jin, R., Ruscio, A. M., Shear, K., & Walters, E. E. (2006). The epidemiology of panic attacks, panic disorder, and agoraphobia in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 63(4), 415–24. (Cited on page 8.)

Kircanski, K., Craske, M. G., Epstein, A., & Wittchen, H.-U. (2009). Subtypes of panic attacks: A critical review of the empirical literature. *Depression and Anxiety*, 26(10), 878–887. (Cited on pages 17, 38, 40, and 107.)

Kirchner, T. R., Cantrell, J., Anesetti-Rothermel, A., Ganz, O., Vallone, D. M., & Abrams, D. B. (2013). Geospatial exposure to point-of-sale tobacco: Real-time craving and smoking-cessation outcomes. *American Journal of Preventive Medicine*, 45, 379–385. (Cited on page 105.)

Kluger, A. N. & Denisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. (Cited on page 118.)

Kozak, M. J., Foa, E. B., Steketee, G., & McNally, R. (1988). Process and outcome of exposure treatment with obsessive-compulsives: Psychophysiological indicators of emotional processing. *Behavior Therapy*, 19, 157–169. (Cited on pages 12, 18, 63, and 85.)

Kristjansson, S. D., Kircher, J. C., & Webb, A. K. (2007). Multilevel models for repeated measures research designs in psychophysiology: An introduction to growth curve modeling. *Psychophysiology*, 44(5), 728–736. (Cited on pages 73 and 94.)

Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological science*, 21(7), 984–991. (Cited on page 115.)

Lacey, J. I. (1956). The evaluation of autonomic responses: Toward a general solution. *Annals of the New York Academy of Sciences*, 67, 127–163. (Cited on pages 78 and 84.)

Lacey, J. I. & Lacey, B. C. (1962). The law of initial values in the longitudinal study of autonomic constitution: Reproducibility of autonomic responses and response patterns over a four year interval. *Annals New York Academy of Sciences*, 98, 1257–1290. (Cited on pages 78 and 84.)

Lang, A. & Craske, M. G. (2000). Manipulations of exposure-based therapy to reduce return of fear: A replication. *Behaviour Research and Therapy*, 38(1), 1–12. (Cited on pages 12, 19, 89, 90, 100, and 101.)

Lang, P. J. (1968). Fear reduction and fear behavior: Problems in treating a construct. In J. Schlien (Ed.), *Research in psychotherapy* (pp. 90–102). Washington, DC: APA. (Cited on page 10.)

Lang, P. J. (1971). The application of psychophysiological methods to the study of psychotherapy and behavior modification. *Handbook of psychotherapy and behavior change*, (pp. 75–125). (Cited on pages 11, 13, 63, and 84.)

Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology*, 16, 495–512. (Cited on pages 13 and 84.)

Lang, P. J. (1993). The three system approach to emotion. In N. Birbaumer & A. Öhman (Eds.), *The organization of emotion* (pp. 18–30). Toronto: Hogrefe-Huber. (Cited on pages 13 and 84.)

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. In P. J. Lang, R. F. Simons, & M. T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes.* (pp. 97–135). Florida, Gainesville: Lawrence Erlbaum Associates Publishers. (Cited on page 12.)

Lang, P. J. & McTeague, L. M. (2009). The anxiety disorder spectrum: Fear imagery, physiological reactivity, and differential diagnosis. *Anxiety, Stress, and Coping*, 22, 5–25. (Cited on pages 12 and 13.)

Lang, P. J., Melamed, B. G., & Hart, J. (1970). A psychophysiological analysis of fear modification using an automated desensitization procedure. *Journal of Abnormal Psychology*, 76(2), 220–234. (Cited on page 63.)

Lang, T., Helbig-Lang, S., Gloster, A. T., Richter, J., Hamm, A. O., Fehm, L., Fydrich, T., Gerlach, A. L., Ströhle, A., Alpers, G. W., Gauggel, S., Kircher, T., Deckert, J., Höfler, M., Arolt, V., & Wittchen, H.-U. (2012). Effekte therapeutenbegleiteter versus patientengeleiteter Exposition bei Panikstörung mit Agoraphobie. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 41(2), 114–124. (Cited on page 64.)

Lederbogen, F., Kirsch, P., Haddad, L., Streit, F., Tost, H., Schuch, P., Wüst, S., Pruessner, J. C., Rietschel, M., Deuschle, M., & Meyer-Lindenberg, A. (2011). City living and urban upbringing affect neural social stress processing in humans. *Nature*, 474(7352), 498–501. (Cited on page 101.)

LeDoux, J. E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences of the United States of America*, 111(8), 2871–2878. (Cited on page 117.)

Lewis, L. E. & Drewett, R. F. (2006). Psychophysiological correlates of anxiety: A single-case study. *Journal of Anxiety Disorders*, 20(6), 829–835. (Cited on pages 40 and 59.)

Leyro, T. M., Zvolensky, M. J., & Bernstein, A. (2011). Distress tolerance and psychopathological symptoms and disorders: A review of the empirical literature among adults. *Psychological Bulletin*, 136(4), 576–600. (Cited on page 41.)

Liddell, A., Bilsbury, C., & Rattenbury, C. (1987). Concordance and discordance of cognitive, behavioural and somatic self-ratings as a function of exposure: A Discan analysis. *Behaviour Research and Therapy*, 25, 425–428. (Cited on page 13.)

Lipp, O. V. (2006). Human Fear Learning: Contemporary Procedures and Measurement. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical implications*, number 2001 (pp. 37–51). Washington, DC: American Psychological Association. (Cited on page 11.)

Loecher, M. (2014). RgoogleMaps. (Cited on pages 94 and 122.)

Lovibond, P. F., Montpetit, M. A., Minard, E., Brady, A., & Menzies, R. G. (2009). Safety behaviours preserve threat beliefs: Protection from extinction of human fear conditioning by an avoidance response. *Behaviour Research and Therapy*, 47, 716–720. (Cited on page 23.)

Maddison, R. & Ni Mhurchu, C. (2009). Global positioning system: A new opportunity in physical activity measurement. *The International Journal of Behavioral Nutrition and Physical Activity*, 6, 73. (Cited on page 23.)

Maheu, M. M., Pulier, M. L., McMenamin, J. P., & Posen, L. (2012). Future of telepsychology, telehealth, and various technologies in psychological research and practice. *Professional Psychology: Research and Practice*, 43(6), 613–621. (Cited on pages 118 and 119.)

Margraf, J. (1990). Ambulatory psychophysiological monitoring of panic attacks. *Journal of Psychophysiology*, 4(4), 321–330. (Cited on page 14.)

Margraf, J., Taylor, C., Ehlers, A., Roth, W. T., & Agras, W. (1987). Panic attacks in the natural environment. *Journal of Nervous and Mental Disease*, 175(9), 558–565. (Cited on page 14.)

Marks, I. (1987). *Fears, phobias, and rituals: Panic, anxiety, and their disorders*. New York: Oxford University Press. (Cited on page 10.)

Marshall, W. (1985). The effects of variable exposure in flooding therapy. *Behavior Therapy*, 16(2), 117–135. (Cited on pages 65 and 83.)

Mausbach, B. T., Moore, R., Roesch, S., Cardenas, V., & Patterson, T. L. (2010). The relationship between homework compliance and therapy outcomes: An updated meta-analysis. *Cognitive Therapy and Research*, 34, 429–438. (Cited on pages 25 and 33.)

Mavissakalian, M. (1988). The relationship between panic, phobic and anticipatory anxiety in agoraphobia. *Behaviour Research and Therapy*, 26(3), 235–240. (Cited on page 15.)

McCambridge, J., de Bruin, M., & Witton, J. (2012). The effects of demand characteristics on research participant behaviours in non-laboratory settings: A systematic review. *PloS one*, 7(6), e39116. (Cited on page 111.)

McCoach, D. B. & Kaniskan, B. (2010). Using time-varying covariates in multilevel growth models. *Frontiers in Psychology*, 1(June), 17. (Cited on pages 73 and 94.)

McNally, R. (2002). Anxiety sensitivity and panic disorder. *Biological Psychiatry*, 52(10), 938–946. (Cited on page 9.)

Mehl, M. R. & Conner, T. S., Eds. (2012). *Handbook of research methods for studying daily life*. New York, NY US: Guilford Press. (Cited on page 3.)

Meltzoff, J. (1998). *Critical thinking about research - psychology and related fields*. Washington, DC, US: American Psychological Association. (Cited on page 110.)

Meuret, A. E., Seidel, A., Rosenfield, B., Hofmann, S., & Rosenfield, D. (2012). Does fear reactivity during exposure predict panic symptom reduction? *Journal of Consulting and Clinical Psychology*, 80, 773–785. (Cited on pages 6, 10, 12, 18, 24, 37, 40, 63, 64, 70, 85, and 108.)

Meuret, A. E., White, K. S., Ritz, T., Roth, W. T., Hofmann, S., & Brown, T. A. (2006). Panic attack symptom dimensions and their relationship to illness characteristics in panic disorder. *Journal of Psychiatric Research*, 40, 520–527. (Cited on pages 7 and 64.)

Michelson, L. (1986). Treatment consonance and response profiles in agoraphobia: The role of individual differences in cognitive, behavioral and physiological treatments. *Behaviour Research and Therapy*, 24(3), 263–275. (Cited on page 38.)

Milad, M. R. & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: Ten years of progress. *Annual Review of Psychology*, 63, 129–151. (Cited on page 9.)

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitive Methods for Psychology*, 4(2), 61–64. (Cited on pages 74 and 95.)

Morris, M. & Guilak, F. (2009). Mobile heart health: Project highlight. *IEEE Pervasive Computing*, 8(2), 57–61. (Cited on page 118.)

Morris, M. E. & Aguilera, A. (2012). Mobile, social, and wearable computing and the evolution of psychological practice. *Professional Psychology: Research and Practice*, 43(6), 622–626. (Cited on pages 23, 118, and 119.)

National Institute for Clinical Excellence (2011). *Generalised anxiety disorder and panic disorder (with or without agoraphobia) in adults: Management in primary, secondary and community care*. Technical report, London. (Cited on page 7.)

Neumann, D. L. (2006). The effects of physical context changes and multiple extinction contexts on two forms of renewal in a conditioned suppression task with humans. *Learning and Motivation*, 37(2), 149–175. (Cited on pages 19, 89, and 100.)

Neumann, D. L., Lipp, O. V., & Cory, S. E. (2007). Conducting extinction in multiple contexts does not necessarily attenuate the renewal of shock expectancy in a fear-conditioning procedure with humans. *Behaviour Research and Therapy*, 45(2), 385–394. (Cited on pages 89 and 112.)

Newman, M. G. (2004). Technology in psychotherapy: An introduction. *Journal of Clinical Psychology*, 60(2), 141–145. (Cited on page 117.)

Ng, S. & McLachlan, G. (2014). Mixture models for clustering multilevel growth trajectories. *Computational Statistics & Data Analysis*, 71, 43–51. (Cited on page 115.)

Ning, L. & Liddell, A. (1991). The effect of concordance in the treatment of clients with dental anxiety. *Behaviour Research and Therapy*, 29(4), 315–322. (Cited on page 13.)

Norton, P. J. & Price, E. C. (2007). A meta-analytic review of adult cognitive-behavioral treatment outcome across the anxiety disorders. *The Journal of Nervous and Mental Disease*, 195(6), 521–531. (Cited on pages 62 and 88.)

Obrist, P. (1981). *Cardiovascular psychophysiology: A perspective.* New York: Plenum Press. (Cited on page 12.)

Odom, J. V. & Nelson, R. O. (1977). Effect of demand characteristics on heart rate during a behavioral avoidance test. *Perceptual and Motor Skills*, 44, 175–183. (Cited on page 114.)

O'Donohue, W. & Fisher, J. E. (2012). *Cognitive behavior therapy: Core principles for practice*. Hoboken, New Jersey: John Wiley & Sons. (Cited on pages 10 and 64.)

Pauli, P., Marquardt, C., Hartl, L., Nutzinger, D., Hötzl, R., & Strian, F. (1991). Anxiety induced by cardiac perceptions in patients with panic attacks: A field study. *Behavior Research and Therapy*, 29(2), 137–145. (Cited on page 65.)

Perna, G., Bertani, A., Politi, E., Colombo, G., & Bellodi, L. (1997). Asthma and panic attacks. *Biological Psychiatry*, 42(7), 625–630. (Cited on page 8.)

Peter, H., Brückner, E., Hand, I., Rohr, W., & Rufer, M. (2008). Treatment outcome of female agoraphobics 3-9 years after exposure in vivo: A comparison with healthy controls. *Journal of Behavior Therapy and Experimental Psychiatry*, 39(1), 3–10. (Cited on pages 2, 17, 63, 88, 100, and 101.)

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Development Core Team (2013). nlme: Linear and Nonlinear Mixed Effects Models. Version 3.1-111. (Cited on pages 45, 74, and 95.)

Pittig, A., Schulz, A. R., Craske, M. G., & Alpers, G. W. (2014). Acquisition of behavioral avoidance: Task-irrelevant conditioned stimuli trigger costly decisions. *Journal of Abnormal Psychology*, 123(2), 314–329. (Cited on pages 23 and 86.)

Porter, M. & Whitton, M. J. (2002). Assessment of driving with the Global Positioning System and video technology in young, middle-aged, and older drivers. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 57, M578–M582. (Cited on page 23.)

Priebe, K., Kleindienst, N., Zimmer, J., Koudela, S., Ebner-Priemer, U. W., & Bohus, M. (2013). Frequency of intrusions and flashbacks in patients with posttraumatic stress disorder related to childhood sexual abuse: An electronic diary study. *Psychological Assessment*, 25(4), 1370–1376. (Cited on page 13.)

Puri, B. & Treasaden, I. (2011). *Textbook of psychiatry: Third edition*. London: Elsevier. (Cited on pages 10 and 64.)

R Development Core Team (2013). R: A Language and Environment for Statistical Computing. (Cited on pages 27, 74, and 95.)

Rabavilas, A., Boulougouris, J., & Stefanis, C. (1976). Duration of flooding sessions in the treatment of obsessive-compulsive patients. *Behaviour Research and Therapy*, 14(5), 349–355. (Cited on page 83.)

Rachman, S. (1984a). Agoraphobia—A safety-signal perspective. *Behaviour Research and Therapy*, 22, 59–70. (Cited on page 110.)

Rachman, S. (1984b). The experimental analysis of agoraphobia. *Behaviour Research and Therapy*, 22, 631–640. (Cited on page 18.)

Rachman, S. (2009). Psychological treatment of anxiety: The evolution of behavior therapy and cognitive behavior therapy. *Annual Review of Clinical Psychology*, 5, 97–119. (Cited on pages 2 and 10.)

Rachman, S. & Hodgson, R. (1974). Synchrony and desynchrony in fear and avoidance. *Behaviour Research and Therapy*, 12(4), 311–318. (Cited on pages 18 and 84.)

Rachman, S. & Lopatka, C. (1986a). Match and mismatch in the prediction of fear - I. *Behaviour Research and Therapy*, 24(4), 387–393. (Cited on pages 16 and 66.)

Rachman, S. & Lopatka, C. (1986b). Match and mismatch of fear in Gray's theory - II. *Behavior Research and Therapy*, 24(4), 395–401. (Cited on page 66.)

Rapee, R. M. & Murrell, E. (1988). Predictors of agoraphobic avoidance. *Journal of Anxiety Disorders*, 2, 203–217. (Cited on page 84.)

Reis, H. T. (2012). Why researchers should think 'real-world': A conceptual rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 3–21). New York: The Guilford Press. (Cited on pages 4, 5, and 109.)

Reiss, S. (1987). Theoretical perspectives on the fear of anxiety. *Clinical Psychology Review*, 7(6), 585–596. (Cited on page 9.)

Reiss, S., Peterson, R. a., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour Research and Therapy*, 24(1), 1–8. (Cited on page 9.)

Rescorla, R. A. (2006). Deepened extinction from compound stimulus presentation. *Journal of Experimental Psychology. Animal Behavior Processes*, 32(2), 135–144. (Cited on pages 9, 12, 19, 90, 99, and 100.)

Rescorla, R. A. & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. Black & W. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Croft. (Cited on pages 84 and 108.)

Richter, J., Hamm, A. O., Pané-Farré, C. C. A., Gerlach, A. L., Gloster, A. T., Wittchen, H.-U., Lang, T., Alpers, G. W., Helbig-Lang, S., Deckert, J., Fydrich, T., Fehm, L., Ströhle, A., Kircher, T., & Arolt, V. (2012). Dynamics of defensive reactivity in patients with panic disorder and agoraphobia: Implications for the etiology of panic disorder. *Biological Psychiatry*, 72(6), 512–520. (Cited on pages 8 and 83.)

Roberson-Nay, R. & Kendler, K. S. (2011). Panic disorder and its subtypes: A comprehensive analysis of panic symptom heterogeneity using epidemiological and treatment seeking samples. *Psychological Medicine*, 41(11), 2411–2421. (Cited on pages 17 and 38.)

Roberson-Nay, R., Latendresse, S. J., & Kendler, K. S. (2012). A latent class approach to the external validation of respiratory and non-respiratory panic subtypes. *Psychological Medicine*, 42(3), 461–474. (Cited on pages 38, 64, and 111.)

Robinson, S., Perkins, S., Bauer, S., Hammond, N., Treasure, J., & Schmidt, U. (2006). Aftercare intervention through text messaging in the treatment of bulimia nervosa–feasibility pilot. *The International Journal of Eating Disorders*, 39(8), 633–638. (Cited on page 118.)

Roth, W. T., Telch, M., & Taylor, C. (1986). Autonomic characteristics of agoraphobia with panic attacks. *Biological Psychiatry*, 21(12), 1133–1154. (Cited on pages 14 and 40.)

Rowe, M. K. & Craske, M. G. (1998). Effects of varied-stimulus exposure training on fear reduction and return of fear. *Behavior Research and Therapy*, 36, 719–734. (Cited on pages 12, 19, and 101.)

Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). *Akaike information criterion statistics*. Dordrecht, Holland: Reidel. (Cited on page 46.)

Salkovskis, P. M., Hackmann, A., Wells, A., Gelder, M. G., & Clark, D. M. (2007). Belief disconfirmation versus habituation approaches to situational exposure in panic disorder with agoraphobia: A pilot study. *Behaviour Research and Therapy*, 45(5), 877–885. (Cited on pages 9 and 110.)

Santangelo, P. S., Ebner-Priemer, U. W., & Trull, T. J. (2013). Experience sampling methods in clinical psychology. In J. S. Comer & P. C. Kendall (Eds.), *The Oxford handbook of research strategies for clinical psychology* (pp. 188–210). New York, NY US: Oxford University Press. (Cited on pages 6, 40, and 113.)

Saul, J., Rea, R. F., Eckberg, D. L., Berger, R. D., & Cohen, R. J. (1990). Heart rate and muscle sympathetic nerve variability during reflex changes of autonomic activity. *The American Journal of Physiology*, 258, H713–H721. (Cited on page 14.)

Schmidt, N., Richey, J. A., Cromer, K. R., & Buckner, J. D. (2007). Discomfort intolerance: Evaluation of a potential risk factor for anxiety psychopathology. *Behavior Therapy*, 38(3), 247–255. (Cited on page 59.)

Schmidt, R. A. & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, 3(4), 207–217. (Cited on page 89.)

Shadish, W. R., Cook, T. D., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin. (Cited on pages 5, 102, 105, 106, 109, and 110.)

Shapiro, D. A. (1995). Finding out how psyhotherapies help people change. *Psychotherapy Research*, 5(June 1994), 1–21. (Cited on page 117.)

Shear, M., Polan, J., Harshfield, G., Pickering, T., Mann, J., Frances, A., & James, G. (1992). Ambulatory monitoring of blood pressure and heart rate in panic patients. *Journal of Anxiety Disorders*, 6(3), 213–221. (Cited on page 14.)

Shiban, Y., Pauli, P., & Mühlberger, A. (2013). Effect of multiple context exposure on renewal in spider phobia. *Behaviour Research and Therapy*, 51(2), 68–74. (Cited on pages 19, 89, 102, and 112.)

Shin, L. M. & Liberzon, I. (2010). The neurocircuitry of fear, stress, and anxiety disorders. *Neuropsychopharmacology*, 35(1), 169–91. (Cited on pages 11, 116, and 117.)

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, 25(3), 289–310. (Cited on page 111.)

Shockley, K. M. & Allen, T. D. (2013). Episodic work-family conflict, cardiovascular indicators, and social support: An experience sampling approach. *Journal of Occupational Health Psychology*, 18(1939-1307 (Electronic)), 262–275. (Cited on page 13.)

Shultz, T. R. & Lepper, M. R. (1996). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review*, 103(2), 219–240. (Cited on pages 65 and 83.)

Sievert, L. L. (2013). Subjective and objective measures of hot flashes. *American Journal of Human Biology*, 25(1520-6300 (Electronic)), 573–580. (Cited on page 13.)

Singer, J. & Willett, J. (2003). *Applied longitudinal data analysis: Modelling change and event occurrence*. New York: Oxford University Press. (Cited on pages 73, 74, 86, 94, and 114.)

Speltz, M. L. & Bernstein, D. A. (1976). Sex differences in fearfulness: Verbal report, overt avoidance and demand characteristics. *Journal of Behavior Therapy and Experimental Psychiatry*, 7, 117–122. (Cited on pages 15, 90, 111, and 114.)

Stemmler, G. (1996). Strategies and designs in ambulatory assessment. In J. Fahrenberg (Ed.), *Ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 257–271). Seattle, WA: Hogrefe & Huber. (Cited on pages 18 and 39.)

Stone, A., Broderick, J., Schwartz, J., Shiffman, S., Litcher-Kelly, L., & Calvanese, P. (2003). Intensive momentary reporting of pain with an electronic diary: Reactivity, compliance, and patient satisfaction. *Pain*, 104, 343–351. (Cited on page 6.)

Sullivan, G. M., Kent, J. M., Kleber, M., Martinez, J. M., Yeragani, V. K., & Gorman, J. M. (2004). Effects of hyperventilation on heart rate and QT variability in panic disorder pre- and post-treatment. *Psychiatry Research*, 125(1), 29–39. (Cited on pages 38 and 111.)

Sung, M., Marci, C., & Pentland, A. S. (2005). Objective physiological and behavioral measures for identifying and tracking depression state in clinically depressed patients (Technical Report 595). (Cited on page 118.)

Tabachnick, B. & Fidell, L. (2007). *Using multivariate statistics.* Boston, MA: Pearson Education, Inc. (Cited on pages 72 and 73.)

Taylor, C., Sheikh, J., Agras, W., Roth, W. T., Margraf, J., Ehlers, A., Maddock, R., & Gossard, D. (1986). Ambulatory heart rate in patients with panic attacks. *American Journal of Psychiatry*, 143(4), 478–482. (Cited on page 14.)

Telch, M., Brouillard M., Telch, C., Agras, W., & Taylor, C. B. (1989). Role of cognitive appraisal in panic-related avoidance. *Behavior Research and Therapy*, 27, 373–383. (Cited on page 28.)

Telch, M., Valentiner, D. P., Ilai, D., Petruzzi, D., & Hehmsoth, M. (2000). The facilitative effects of heart-rate feedback in the emotional processing of claustrophobic fear. *Behaviour Research and Therapy*, 38(4), 373–387. (Cited on page 12.)

Thayer, J. F., Friedman, B., & Borkovec, T. D. (1996). Autonomic characteristics of generalized anxiety disorder and worry. *Biological Psychiatry*, 39(4), 255–266. (Cited on page 14.)

Thayer, J. F. & Lane, R. D. (2009). Claude Bernard and the heart-brain connection: further elaboration of a model of neurovisceral integration. *Neuroscience and Biobehavioral Reviews*, 33(2), 81–88. (Cited on page 14.)

Trudel, G. (1979). The effects of instructions, level of fear, duration of exposure and repeated measures on the Behavioral Avoidance Test. *Behaviour Research and Therapy*, 17, 113–118. (Cited on pages 15, 65, 83, and 90.)

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. *Psychological Assessment*, 21, 457–462. (Cited on pages 13 and 23.)

Tryon, W. W. (2005). Possible mechanisms for why desensitization and exposure therapy work. *Clinical Psychology Review*, 25(1), 67–95. (Cited on pages 65, 83, 84, 86, and 108.)

Vermunt, J. K. (2003). Multilevel latent class models. *Sociological Methodology*, 33, 213–239. (Cited on page 115.)

Vermunt, J. K. & Magidson, J. (2002). Latent class cluster analysis. In J. Hagenaars & A. McCutcheon (Eds.), *Applied latent class analysis* (pp. 89–106). Cambridge: Cambridge University Press. (Cited on pages 18 and 39.)

Vervliet, B., Baeyens, F., Van den Bergh, O., & Hermans, D. (2012). Extinction, generalization, and return of fear: A critical review of renewal research in humans. *Biological Psychology*, 92, 51–58. (Cited on pages 11 and 89.)

Vervliet, B. & Raes, F. (2013). Criteria of validity in experimental psychopathology: Application to models of anxiety and depression. *Psychological Medicine*, 43(11), 2241–2244. (Cited on page 89.)

Vila, J., Guerra, P., Muñoz, M. A., Vico, C., Viedma-del Jesús, M. I., Delgado, L. C., Perakakis, P., Kley, E., Mata, J. L., & Rodríguez, S. (2007). Cardiac defense: From attention to action. *International Journal of Psychophysiology*, 66(3), 169–182. (Cited on page 14.)

West, B. T., Welch, K. B., Ga, A. T., & Crc, H. (2007). *Linear mixed models: A practical guide using statistical software*. Boca Raton, FL: Chapman & Hall. (Cited on pages 74 and 114.)

Westra, H. A., Dozois, D. J. A., & Marcus, M. (2007). Expectancy, homework compliance, and initial change in cognitive-behavioral therapy for anxiety. *Journal of Consulting and Clinical Psychology*, 75(3), 363–373. (Cited on pages 66 and 114.)

White, A. J., Neben, T., von der Trenck, A., Heinzl, A., & Alpers, G. W. (2013). Personalized tracking of goals and gains after psychotherapy using behavioral data. In *Personal Informatics in the Wild: Hacking Habits for Health & Happiness—CHI 2013 Workshop* Paris, France. (Cited on pages 101 and 118.)

White, A. J., Umpfenbach, K., & Alpers, G. W. (2014). Where have they gone? Tracking movement patterns to document the process of situational exposure in agoraphobia. *Professional Psychology: Research and Practice*, 45(3), 171–179. (Cited on pages 3, 82, 91, and 102.)

Whittal, M. & Goetsch, V. (1997). The impact of panic expectancy and social demand on agoraphobic avoidance. *Behaviour Research and Therapy*, 35(9), 813–821. (Cited on page 66.)

Whittal, M. L., Goetsch, V. L., & Eifert, G. H. (1996). Introduction of a dynamic, idiographic model for identifying panic. *Journal of Anxiety Disorders*, 10(2), 129–144. (Cited on pages 17 and 101.)

Wiehe, S. E., Hoch, S. C., Liu, G. C., Carroll, A. E., Wilson, J. S., & Fortenberry, J. D. (2008). Adolescent travel patterns: Pilot data indicating distance from home varies by time of day and day of week. *The Journal of Adolescent Health*, 42, 418–420. (Cited on page 23.)

Wilhelm, F. H., Alpers, G. W., Meuret, A. E., & Roth, W. T. (2001). Respiratory pathophysiology of clinical anxiety outside the laboratory: Assessment of end-tidal pCO2, respiratory pattern variability, and transfer function RSA. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment* (pp. 313–343). Göttingen: Hogrefe. (Cited on page 33.)

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. (Cited on pages 3, 5, 7, and 60.)

Wilhelm, F. H., Grossman, P., & Müller, M. (2012a). Bridging the gap between the laboratory and the real world: Integrative ambulatory

psychophysiology. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 210–234). New York: The Guilford Press. (Cited on pages 3, 5, and 113.)

Wilhelm, F. H. & Roth, W. T. (1998). Taking the laboratory to the skies: Ambulatory assessment of self-report, autonomic, and respiratory responses in flying phobia. *Psychophysiology*, 35(5), 596–606. (Cited on pages 6, 7, and 23.)

Wilhelm, P., Perrez, M., & Pawlik, K. (2012b). Conducting research in daily life: A historical review. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 62–86). New York: The Guilford Press. (Cited on pages 3, 5, and 6.)

Wittchen, H.-U., Gloster, A. T., Beesdo-Baum, K., Fava, G., & Craske, M. G. (2010). Agoraphobia: A review of the diagnostic classificatory position and criteria. *Depression and Anxiety*, 27, 113–133. (Cited on pages 7, 8, 18, and 61.)

Wittmann, W. W. (2012). Principles of symmetry in evaluation research with implications for offender treatment. In T. Bliesener, A. Beelmann, & M. Stemmler (Eds.), *Antisocial behavior and crime. Contributions of developmental and evaluation research to prevention and intervention*, number 2011 (pp. 357–368). Cambridge: Hogrefe. (Cited on pages 5 and 112.)

Wittmann, W. W. & Klumb, P. (2006). How to fool yourself with experiments in testing theories in psychological research. In R. R. Bootzin & P. E. McKnight (Eds.), *Strengthening research methodology: Psychological measurement and evaluation* (pp. 185–211). Washington, DC, US: American Psychological Association. (Cited on pages 5, 111, and 112.)

Wolff, J. J. & Symons, F. J. (2013). An evaluation of multi-component exposure treatment of needle phobia in an adult with autism and intellectual disability. *Journal of Applied Research in Intellectual Disabilities*, 26(4), 344–348. (Cited on page 110.)

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press. (Cited on page 10.)

Woods, A. M. & Bouton, M. E. (2007). Occasional reinforced responses during extinction can slow the rate of reacquisition of an operant response. *Learning and Motivation*, 38, 56–74. (Cited on pages 84 and 108.)

Zuzanek, J. (2013). Does being well-off make us happier? Problems of measurement. *Journal of Happiness Studies*, 14(3), 795–815. (Cited on page 13.)

EIDESSTATTLICHE VERSICHERUNG

Eidesstattliche Versicherung gemäß § 9 Absatz 1 Buchstabe e) der Promotionsordnung der Universität Mannheim zur Erlangung des Doktorgrades der Sozialwissenschaften:

1. Bei der eingereichten Dissertation mit dem Titel, "Methodological Approaches to the Ambulatory Assessment of Anxiety During Situational Exposure", handelt es sich um mein eigenständig erstelltes Werk.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtliche Zitate aus anderen Werken als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bisher nicht an einer Hochschule des In- oder Auslandes als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärung bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Mannheim, July 2014 Andrew J. White