Ecological Momentary Assessment applied in Patient-Focused Psychotherapy Research
- Challenges and Chances

Inauguraldissertation zur Erlangung der Doktorwürde (Dr. rer. Nat.)
im Fach Psychologie, Fachbereich I an der Universität Trier

Vorgelegt im
März 2016

Von
Kristin Husen

Gutachter
Prof. Dr. Wolfgang Lutz
Prof. Dr. Martin grosse Holtforth
DANKSAGUNG

Mein besonderer Dank gilt meinem Chef Prof. Wolfgang Lutz für die Gelegenheit, diese Arbeit zu planen und umzusetzen und die vielfältigen Möglichkeiten für ein sehr interessantes und herausforderndes Arbeiten über die letzten Jahre.

Weiterhin danke ich Prof. Martin grosse Holtforth für seine spontane Anpassung an lebensweltliche Ereignisse, die Übernahme der Zweitbetreuung und seine literarischen Empfehlungen.

A special thanks goes to Prof. Eshkol Rafaeli for his valuable and cordial support with expert knowledge on EMA applications throughout the entire process of this thesis.

Ich danke Kaitlyn Boyle für den sprachlichen Feinschliff dieser Arbeit.


Meiner Familie danke ich für Ihr Interesse und Ihr Vertrauen in meine Fähigkeiten und in mein Glück.

Ganz besonders danke ich meinem Mann Onno, für das Lesen dieser Arbeit, der Eröffnung neuer Blickwinkel auf dem gesamten Entstehungsweg dieser Arbeit und deine humorvolle Unterstützung durch alle Höhen und Tiefen. Ich freue mich auf alles, was kommt!
# Table of Contents

Danksagung .......................................................................................................................... 2  
Table of Contents ............................................................................................................... 3  
0 ABSTRACT ......................................................................................................................... 6  
1 THEORETICAL BACKGROUND ..................................................................................... 8  
   1.1 Paradigms of Psychotherapy Research ....................................................................... 8  
      1.1.1. Efficacy and Effectiveness Research .................................................................. 8  
      1.1.2. Patient-Focused Psychotherapy Research ............................................................. 9  
   1.2 Ecological Momentary Assessments (EMA) .............................................................. 12  
      1.2.1. EMA in Clinical Psychology and Psychotherapy Research ............................... 16  
      1.2.2. EMA in Patient-Focused Psychotherapy Research ................................................ 18  
2 METHODS ......................................................................................................................... 21  
   2.1. Recruitment and Sample Description ...................................................................... 21  
   2.2. EMA – Procedure and Design ................................................................................. 23  
   2.3. Measures .................................................................................................................. 24  
      2.3.1. Regular Questionnaires ...................................................................................... 25  
         Brief Symptom Inventory (BSI) ............................................................................... 25  
         Response Styles Questionnaire (RSQ) ..................................................................... 26  
         Hopkins Symptom Checklist-11 (HSCL-11) ............................................................ 27  
         EMA Evaluation Form .............................................................................................. 27  
      2.3.2. EMA Measures .................................................................................................... 27  
         Global Assessment .................................................................................................. 27  
         Affective states ........................................................................................................ 28  
         State Rumination and Accompanied Sensory Experiences ...................................... 28  
         Social Support and Self-Efficacy ............................................................................. 29  
         Life Events ............................................................................................................... 29  
   2.4. Statistics: Hierarchichal Linear Modeling ................................................................. 29  
3 STUDY I: Feasibility of an EMA – Application in an Outpatient Setting ......................... 31  
   3.1. Introduction ............................................................................................................... 31  
   3.2. Methods .................................................................................................................... 38  
      3.2.1. Sample ............................................................................................................... 38  
      3.2.2. Instruments ....................................................................................................... 40  
      3.2.3. Statistical Analysis ........................................................................................... 40
III: Daily affect dynamics predict early response in CBT

3.3. Results .................................................................................................................. 44
  3.3.1. Sampling........................................................................................................ 44
  3.3.2. Technical Aspects .......................................................................................... 45
    Unconditional Model and Single Predictor Models Technical Missings ............... 46
  3.3.3. Reactivity and Compliance ............................................................................. 47
    Unconditional Model and Single Predictor Models Duration ............................... 47
    Combined Model .................................................................................................. 48
  3.3.4. Acceptance and Burden for Patients .............................................................. 51
  3.4. Conclusions ........................................................................................................ 59

4 STUDY II: The rumination-affect link in everyday life ............................................. 67
  4.1. Introduction ........................................................................................................ 67
  4.2. Methods ............................................................................................................ 70
    4.2.1. Sample ....................................................................................................... 70
    4.2.2. Instruments ............................................................................................... 70
    4.2.3. Statistical Analysis .................................................................................... 71
  4.3. Results ................................................................................................................ 73
    4.3.1. Descriptive Results .................................................................................... 73
    4.3.2. Relation between State Rumination and Positive-Negative Affect .......... 74
    4.3.3. Moderator Effect of Dispositional Rumination ........................................... 77
  4.4. Conclusions ........................................................................................................ 78

5 STUDY III: Daily affect dynamics predict early response in CBT ......................... 83
  5.1. Introduction ........................................................................................................ 83
  5.2. Methods ............................................................................................................ 87
    5.2.1. Sample ....................................................................................................... 87
    5.2.2. Instruments ............................................................................................... 87
    5.2.3. Statistical Analysis .................................................................................... 88
  5.3. Results ................................................................................................................ 90
    5.3.1. Descriptive Statistics ................................................................................ 90
    5.3.2. Unconditional Model and Separate Predictor Models ............................. 91
    5.3.3. Combined Model ....................................................................................... 93
  5.4. Conclusions ........................................................................................................ 96

6 GENERAL DISCUSSION ......................................................................................... 101
  6.1. Summary and discussion of study results ............................................................ 101
6.1.1. Study I .................................................................................................................. 101
6.1.2. Study II ............................................................................................................... 106
6.1.3. Study III ............................................................................................................. 108
6.2. Conclusions and Future Prospects ....................................................................... 110

7 REFERENCES .............................................................................................................. 115
8 FIGURES .................................................................................................................... 139
9 TABLES ....................................................................................................................... 141
10 APPENDICES .......................................................................................................... 142
ABSTRACT

The efficacy and effectiveness of psychotherapeutic interventions have been proven time and again. We therefore know that, in general, evidence-based treatments work for the average patient. However, it has also repeatedly been shown that some patients do not profit from or even deteriorate during treatment. Patient-focused psychotherapy research takes these differences between patients into account by focusing on the individual patient. The aim of this research approach is to analyze individual treatment courses in order to evaluate when and under which circumstances a generally effective treatment works for an individual patient. The goal is to identify evidence based clinical decision rules for the adaptation of treatment to prevent treatment failure. Patient-focused research has illustrated how different intake indicators and early change patterns predict the individual course of treatment, but they leave a lot of variance unexplained. The thesis at hand analyzed whether Ecological Momentary Assessment (EMA) strategies could be integrated into patient-focused psychotherapy research in order to improve treatment response prediction models. EMA is an electronically supported diary approach, in which multiple real-time assessments are conducted in participants’ everyday lives. We applied EMA over a two-week period before treatment onset in a mixed sample of patients seeking outpatient treatment. The four daily measurements in the patients’ everyday environment focused on assessing momentary affect and levels of rumination, perceived self-efficacy, social support and positive or negative life events since the previous assessment. The aim of this thesis project was threefold: First, to test the feasibility of EMA in a routine care outpatient setting. Second, to analyze the interrelation of different psychological processes within patients’ everyday lives. Third and last, to test whether individual indicators of psychological processes during everyday life, which were assessed before treatment onset, could be used to improve prediction models of early treatment response.
Results from Study I indicate good feasibility of EMA application during the waiting period for outpatient treatment. High average compliance rates over the entire assessment period and low average burdens perceived by the patients support good applicability. Technical challenges and the results of in-depth missing analyses are reported to guide future EMA applications in outpatient settings.

Results from Study II shed further light on the rumination-affect link. We replicated results from earlier studies, which identified a negative association between state rumination and affect on a within-person level and additionally showed a) that this finding holds for the majority but not every individual in a diverse patient sample with mixed Axis-I disorders, b) that rumination is linked to negative but also to positive affect and c) that dispositional rumination significantly affects the state rumination-affect association. The results provide exploratory evidence that rumination might be considered a transdiagnostic mechanism of psychological functioning and well-being.

Results from Study III finally suggest that the integration of indicators derived from EMA applications before treatment onset can improve prediction models of early treatment response. Positive-negative affect ratios as well as fluctuations in negative affect measured during patients’ daily lives allow the prediction of early treatment response. Our results indicate that the combination of commonly applied intake predictors and EMA indicators of individual patients’ daily experiences can improve treatment response predictions models. We therefore conclude that EMA can successfully be integrated into patient-focused research approaches in routine care settings to ameliorate or optimize individual care.
1 THEORETICAL BACKGROUND

1.1 Paradigms of Psychotherapy Research

Psychotherapy research has profited from explosive developments within the past 20 – 25 years. These developments were decidedly promoted by continuous efforts in the field of quality assurance and quality management (Lutz, & Grawe, 2005). The goals of psychotherapy research are (at least) twofold: on the one hand, it attempts to test the efficacy and effectiveness of different treatment approaches and on the other hand, it attempts to understand how empirically supported treatments promote change, i.e., which processes take place in successful treatments (Orlinsky, Grawe, & Parks, 1994). The following sections describe current developments from treatment- to patient-focused psychotherapy research, omitting a detailed description of process research, as it is not essential to the analyses conducted in the context of the present dissertation project.

1.1.1. Efficacy and Effectiveness Research

Efficacy and effectiveness research have continued the tradition that was initiated when psychotherapeutic interventions were required to generate proof of their general effectiveness, i.e., legitimate their existence in the health care system (Grawe, 1992, 1997). Two research arms were established, the one focusing on efficacy under experimental conditions (Howard, Orlinsky, Brill, Martinovich, & Lutz, 1996; Lambert, & Ogles, 2004; Lutz, & Grawe, 2007) and the other testing effectiveness under routine conditions (Howard et al., 1996).

Efficacy research is applied in clinical studies that test whether a specifically defined intervention achieves the desired outcome in a specific and clearly defined patient subgroup under clinically controlled conditions. The design and the statistical assumptions of these randomised controlled trials (RCT) focus on internal validity and allow an estimation of the
maximum effects an intervention can provoke under the best possible conditions (e.g., Krause, & Howard, 2003).

After empirical support has been generated for a specific intervention in clinically controlled settings, the question of whether this intervention is efficacious in routine practice - and if so, to what extent - is still unanswered. Therefore, efficacy research is extended by effectiveness research under naturalistic conditions, with a focus on external validity. This next step is indispensable, as the differences between treatments applied under controlled vs. naturalistic conditions are manifold. Among these differences are the following: in naturalistic settings, interventions are rarely applied strictly following a manual; the duration of psychotherapy is often dictated by health insurance systems; and patients often suffer from comorbidity (Barkham et al., 2008; Böhnke, 2012; Lutz, & Böhnke, 2010; Lutz, & Grawe, 2007). This discrepancy led to a gap between research and practice, which effectiveness research attempted to bridge.

1.1.2. Patient-Focused Psychotherapy Research

Even after the implementation of effectiveness research, many practitioners and some researchers still saw a gap between research and practice (e.g., Castonguay, Barkham, Lutz, & McAleavey, 2013, Howard et al., 1996; Lutz, 2002). Practitioners are often faced with the question, whether an intervention, which has been shown to be efficacious and effective on average, also works for the individual patient they attempt to treat. This essential question, which could neither be answered by efficacy nor effectiveness research, has fostered a paradigm shift from treatment- to patient-focused research. Patient-focused psychotherapy research aims to close the research-practice gap by evaluating whether generally successful treatments also show positive effects for an individual patient and by analyzing how this treatment can be adapted to improve individual treatment outcome. To achieve this goal,
information from the treatment process is used to optimize clinical decision-making at the beginning, during, and at the end of treatments. In doing so, differential adaptations of the treatment process can be provided to avoid treatment failure and to guarantee the best possible treatment outcome for an individual patient (Lambert, Hansen, & Finch, 2001; Lutz, & Böhnke, 2010). Patient-focused psychotherapy research attempts to broaden nomothetic research results by integrating idiographic elements and therefore adding to the research field’s practice orientation. Statistical developments, which allow the combination of nomothetic and idiographic elements when analyzing treatment data (e.g., hierarchical linear models), open up new possibilities in this area of research.

Patient-focused psychotherapy research requires close monitoring of the developments an individual patient undergoes during different phases of the treatment process and the feedback of this information to the therapist (Lambert, 2007; Lutz, 2002). This monitoring process involves repeated assessments (usually at every therapy session) and allows the investigation of typical change patterns over the entire treatment period. These typical change patterns are essential to therapists as a reference point for the decision, whether the change observed for an individual patient corresponds to expected change trajectories, or whether it deviates from the expected course. Feedback studies have provided clear evidence of the fact that treatments showed enhanced effects, when therapists were provided with feedback of their patients’ progress (e.g., Lambert, & Shimokawa, 2011) – especially when therapists were committed to actually use the feedback provided (De Jong, Van Sluis, Nugter, Heiser, & Spinhoven, 2012). A possible explanation of these findings could be the repeatedly detected result that therapists overestimate their patient’s progress until a certain point during and until the end of treatment (e.g., Breslin, Sobell, Sobell, Buchan, & Cunningham, 1997; Hannan et al., 2005; Hatfield, McCullough, Frantz, & Krieger, 2010; Walfish, McAlister, O’Donnel, & Lambert, 2012). This overestimation bias is especially harmful to patients who do not profit.
from or even deteriorate during treatment. Especially for these cases, empirically derived decision rules can help the therapist to detect patients showing negative developments early during treatment and to adapt treatment strategies accordingly (Lambert, 2007; Lambert, & Shimokawa, 2011).

Different approaches have been applied to generate expected treatment course predictions, i.e., identify typical change patterns (e.g., Lutz et al., 2013; Stulz, Lutz, Lucock, & Barkham, 2007; Tang, & DeRubeis, 1999) in order to develop reliable decision rules for therapists’ evaluations of their patients’ treatment courses. Expected treatment response methods use patient intake characteristics to predict the following course of treatment. Similarly, nearest neighbor approaches (Lutz et al., 2010) use patient intake data to identify a number of alike patients, in order to predict the patient’s treatment course by using the progress shown by the patient’s nearest neighbors (Lutz, 2002; Lutz, & Böhnke, 2010).

Among the patient intake data used to identify nearest neighbors and predict treatment response have been demographic variables (e.g., sex and age) and intake scores on different outcome measures (e.g., BDI-II, BAI, IIP-32, CORE-OM, EMI, subscales from the SCL-90) (Lutz et al., 2005, Lutz et al., 2006).

The present dissertation project was developed in the context of and with data derived from patient-focused psychotherapy research. As described above, the fundament of patient-focused research is the continuous monitoring of outcome variables over the entire course of treatment, in order to use the information gained for ongoing treatments via feedback to the therapist (Castonguay et al. 2013; Howard et al., 1996; Lambert, Hansen, & Finch, 2001; Lutz, 2002; Lutz, De Jong, & Rubel, 2015). This continuous monitoring process employs psychometric questionnaires, providing researchers with large databases on treatments, which are conducted under the conditions of routine care. These databases present multifaceted opportunities to close the research-practice gap and to provide practitioners with research
results that are applicable to the individual patient. The objective of this dissertation project was to test whether the highly frequent assessments within the monitoring process of patient-focused research could be extended by a research method applying even higher frequency assessments in order to use the information gained to improve predictions on expected treatment courses and decision-making derived from patient-focused research applications. We expanded the assessment period to the time period before the onset of treatment and applied ecological momentary assessment (EMA, Stone, & Shiffman, 1994) strategies to assess data from the individual patient before the regular treatment monitoring process began. As research applications of EMA are still not commonly applied in psychotherapy research - although the number of applications is growing (Trull, Ebner-Priemer, Brown, Tomko, & Scheiderer, 2012) - the method is described in more detail below.

1.2 Ecological Momentary Assessments (EMA)

EMA strategies are typically applied in daily life research. This kind of data assessment is characterized by repeated event and/or time contingent inquiries of persons in their natural environment, during which information is given on current situations, behaviors, and experiences. These assessments are often carried out using electronic data assessment instruments (Fahrenberg, & Myrtek, 2001). Within daily life research, ambulatory assessment methods are broadly classified in two categories: the first category encompasses self-reports of behaviour, affect, and cognition, collected repeatedly over a number of days, either once a day (labelled daily diaries) or sampled several times a day. EMA strategies (also known as experience sampling method; ESM, Csikszentmihalyi, & Larson, 1987) fall into this category. The second category includes technically sophisticated methods that allow not only self-reports but also diverse, non-self-reported aspects of daily experiences, such as the auditory environment, psychophysiological status, physical location or proximity to a
particular other person to be assessed (Reis, 2012). In the National Institutes of Health’s Healthy People 2020 initiative, Bachrach (2010) described EMA methods as “tools that can revolutionize the behavioural and social sciences” albeit noting that “researchers are still in the earliest stages of tapping into [their] vast potential” (as cited in Reis, 2012).

There are several rationales that justify the application of EMA methods. The three most valuable arguments are briefly described in the following paragraphs: real-life, real-time and within-person. Daily life protocols intend to capture “life as it is lived” (Bolger, Davis, & Rafaeli, 2003, p. 580). They do so by providing extensively detailed data on behaviour as it occurs within its natural setting, justified by the assumption that the context in which the assessed processes unfold matters. For daily life researchers, a proper understanding of real-life behaviour can only succeed when contextual factors are taken into account in order to ensure ecological validity.

Another rationale behind EMA is defined in methodological terms: real-time assessments eliminate retrospection bias and minimize selectivity in describing experiences (Schwarz, 2012). Self-report-assessments rely on the assumption that people know their thoughts, feelings and behaviours and are able to report them accurately. A large body of literature has addressed concerns with this assumption, especially in research contexts where participants are faced with questions that are ambiguous, difficult to understand or merely exceed the participants’ knowledge or memory (e.g., Kane, Brown, Little, Silvia, Myin-Germey, & Kwapi, 2007; Redelmeier, & Kahnemann, 1996; Shiffman, 2009; Trull, & Ebner-Priemer, 2013). Reconstruction of experiences shortly after the experience can result in accurate reports (Kahneman, Krueger, Schkade, Schwarz, & Stone, 2004), but as time passes, general knowledge infers with past experiences. In this process, prediction, intention, choice, global memories, attitudes, and preferences converge and influence the reconstruction of experiences at the expense of the mere facts. By decreasing the latency periods between
events and reports, EMA minimizes recall biases, leading to higher data validity compared to conventional questionnaires. Covariations of selected variables (e.g., emotional state and interpersonal experiences or rumination) can be modelled statistically, which is of high relevance, as these contingencies are often misrepresented in retrospective self-reports (Gloster et al., 2008; Nisbett, & Wilson, 1977; Piasecki, Hufford, Solhan, & Trull, 2007). Studies that directly compare EMA and retrospective questionnaires point to discrepancies between the two data sources, especially regarding chronological patterns (Samo, Tucker, & Vuchinich, 1989; Shiffman et al., 1997; Stone, Broderick, Shiffman, & Schwartz, 2004). If researchers are interested in undistorted images of real-time experiences, then in situ measurements can shed light on the underlying dynamics from the participant’s perspective (Schwarz, 2012).

After the real-life and real-time arguments discussed in the preceding paragraphs, the last and maybe most important aspect that distinguishes EMA methods from other methods of data collection follows: the opportunity to study within-person dynamics and relations. Psychological research is – in most cases – concerned with analyzing large sample sizes in order to be able to generalize the results on a population level. This nomothetic kind of research uses summary statistics to describe the distribution of variables at the population level. With this approach, conclusions can be drawn about what applies to the aggregate, but still very little is known about what applies in general, i.e., what is true for each and every individual in the population. Especially when research interests lie in processes that unfold within individuals over time, the large-sample approach is limited. This touches on practitioners’ and some researchers’ criticism that population results are not necessarily reflective of within-person phenomena, as already discussed in section 2.1.2 (Hamaker, 2012). Intraindividual variations in behaviour and experience and how these trigger other intraindividual processes across real-life conditions can only be studied within individuals.
For decades, ambulatory assessment strategies have been used and are now indispensable as routine methods to inform diagnostic and treatment decisions in medicine (e.g., 24h blood pressure monitoring or electrocardiograms). Ambulatory psychological assessments have been applied as a research method in Germany since the early 1980s. The introduction was triggered by the emergence of computer-assisted assessments as an alternative to stationary diagnostics. In the United States, ambulatory assessments in psychological research were first introduced in paper-pencil format to study changes in mood states in representative samples of situations (Fahrenberg, Myrtek, Pawlik, & Perrez, 2007). The introduction of modern technical equipment such as personal digital assistants, mobile phones, and online questionnaires has facilitated the collection of daily life data. Additionally, the increased computational power of computers and the development of new statistical methods, which allow the simultaneous analysis of between- and within-person associations, further ease the evaluation of such extensive longitudinal data sets (Hamaker, 2012). Today, a number of programs exist that facilitate the implementation of EMA (Conner, & Feldman Barrett, 2013; Perrez, & Reicherts, 1996) using programs specifically developed for the implementation of EMA on mobile phones (Conner, 2013). Besides the possibility of launching audible signals, many electronic devices add time stamps, indicating when entries are made and how much time a participant dedicates to the items on the questionnaire – both indicators for ecological validity that cannot be assessed by conventional paper diaries (Bolger et al., 2003). Studies have shown that even though paper diaries regularly show high compliance rates, entries are often made before or after the time point aimed to be assessed (Broderick, Schwartz, Shiffman, Hufford, & Stone, 2003; Stone, Shiffman, Schwartz, & Hufford, 2002). According to those results the instruction-conform compliance must be rated considerably lower. On average, EMA studies report predominantly high compliance rates,
varies between 80 and 95%, although compliance decreases when only short time slots are defined to answer the questionnaire (e.g., Broderick, et al., 2003, & Jamison et al., 2001).

Today, the fields of research applications for ambulatory assessment methods in psychology are diverse, as indicated by the current volume of The Handbook of Research Methods for Studying Daily Life (Mehl, & Conner, 2012), in which 10 chapters on perspectives from different fields are illustrated. The following section summarizes research applications in clinical psychology.

1.2.1. EMA in Clinical Psychology and Psychotherapy Research

EMA data can shed light on an individual’s experience of problems in daily life. The detailed accounts gained by EMA can both inform and enhance clinical practice. EMA applications have been applied in various forms within clinical research. For instance, this approach is used to characterize clinical symptoms and features of psychopathology by tracking problematic behaviours, mood states or cognitions in daily life. Many research projects have targeted the core features of a specific disorder that manifest behaviourally (e.g., instances of substance abuse or binge and purge episodes) and tracked these during daily life (e.g., Hufford, 2007; Shiffman, 2009, Smyth et al., 2009). Others have focused on problematic mood states in those forms of psychopathology, which are characterized by elevated or extreme mood states (e.g., depression, anxiety, mania) or extreme fluctuations in mood states (e.g., borderline personality disorder) (Ebner-Priemer, & Trull, 2009). These study designs allow the assessment and analysis of the dynamic interplay between other symptoms that are characteristic for the specific disorder (e.g., depressive symptoms in depression) and momentary affect on a between- and within-subject level (e.g., Ebner-Priemer, & Trull, 2009; Moberly, & Watkins; 2008Vranceanu, Gallo, & Bogart, 2009). Similarly, problematic cognitions, expectancies, urges and their consequences can and have been targeted in EMA
1 THEORETICAL BACKGROUND

studies (e.g., the relationship between cravings and actual cocaine use; Preston et al., 2009). Importantly, for the analysis of this dissertation project, cognitive biases associated with certain disorders (e.g. depression) can be assessed alongside mood states, providing a functional analysis of increases in negative or positive mood states (Trull et al., 2012).

The previous examples show how using EMA applications to analyze specific symptoms of various disorders can enrich our clinical knowledge. The need for EMA application is well founded by the limitations of conventional diagnostic procedures. Conventional diagnostic procedures aim to create an image of reality by recapitulation, often using categorical statements (Helbig, Lang, Swendsen, Hoyer, & Wittchen, 2009). In order to gain improved and deeper knowledge of processes that cause disorders, research attempts that take the situational and context specific cognitive, emotional and behavioral reactions into account in a dimensional way are required. Therefore, the timeframe of assessments must be expanded. Ambulatory assessment strategies “have the potential to enhance the clinical assessment enterprise and deepen clinicians’ understanding of their clients’ symptoms, motives, and life circumstances” (Piasecki, Hufford, Solhan, & Trull, 2007, p. 25).

In addition to the description and characterization of psychopathology and its correlates, EMA methods can be used to monitor treatment progress (Trull et al., 2012). Treatment progress is usually monitored by weekly assessments (as in our patient-focused research setting described in section 2.1.). These weekly assessments are equally prone to retrospective biases; they are notoriously context-dependent (Schwarz, 2012) and highly influenced by momentarily accessible information (e.g. Fredrickson, 2000). Treatment studies using real-time data collection allow the comparison of treatment responses to different treatment modalities or interventions (e.g., Barge-Schaapveld, & Nicolson, 2002; Bauer et al., 2005; Klosko, Barlow, Tassinari, & Cerny, 1990). Treatment studies applying ambulatory
assessment strategies are still relatively infrequent, even though it has been shown that they reflect therapeutic effects more quickly than standard weekly assessments (Trull et al., 2012).

The final field of EMA application in clinical psychology, which is mentioned for the sake of completeness, has been broadly termed interactive assessment (Ebner-Priemer, & Trull, 2009; Fahrenberg, 1996; Shiffman, 2007) or ecological momentary interventions (Heron, & Smyth, 2010). The former describes EMA that includes real-time feedback to the patients, whereas the latter involves electronically-mediated interventions (Trull et al., 2012). According to the responses logged by the patient, feedback, guidance or treatment components are provided. Although research has not yet provided sufficient evidence on the effectiveness of interactive ambulatory assessments compared to treatment as usual, there are some promising results from studies that have applied electronic treatment to a variety of different disorders (e.g., Alpers, 2009; Intille, 2007; Kenardy et al., 2003).

EMA applications in clinical psychology and psychotherapy research help to broaden nomothetic research approaches by integrating idiographic elements. This approach is therefore easily brought in line with patient-focused psychotherapy research, as described in section 1.1.2. Therefore, the next paragraph highlights EMA applications, which can be linked to patient-focused research.

1.2.2. EMA in Patient-Focused Psychotherapy Research

Although not specifically in the context of patient-focused research, some studies have already analyzed the possibilities of using information gained from EMA data before treatment onset in order to optimize clinical decision-making or predict treatment courses. The following examples focus on studies conducted in the context of psychotherapeutic treatments, notwithstanding a number of studies conducted in the context of pharmacotherapy (e.g., Barge-Schaapveld, Nicolson, van der Hoop, & deVries, 1995; Barge-Schaapveld, &
Nicolson, 2002; Klosko et al., 1990; Wichers, Barge-Schaapveld, Nicolson, Peeters, de Vries, et al., 2009).

For example, Peeters, Berkhof, Rottenberg, and Nicolson (2010) analyzed whether emotional reactivity to daily life events functions as a predictor of a) treatment response within the first month of psychotherapy (combined with pharmacotherapy if indicated) and b) remission rates within 18 months in a sample of depressed patients. EMA was applied over 6 consecutive days before treatment onset 10 times daily to assess emotions and daily life events. They found that a) less emotional reactivity to negative and positive life events predicted higher depressive symptom severity after the first month of treatment and b) patients with less negative emotional reactivity to negative life events were less likely to recover from depression over the 18-month follow up.

Wichers, Lothmann, Simons, Nicolson, and Peeters (2011) found that reductions in negative affect following the maximum daily increase of positive affect provided a means to discriminate between treatment responders (assessed after 8 weeks of treatment, defined as a 50% reduction in the Hamilton Depression Rating Scale, HDRS) and non-responders in a sample of depressed patients and that higher negative affect following maximum increases in positive affect was associated with more depressive symptomatology at six-month follow-up.

Forbes et al. (2012) sought to predict the course and outcome of an eight-week open trial of cognitive behavioral therapy (CBT), pharmacotherapy, or a combination of the two in children and adolescents suffering from depression or anxiety, by social interaction and affective dynamics in daily life. EMA was applied over four days before treatment onset. The results of this study showed that higher positive affect levels, lower negative affect levels, higher positive : negative affect ratios and more time spent with fathers predicted lower post-treatment severity of depressive and anxiety symptoms.
Fisher (2015) analyzed a sample of generalized anxiety disorder (GAD) patients to determine the latent syndrome structure between avoidance, worry and anxiety for each individual before treatment onset in order to personalize intervention strategies. Fisher used information on GAD symptoms gathered from each patient by EMA (end of day diaries) over 60 consecutive days to inform treatment decisions regarding the timing of intervention strategies. Results from his exploratory analyses ($N = 10$) suggest that the interaction of different symptom dimensions varies across individuals and that increased knowledge about which process influences the next in a specific individual can help to decide which symptom should be targeted first during treatment to catalyze therapeutic effects.

The aforementioned examples illustrate how the implementation of EMA opens up new perspectives in patient-oriented psychotherapy research. Besides the possible exploration of context-specific intra- and interindividual differences in experiences and behavior, the frequency and intensity of activities and psychological processes can be captured with high ecological validity. Repeated measurements enable the exploration of chronological patterns and therefore allow for conclusions regarding antecedents, consequences and moderators of specific events (e.g., Greeno, Wing, & Shiffman, 2000; Shiffman, Hufford, Hickcox, Paty, Gnys, & Kassel, 1997; Shiffman, & Waters, 2004; Swendsen et al., 2000).

The research project in which we apply EMA aims to broaden our knowledge about individual patterns of change and patient characteristics that allow the prediction of treatment response and course. Therefore, we assess cognitive state and emotional experience variables in daily life four times a day from patients over a two-week period, before the onset of outpatient treatment. Before EMA should be applied routinely in the context of patient-focused psychotherapy research, a thorough analysis of its feasibility is indispensable. Only if a low-burden implementation (for researchers and patients) can be achieved, can the next step be to test whether EMA data provides additional information above and beyond prediction
models of treatment response and outcome derived from conventional self-report instruments. Therefore, Study I focuses on the feasibility and acceptance (patient perspective) of the implementation of an EMA application in an outpatient setting. Study II focuses on the EMA period itself, analyzing the well-studied link between rumination and affect (see section 4.1.) in our current sample in order to test whether results from earlier studies can be replicated and thereby secure the validity of the assessed data. Study III finally analyzes whether the EMA data that was assessed before treatment onset can be used to predict early treatment response and whether this data provides information beyond commonly applied data from self-report questionnaires. Early treatment response was chosen as an outcome variable, as early change patterns have been shown to soundly predict later change patterns (see section 5.1.).

2 METHODS

2.1. Recruitment and Sample Description

All three studies were based on an EMA project that was conducted at the University of Trier’s outpatient clinic. The study was approved by the University’s ethics committee and written informed consent was obtained. Patients were assessed via EMA before the onset of their treatment. Recruitment for EMA was conducted between October 2013 and April 2015. When patients registered for treatment, they were asked to indicate whether they were interested in participating in research studies before the onset of treatment. Patients were also asked to fill out a range of questionnaires, which are part of the clinic’s regular quality assurance and monitoring system. These questionnaires were screened for exclusion criteria. Exclusion criteria included suicidality, psychosis, and mania. Patients who were interested in study participation and did not fulfill any exclusion criteria were contacted by phone by the principal investigator of the EMA project, who introduced the aim of the study and explained the conditions of participation. When patients were interested in participation, screening for
eligibility was carried out via the German Version of the Mini-International Neuropsychiatric Interview 5.0.0. (M.I.N.I.; Ackenheil, Stotz-Ingenlath, & Dietz-Bauer, 1999). The M.I.N.I. is a brief structured interview for the major Axis I psychiatric disorders in DSM-IV and ICD-10. The interview was carried out via phone by trained psychologists attending the post-graduate psychotherapy training program. If the patients proved eligible for the study, they were invited to take part.

Figure 2.1 displays the recruitment flow chart. Within the recruitment period, 63 patients were screened eligible. Of the 63 patients, 61 started EMA and 60 completed the entire 2-week period. One patient dropped out during the EMA period and data from two patients was lost at transfer after the 2-week period. Thus, EMA data from 58 patients could be analyzed for this dissertation project.

![Recruitment flow chart](image)

**Figure 2.1:** Recruitment flow chart
63.80% of EMA participants were female, age ranged from 19 – 59 years (mean = 35.40, SD = 11.45) and impairment levels measured by the Global Severity Index of the Brief Symptom Inventory (BSI; Franke, 2000; German translation by Derogatis, 1975) ranged from .11 to 2.79 (mean = 1.28, SD = .65).

48 patients were diagnosed after treatment onset. Diagnoses were based on the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV) Axis I Disorders (First, Spitzer, Gibbon, & Williams, 2002). Most patients in the sample were diagnosed with an affective disorder (46.0%), followed by post-traumatic stress disorder (PTSD; 23.0%) and anxiety disorder (13.2%) as the primary diagnosis. Additional primary diagnoses were eating disorders (6.3%), substance-related and addictive disorders (5.2%), and dissociative disorder (2.1%). For the diagnosis of personality disorders, the International Diagnostic Checklist for Personality Disorders (IDCL-P; Bronisch, Hiller, Mombour, & Zaudig, 1996) was adopted, which identified 4.2% of the sample as having a personality disorder as the primary diagnosis.

2.2. EMA – Procedure and Design

The EMA period was integrated into the clinic’s regular care process and took place during the time between registration and treatment onset. We applied iOS-based iDialogPad software, Version 1.922 (Mutz, unpublished) for the assessments. The EMA period started with a training session, instructing the participant on the usage of the iPod, which was used for EMA data collection. Written informed consent was obtained at this session and self-report questionnaires were answered (see section 2.3.1.). The two-week EMA period started immediately after the training session and was part of the regular waiting period before the onset of treatment. Patients were signaled 4 times a day for 14 consecutive days (handover and return days not included). Audio signals followed a time contingent sampling plan, with
signals every four hours between 08.00 am and 08.00 pm on weekdays and between 10.00 am and 10.00 pm on weekends. Unanswered signals were repeated three times in fifteen-minute intervals. That way, participants had the chance to postpone entries up to a maximum of one hour (in total four chances to fill out the questionnaire). For instance, if the entry was made at the second signal, no more signals were given for this entry. If there was still no response to the signal after the fourth beep, it was coded as a missing recording. Participants were contacted after the first two days of the EMA period to ensure that no problems arose. Additionally, they received a phone number they could call in case of questions or problems. After 14 days of data collection, participants returned the iPods. Participants were compensated with €80 for completing the 14-day EMA period. After this session, patients routinely continued their waiting period before the onset of treatment. After the onset of treatment, the routine monitoring and diagnostic process began, during which patients report weekly on well-being and impairment. Treatment course was followed until session 5 for each study participant.

2.3. Measures

The following section gives an overview of all relevant measures included in the study. They are described in the order of application; regular questionnaires are described separately from the EMA measures. Figure 2.2 gives an overview of the time point of assessment for each measure or construct during the course of the study.
2 METHODS

**Figure 2.2:** Study flow chart with relevant measures and assessed constructs

### 2.3.1. Regular Questionnaires

**Brief Symptom Inventory (BSI)**

To assess overall impairment levels, the BSI was administered at registration for treatment (see Figure 2.2). The BSI is a 53-item short form of Derogatis’ Symptom Checklist that assesses general impairment levels of patients on nine symptomatic subscales on a five-point Likert scale ranging from 0 (not at all) to 4 (extremely). In this study, only the BSI's Global Severity Index (GSI) was used. The internal consistency of the BSI is \( \alpha = 0.92 \) and the retest-reliability is \( r_{tr} = 0.90 \) (Franke, 2000).

**Beck Depression Inventory (BDI-II)**

The BDI is a self-report instrument, which assesses depressed mood (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). The items relate to depressive symptoms such as mood and hopelessness, cognitions such as guilt or feelings of being punished as well as physical symptoms such as loss of appetite or loss of libido. In this study, the BDI-II, a revised version
of the BDI, was applied (Beck, Steer, Ball., & Ranieri, 1996). The BDI-II contains 21 questions, of which some have been changed due to adapted diagnostic criteria for major depressive disorder (MDD) in the DSM-IV (American Psychiatric Association, 2000). Each item can be answered with scores between 0 and 3 whereby higher total scores indicate more severe depressive symptoms. The BDI-II’s psychometric properties range between acceptable and excellent ($\alpha = .76\text{--}.95$; $r_{tt} = .90$; Beck et al., 1988; Beck et al., 1996). In the current study, the BDI total score was used to measure depressive symptom changes over the EMA-period. Therefore, it was administered before EMA onset and immediately after its termination after two weeks (see Figure 2.2).

**Response Styles Questionnaire (RSQ)**

The RSQ (Nolen-Hoeksema, & Morrow, 1991) is designed to measure dispositional cognitive and behavioral responses to dysphoric mood by asking patients what they generally do, when they feel sad, down or depressed. We used the German short version of the RSQ (RSQ-D; Kühner, Huffziger, & Nolen-Hoeksema, 2007). The short version consists of 23 items, which are classified on three subscales assessing symptom-related and self-related rumination as well as distraction. Items are rated on a 4-point frequency scale. Symptom-related rumination is assessed by 8 items, self-related rumination by 7 items and distraction by 8 items. Sum scores reflect the values a participant scores on the corresponding subscale. The internal consistency of the three subscales ranges between $\alpha = .76$ and $\alpha = .88$, the retest-reliability after 5 months ranges between $r_{tt} = .51$ and $r_{tt} = .70$ (Kühner et al., 2007). The RSQ-D was administered before EMA onset and immediately after its termination after two weeks (see figure 2.2).
Hopkins Symptom Checklist-11 (HSCL-11)

The HSCL-11 (Lutz, Tholen, Schürch, & Berking, 2006) was administered at the beginning of each session. This 11-item self-report inventory assesses symptomatic distress. It is a brief version of the SCL-90-R (Derogatis, 1992). The items are answered on a 4-point Likert scale ranging from 1 (not at all) to 4 (extremely). The mean of the 11 items represents the client’s level of global symptomatic distress in the preceding week. It is highly correlated with the GSI (r = .91) and has a high internal consistency (α = .92; Lutz et al., 2006).

EMA Evaluation Form

After termination of the EMA period, participants routinely answered a set of questions regarding the perceived effects and consequences of the regular assessments and the perceived associated burden. The questionnaire consists of 18 items (e.g., “Did you perceive the daily assessments as a burden?”; see Appendix A6 for all items) that were assessed on a five-point Likert scale ranging from 0 (not at all) to 4 (extremely). Additionally, participants could give further written feedback in an open response space.

2.3.2. EMA Measures

Patients answered 23 items at each of the four daily assessments. The items measure the following constructs:

Global Assessment

Each measurement started with an item that assessed global functioning and well-being. We decided to use the item: “How well did you get along within the past 3-4 hours?” from the Compass Tracking System (Howard et al., 1996).
2 METHODS

Affective states

Positive (PA) and negative affect (NA) were assessed during the EMA procedure. Participants rated eight momentary affective states on 5-point Likert scales (ranging from 1 - “a little or not at all” to 5 - “very”). Our choice of the EMA affect items was guided by the Positive and Negative Affect Schedule (PANAS) questionnaire and by results of previous studies (factor analytic selection of items with high loadings on NA and PA factors and sufficient item difficulty). Ratings of the items ‘excited’, ‘determined’, ‘alert’, and ‘active’ were averaged to form the PA scale. The average ratings of ‘depressed’, ‘ashamed’, ‘anxious’, and ‘nervous’ formed the NA scale (see Appendix for item list). Mean PA, mean NA and mean PA/NA scores, as well as mean squared successive differences (MSSD) in PA and NA to assess temporal fluctuation in affective states, were computed for each participant. MSSD is the average of the squared differences between successive observations at occasion $i + 1$ and $i$ ($\text{MSSD} = \frac{1}{N-1} \sum_{i=1}^{N-1} (x_{i-1} - x_i)^2$). The MSSD is a preferred index of affective fluctuation in EMA studies, because it captures both variability and temporal dependency in a time series (Jahng, Wood, & Trull, 2008).

State Rumination and Accompanied Sensory Experiences

State rumination was assessed by 4 modified items from the RSQ, which have already been used in a diary study by Genet and Siemer (2012). For example: “Within the past three hours, I continued to think about a situation, wishing it had gone differently.” (See Appendix A5 for all items). The selected items assessed symptom- and self-focused rumination. Items were answered on a 9-point scale (1 = not at all, 9 = extremely). Sum scores of the four items were used as an indicator of state rumination at each measurement occasion.

Sensory properties were assessed by three adapted items from the Sensory Properties of Depressive Thoughts Questionnaire (SPD; Moritz et al. 2013). Participants were asked to
indicate whether or not their ruminative thoughts were accompanied by sensory experiences, including auditory, visual, bodily, tactile, and olfactory sensations. Moritz et al. (2013) report good internal consistency (α = .76) and associations to several indexes of psychopathology.

**Social Support and Self-Efficacy**

Five adapted items from the Bern Inventory for the Assessment of Personal and Interpersonal Resources and Resource Realization (RES; Trösken, & Grawe, 2004) were administered to assess social support (e.g., “Within the past three hours, someone offered me help”) and self-efficacy (e.g., “Within the past three hours, I could experience my own capabilities and potential”). Items were answered on a 9-point scale (1 = not at all, 9 = extremely). See Appendix A5 for all items.

**Life Events**

By means of two single items, patients were asked to indicate whether or not something pleasant or unpleasant had happened to them within the past three hours.

**2.4. Statistics: Hierarchical Linear Modeling**

EMA data exhibits a nested structure, where measurement occasions (Level 1) are nested within persons (Level 2). In our analyses, different nested data structures are relevant: For Project I and II, the aforementioned structure of measurement occasions nested in persons is analyzed. For project III, information from EMA data is extracted and included in prediction models, in which therapy sessions (Level 1) are nested in patients (Level 2). All analyses require a statistical method that takes this multilevel data structure into account. Hierarchical linear models (HLM) fulfill these requirements and are therefore currently the recommended statistical approach. All analyses were done using HLM (Version 7, Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011). HLMs applied to daily assessment data allow not only
the analysis of how mean levels of daily observations (e.g., mood) vary across individuals, but also of how within-person relationships between daily observations (e.g., rumination and mood) vary across individuals (Nezlek, 2012). If there is no theoretical or methodological justification for an alternate procedure, random coefficient models are used, starting with unconditional or null models, which are successively augmented to incorporate predictors at Level 1 and Level 2. The focus of analysis varies between intercepts as outcomes, slopes as outcomes and cross level interactions. Specific models for the separate analyses are described in the methods sections of each of the three studies (see sections 3.2.3., 4.2.3. and 5.2.3.).
3 STUDY I:FEASIBILITY OF AN EMA – APPLICATION IN AN OUTPATIENT SETTING

3.1. Introduction

EMA approaches are growing in number, with various fields of application in clinical psychology. EMA is used to assess problematic behavior, mood, cognitions, expectancies, and urges. It is applied to monitor responses to treatment and to give real-time feedback to patients as part of the treatment strategy (Trull et al., 2012). However, the applications in patient-oriented psychotherapy research – especially with regard to the prediction of treatment response and outcome - are still limited in number. Possible causes may be open questions with regard to feasibility - especially in outpatient settings. In order to test feasibility and acceptance of EMA in outpatient populations, different aspects of its application must be considered. In the introduction to this study, some problematic aspects that may arise when applying EMA approaches and how they can be assessed are discussed. They are listed in order of potential occurrence.

First, systematic sampling selection effects might occur due to person-variables (e.g., age). Although up to now no sampling selection effects with regard to education or age that systematically influence who participates in an EMA study have been reported (Finkelstein et al., 2000; Shiffman et al., 2007), some critics of ambulatory assessment strategies argue that the method bears a high risk of self-selection bias in favor of technically-experienced individuals (Bussmann, & Ebner-Priemer, 2013). Age could be a person-variable, which may hint at a self-selection bias caused by familiarization in the use of electronic devices. A thorough analysis of study flow charts helps to reveal whether systematic sampling selection (and dropouts) occurs, which discriminates certain individuals. Additionally, control group comparisons allow the testing of sampling effects with regard to different patient characteristics such as age, education, employment etc.
Another topic to be addressed with regard to feasibility is measurement reactivity. In total, the field of diary methods has paid little attention to measurement reactivity. Measurement reactivity refers to “the systematically biasing effects of instrumentation and procedures on the validity of one’s data” (Barta, Tennen, & Lit, 2012, p. 108). Even though almost any measurement method is likely to generate reactivity, high frequent self-monitoring of one’s own behavior, as it is required in EMA applications, raises the question, whether diary methods result in particular reactivity. Reactivity appears in different forms. For instance in reactance – the change in participants’ experience or behavior as a result of participation in the study, or in habituation effects – which can lead to the development of a habitual response style when making diary entries (Bolger et al., 2003). An important topic in research applications of EMA is the possibility that completing diary style assessments several times a day could itself alter the perception of the assessed construct or the elaborateness of completing assessments (i.e., become a reactive task), which again would influence the validity of the data. Although a consensus exists that the accuracy of records can be increased, if the behavior is recorded at the time and place of its occurrence, since early behaviorism it is known that self-monitoring also has the potential to alter the behavior under observation. This might, for instance, be due to an increase in awareness and reflection caused by mere self-monitoring. Never the less, research on this particular part of reactivity has shown that behavior changes are most likely among individuals who are motivated to change their behavior (Korotitsch, & Nelson-Gray, 1999). One example of alteration of the assessed behavior or experiences could be systematic mood changes when the affective experiences of the participants are under observation. According to Nolen-Hoeksema’s Response Styles Theory (1991), being overly engaged in thinking about one’s own feelings (a process, which could be triggered when participants are asked to reflect on their affective experiences - e.g., 4 times a day) can lead to a vicious circle that promotes maintenance and increase of depressive
symptoms, when the focus is mainly on negative emotions. Following this line of argumentation, the regular assessments of EMA applications might lead to an increase in negative affect in some participants.

One other form of reactivity could result in positive effects of EMA application for participants. Especially when EMA applications are implemented in the waiting period for outpatient treatment, it could influence the psychological symptoms of the participants. In psychotherapy research, decreases in psychological symptoms during the waiting period for outpatient psychotherapy, the so-called waiting period effect, could be observed repeatedly (e.g., Greenberg, Constantino, & Bruce, 2006, & Huckert, Hank, & Krampen, 2012). It is explained by an increase of positive expectations as soon as patients have applied for treatment and thus have overcome the first obstacle on their individual path to increased well-being. It is arguable that participation in an EMA study during the waiting period could further increase positive expectations (e.g., by the mere fact of having the impression of starting to do something for their psychological health or by the first personal contact with a health professional). If such effects occur, they should be measurable by the same instruments that usually assess waiting period effects (in our case, the BSI). As suggested by the Handbook of Research Methods for Studying Daily Life (Mehl, & Conner, 2012), control group comparisons open up the possibility of comparing changes on outcome measures (e.g., BSI) between the EMA group and a matched control group, to identify whether there are systematic method effects that exceed the usual improvement of well-being, which occurs between registration for treatment and its onset (waiting period effect).

Another aspect that highly influences reactivity is the social desirability of the behavior under observation. It can be hypothesized that socially undesirable behavior produces negative affect and is therefore negatively reinforcing, whereas socially desirable behavior produce positive affect and positive reinforcements (Kazdin, 1980). The framing of
the study’s contents additionally influences the extent of reactivity, as it has been shown that
the behavior changing effects of self-monitoring are greater, when presented to the
participants as a treatment modality (Abrams, & Wilson, 1979). Reactivity is decreased when
more than one behavior is monitored. Especially when it comes to sensitive questions, self-
administered questionnaires - as they are employed in computerized data collection methods -
are less reactive than face-to-face interviews, because of the perceived anonymity they offer
(Barta et al., 2012). Although computerized data collection methods are associated with
perceived anonymity (DiLillo, DeGue, Kras, Di Loreto-Colgan, & Nash, 2006), EMA
applications include the regular necessity of making entries on technical devices such as
PDAs, iPods or smartphones, which are additionally announced by auditory signals, might
evoke reactions from others. There is a trend, which can be described as self-tracking and has
been observed with regard to different aspects of the body and mind, reaching from diets to
physical and psychological activities and has become a quickly disseminated topic in mass
media (e.g., “Trend zur Selbstvermessung”, 2015; Laaff, 2011). This trend is promoted by an
ever increasing multitude of self-tracking apps, which are available to the modern smartphone
user (Armstrong, 2015). It also initiates research projects that analyze different aspects of self-
tracking e.g., the motives behind and effects of the quantification of the self (e.g., Digitale
Selbstvermessung [Digital Self-Tracking], Zilien, Fröhlich, Kofahl, & Spengler; project
funded by the German Research Foundation [Deutsche Forschungsgemeinschaft; DFG]) in
2014). This trend facilitates EMA applications for psychological constructs, as more and more
people become used to others filling out questionnaires on electronic devices, however the
high frequency of entries and the auditory signals in EMA applications might still provoke
reactions. These reactions can be positive or negative with corresponding effects for the
participant. Negative reactions might have a negative impact on compliance rates, whereas
compliance rates might benefit from positive reactions.
Satisficing, defined as “the limits a respondent imposes on the amount of effort she or he is willing to apply to answer a question or a set of questions” (Barta et al., 2012, p. 110), is another aspect of measurement reactivity. To prevent the risk carried by satisficing, time pressure to respond quickly and the burdening of participants by means of lengthy questionnaire protocols should be minimized. Participants should be thoroughly familiar with the contents of the administered questionnaires and the handling of the electronic device being used. Although the elements of optimal diary design are easily accessible for researchers studying daily life, they have yet to be evaluated. Inattentive responding because of fatigue, distraction or hurry might potentially occur within the timeframe of our EMA-application (14 consecutive days), in which participants are asked to answer the questionnaire four times daily. Fatigue effects in their most pronounced forms can lead to a decrease in the amount of effort a participant is willing to apply and thus elicit poorer compliance (defined as the beep-wise response rates). Nevertheless, compliance rates in EMA applications using electronic devices are usually very high compared to paper-pencil diaries (Stone, Broderick, Shiffman, & Schwartz, 2004). Studies that involve multiple reports per week typically run from 3 days to 3 weeks (Conner, & Lehman, 2012). With a length of two weeks, our protocol is located in the upper third of typical assessment duration and fatigue effects may occur, especially at the end of the individual assessment periods.

In summary, there is a risk that reactivity potentially distorts data derived from EMA in ways that could undermine the advantages of the research technique. Although influential research in the field of diary applications was able to show that reactivity effects are minimal in specific patient populations (e.g., Stone, Broderick, Schwartz, Shiffman, Litcher-Kelly, & Calvanese, 2003), every application of EMA requires thorough analyses of reactivity caused by the assessment strategy.
Furthermore, technical aspects of an EMA application are also involved in missing data rates. As different EMA programs exist for different electronic devices, the rates of malfunctioning and data loss vary significantly. Peters, Sorbi, Kruise, Kerssens, Verhaak and Bensing (2000) conducted a study with patients with unexplained pain symptoms, which was similar in design to our study, as they also assessed psychological variables four times a day. Nevertheless, the assessment period was twice as long (four weeks). Peters et al. reported 5.1% of missing data due to technical problems. To date, no systematic influence of day of week on the measurement results could be identified (Ebner-Priemer, & Trull, 2009).

The Handbook of Research Methods for Studying Daily Life (Mehl, & Conner, 2012) suggests testing for reactivity in systematic missing analyses using hierarchical linear modelling (HLM) and by control group comparisons. Missing analyses help to identify whether data is missing completely at random or whether there are systematic effects. Missing data rates should be reported, as they can provide important information for the future planning and implementation of EMA in clinical settings.

In addition, when analyzing EMA feasibility, the burden for participants possibly caused by intensive assessment schemes must be considered. In general, strategies to keep the burden low include: limiting the length and difficulty of diary assessments, using simple language, limiting the clauses within a sentence as well as the number of response options per question and avoiding subjective items in favor of concrete, objective items (Barta et al., 2012). Even when following those guidelines, the acceptance and perceived burden for the participants should still be assessed after the EMA period to make reliable statements about the participants’ burden in the specific study. Especially when facing the fact that EMA strategies, despite their numerous advantages, are still not a regularly applied method in clinical psychology, insights on the perceived burden and participants’ acceptance can inform future implementations, particularly in outpatient settings. Feasibility in terms of patients’
acceptance can be analyzed by rates of compliance and patients’ subjective reactions after EMA application.

Considering the benefits of EMA, which were reported in the theoretical background of this thesis and could ameliorate patient-oriented care, in the following, open questions regarding the feasibility of its implementation in an outpatient setting are addressed. Thus, Study I focuses on the following research questions:

1.) With regard to sample selection and daily assessments, can context variables be detected, which impede the successful implementation of EMA?

   We hypothesize that:
   a) There are no systematic sampling selection effects and
   b) data loss due to technical problems does not impede application and does not appear systematically.

2.) Can systematic missing analyses and control group comparisons detect reactivity effects that undermine the advantages of EMA?

   We hypothesize that:
   a) Participants become faster in answering the EMA protocol over time, but reductions in time to answer only appear within the first days (training effect).
   b) Overall compliance is high and does not systematically change over the 14-day period, no patients need to be excluded because of lack of compliance and there are no effects of patient characteristics on compliance. The number of dropouts is low.
   c) The improvement in impairment levels from registration for treatment to the onset of treatment does not systematically differ between the EMA sample and a matched control group, meaning that no intervention effects of EMA on impairment levels can be detected.
3.) What do patients say? Do patients’ ratings of acceptance and perceived burden after the EMA period support or oppose its implementation?

3.2. Methods

3.2.1. Sample

For the Study I analyses, the entire EMA sample, as described in section 2.1, was included.

For research question 1a we compared our EMA sample to those patients from the outpatient clinic who registered for therapy in the same time period (October 2013 – April 2015), but did not take part in the study, although they also did not fulfill exclusion criteria. This group (control group 1) consisted of 877 patients.

For research question 2c, a second control group was generated by means of propensity score matching (for an example of propensity score matching, PSM, in clinical psychology see Lutz, Schiefele, Wucherpfennig, Rubel, & Stulz, 2016). Control Group 2 consisted of 58 patients who registered for therapy in the same time period as the EMA group.

Matching was conducted with the R Package MatchIt (Ho, Imai, King, & Stuart, 2011). Matches were identified from a sample of 725 patients from the University of Trier’s outpatient clinic. We used the following covariates for the matching procedure: initial impairment measured by the BSI’s Global Severity Index at the time of registration for treatment, gender, age, family status, education, professional education, and employability. Informed by Lutz et al., 2015, we applied the nearest neighbor (NN) matching procedure. The goodness of the matching procedure is indicated by the degree to which both samples (in our case, EMA sample and matched control sample; control group 2) result in similar distributions of covariates. A widely used method to check covariate balance between samples is the standardized mean difference (smd) technique. The smd method is similar to Cohen’s d and allows the comparison of differences in matched and unmatched conditions for
each covariate. A smd < .25 indicates an acceptable match between samples on the respective covariate (Rubin, 2001).

**Table 3.1: Demographics of EMA sample and matched controls**

<table>
<thead>
<tr>
<th></th>
<th>EMA Sample</th>
<th>Control Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex: female</strong></td>
<td>N (%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>37 (63.8)</td>
<td>40 (69.0)</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>Average (SD)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>35.00 (11.35)</td>
<td>35.74 (12.04)</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19 – 59</td>
<td>18 - 60</td>
</tr>
<tr>
<td><strong>Family Status</strong></td>
<td>N (%)</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>33 (56.9)</td>
<td>30 (51.7)</td>
</tr>
<tr>
<td>Married</td>
<td>09 (15.5)</td>
<td>8 (13.8)</td>
</tr>
<tr>
<td>living separately/divorced</td>
<td>14 (24.1)</td>
<td>19 (32.7)</td>
</tr>
<tr>
<td>widowed</td>
<td>1 (1.7)</td>
<td>1 (1.7)</td>
</tr>
<tr>
<td>Re-married</td>
<td>1 (1.7)</td>
<td>0</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td>N (%)</td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>1 (1.7)</td>
<td>0</td>
</tr>
<tr>
<td>&quot;Hauptschule&quot;*</td>
<td>11 (19.0)</td>
<td>8 (13.8)</td>
</tr>
<tr>
<td>&quot;Realschule&quot;*</td>
<td>20 (34.5)</td>
<td>21 (36.2)</td>
</tr>
<tr>
<td>&quot;Fachabitur/Abitur&quot;*</td>
<td>24 (41.4)</td>
<td>28 (48.3)</td>
</tr>
<tr>
<td>Other</td>
<td>2 (3.4)</td>
<td>1 (1.7)</td>
</tr>
<tr>
<td><strong>Professional Education</strong></td>
<td>N (%)</td>
<td></td>
</tr>
<tr>
<td>Trainee/University Student</td>
<td>14 (24.1)</td>
<td>12 (20.7)</td>
</tr>
<tr>
<td>Traineeship/Polytechnic Degree</td>
<td>29 (50.0)</td>
<td>29 (50.0)</td>
</tr>
<tr>
<td>University Degree</td>
<td>8 (13.8)</td>
<td>12 (20.7)</td>
</tr>
<tr>
<td>None</td>
<td>4 (6.9)</td>
<td>1 (1.7)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (5.2)</td>
<td>4 (6.9)</td>
</tr>
<tr>
<td><strong>Employable</strong></td>
<td>N (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>44 (75.9)</td>
<td>47 (81.0)</td>
</tr>
<tr>
<td>No</td>
<td>12 (20.7)</td>
<td>10 (17.2)</td>
</tr>
<tr>
<td>Unknown</td>
<td>2 (3.4)</td>
<td>1 (1.7)</td>
</tr>
<tr>
<td><strong>Initial Impairment</strong></td>
<td>Average (SD)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.26 (0.65)</td>
<td>1.35 (0.69)</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.11 – 2.79</td>
<td>0.13 – 3.62</td>
</tr>
</tbody>
</table>

*Note.* * Correspond to the three common school leaving certificates in the German education system. The three graduation levels rank in the following order (from lowest to highest): Hauptschule, Realschule, Fachabitur/Abitur.

Table 3.1 gives an overview of demographic variables for the EMA sample and the control group 2. The vast majority of smd scores indicated acceptable balance for all
covariates. Only the smd score for the subcategory \textit{divorced} from the family status covariate exceeded the recommended score with a value of -0.31.

\subsubsection*{3.2.2. Instruments}

Data from the EMA evaluation form, as described in section 2.3, was analyzed for Study I. Furthermore, systematic missing analyses were carried out, which incorporated sample selection data (flow chart of recruitment and dropouts) as well as the full EMA dataset from all 58 patients.

\subsubsection*{3.2.3. Statistical Analysis}

With regard to research question 1a), we first compared those patients who agreed to take part in the EMA study with all other patients who registered for therapy in the same time period and did not take part (disregarding those patients who fulfilled exclusion criteria). Comparisons where made with regard to demographic variables (gender, age, education [highest school leaving certificate] and occupation [currently employed or not]) and with regard to impairment levels at application (measured by BSI). We conducted chi-square tests for independence for all categorical variables and independent t-tests for continuous variables.

In a second step of analyses for research question 1b), we tested whether the number of technical missings systematically changed over the EMA period and whether certain patient characteristics significantly influenced the number of technical missings. Therefore, we first calculated the sum of signals per day (minimum 1, maximum 4) within each participant. We entered the sum of signals per day as the outcome variable in a random intercept and slope HLM, because of the nested structure of measurement occasions (Level 1) within persons (Level 2). Number of days was entered as the slope variable to account for time effects. We applied the following model:
**Level 1 Model**

Sum of daily signals\(_{ti}\) = \(\pi_{0i} + \pi_{1i} \times (\text{Number of days}_{ti}) + e_{ti}\)  \(\text{(1)}\)

Sum of daily signals\(_{ti}\) is the observed sum of signals on a particular day. The \(\pi_{0i}\) parameter (intercept) is a patient’s expected sum of signals on the first day. The \(\pi_{1i}\) parameter (slope) is the expected change in the number of signals per day. The random error term (\(e_{ti}\)) refers to normally distributed deviations from expected values for patient \(i\) at assessment \(t\). This model is referred to as the Level 1 model.

Individual differences between patients in the number of daily signals (intercept) and individual change in the number of signals from the first to the last day of assessment (slope) are predicted with successive Level 2 models, in which the Level 1 intercept and slope coefficients are treated as the dependent variable in random coefficient models. We started with an unconditional model (Model 1), in which the variation of patient intercepts and slopes was modeled as a constant plus random effect:

**Level 2 Model**

\(\pi_{0i} = \beta_{00} + r_{0i}\)  \(\text{(2)}\)

\(\pi_{1i} = \beta_{10} + r_{1i}\)  \(\text{(3)}\)

This model served as our unconditional base model (Model 1). We then successively augmented equations 2 and 3 to test the cross-level interaction of different level 2 predictors with regard to their predictive values for the intercept and slope factor. We tested the following patient characteristics as predictors for differences in intercept and slope: age, gender, and initial impairment. Initial impairment was grand-mean centered, whereas age and gender were entered uncentered. We first entered all predictors separately in single predictor
models. In a second step, we entered all significant predictors from the single predictor models into one combined model.

With regard to research question 2a, in a first step we tested whether the number of days of assessment had an effect on the duration of the single assessments (e.g., participants become faster over time) and whether certain patient characteristics significantly influenced the duration. We modelled duration as the outcome variable in a random intercept and slope HLM. Number of assessment was entered as the slope variable to account for time effects. The combined equation for level 1 and level 2 looks as follows:

$$DURATION_{ti} = \beta_{00} + \beta_{10} (\text{Number of assessment}_{ti}) + r_{0i} + r_{1i}$$

DURATION_{ti} is the observed duration at a particular assessment. The \(\beta_{00}\) parameter (intercept) is the average expected duration at the first assessment. The \(\beta_{10}\) parameter (slope) is the expected change in duration per assessment number. The random error term \((r_{0i})\) refers to the Level 1 residuals (i.e., deviations of the observed duration scores from the expected duration scores at each measurement occasion). The random error terms \((r_{1i} \text{ and } r_{1i})\) refer to the Level 2 residuals of intercepts and slopes (i.e., individual differences in duration mean levels and individual differences in the within-person regression slope). As in the previous model, we started with this unconditional base model and then successively entered different Level 2 predictors to test their predictive value for the intercept and slope factor. We tested the following patient characteristics as predictors for differences in intercept and slope: age, gender, and initial impairment. Initial impairment was grand-mean centered, whereas age and gender were entered uncentered. We first entered all predictors separately in single predictor models. In a second step, we entered all significant predictors from the single predictor models into one combined model.
With regard to the missing analyses, we attempted to distinguish between missings due to technical problems and missings due to participants’ lack of compliance. Technical problems were considered those instances when the iPod did not signal a beep, although an entry should have been made. We assessed compliance as the percentage of full entries measured by the total number of occurred signals. To test whether compliance systematically changed over the EMA period and whether certain patient characteristics significantly influenced the number of missings, we first calculated the sum of missed entries per day (minimum zero, maximum 4) within each participant. We entered the sum of missed entries per day as the outcome variable in a random intercept and slope HLM. Number of days was entered as the slope variable to account for time effects.

\[
\text{Sum of daily missings}_{ti} = \beta_{00} + \beta_{10} (\text{Number of days}_{ti}) + r_{0i} + r_{1i}
\]

Sum of daily missings\(_{ti}\) is the observed sum of missings on a particular day. The \(\beta_{00}\) parameter (intercept) is the average expected sum of missings on the first day. The \(\beta_{10}\) parameter (slope) is the expected change in missings per day. The random error term \(r_{0i}\) refers to the Level 1 residuals (i.e., deviations of the observed missing scores from the expected missing scores at each measurement occasion). The random error terms \(r_{1i}\) and \(r_{1i}\) refer to the Level 2 residuals of intercepts and slopes (i.e., individual differences in mean levels of missings and individual differences in the within-person regression slope). Again, we started with this unconditional base model, which we augmented with successive Level 2 models to account for individual differences between patients in the number of missings (intercept) and individual change in missings from the first to the last assessment (slope). We tested the following patient characteristics as predictors for differences in intercept and slope: age, gender, initial impairment, and perceived burden of daily assessments. The latter predictor was derived from the single item “Did you feel stressed by the daily surveys?” from
the EMA evaluation form. We added this Level 2 predictor to our missing analyses due to earlier studies, which found that the number of daily prompted diary entries correlated significantly with perceived burden of electronic diaries (Stone et al., 2003). Initial impairment was grand-mean centered, whereas age, gender, and perceived burden were entered uncentered. First, we entered all predictors separately in single predictor models.

To test whether the EMA period had a significant effect on impairment levels and should therefore rather be regarded as an intervention (research question 2c), we compared changes in impairment levels, measured by the BSI from the time of registration to the time of treatment onset, between the EMA sample and control group 2. Therefore, we first calculated changes in impairment levels by subtracting BSI pre-treatment scores from BSI scores at registration. Positive scores indicate drops in impairment levels. We then compared these scores via independent sample t-tests.

With regard to research question 3, we descriptively analyzed data from the EMA evaluation form.

3.3. Results
The following section displays data analytic results in the order of the research questions.

3.3.1. Sampling

*Research question 1*)

With regard to sample selection and daily assessments, can context variables be detected, which impede the successful implementation of EMA?

*Hypothesis a:* There are no systematic sampling selection effects.
In order to test whether systematic sampling effects occurred, we first compared our EMA sample to control group 1. Chi-square analyses showed no significant relation between group membership (EMA or control group 1) and gender, $X^2 (1, N = 933) = .03, p = .87$. Furthermore, no significant relationship between group membership and education $X^2 (6, N = 927) = 3.55, p = .74$ and no significant relationship between group membership and employment $X^2 (3, N = 933) = 1.75, p = .63$ could be found. With regard to the continuous variables, the results of our analyses showed that no significant differences in age occurred between the EMA sample ($M = 35.35, SD = 11.52$) and control group 1 ($M = 34.93, SD = 13.08$); $t(923) = -.24, p = .81$. Nevertheless, there were significant differences between the EMA sample ($M = 1.28, SD = .65$) and control group 1 ($M = 1.57, SD = .74$) with regard to initial impairment levels, $t(921) = 2.88, p = .004$. The EMA sample showed significantly lower impairment levels at registration compared to control group 1.

3.3.2. Technical Aspects

**Hypothesis b:** Data loss due to technical problems does not impede application and does not appear systematically.

With regard to hypothesis b, descriptive results indicated that across all participants and all days, 3063 signals for data entry were given. With 58 patients, a maximum number of 3248 assessments could have been reached without any technical problems and a compliance of 100%. With 3063 signals, 94.30% of the maximum number of signals was obtained, meaning that 185 prompts (5.70%) were not signaled, either because of technical problems (the iPod was fully charged and turned on but still no prompt was given) or because the iPod was turned off (actively by the patient or because of low battery). On the majority of days (90.0%), 4 prompts were signaled. 7.6% of the days had only 3 signals, 2.0% had 2 signals...
and only 0.4% of the days had only 1 signal. Descriptive values for the number of prompts per assessment within days did not reveal any systematic missings: 1st assessment of the day = 755 (24.6%), 2nd assessment of the day = 779 (25.4%), 3rd assessment of the day = 769 (25.1%), 4th assessment of the day = 760 (24.8%). A chi-square test of independence was performed to examine whether certain measurement points within days were missing systematically more often than others. No significant relationship was found $X^2 (4, N = 3063) = 4.53, p = 1.00$.

**Unconditional Model and Single Predictor Models Technical Missings**

To test whether the number of technical missings changed over time and can be predicted by participant variables, we first modeled the course of the daily number of signals over time in an unconditional model. The fixed effect estimates indicated an average number of signals of 3.72 on the first day across all participants. A non-significant fixed effect for the time slope indicated that no significant changes in the number of signals occurred from day one to day fourteen. Participants differed in their mean number of signals at the first session, as indicated by a significant random effect for intercept variation. Participants differed in the amount of change in the number of signals over time, as indicated by a significant random effect for slope variation.

**Table 3.2: Results of multilevel analyses predicting the daily number of signals with an unconditional random slope model (Model 1)**

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Fixed Effects</th>
<th>Random Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
</tr>
<tr>
<td>Daily # of signals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>3.72</td>
<td>0.06</td>
</tr>
<tr>
<td>Time Slope</td>
<td>.01</td>
<td>.01</td>
</tr>
</tbody>
</table>

*Note. N (Level 2) = 58. N (Level 1) = 808. Time as the number of days was entered uncentered as a Level 1 predictor. Coeff. = Regression Coefficient.* $p < .05$, *** $p < .001$. 

46
Results from the single predictor models indicated that initial impairment neither predicted intercept \((B = -0.11, p = .18)\) nor slope variance \((B = -0.00, p = .99)\) of the daily number of signals, when entered as a single predictor. The same holds for gender; neither intercept \((B = 0.11, p = .36)\) nor slope variance \((B = -0.00, p = .82)\) could be explained. The same result was found for age: no intercept effect could be detected \((B = 0.01, p = .37)\) and no slope effects could be detected \((B = 0.00, p = .39)\). Of the patient variables analyzed, none could explain differences in the number of technical missings between patients.

### 3.3.3. Reactivity and Compliance

*Research question 2)*

Can systematic missing analyses and control group comparisons detect reactivity effects that undermine the advantages of EMA?

**Hypothesis a:** Participants become faster in answering the EMA protocol over time, but reductions in time to answer only appear within the first days (training effect).

The average duration to answer a single assessment was 1 minute and 58 seconds (SD: 1 minute and 14 seconds).

*Unconditional Model and Single Predictor Models Duration*

To test whether the duration to answer the single assessments changed over time and can be predicted by participant variables, we first modeled the course of duration in an unconditional model. The fixed effect estimates indicated an average duration of 2 minutes and approximately 18 seconds at the first session and a mean rate of change of 0.77 seconds per assessment (see table 3.3). Participants differed in their mean duration rates at the first
session, as indicated by a significant random effect for intercept variation. Additionally, participants differed in the amount of decrease in duration over time, as indicated by a significant random effect for slope variation.

**Table 3.3:** Results of multilevel analyses predicting duration with an unconditional random slope model

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Fixed Effects</th>
<th>Random Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
</tr>
<tr>
<td>DURATION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>138.14</td>
<td>6.37</td>
</tr>
<tr>
<td>Time Slope</td>
<td>-.77</td>
<td>.08</td>
</tr>
</tbody>
</table>

*Note. N (Level 2) = 58, N (Level 1) = 3063. Time as the number of assessment was entered uncentered as a Level 1 predictor. Coeff. = Regression Coefficient.* p < .05, *** p < .001.*

Results from the single predictor models indicated that initial impairment neither predicted intercept \( (B = 2.02, p = .84) \) nor slope variance \( (B = -.12, p = .38) \) of duration, when entered as a single predictor. Gender could not explain intercept \( (B = 1.39, p = .92) \), but slope variance \( (B = -.39, p = .03) \). As we coded female participants with 0 and male participants with 1, this result indicates that the decrease in time that participants needed to answer the EMA questionnaire was greater for male participants compared to female participants. A similar result was found for age: no intercept effect could be detected \( (B = .68, p = .29) \), but differences in slope variance could be explained by age \( (B = -.02, p = .02) \). Older participants showed greater decreases in time to answer the EMA questionnaire over time.

**Combined Model**

We then entered age and gender – the significant predictors from the single models - into one full model (Model 2). The results indicate that when controlled for age, gender does not explain additional slope variance in duration (see Table 3.4).
Table 3.4: Results of multilevel analyses predicting duration to answer EMA assessments by demographic variables (Model 2)

<table>
<thead>
<tr>
<th>Model 2</th>
<th>Fixed Effects</th>
<th>Random Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
</tr>
<tr>
<td>Duration</td>
<td>Intercept</td>
<td>113.76</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>-3.48</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Time Slope</td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>-0.29</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

* Based on the assumption of normally distributed slope coefficients, this value indicates the estimated percentage of slope coefficients that are in the direction indicated by the algebraic sign of the regression coefficient for the significant predictors (Hox, 2010).

Table 3.4 indicates that older participants were estimated to show greater reductions in duration over time. When controlled for age, participants did not differ in reductions of durations over time, indicated by a non-significant slope intercept in Model 2. All differences between participants were explained by age differences. To examine the pattern of individual differences in more detail, we calculated the estimated percentage of slope coefficients that are negative (Hox, 2010, p. 19; see last column of Table 3.4). For 51.20% of the participants, age and the duration of filling out the single assessments were negatively related. To further analyze the reductions in duration from the first to last assessment, we conducted additional repeated measurement ANOVAs, comparing differences in mean durations per day. There was a significant effect of days of the EMA period, Wilks’ Lambda = 0.33, \(F(13.40) = 6.14\), \(p < .001\). Repeated contrasts between successive days revealed a significant reduction in duration from day 1 to day 2 (\(F(1) = 18.09\), \(p < .001\)) and from day 2 (\(F(1) = 11.11\), \(p = .002\)) to day 3. No significant reductions in duration could be revealed for the following days (3 to 4 = (\(F(1) = .06\), \(p = .81\)).
4 to 5 = (F (1) = 3.47, p = .07); 5 to 6 = (F (1) = .29, p = .59); 6 to 7 = (F (1) = .01, p = .92); 7 to 8 = (F (1) = .17, p = .68); 8 to 9 = (F (1) = .27, p = .60); 9 to 10 = (F (1) = .42, p = .52); 10 to 11 = (F (1) = 1.03, p = .31); 11 to 12 = (F (1) = .16, p = .70); 12 to 13 = (F (1) = .10, p = .75); 13 to 14 = (F (1) = .01, p = .93). Results are displayed in Figure 3.1.

**Figure 3.1:** Reductions in the duration to answer the EMA protocol from day 1 to day 14 of the EMA period. Duration is displayed as daily averages in seconds over all participants. **p < .005, *** p < .001.

**Hypothesis b:** Overall compliance is high and does not systematically change over the 14-day period, no patients need to be excluded because of lack of compliance and there are no effects of patient characteristics on compliance. The number of dropouts is low.

Overall compliance was high (92.89%). The individual range in compliance was 53.57 – 100%. Average number of missings due to participants over the entire EMA period was 3.98 (range 0– 26). No participant exceeded 50% of missings (recommended value for exclusion of
multivariate data analyses; Hair et al., 2010) and therefore no data set needed to be excluded from analyses. The average number of missings due to patients per day was .29 across all participants. Figure 3.2 shows the distribution of mean number of missings per day from day 1 – day 14.

**Figure 3.2:** Average number of missings per day from day 1 to day 14 of the EMA period

**Unconditional Model and Single Predictor Models Compliance**

To test whether compliance changed over time and can be predicted by participant variables, we first modeled the course of the sum of missed entries per day in an unconditional model. Table 3.5 illustrates the results of multilevel analyses predicting participant-induced missings.

<table>
<thead>
<tr>
<th>Table 3.5. Results of multilevel analyses predicting missings due to participants with an unconditional random slope model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MISSINGS</strong></td>
</tr>
<tr>
<td>Model 1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Day Slope</td>
</tr>
</tbody>
</table>

*Note. N (Level 2) = 58. N (Level 1) = 3063. Time as the number of days was entered uncentered as a Level 1 predictor. Coeff. = Regression Coefficient. * p < .05, *** p < .001.*
The fixed effect estimates indicated an average number of missings of .20 on the first day. Participants significantly differed in the number of missings, indicated by a significant random effect. The number of missings did not significantly change over the 14-day period, as indicated by a non-significant day slope.

Results from the single predictor models indicated that neither gender ($B = .01, p = .89$) nor age ($B = -.01, p = .08$), nor initial impairment ($B = .00, p = .99$), nor perceived burden of the daily assessments ($B = -0.05, p = .28$) explained intercept variance of missed entries when entered as single predictors.

**Figure 3.3:** Flow chart of recruitment and study participants

The study flow chart (Figure 3.3) indicates that one person dropped out during the EMA period. The reason for dropout given was “too high stress levels due to worrying about missing an assessment”. 

52
**Hypothesis c:** The improvement in impairment levels from registration for treatment to the onset of treatment does not systematically differ between the EMA group and a control group, meaning that no intervention effects of EMA on impairment levels can be detected.

Changes in impairment levels from the time of registration to the time of treatment onset could be assessed for 90 patients (41 from control group 2 and 49 from the EMA group). 26 patients did not start treatment and therefore no pre-treatment assessment was conducted. The average time between the time of registration and the onset of treatment was 140 days (SD approximately 19 weeks / four months). For the EMA group, the time between registration and onset was 150 days on average (SD 47.38). This corresponds to approximately 21 weeks or 4 months. The time between registration and onset was 128 (SD 53.28) days for control group 2. This corresponds to approximately 18 weeks, 4 months. On the BSI, the average change in impairment levels from registration to the pre-treatment assessment was .21 (SD .52). This corresponds to the waiting period effect scores usually found in the outpatient clinic (N = 1821, Mean = .23, SD = .55) For the EMA group, changes averaged out at .24 (SD .53) and for control group 2 at .19 (SD .51). Descriptively, this indicates a small difference in impairment drops in favor of the EMA group. Results from independent sample t-tests indicated no significant difference in impairment change between the EMA group and control group 2 from time of registration to treatment onset; t(88) = -.51, p = .62.

### 3.3.4. Acceptance and Burden for Patients

*Research question 3)*

What do patients say? Do patients’ ratings of acceptance and perceived burden after the EMA period support or oppose its implementation?
Figure 3.4 to 3.11 (page 55-56) display descriptive results of selected items from the EMA evaluation form (for all results, see Appendix B). Data from all 58 participants who completed the EMA period were analyzed.

Figure 3.4 shows that 91.38% of participants indicated that the EMA assessments helped them in some way (between slightly and very much). The majority of participants indicated that the assessments helped them either slightly (37.93%) or moderately (37.93%). When we take a closer look at the positive effects participants reported as a result of the regular assessments, it becomes clear that most participants profited from a more conscious perception of their emotions as a result of the assessments (89.68%, accumulated from slightly to very much). The distribution of participants who indicated having profited slightly (32.76%), moderately (29.31%) and quite a bit (22.41%) lies close together. 5.17% indicated having profited very much with regard to a more conscious perception of their emotions. 68.97% indicated that the assessments encouraged them between slightly and very much to pay more attention to their own interests and needs. Most participants (32.76%) indicated that they were slightly encouraged to pay more attention to their interests and needs. 53.45% indicated that the regular assessments had a positive effect on their mood (accumulated from slightly to quite a bit). Here, the range of chosen answers (from slightly to quite a bit) already indicates that the positive effects on mood are not as high as the aforementioned positive effects of the assessments.

When we take a look at the negative effects of the EMA application in our sample, the following picture forms: 60.34% indicated that they were stressed by the assessments (accumulated from slightly to very much). Nevertheless, most of these patients (46.55%) indicated that they were only slightly stressed. Only one participant (1.72%) indicated that he was stressed very much by the assessments. (Further analyses of this patient revealed that he had 0.00% missing data and – as the only participant in the sample - did not make use of the
postponement option, meaning that he answered every signal at the first beep.) 75.86% indicated that the beep signaling the next entry sounded at inconvenient times (accumulated from slightly to very much). For 32.76% of these 75.86%, the signals only sounded at slightly inconvenient times, compared to only 3.45% who indicated that signal times were very much inconvenient for them. The burden caused by inconvenience of signal times therefore appears rather low. Patients were asked whether they had worried about missing the beep, about forgetting the iPod and about handling the iPod incorrectly. Most worries were indicated about missing the beep. 77.59% indicated that they worried between slightly and very much. Still the percentage of participants who only worried slightly about missing the beep (32.76%) outweighs the percentage of people who worried a lot about missing a beep (5.17%). Only 20.69% of the participants did not worry at all about missing a beep. Figure 3.11 displays that only 37.93% indicated that the daily self-observations had a slight, moderate or quite a bit negative effect on their mood. This number must be interpreted compared to the aforementioned 53.45% of the participants who indicated that the assessments had a positive effect on their mood (also accumulated from slightly to quite a bit; see Figure 3.5). 62.07% of the participants indicated no negative effect on their mood. Nevertheless, the percentage of participants who neither indicated a positive nor a negative effect on their daily mood still appears comparably high (46.55% for positive effects, 62.07% for negative effects). In total these results suggest that if effects on daily mood occurred in our EMA application, then the positive effects outweighed the negative effects.
Figure 3.4: To what extent did the daily self-observation help you?

Figure 3.5: Did the daily self-observation have a positive effect on your mood?

Figure 3.6: Did the surveys lead to a more conscious perception of your emotions?

Figure 3.7: Did the surveys encourage you to pay more attention to your own interests and needs?
Figure 3.8: Did you feel stressed by the daily surveys?

Figure 3.9: Did the beep sound at inconvenient times?

Figure 3.10: Did you worry about missing the beep?

Figure 3.11: Did the daily self-observation have a negative effect on your mood?
No significant effects of reactions from others on compliance rates could be found in our sample. Whether participants indicated that others reacted positively \((r = -.18, p = .18)\), or negatively to the EMA \((r = -.11, p = .41)\), or that they had difficulties giving others an explanation for the regular assessments \((r = .03, p = .85)\) did not correlate significantly with compliance rates. Also, perceived stress level due to the assessments did not correlate significantly with compliance rates \((r = -.12, p = .38)\) (see also results from the multilevel analyses with regard to compliance rates).

The analysis of the open question answers on the EMA evaluation form reflects a diverse picture. 43 participants \((74.14\%)\) made use of the open question option in the evaluation form. 17 participants \((29.31\%)\) indicated technical problems with the iPod or problems in its handling (battery capacity, missing signals, shifting signal times, missing assessments and missing reminders after 15 minutes). 10 participants \((17.24\%)\) reported difficulties integrating the assessments into their daily routines (occupation, family or friends). 8 participants \((18.80\%)\) reported additional positive effects of the assessments (better identification of stress causing situations, opportunity for reflection, a better perception of small pleasant situations/things, support in structuring everyday routines, insights in associations between events and emotions, a more pronounced perception of emotions). 8 participants \((18.80\%)\) recommended adaptations of assessment times, for which reasons were varying (general wish for more flexible timing, dissatisfaction with the applicability of some items at the 8.00 am assessment, worries that the intervals of assessments did not fit to daily fluctuations, the wish for less but longer assessments and the wish for random interval assessments). 3 participants \((5.17\%)\) indicated difficulties with the formulation of the items or the applicability of the items to their daily experiences. 3 participants \((5.17\%)\) considered the financial compensation unnecessary. 2 participants \((3.45\%)\) indicated that they would have profited from an option for open feedback or comments at every assessment.
3.4. Conclusions

Study I served as an exploratory cost-benefit analysis of EMA applications in routine outpatient settings. The findings generally support its feasibility during the waiting period for outpatient treatment. In summary our analyses revealed the following results: (a) EMA participants were significantly less impaired compared to a matched control group, (b) the rate of technical missings was estimated at 5.7%, (c) participants became faster in answering the protocol within the first three days, (d) compliance rates were high with about 93% on average, (e) EMA participation had no significant effects on impairment levels and (f) patients’ evaluation of their EMA participation revealed a general satisfaction even though some critical aspects appeared, which can inform future applications. In the following section, the findings from Study I will be reiterated and conclusions will be drawn.

Results from the sampling selection analyses revealed that the participants of the EMA application showed significantly lower impairment levels at application for treatment compared to those patients that applied for treatment during the same time period and did not take part in the study (control group 1). This result suggests that a certain level of functioning might be required to agree to take part in an EMA study in an outpatient setting, where patients cannot be closely supervised during data collection. Since EMA strategies intend the opposite, namely not to personally supervise data collection, as this would mean an interference with the natural environment of the patient’s everyday life, the analyses indicate that one faces the risk of missing out on those patients who are more severely impaired. Further analyses on different outcome measures, which assess impairment levels, are needed to test whether our results can be replicated. If this is the case, then future studies conducted in outpatient settings must take different sampling strategies into account to avoid the risk of having a limited sample with regard to initial impairment level.
Analyses for technical missings indicate that 5.7% of prompts were not signaled. The interpretation of this number is complicated by the fact that the applied software (iDialogPad) did not allow the distinction of missing signals due to technical problems and missing signals due to the fact that the iPod was turned off (which can be participant initiated). The 5.7% must therefore be regarded as a mix of both causes. Importantly, missing signals were randomly distributed across time of day and day of assessment period. Participants differed in their number of technical missings, suggesting that handling was differentially manageable for the participants. This interpretation is supported by the results of the EMA evaluation form, in which the majority of participants indicated that they had no difficulties in handling the iPod (86.21%) but 12.06% indicated that they had slight or moderate difficulties. This is not surprising, given that we cannot distinguish between mere technical missings and missings caused by participant variables, we must assume that a certain amount of the 5.7% is caused by a lack of technical know-how. The sample selection procedure did not include variables that assessed the level of technical know-how. The age range, which covered a span of 40 years, hints at variations regarding the level of familiarity with newer technical equipment such as the iPod. A common concern in EMA applications is that older adult participants are unable or unwilling to use high-tech devices. However, there are several studies that show good results concerning compliance rates, even with participants over 80 years of age, when movement behavior was assessed (Bussmann, & Ebner-Priemer, 2013). Also, our results suggest that age itself, as a proxy for the level of familiarity with the handling of modern technical equipment, did not explain differences in the number of technical missings, making it is advisable to assess familiarity or technical know-how separately. When we compare our 5.7% to the percentage of missings due to technical problems reported by earlier studies with similar designs (e.g. Peters et al., 2000), it seems acceptably low and does not outweigh the advantages of the EMA application. It can still be
argued that over the years, technical developments should have improved software applications with accompanying decreases in the amount of missing data caused by technical problems, but reality rather reflects the fact that electronic devices are never entirely error proof and – as supported by our results – data may be lost due to technical problems.

One form of reactivity that might influence the number of answered signals, is the time that participants need to answer a single assessment. Reductions in time to answer can be interpreted in at least two different ways: on the one hand training effects can cause reductions in time to answer the questionnaire – participants get familiar with the handling and the formulation of the questions – on the other hand reductions in time can also be caused by fatigue effects – participants become tired of answering the questionnaire and begin checking random numbers. In our case, training effects would be expected within the first few days of the assessment period, whereas fatigue effects can occur over the entire assessment period, especially in latter stages of the assessment period. Our results indicate that reductions in time to answer the questionnaire occurred, but only within the first 3 days of the 14-day period. Following the aforementioned line of argumentation, reductions in duration from day one to day three without significant reductions on the following days support the occurrence of training effects. Older and male participants showed greater trainings effects indicated by drops in time to answer EMA questionnaires, although the effects for gender disappeared when controlled for age. This result highlights the need for adequate training and introduction to the handling of the electronic device before starting the EMA period. The average duration of approximately 2 minutes to answer the questionnaire seems reasonable. This conclusion is supported by participants’ opinions on whether or not it took too long to answer the questionnaire: the majority (89.66%) indicated that the input times where not at all too long (see Appendix B).
Compliance rates in EMA applications – measured as beep-wise response rates – are often very high at around 90% (e.g., Broderick et al., 2003; Jamison et al., 2001; Sokolovsky, Mermelstein, & Hedecker, 2014). Other studies report common compliance rates between 65 – 85% (Silvia, Kwapil, Eddington, & Brown, 2013). Our results replicate high compliance rates with an average of 92.89%. As opposed to other studies, which found that compliance rates drifted across the days of the study (e.g., Courvoisier, Eid, & Lischetzke, Pfeifer, Crayen, & Eid, 2012; Silvia et al., 2013) our results do not indicate significant changes in compliance rates over the 14-day period. Nevertheless, the individual range of compliance rates was high. However, this could not be explained by the analyzed person-variables, which had been shown to have a significant influence on or be correlated with compliance rates in earlier studies (e.g., gender in Messiah, Grondin, & Encrenaz, 2011 and Silvia et al., 2013; or perceived burden in Stone et al., 2003). We replicated the result from earlier studies (e.g., Courvoisier et al., 2012 and Silvia et al., 2013) that age did not have a significant influence on compliance rates, which alleviates concerns about limited applicability regarding the age of participants. Also, initial impairment levels did not influence compliance rates in our sample. As control group comparisons regarding sampling selection revealed a significantly lower impairment level of the study participants compared to other patients from the outpatient clinic (even though the range of impairment levels in the sample was reasonable), future studies in outpatient settings should reconsider impairment levels as a potential predictor or moderator of compliance rates. The case example in our study with no missing data and no use of the postponement function also suggests that other variables, which should be taken into account in future studies are personality characteristics such as perfectionism or accentuations in compulsivity. Levels of perfectionism, which are too pronounced, can increase the perceived burden for participants to an unbearable level, which is supported by the single case in our sample, who dropped out from the study because of this reason.
Apart from mere assessments, EMA strategies also allow electronically-mediated interventions (interactive assessments or ecological momentary interventions, see theoretical background). These applications involve either moment-specific feedback to the participant or treatment components (Trull et al., 2012). Even though our application of EMA did not involve either of the two, one could argue that the intensive self-focus triggered throughout the 14-day assessment period and also the personal contacts with the principal investigator in the outpatient clinic could have intervention-like effects for the participants. As we argued in the introduction, one consequence could be an increase of the waiting period effect. Our results indicate that even though we can descriptively observe differences in the waiting period effect between the study sample and a matched control group, these differences do not hold up against inferential testing. Future studies should investigate whether this result can be replicated on different outcome measures and with specific diagnostic subgroups. In general, this result supports our application as a mere assessment strategy, however the high percentage of patients who indicated that the EMA application has helped them in some way (91.38%) and the individual feedback from patients after the end of the EMA-period suggests that certain patients experienced positive effects of the assessments, which are comparable to effects reported from cognitive behavioral therapy (e.g., “helped me structure my everyday routines”, “better identification of stress causing situations”, “opportunity for reflection” or “more conscious perception of emotions”). Further research on the question: “What patient profits in which way from an EMA-application?” could help design EMA applications that can be integrated as a primary intervention in stepped care treatment programs (e.g., Bower, & Gilbody, 2005).

The fact that participants reported satisfaction with regard to their study participation supports EMA applications during the waiting period for outpatient treatment. In-depth analyses of the positive and negative consequences of the application suggest that the positive
effects outweigh the negative effects and seem to compensate the participants for the burden of the high frequency assessments. On a global level, the percentage of participants who indicated that the assessments helped them (91.38%) outweighs the percentage of participants who indicated that they were stressed by the assessments (60.34%). However, the range of perceived stress level between different participants was greater (from *not at all* to *very much*) than the range of the level of perceived help (from *not at all* to *moderately*). This result again suggests that more research on specific patient characteristics, which facilitate or impede EMA applications in specific patient subgroups, is needed.

Among the study’s important findings, which can inform the future design of EMA applications in outpatient settings, are the correlation (and regression) results of perceived burden due to the assessments (e.g., by perceived stress levels or negative reactions from others) and compliance rates. We did not find significant effects of perceived burden on compliance rates, which suggests that the number of daily assessments did not overstrain participants. Four daily assessments seem to be an advisable number of assessments for patients during the waiting period for outpatient treatment. Still, more importantly, the timing of the assessments must reflect the nature of the assessed construct (Conner, & Lehman, 2012) so that fluctuations can be captured. The chosen number of daily assessments must always be a result of thorough balancing between reasonable levels of burden for participants and the degree of micro-level insights in the phenomenon of interest.

Our results highlight the importance of assessing participants’ evaluation after an EMA application. The high number of participants who made use of the opportunity to give an open feedback at the end of the evaluation form also supports the integration of an open feedback question in post EMA evaluation forms. Important information on unexpected positive effects (e.g., “helped me structure my everyday routines”) would have otherwise been missed and can now help guide future applications in outpatient settings.
When interpreting the results of our study, a few limitations should be taken into account: first, the small sample size causes a lack of power, so that our results can only serve as exploratory results and should be tested in future studies. Still, our results do not deviate from other studies that tested feasibility in outpatient settings (Helbig et al., 2009) with a significantly smaller sample size. Second, our sample consisted of patients with different primary diagnoses (mainly affective or anxiety disorders). Even though these are the most common psychological disorders in Germany (Jacobi, Klose, & Wittchen, 2004), we do not know whether our results generalize to other patient groups. Third, the study design may have influenced the participants’ evaluation of their study participation. Aside from screening before study onset, the principal investigator of the study was the only contact person for study participants, meaning that the same person introduced them to the study and the handling of the iPod, had phone contact during the assessment period, ended the assessment period, was therefore the recipient of the evaluation form and paid them their financial compensation for participating in the study. Although the compensation was paid at the very end, after participants had completed the evaluation form, the fact that participants knew that they would receive the money in the end might have resulted in a more positive evaluation. Fourth, we have only investigated feasibility from the participants’ point of view. The other point of view we have not highlighted is feasibility for the researcher who applies EMA. There are several obstacles that might prevent a researcher from applying the method. First, the costs of electronic devices must be considered. Although technical developments are rapid, lowering the cost of smartphones, handheld devices or other equipment that can be used for data assessment, they are all still more expensive than paper-pencil solutions. Second, the application of the method requires statistical knowledge and experience in sophisticated strategies of data analysis (e.g., multilevel analysis). Third, ethical considerations may arise. Privacy could be a concern, as with all electronic devices and communications used to
document health status. In applications where participants actively contribute to data collection, most concerns can be allayed by using password-protected devices and protocols, data encryption, secure servers to house data and deactivating real-time transfer from the electronic device to the server (Trull et al., 2012). Fourth, the number of validated item sets of different constructs for EMA applications is very limited, leaving the researcher with the unresolved question of which item sets to apply which leads to a high diversity which again impedes comparability between studies (also refer to the general discussion for this aspect). All these aspects must be considered and might hinder researchers in using EMA applications, despite the vast potential they offer.

In summary, the results of our analyses support the feasibility of EMA applications in outpatient settings. Application during the waiting period helps patients to fill the gap between registration for treatment and its actual onset. Additionally it might open a window in which important pieces of information can be collected to inform and enhance clinical practice (see study II and III). Our results can inform the planning, design, sampling strategies and implementation of future studies that apply EMA in outpatient settings.
4 STUDY II: THE RUMINATION-AFFECT LINK IN EVERYDAY LIFE

4.1. Introduction

While rumination has long been considered a causal and maintaining factor for depression, more recent research suggests that it can equally be associated with other disorders, such as anxiety (e.g., Nolen-Hoeksema, 2000; Ruscio, Gentes, Jones, Hallion, Coleman, & Swendsen, 2015). A meta-analytic review of emotion-regulation strategies across psychopathology showed that rumination was not only linked to depression and anxiety but also to eating, and substance-related disorders (Aldao, Nolen-Hoeksema, & Schweizer, 2009). Rumination has been linked to negative affect (NA) in different ways (Martin, & Tesser, 1996). Following Nolen-Hoeksema’s definition of rumination as “repetitive and passive thinking about one’s symptoms […] and the possible causes and consequences of these symptoms” (2004, p. 107), rumination can be broadly defined as repetitive negative thinking, especially since recent elaborations on rumination have extended the concept to include attention to negative life events an individual has experienced (Alloy et al., 2000; Smith, & Alloy, 2009). Repetitive negative thinking seems to be causally involved in the maintenance of several emotional disorders and some researchers even consider it to be a transdiagnostic process linked to nearly all Axis I disorders (Ehring, & Watkins, 2008).

Rumination has been studied in the context of stress, where it leads to heightened physiological response, delayed recovery from the stressor, increased reactivity to subsequent stressful events and increased general stress sensitivity (e.g., Watkins, 2004; Watkins, Moberly, & Moulnds, 2008; Zoccola, Quas, & Yim, 2010; Ruscio et al., 2015). The interaction of rumination and stress prospectively predicts psychological distress (e.g., dysphoria and hopelessness) (Morrison, & O’Connor, 2008). Studied in the context of depression and anxiety, rumination has been shown to predict negative affect and depressive symptoms (Nolen-Hoeksema, 2000). Furthermore, rumination has been found to influence the duration
of symptoms, the characteristics of and recovery from depressive episodes and has been linked to the emergence of new depressive episodes in many studies (e.g., Just, & Alloy, 1997; Nolan, Roberts, & Gotlib, 1998; Nolen-Hoeksema, & Morrow, 1991; Nolen-Hoeksema, Morrow, & Fredrickson, 1993; Nolen-Hoeksema, Parker, & Larson, 1994; Nolen-Hoeksema, 2000; Kuehner, & Weber, 1999). Rumination seems to be linked to an array of different negative emotions, such as sadness (Lyubomirsky, & Nolen-Hoeksema, 1993), anxiety (McLaughlin, & Nolen-Hoeksema, 2011) and irritation (Thomsen, Mehlsen, Christensen, & Zachariae, 2003).

The concept of rumination has long been studied in laboratory or cross-sectional settings. These approaches are limited to research questions that explore between-person associations of rumination and psychological symptoms (e.g., Is trait rumination related to depressive symptoms?), leaving out important within-person questions (e.g., Can increased rumination following a stressful life event predict mood changes on subsequent measurement occasions?). The intrapersonal association between rumination and affect and how this relation unfolds in everyday life can lead to important insights concerning individual etiological models of disorders. EMA applications allow the exploration of the day-to-day associations of rumination and affect over extended time periods. Several EMA studies have already explored different aspects of rumination in daily life. An early and influential study by Moberly and Watkins (2008) for instance, showed that ruminative self-focus was positively related to negative affect and also predicted negative affect at subsequent measurement occasions. Another EMA study by Genet and Siemer (2012) found that rumination moderates the relation between unpleasant daily events and negative mood. On days with increased rumination, higher levels of unpleasant daily events predicted higher levels of negative mood. Similarly, Brans, Koval, Verduyn, Lim, & Kuppens (2013), who studied the use of different emotion regulation strategies and their consequences for positive
and negative affect, found that rumination was associated with increases in NA and decreases in PA within-persons. The results from this study suggest that the effects of rumination on mood are not limited to negative affect but also incorporate changes in positive affect. Huffziger et al. (2013) also found that induced rumination in daily life led to mood deterioration (measured by valence and calmness). Importantly the results of this association were independent of depressive symptoms. A study by Ruscio et al. (2015), which focused on individuals diagnosed with major depressive disorder and generalized anxiety disorder, found that increased rumination in response to a stressful event predicted poorer affect, more maladaptive behavior and more disorder specific symptoms at the following measurement occasion.

The reported findings support a strong relation between rumination and affect in daily life and therefore underline the negative effects of this maladaptive emotion regulation strategy. A shortcoming of previous studies is that they usually apply EMA in specific samples – either non-clinical or diagnosis specific samples (e.g., depression or anxiety). To our knowledge, rumination has not yet been studied in a mixed outpatient sample of individuals that sought treatment. Additionally, only few studies have assessed rumination both retrospectively with commonly applied questionnaires and real-time via EMA. A combination of both assessment approaches allows differentiated analyses of moderation effects of trait rumination on the relation between state rumination and affect. Moberly and Watkins (2008), who applied both assessment strategies, found that dispositional rumination was associated with mean levels of momentary ruminative self-focus. How dispositional rumination influenced the association between rumination and affect via cross-level interactions was, however, not investigated. Study designs that combine retrospective and real-time assessments of rumination not only allow the analysis of differences between trait and state aspects of the construct, but also allow testing whether retrospective and real-time
measures come to similar conclusions regarding the dependent construct. While many EMA studies have focused on the effects of rumination on negative affect, we want to capture the effects on both negative and positive affect simultaneously (as suggested by study results from Brans et al., 2013). Therefore we applied the composite positive-negative affect ratio of positive (PA) over negative (NA) affect (i.e., PA/NA) as our outcome variable.

The following analyses attempt to extend the research on rumination to the context of different diagnostic subgroups in an outpatient setting. We attempt to add to the findings on how rumination concurrently and subsequently affects affect within-person and whether these relations are moderated by dispositional rumination. Consequently the present study tests:

1. Whether state rumination and concurrent positive – negative affect are negatively related within-person.

2. Whether state rumination predicts a decrease in subsequent positive – negative affect.

3. Whether dispositional rumination moderates the within-person relationship between state rumination and positive-negative affect.

4.2. Methods

4.2.1. Sample

The data from all 58 study participants (described in section 4.1.) could be analyzed for research questions 1 and 2. One participant needed be excluded from the analyses of research question 3, as we did not acquire RSQ-D data from this participant.

4.2.2. Instruments

Dispositional rumination

We analyzed data from the two subscales self-focused and symptom-focused rumination from the RSQ-D, as described in section 2.3.1.
State rumination

State rumination was assessed by all four EMA rumination items, which are described in section 2.3.2. High values indicate high state rumination.

Positive-negative affect

As a composite indicator of positive-negative affect, we computed the quotient of positive and negative affect. We therefore averaged the sum of PA items and the sum of NA items (PA/NA). Scores above 1 indicate that, at a certain measurement occasion, PA outweighs NA. Higher scores indicate better (more positive) affect.

4.2.3. Statistical Analysis

Our data exhibited a nested structure, in which measurement occasions (Level 1) were nested within persons (Level 2). To analyze within-person relationships, we again adopted a multilevel modeling strategy. To test whether state rumination is related to positive-negative affect, we constructed the following two-level model (see also Figure 4.1, Model 1):

\[ Affect_{it} = \beta_{00} + \beta_{10}Rumi_{it} + \beta_{20}Day_{it} + r_{0i} + r_{1i}, \]

where \( Affect_{it} \) is individual \( i \)’s positive-negative affect score at measurement occasion \( t \); \( \beta_{00} \) is the mean intercept; \( Rumi_{it} \) is individual \( i \)’s momentary rumination score at measurement occasion \( t \) with corresponding slope coefficient \( \beta_{10} \). To account for time effects, we added days within the assessment period as an additional predictor. The random effects comprised the Level 1 residuals, (i.e., deviations of the observed affect scores from the expected affect scores at each measurement occasion), the Level 2 residuals of intercepts (i.e., individual differences in mean affect levels) and – in case of a significant deviance test that compared the random slope model with a more restrictive random intercept model – the Level 2 residuals of slopes (i.e., individual differences in the within-person regression slope). We
group-mean centered the continuous predictor variables in all analyses, as our interest was in “pure” within-person relationships (Enders, & Tofighi, 2007).

Figure 4.1: Path diagrams of the analyzed Level 1 relationship between state rumination and positive-negative affect. Time effects are not depicted. The line indicates that the measure of state rumination refers to the time period since the last measurement occasion. Arrows without circles represent fixed effects, arrows with circles represent random effects, an arrows not originating from a variable indicates Level 1 residuals. $Rumi_{it}$: State rumination of individual $i$ at measurement occasion $t$. $Affect_{it}$: Positive-negative affect of individual $i$ at measurement occasion $t$. $Affect_{it-1}$: Lagged positive-negative affect. $Rumi_{it-1}$: Lagged state rumination.

To analyze whether state rumination is related to a decrease in positive-negative affect in successive measurement occasions, we added lagged positive-negative affect to the model as a group-mean centered Level 1 predictor (see Figure 4.1, Model 2). We only included within-day lags, since overnight lags represent structurally different lags with longer time
periods and intervening night’s sleep. We also excluded days with only one measurement occasion, as they did not allow the creation of lags. To analyze whether state rumination since the last signal also predicted mood changes in the subsequent time interval, we analyzed a third model (see Figure 4.1, Model 3). In this model, lagged state rumination and lagged positive-negative affect were entered as predictors of positive-negative affect. Again, only within-day lags were analyzed.

To test whether dispositional rumination moderates the within-person relation between state rumination and affect, we added grand-mean centered dispositional rumination (self-focused and symptom-focused rumination) as Level 2 predictors of the intercept and the rumination slope coefficients.

4.3. Results

4.3.1. Descriptive Results

Descriptive statistics for all variables can be found in table 4.1.

### Table 4.1: Descriptive Statistics

<table>
<thead>
<tr>
<th>Measure</th>
<th>Items</th>
<th>Scale</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training Session</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dispositional self-focused rumination</td>
<td>7</td>
<td>1-4</td>
<td>15.49</td>
<td>4.46</td>
<td>7</td>
<td>28</td>
</tr>
<tr>
<td>Dispositional symptom-focused rumination</td>
<td>8</td>
<td>1-4</td>
<td>19.12</td>
<td>5.25</td>
<td>9</td>
<td>28</td>
</tr>
<tr>
<td><strong>EMA period</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State rumination</td>
<td>4</td>
<td>1-9</td>
<td>12.66</td>
<td>8.85</td>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>Positive-negative affect</td>
<td>8</td>
<td>1-5</td>
<td>1.69</td>
<td>1.03</td>
<td>.21</td>
<td>5</td>
</tr>
</tbody>
</table>

*Note. For EMA period variables, means represent averaged momentary scores across individuals and measurement occasions. Items = Number of items in the scale. Scale = potential range of scale scores.*

The mean score of approximately 15 on the self-focused rumination scale corresponds to a percentile rank of 40 in a depressive sample (i.e., 60% of depressed patients have the same or higher values in self-focused rumination), whereas it corresponds to a percentile rank of 60 in a non-clinical sample (i.e., 40% of non-clinical participants have the same or higher values in self-focused rumination). 52.6% of the sample received a percentile rank of 50 (depressed
sample) or above on self-focused rumination. The mean score of approximately 19 on the symptom-focused rumination scale corresponds to a percentile rank of 40 in a depressive sample (i.e., 60% of depressed patients have the same or higher values in symptom-focused rumination), whereas it corresponds to a percentile rank between 80 and 90 in a non-clinical sample (i.e., between 10 and 20% of non-clinical participants have the same or higher values in symptom-focused rumination). 50.9% of the sample received a percentile rank of 50 (depressed sample) or above on symptom-focused rumination.

Descriptive results for state rumination indicate that the full range of state rumination scores was obtained within the EMA period. Mean values for positive-negative affect indicate that PA outweighs NA across all participants. This is true for 74.8% of the valid measurement occasions across all participants.

State rumination and the composite positive-negative affect correlate significantly within measurement occasions ($r = -.45$, $p = .00$). The more PA outweighs NA, the less rumination on average.

### 4.3.2. Relation between State Rumination and Positive-Negative Affect

To test whether state rumination and positive-negative affect were negatively related within participants, we constructed a two-level model predicting affect by time and state rumination (depicted in Figure 4.1, Model 1). The results for this model (Model 1) can be found in Table 4.2.
### Table 4.2: Results of multilevel analyses predicting positive-negative affect by state rumination

<table>
<thead>
<tr>
<th>Model</th>
<th>Dependent variable</th>
<th>Predictor</th>
<th>Fixed Effects</th>
<th>Random Effects</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Coeff.</td>
<td>SE</td>
<td>t-ratio</td>
</tr>
<tr>
<td>Model 1</td>
<td>Positive-negative affect</td>
<td>Intercept</td>
<td>1.70</td>
<td>.10</td>
<td>17.84***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>State rumination</td>
<td>-.05</td>
<td>.01</td>
<td>-9.34***</td>
</tr>
<tr>
<td>Model 2</td>
<td>Positive-negative affect</td>
<td>Intercept</td>
<td>1.75</td>
<td>.10</td>
<td>18.12***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>State rumination</td>
<td>-.04</td>
<td>.01</td>
<td>-9.06***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lagged positive-negative affect</td>
<td>.27</td>
<td>.03</td>
<td>8.32***</td>
</tr>
<tr>
<td>Model 3</td>
<td>Positive-negative affect</td>
<td>Intercept</td>
<td>1.75</td>
<td>.10</td>
<td>18.12***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lagged state rumination</td>
<td>-.01</td>
<td>.00</td>
<td>-2.51*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lagged positive-negative affect</td>
<td>.33</td>
<td>.03</td>
<td>10.13***</td>
</tr>
</tbody>
</table>

*Note. N (Level 2) = 58 for Model 1, 2, & 3. N (Level 1) = 3019 for Model 1 and N (Level 1) = 3009 for Model 2, & 3. Time was entered as a group-mean centered predictor on Level 1. All continuous Level 1 predictors were group-mean centered. Coeff. = Unstandardized regression coefficient. Stand. Est. = standardized estimate (standardized regression coefficient).

* Based on the assumption of normally distributed slope coefficients, this value indicates the estimated percentage of slope coefficients that are in the direction (positive or negative) indicated by the coefficient (Hox, 2010).

* p < .05. *** p < .001.
The overall intercept (1.70) represents the predicted affect at measurement occasions in the middle of the assessment period, at which individual’s rumination was average (because the continuous predictors were centered). As expected, state rumination negatively predicted positive-negative affect. Time (days within the assessment period) did not turn out to be a significant predictor for the outcome variable in any of the models and is therefore not included in Table 4.2. The within-person relation between state rumination and affect varied across participants, as indicated by a significant random effect for state rumination (see Table 4.2). To examine the pattern of individual differences in more detail, we calculated the estimated percentage of slope coefficients that are negative (Hox, 2010, p. 19; see last column of Table 4.2). For 95% of the participants, state rumination and positive-negative affect were negatively related within the same time interval.

We extended Model 1 to test whether the relation between state rumination and positive-negative effect held, when positive-negative affect at the previous measurement occasion was controlled for (see Figure 4.1, Model 2). We therefore added lagged positive-negative affect to the model as a group-centered Level 1 predictor. As state rumination represents the extent to which an individual showed ruminative behavior since the last signal, this analysis allowed us to examine, whether an increase in state rumination was associated with a change in positive-negative affect. This analysis was based on 3009 measurement occasions (only within-day lags). Results can be found in Table 4.2 (Model 2). Even when lagged positive-negative affect was controlled for, the average slope coefficient of state rumination remained almost the same. This means that state rumination is related to a negative change in positive-negative affect (i.e., a worsening of affect).

Model 3 (Table 4.2) depicts results from the analyses, which tested whether rumination since the last signal also predicted affect changes in the subsequent time interval. As the results indicate, state rumination predicted subsequent changes in positive-negative
affect beyond lagged positive-negative affect. There were no individual differences in this within-person association, as indicated by a non-significant random effect for the state rumination slope. This means that state rumination since the last signal predicted a negative change in positive-negative affect (i.e., worsening of affect) from one measurement occasion to the next, independent of preceding affect ratings. This effect was, however, somewhat smaller than in Model 2, where concurrent affect change (across the same time period) was analyzed.

4.3.3. Moderator Effect of Dispositional Rumination

To analyze whether dispositional rumination moderated the within-person association between state rumination and positive-negative affect, we extended Model 1 and added dispositional self-focused and symptom-focused rumination as Level 2 predictors for the varying intercepts and varying rumination slope coefficients. Dispositional self-focused rumination neither predicted the intercept ($B = -.01, p = .60$), nor the rumination slope ($B = -.01, p = .60$). Dispositional symptom-focused rumination negatively predicted the intercept ($B = -.05, p < .05$), indicating that individuals who retrospectively indicated having ruminated less on their symptoms, experienced better affect (more PA compared to NA) across the EMA period. However, dispositional symptom-focused rumination had no significant influence on the rumination slope ($B = .00, p = .09$).

We reran the same model, replacing dispositional rumination (measured by the RSQ-D) with aggregated state rumination (i.e., mean state rumination for each participant across the EMA period) as a Level 2 predictor on the varying intercepts and rumination slope. We obtained similar results with regard to the intercept. Aggregated state rumination negatively predicted the intercept ($B = -.05, p < .05$), indicating that participants who ruminated less across the EMA period also experienced better affect (more PA compared to
NA) across that time period. Aggregated state rumination, however, also generated significant cross-level interactions. Aggregated state rumination positively predicted the rumination slope \((B = .01, p = .03)\). Aggregated state rumination dampens the influence of state rumination on positive-negative affect, i.e., participants who ruminated more across the EMA period displayed decreased effects of state rumination on positive-negative affect. When both person-variables (symptom-focused rumination from the RSQ-D and aggregated state rumination) were entered simultaneously as Level 2 predictors, both negatively predicted the varying intercepts (symptom-focused: \(B = -.04, p < .05\); aggregated rumination: \(B = -.05, p < .001\)), and again, the cross-level interaction term was only significant for aggregated state rumination \((B = -.01, p = .04)\).

### 4.4. Conclusions

Our analyses replicated and extended findings from previous research on the within-person associations between state rumination and affect. In summary our analyses revealed the following results: (a) a negative within-person relationship between state rumination and positive-negative affect in everyday life for 95% of the sample, (b) state rumination predicts negative changes in affect independent of previous affect concurrently and subsequently (c) dispositional symptom-focused rumination predicts mean levels of positive-negative affect whereas aggregated levels of rumination as a group-mean centered level 2 predictor not only predicts mean levels of positive-negative affect but also displays cross-level effects for the rumination slope. The latter result indicates a dampening effect of aggregated rumination on the state rumination–affect association. (d) Individual restrospective indicators of self-reported rumination and real-time indicators of rumination both explain variance in affect independent of each other.
Our analyses extended findings from previous research in four important ways. First, we included a mixed sample with different diagnoses and comorbidity, as the literature suggests that rumination is an important factor (or possibly mechanism) for the maintenance and the course of different Axis-I disorders. Second, as preliminary results suggested that rumination is not only linked to negative, but also to positive affect (Moberly, & Watkins, 2008), we used a composite measure of positive-negative affect as our outcome variable. Third, our design allowed the additional analysis of the effects of dispositional rumination on the within-person relationship between state rumination and affect. Fourth, in-depth analyses of the individual differences of within-person associations between state rumination and affect allowed a statistical evaluation of the dimension of the influence of rumination on affect. In the following section, the findings from Study II will be reiterated and conclusions will be drawn.

Our analyses revealed a negative within-person relationship between state rumination and positive-negative affect in everyday life, thereby conceptually replicating findings from previous EMA studies (e.g., Brans et al., 2013; Huffziger et al., 2013; Moberly, & Watkins, 2008; Ruscio et al., 2015). Significant random effects for the slope term however indicated that individuals differed in the extent to which state rumination influenced positive-negative affect. Our results revealed that for 95% of the sample, the association between state-rumination and affect was negative. That means that for 5% of the sample higher state-rumination corresponded to better affect (more PA than NA). This result potentially has critical implications for the diagnostic of individual psychopathology but also for delivering personalized interventions. A standardization of treatment content and delivery can result in interventions that are both incomplete and/or counterproductive – as it would have been the case for 5% of our sample if interventions were applied targeting rumination in order to increase PA (see also section 6.2. in the general discussion for this aspect).
As previous EMA studies have suggested that a reciprocal relationship may exist between rumination and affect – in which rumination predicts deterioration in affect, but negative affect also predicts increases in rumination (Moberly, & Watkins, 2008) – we controlled for positive-negative affect at previous measurement occasions to extract pure influences of rumination on affect. Our results suggest that state rumination alone is associated to a change in positive-negative affect, i.e., decreases in PA and concurrent increases in NA. This result is in accordance with earlier findings, which have shown that the within-person relation between rumination and affect is independent of other depressive symptoms (Huffziger et al., 2013). Importantly, our analyses show that state rumination predicted subsequent changes in positive-negative affect beyond lagged positive-negative affect. Independent of previous affect ratings, rumination predicts negative changes in affect in subsequent time intervals. As we analyzed time intervals of a maximum of 5 hours, this association suggests that rumination not only shows immediate effects on affect, but also effects, which last for about one third of an adult’s daily waking hours. Furthermore, our results suggest that the effects of this association do not differ across individuals with different Axis-I disorders.

Results from the moderator analyses of dispositional rumination indicated differential effects of self- vs. symptom-focused rumination. Only dispositional symptom-focused rumination positively predicted mean levels of positive-negative affect. Symptom-focused rumination covers rumination about symptoms and their consequences, whereas self-focused rumination comprises e.g., self-reproaches, withdrawal or brooding over the question, why one suffers from mental illness (Kuehner et al., 2007). This result contradicts earlier findings, which suggest that brooding is associated with higher mean levels of negative affect (Moberly, & Watkins, 2008). We therefore reran the same model with aggregated state rumination. Similarly, person-specific aggregated levels of rumination negatively predicted
mean levels of positive-negative affect. Additionally, in this model, we found cross-level effects for the rumination slope, indicating a dampening effect of aggregated rumination on the state rumination – affect association. The negative effects of state rumination on positive-negative affect seem to be less pronounced for individuals who rather describe themselves as the ruminating type (almost like a habituation effect). As a personal characteristic, rumination may therefore be a variable, which can explain individual differences in the extent to which state rumination influences affect within a certain time interval.

Results from the combined model with aggregated rumination and dispositional rumination as Level 2 predictors showed that both predictors significantly explained variance independent of each other. This is an interesting result that can be interpreted in different ways: either each of the Level 2 predictors are different indicators, capturing different aspects of the same underlying construct, or they capture different underlying constructs. The latter assumption is related to the phenomena of repeatedly observed discordance between retrospective questionnaires and real-time measures. Conner and Feldman Barrett (2013) argue that retrospective self-reports and ambulatory self-reports capture qualitatively different aspects – one focusing on memories and beliefs and the other on experiences. Depending on the research objectives, one must decide whether the focus should rather lie on the experiencing self (ambulatory self-reports) or the remembering and believing self (retrospective self-reports), to obtain maximum relevance to one’s own research question.

Several limitations of our study should be mentioned: First, our sample consisted of individuals seeking outpatient treatment; therefore results cannot be generalized to non-clinical or inpatient samples. Additionally, control group comparisons indicated that our sample was slightly less impaired than the usual patient treated at our outpatient clinic. Future research in this area would profit from collecting more diverse samples with regard to initial
impairment levels. Furthermore, the sample size did not allow comparisons between different diagnoses – our results suggest that the rumination - affect link applies across 95% of the individuals of our sample diagnosed with different Axis-I disorders, but future studies should analyze whether a moderator effect of different diagnoses can be identified.

To summarize, the results of our analyses added to previous studies showing a strong within-person association between rumination and affect. It extended previous findings, which were limited to specific diagnostic subgroups, supporting the assumption that rumination may be considered a transdiagnostic factor (Ehring, & Watkins, 2008). Additionally, results from the moderator analyses helped elucidate personal characteristics that explain which individuals are more prone to an affect-deteriorating effect of rumination on momentary and subsequent affect. Clinical implications drawn from our results highlight the importance of targeting rumination as an intervention in treatment approaches for varying Axis-I disorders; but at the same time Study II underlines that – even though it applies to the majority of the sample – the within-person link between rumination and affect varies across individuals. The latter finding highlights the importance of idiographic diagnostic approaches in order to realize the most effective individualized treatments.
5 STUDY III: Daily affect dynamics predict early response in CBT

5.1. Introduction

The effectiveness and efficacy of empirically based treatments have been shown time and again (Lambert, 2013). Recently, a new topic of research has emerged in the literature, which addresses individual patterns of change in response to empirically validated treatments. There are replicated findings, which support that the general pattern of treatment response can be described – as postulated by the dose-effect model (Howard, Kopta, Krause, & Orlinsky, 1986) – as a log-linear decrease in impairment levels with increasing number of sessions (e.g., Lambert, Hansen, & Finch, 2001; Stulz, Lutz, Kopta, Minami, & Saunders, 2013). However, even if we assume this general pattern, we know that individual patients differ in their change trajectories and deviate from the general trend. Studying the phenomenon of change, as it occurs at the beginning of treatment (early change), has gained increasing interest. Within this research approach, target behaviors, e.g., depressive symptoms over time, are used to predict treatment outcome (Cuijpers, van Lier, van Straten, & Donker, 2005; Lutz, Stulz, & Köck, 2009). Early change patterns have been shown to be associated with outcome across different diagnoses (Bradford et al., 2011; Hayes, Feldman, Beevers, Laurenceau, & Cardaciotto, 2007; Lewis, Simons, & Kim, 2012), different treatment approaches (Crits-Christoph et al., 2001; Gunlicks-Stoessel, & Mufson, 2011) and different measures (Hunter, Muthén, Cook, & Leuchter, 2010). Rubel, Lutz, and Schulte (2015) were able to identify different patient subgroups with regard to their early change patterns in an outpatient sample with varying diagnoses. They could also show that most change in patients’ progress scores took place in an early phase of the treatment. Lutz and colleagues (2014) investigated early change patterns (first 5 sessions) in patients with panic disorder (N=326) who underwent cognitive behavioral therapy. Using growth mixture modeling (GMM), one group could be identified, which was characterized by high initial impairment and rapid early improvement. Class membership
predicted outcome and early treatment termination. Although we know that early response is a critical predictor for treatment outcome (Lutz et al., 2009; Lutz et al., 2014; Nordberg, Castonguay, Fisher, Boswell, & Kraus, 2014) and the described findings show that patients differ with regard to early change patterns, we cannot satisfactorily explain how these differences in early response arise (e.g., Lutz et al., 2014; Rubel, Lutz, Kopta, Köck, Minami, Zimmermann, & Saunders, 2014; Lambert, Whipple, Bishop, Vermeersch, Gray, & Finch, 2002).

What the previous studies, which explored the effects of patient characteristics and change trajectories have in common, is that they use intake measures as predictors of early response. Research, which takes the situational and context-specific cognitive, emotional, and behavioral reactions into account in a dimensional way, may improve and deepen our knowledge of processes that cause disorders. Thus, assessment timeframes must be expanded.

There is a contradiction between prevailing diagnostic criteria and our standard methods of clinical assessment: Diagnostic criteria often describe dynamic symptomatology over a certain period of time (e.g., manifest symptoms over a two-week period in depressed patients), whereas the assessments of these criteria typically apply clinical interviews or self-report questionnaires, which ask patients to rate symptom distress retrospectively. This discordance has already caused researchers to call for dynamic models of psychological assessment, in order to improve personalized care (e.g., Fisher, 2015; Trull et al., 2012). EMA applications allow dynamic and longitudinal assessments of different aspects of psychological disorders. The various advantages of the method with regard to the exploration of context specific intra- and interindividual differences in experience and behavior, the frequency and intensity of activities and the accompanied dynamic psychological processes have already been illustrated in the theoretical background of this thesis. It was argued that gaining deeper insights into the daily routines and experiences of an individual patient may help to better
understand both health and pathology and guide decision-making regarding treatment intervention options (Shiffman, Stone, & Hufford, 2008). Following this line of argumentation, data from EMA gathered before treatment onset may enhance individualized prediction models with regard to treatment response and outcome.

Hofmann, Sawyer, Fang, and Asnaani (2012) present a transdiagnostic emotion dysregulation model of mood and anxiety disorders, in which they propose that both disorders are the result of emotion dysregulation of negative affect coupled with deficiencies in positive affect. Among the model assumptions regarding the most effective ways to treat mood and anxiety disorders, the authors list “decreasing negative affect and increasing positive affect” (Hofmann et al., 2012, p. 409). The theoretical background behind these treatment strategies encompasses the broaden-and-build model by Fredrickson (2000), in which negative affect is assumed to be associated with a limited behavioral repertoire in a given situation and positive affect is assumed to loosen the influence of negative affect on the person and to broaden the behavioral repertoire. Additionally positive affect is associated with approach, whereas negative affect is associated with withdrawal tendencies. Hofmann et al. (2012) underline the importance of the two primary dimensions of negative and positive affectivity for the onset, overlap and maintenance of anxiety and depression and assume that the impact of these constructs have been underestimated.

There are some studies that have previously applied EMA methods to investigate different aspects of affective dynamics in everyday life with regard to treatment response, outcome and relapse in remitted patients for specific diagnostic subgroups (Peeters, Berkhof, Rottenberg, & Nicolson, 2010; Wichers et al., 2010; Wichers, Lothmann, Simons, Nicolson, & Peeters, 2011). The results of this research indicate that the reactivity of negative (NA) and positive affect (PA) in response to positive and negative daily life events allows the distinction between treatment responders and non-responders for depressed patients.
Furthermore, it indicates that negative affective variability allows the prediction of relapse in remitted depressed patients, independent of conventional questionnaire measures and other dynamic emotional patterns. Thompson et al. (2012) showed that clinical samples display greater fluctuation in NA compared to healthy controls, which indicates that negative affect fluctuation might be an indicator for psychological functioning. Regarding affective states, but independent of their fluctuation, Forbes et al. (2012) recently showed that the PA/NA ratio in daily life before treatment onset predicted treatment outcome in children and adolescents with depression and anxiety. To our knowledge, the effects of daily life affective states and dynamics on early response have not yet been investigated.

The research project in which we have applied EMA aims to broaden our knowledge about patient characteristics, which allow the prediction of early treatment response. We apply EMA methods to collect real-time affective dynamics of patients who are – at the time of assessments - waiting to be treated in our outpatient clinic. In doing so, we test whether EMA methods can help answer the question of how differences in early response arise and which role patient characteristics play. Integrating the findings stated above, this study focuses on the following research questions: Do affective states and their temporal dynamics (fluctuation), assessed via EMA before treatment onset, allow the prediction of early treatment response? More specifically, we hypothesize that (a) not mean levels of NA or PA but PA/NA ratios predict early treatment response. We expect that Forbes et al.’s findings (2012) that higher PA/NA ratios correspond to faster treatment responses in children and adolescents with depression or anxiety disorders generalize to adult patients with mixed diagnoses. Furthermore, with regard to Wichers’ (2010) findings that NA variability predicts future negative affective symptoms in remitted depressed patients and also taking Thompson’s (2012) findings that clinical samples display greater instability in NA into account, we hypothesize that (b) fluctuations in NA (when controlled for initial impairment)
predict early response (patients with high fluctuations in NA in daily life are less likely to show early response), whereas fluctuations in PA have no impact on early treatment response. Furthermore, given the advantages of EMA in eliminating retrospective biases, we hypothesize (c) that EMA measures of real-time affect show additional predictive power beyond predictors derived from intake measures.

5.2. Methods

5.2.1. Sample

The analyses were based on the subsample of patients who took part in the EMA application during the waiting period, began treatment afterwards and did not drop out before session 6. In total, this sample consisted of 39 patients who were treated with CBT at the University’s outpatient clinic. The majority of the sample was diagnosed with an affective (46.2%) or anxiety (38.5%) disorder as the primary diagnosis. Additional primary diagnoses were eating disorders (5.1%) as well as substance-related and addictive disorders (5.1%). For the diagnosis of personality disorders, the International Diagnostic Checklist for Personality Disorders (IDCL-P; Bronisch, Hiller, Mombour, & Zaudig, 1996) was adopted, which identified 5.1% of the sample as having a personality disorder. 59.0% of the sample were female, age ranged from 19 – 59 years (mean = 35.69, SD = 11.48) and impairment levels measured by the Global Severity Index of the Brief Symptom Inventory (BSI; Franke, 2000; German translation by Derogatis, 1975) ranged from .11 to 2.55 (mean = 1.24, SD = .59).

5.2.2. Instruments

Figure 5.1 displays all relevant measures that were included in the analyses for Study III. Descriptions of the measures can be found in section 2.3.
Figure 5.1: Study flow chart with relevant measures or assessed constructs

5.2.3. Statistical Analysis

We analyzed treatment response from session 1 to 5. Due to the nested structure of the data (sessions nested within patients), we conducted hierarchical linear models. Following prior research on change trajectories in psychotherapy, a log-linear (base 10) transformation of the time scale was used for these analyses (Lutz et al., 2009; Lutz et al., 2014; Stulz et al., 2007). In most analyses of rapid response, dose-effectiveness research has shown a consistent pattern (e.g., Lambert 2007). A log-linear transformation of session numbers, as presented in this case, can parsimoniously approximate this consistent pattern, which is why it is widely applied in this area of research (e.g., Gibbons et al., 1993). Each patient’s HSCL is modeled as a function of session number (S) as follows:

\[
HSCL(S) = \beta_{00} + \beta_{10}\log_{10}(S) + r_{0i} + r_{1i}
\]

HSCL (S) is the observed HSCL score at a particular session. The \( \beta_{00} \) parameter (intercept) is the mean expected HSCL score at the first session. The \( \beta_{10} \) parameter (slope) is the expected change in HSCL scores per \( \log_{10} \) of session number. The random effects comprised the Level 1 residuals, (i.e., deviations of the observed affect scores from the expected affect scores at each measurement occasion), the Level 2 residuals of intercepts and
the Level 2 residuals of slopes (i.e., individual differences in HSCL mean levels and individual differences in the within-person regression slope). Individual change from session one to five (slope) is predicted with successive Level 2 models, in which the Level 1 slope coefficient was treated as the dependent variable in random coefficient models.

As the HSCL is a short version of the BSI and both are highly correlated, we first entered the Global Severity Index of the BSI as an intercept predictor. Because $\beta_{0i}$ represents the overall HSCL at the first session, the BSI scale scores extract all reliable variation when entered into the Level 2 model predicting intercept; therefore no random effect is included. This model served as our unconditional base model (Model 1). We then successively augmented our equation to test the cross-level interaction of different level 2 predictors with regard to their predictive value for the slope factor (i.e., change from session one to five).

First, we entered the Global Severity Index of the BSI as an indicator for initial impairment. As initial impairment has been repeatedly found to be a predictor of change, we tested all additional predictors against the GSI (Lambert et al., 2002; Lutz et al., 1999; Rubel et al., 2014). Of the seven additional predictors we added, three were patients’ ratings of the following questions: (a) Treatment expectation = “How convinced are you that psychotherapy will help you deal with your problems?” (responses range from 1 = “not at all” to 4 = “very much”) (b) Prior Psychotherapy = “How much psychotherapy have you had in the past?” (responses range from 1 = “none” to 6 = “more than a year”) (c) Chronicity = “How long has the problem for which you are presently seeking treatment been of concern to you?” (responses range from 1 = “less than a month” to 6 = “more than 2 years”). As a fourth predictor, we used the therapist-rated Global Assessment Scale (GAS), which indicates overall functioning of an individual on a continuum from psychological or psychiatric illness to health (Endicott, Spitzer, Fleiss, & Cohen, 1976). The decision for these first four indicators was informed by Lutz, Martinovich and Howard (1999), who identified them as
significant intake predictors for slope variance. The remaining three predictors were derived from the EMA period: alongside entering the PA/NA ratio (like Forbes et al., 2012), we also included fluctuation indices of PA and NA (measured by MSSD). Mean levels of PA and NA were also included to control for these variables.

We followed a two-step approach: In a first step, in addition to the GSI, we entered all aforementioned predictors into single models separately, to test whether we could replicate Lutz et al.’s findings with regard to the intake measures and to test whether the indicators of affective states and dynamics collected via EMA predicted significant variance on the slope factor, when controlled for initial impairment (hypotheses a and b). In a second step, we incorporated all significant predictors from the separate models in a combined model to test whether – if predictive – EMA measures explained slope variance beyond intake measures (hypothesis c). All Level 2 predictors were grand-mean centered in all analyses.

5.3. Results

5.3.1. Descriptive Statistics

Descriptive statistics for all variables can be found in Table 5.1. The GSI of 2.08 represents impairment levels of a common outpatient sample. The mean value of 3 for treatment expectation indicates that the average patient in the sample was rather convinced that treatment would help to overcome the problems he/she sought help for. The average patient in the sample had experienced one to three months of prior psychotherapy and had dealt with the presenting problem between one and two years before starting treatment. The average of therapist-rated global levels of functioning indicates moderately pronounced symptoms.

Mean levels of PA and NA indicate that, on average, the amount of PA outweighs the amount of NA in the sample ($t (38) = -2.65, p < .05$). This relation of PA to NA is also
indicated by an average PA/NA score of above one. Fluctuation in affect is significantly higher for PA than for NA ($t (22) = -5.13, p < .01$).

Table 5.1: Descriptive statistics

<table>
<thead>
<tr>
<th>Measures</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intake</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Impairment</td>
<td>2.08</td>
<td>0.63</td>
<td>1.08</td>
<td>3.33</td>
</tr>
<tr>
<td>Treatment Expectation</td>
<td>3</td>
<td>0.56</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Prior Psychotherapy</td>
<td>2.69</td>
<td>1.76</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Chronicity</td>
<td>5.54</td>
<td>0.88</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Global Assessment</td>
<td>59.69</td>
<td>8.42</td>
<td>45</td>
<td>78</td>
</tr>
<tr>
<td><strong>EMA-Period</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA (4 items)</td>
<td>2.31</td>
<td>1.03</td>
<td>1.00</td>
<td>4.50</td>
</tr>
<tr>
<td>NA (4 items)</td>
<td>1.69</td>
<td>0.69</td>
<td>1.00</td>
<td>3.25</td>
</tr>
<tr>
<td>PA/NA</td>
<td>1.72</td>
<td>0.77</td>
<td>0.56</td>
<td>3.50</td>
</tr>
<tr>
<td>MSSD PA</td>
<td>0.65</td>
<td>0.49</td>
<td>0.06</td>
<td>2.04</td>
</tr>
<tr>
<td>MSSD NA</td>
<td>0.28</td>
<td>0.20</td>
<td>0.02</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Note. For EMA-Period variables, means represent averaged momentary scores across individuals and across measurement occasions. MSSD scores were calculated within days.

Correlation results revealed low correlations between all Level 2 predictors. All but two predictors correlated clearly below +/- .5 (range: .01 to -.44). Only PA/NA ratios correlated by -.66 with mean levels of NA and by .58 with mean levels of PA, which is not surprising, as the indicator is derived from mean levels of PA and NA at each measurement occasion.

5.3.2. Unconditional Model and Separate Predictor Models

To test whether real time affective states and temporal dynamics can improve predictions of early response, we first modeled the course of treatment response in an unconditional model. The fixed effect estimates indicate an average HSCL of 2.08 at the first session and a mean rate of change of .35 HSCL scores per $\log_{10}$ of session number. This corresponds to a mean change of over 0.6 $SD$ over the first five sessions. Participants differed in their response rates
over time, as indicated by a significant random effect for slope variation. As expected, there was no random effect for intercept variation, as indicated by a non-significant $\chi^2$ value (see Table 5.2).

**Table 5.2: Results of multilevel analyses predicting treatment response with an unconditional random slope model (Model 1)**

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Fixed Effects</th>
<th>Random Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
</tr>
<tr>
<td>HSCL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.08</td>
<td>.02</td>
</tr>
<tr>
<td>GSI</td>
<td>0.84</td>
<td>.04</td>
</tr>
<tr>
<td>Time Slope</td>
<td>-.35</td>
<td>.11</td>
</tr>
</tbody>
</table>

*Note. N (Level 2) = 39. N (Level 1) = 195. Time as the common logarithm of session number was entered uncentered as a Level 1 predictor. Coeff. = Regression Coefficient. * $p < .05$, *** $p < .001$.*

As expected, initial impairment significantly predicted slope variance when entered as a single predictor ($B = -0.51, p = .008$). Then we tested all other predictors separately with regard to the slope factor, while controlling for initial impairment. The following predictors explained slope variance beyond initial impairment: mean PA/NA ($B = -0.36, p < .001$), mean levels of NA ($B = 0.23, p = .04$) and fluctuation in NA ($B = 1.00, p = .03$). Controlled for initial impairment, these Level 2 predictors accounted for differences in the slope factor. The following predictors: treatment expectation ($B = -0.03, p = .89$), prior psychotherapy ($B = 0.10, p = .09$), chronicity ($B = 0.08, p = .31$), global assessment ($B = -0.02, p = .05$), mean levels of PA ($B = -0.08, p = .33$) and fluctuation in PA ($B = -0.14, p = .43$) did not turn out to be significant predictors of the slope factor and were therefore not integrated into the full model.

As the PA/NA ratio is correlated with mean levels of NA and PA, we conducted additional analyses to test whether the predictive value of the indicator was confounded with mean levels of PA or NA. We therefore first incorporated mean PA and mean PA/NA into
one model (doing the same for NA), and then added the interaction of mean PA * mean PA/NA (and NA respectively), to test whether the effects of mean PA/NA on the slope were moderated by mean levels of PA (or NA). PA/NA ratios still significantly predicted the slope coefficient, when mean levels of PA \((B = -0.42, p = .002)\), mean levels of NA \((B = -0.38, p = 0.01)\), and their interaction (mean PA * mean PA/NA: \(B = -0.48, p = .005\); mean NA * mean PA/NA: \(B = -0.35, p < .001\)) were controlled for. Therefore, the predictive value of PA/NA ratios on the slope coefficient seems independent of mean levels of PA and NA.

5.3.3. Combined Model

We then entered all significant predictors from the single models into one full model (Model 2). The results indicate that besides mean levels of NA, all predictors from the previous models significantly explain variance in early change slopes (see Table 5.3). As expected, because of the high correlations between PA/NA ratios and NA, NA did not have any additional explanatory value beyond PA/NA ratios and was excluded from the full model.

**Table 5.3:** Results of multilevel analyses predicting treatment response by EMA affective states and temporal dynamics (Model 2)

<table>
<thead>
<tr>
<th>Model 2</th>
<th>Fixed Effects</th>
<th>Random Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.08</td>
<td>.09</td>
</tr>
<tr>
<td>GSI</td>
<td>.91</td>
<td>.04</td>
</tr>
<tr>
<td>Time Slope</td>
<td>-.35</td>
<td>.09</td>
</tr>
<tr>
<td>GSI</td>
<td>-.74</td>
<td>.20</td>
</tr>
<tr>
<td>PA/NA</td>
<td>-.36</td>
<td>.14</td>
</tr>
<tr>
<td>MSSD_NA</td>
<td>.78</td>
<td>.36</td>
</tr>
</tbody>
</table>

Note. \(N\) (Level 2) = 39. \(N\) (Level 1) = 195. Time as the logarithm of session number was entered as a Level 1 predictor. All Level 2 predictors were grand-mean centered. Coeff. = Regression Coefficient.

\(^a\) Based on the assumption of normally distributed slope coefficients, this value indicates the estimated percentage of slope coefficients that are in the direction indicated by the algebraic sign of the regression coefficient for the significant predictors (Hox, 2010).
Table 5.3 indicates that patients with higher impairment levels at intake (measured by GSI scores) were estimated to improve faster over the first five sessions. Higher PA/NA ratios and less fluctuation in NA corresponded to faster rates of change over the first five sessions. The relationship between the significant EMA predictors and rate of change is illustrated in Figures 5.2a and 5.2b. Figure 5.2a indicates that patients with an average PA/NA ratio or a PA/NA ratio one standard deviation above the mean (plotted are standardized estimates, as the predictor was grand-mean centered) improve during the early treatment period, whereas patients with PA/NA ratios one SD below the mean do not. The opposite is true for the impact of fluctuations in NA on early response: higher fluctuations in NA (measured by MSSD scores) correspond to slower rates of early response. Figure 5.2b illustrates that patients with fluctuation scores one SD above mean show increased impairment levels in an early treatment phase, whereas patients with mean fluctuation levels and fluctuation scores one SD below the mean seem to improve.

For these standardized estimates, the regions of significance - which define the specific values of the moderator, at which the slope of the regression of y on the predictor transitions from non-significance to significance - for PA/NA ratios are -6.16 at the lower bound and -0.34 at the upper bound (simple slopes are significant outside this region). The upper bound corresponds to a value of 1.46 on the PA/NA predictor, indicating that all patients with a PA/NA ratio above 1.46 show a significant reduction in their impairment levels over the first five sessions. The regions of significance for fluctuations in NA (MSSD_NA) are -2.26 at the lower bound and 0.14 at the upper bound (simple slopes are significant inside this region). The upper bound corresponds to a value of 0.31 on the MSSD_NA predictor, indicating that all patients with fluctuation scores in NA below 0.31 show a significant reduction in their impairment levels over the first five sessions.
Rates of early response depending on daily life PA/NA ratios (Figure 5.2a, left) and daily life fluctuation in NA measured by MSSD (Figure 5.2b, right). Simple slopes are depicted for the grand-mean of the predictors (0) as well as one SD below (-1) and above (+1).

In order to examine the pattern of individual differences in more detail, the percentages of slope coefficients that are in the direction indicated by the algebraic sign for the significant slope predictors were estimated (following Hox, 2010, see Table 3). Based on the assumption of normally distributed slope coefficients, approximately 76.25% of the sample showed early response indicated by a drop in HSCL scores from session 1 – 5 (see Table 5.3). For 93.45% of the sample high initial impairment scores corresponded to high response rates over the first five sessions (as indicated by a negative coefficient, when the GSI score was entered as a predictor on the slope, see table 3). That means that only 6.55% displayed either high initial impairment scores and corresponding low response rates over the first five sessions or low initial impairment scores and high response rates over the first five sessions. For 76.87% of the sample high PA/NA ratios corresponded to high response rates over the first five sessions (as indicated by a negative coefficient, when the PA/NA score was entered as a predictor on the slope, see table 5.3 and Figure 5.2a). That means that only 23.13% of the sample displayed either high PA/NA ratios and corresponding low response
rates over the first five sessions or low PA/NA ratios and high response rates over the first five sessions. For 94.43% of the sample high fluctuation scores in NA corresponded to low response rates over the first five sessions (as indicated by a positive coefficient, when the PA/NA score was entered as a predictor on the slope, see table 5.3 and Figure 5.2b). That means that only 5.57% of the sample displayed either high fluctuation scores in NA and corresponding high response rates over the first five sessions or low fluctuation scores in NA and low response rates over the first five sessions.

In order to compare the additional value of real-time EMA measures with regard to the explanation of early response, we calculate the proportional reduction of residual slope variance (Raudenbush, & Bryk, 2002). First, we integrated only the significant intake predictor (intake impairment levels measured by GSI) and compared this model to the unconditional model. By integrating GSI, the slope variance was reduced by 16.10%. We then compared our full model with the more restrictive model encompassing only the intake predictor, to test how much additional variance was explained by the EMA predictors. By entering the EMA predictors, slope variance was reduced by an additional 21.73%.

5.4. Conclusions

In this study, we found that the assessment of affective states and dynamics in daily life allows the prediction of early treatment response. In summary our analyses revealed the following results: (a) PA/NA ratios predict early treatment response, (b) fluctuations in NA predict early treatment response beyond PA/NA ratios and (c) both EMA measures named in (a) and (b) predict early treatment response beyond initial impairment levels. The findings from Study III underline how EMA measures can be integrated in prediction models of treatment response to decrease the proportion of unexplained variance. In the following section, the findings from Study III will be reiterated and conclusions will be drawn.
As summarized in the introduction to this study, there is a body of literature on how intake indicators predict course of treatment (Lambert et al., 2002; Lutz et al., 1999; Rubel et al., 2014), but they leave a lot of variance to be explained. This may be due to floor effects for those patients who are less distressed or more room for improvement for those who are more distressed. As this possible explanation can only be regarded as a hypothesis, more research is needed to explain the remaining variance between patients in their response to treatment. This research question becomes increasingly important, as treatment approaches are pushed to become increasingly individualized (Insel, 2009). Attempts to individualize treatments early in the therapeutic process can be supported by methods, which allow high resolution longitudinal data collection before the onset of treatment.

The results from this study show that EMA, applied before treatment onset, can help to explain a considerable proportion of the remaining variance in treatment response between different patients, even in an early stage of treatment. Specifically, and with regard to our hypotheses, we showed that not mean levels of PA and NA but the interplay of PA and NA at a certain time - measured by PA/NA ratios - predicted early treatment response. Patients with higher PA/NA ratios, indicating preponderance of PA over NA, were more likely to show early response. Secondly we showed that fluctuations in NA, when measured in daily life, improve predictions of early response beyond PA/NA ratios, whereas fluctuations in PA do not explain additional variance in early response. Patients with lower fluctuations in NA, measured in daily life before treatment onset, were more likely to show early response. Furthermore, the results of this study indicate that EMA assessments of real-time affect have additional predictive power beyond intake measures. PA/NA ratios and NA fluctuation explain early treatment response beyond initial impairment, therefore improving prediction models of early response.
The results from the separate predictor models indicate that initial impairment levels predict differences in treatment response. Patients with higher impairment levels at intake were more likely to show early response. Therewith, we replicated findings from previous studies that found that initial impairment predicts rates of change (Lambert et al., 2002; Lutz et al., 1999; Rubel et al., 2014).

As we expected of the EMA predictors tested, neither mean levels of NA or PA predicted early slope variation but PA/NA ratios and fluctuation in NA did. Notably, PA/NA ratios predicted slope variation independent of mean levels of PA and NA. This result supports the interrelation of PA and NA described by Hofmann et al. (2012). PA loosens the influence of NA on the person and broadens the behavioral repertoire by enhancing physical, social, and intellectual resources (Fredrickson & Branigan, 2005). Treatment strategies especially for depression and anxiety should therefore not only focus on the reduction or elimination of patients’ suffering caused by negative affective states (Hofmann et al., 2012) but also focus on resource activation as proposed by Flückiger and others (e.g., Flückiger, Caspar, grosse Holtforth, & Willutzki, 2009; Flückiger, Wüsten, Zinbarg, & Wampold, 2010; Flückiger, Zinbarg, Znoj, & Ackert, 2014) to promote positive affective states. Resource activation to increase patients’ well-being and reduce distress is realized by emphasizing patients’ strengths and by focusing on coping and success rather than on symptoms and distress. This treatment approach, which reminds of the therapeutic mindset of for instance solution focused (brief) therapy (Berg & de Shazer, 1993), is compatible with evidence based treatment approaches and many strategies and exercises are described in depth by Flückiger et al. (2010) so that therapists can easily integrate them during treatments.

Although fluctuation scores were significantly higher for PA than for NA across the sample, fluctuations in NA proved to be a more informative indicator with regard to early response. This shows that fluctuation in NA can not only predict relapse in remitted depressed
patients, as shown by Wichers et al. (2010), but also early response. Fluctuation in NA does not only differentiate between clinical samples and healthy controls (Thompson et al., 2012) and not only proves to be a good indicator for relapse in specific diagnostic subgroups, namely depression (Wichers et al., 2010), but also shows to be a significant predictor of treatment response with transdiagnostic relevance. The assessment of NA fluctuation before the onset of treatment may help to elucidate an important aspect of psychological functioning in patients possibly regardless of their diagnoses.

Some limitations of this study should be acknowledged. This study’s sample size did not allow the comparison between different diagnostic subgroups. The majority of the sample was diagnosed with an affective or anxiety disorder. In separate analyses, we controlled for depressive symptoms to ensure the relation between PA/NA dynamics and early treatment response was not limited to depressed patients, which it was not. However, we cannot ensure that the findings generalize to all other psychological disorders. The small sample size, the accompanied lack of power and lack of full range of scores on the respective predictors might also be a possible explanation for the result that the intake predictors, which explained slope variation in the study by Lutz et al. (1999), did not turn out to be significant in our analyses. It is possible that intake indicators identified to predict treatment response in cross-sectional studies provide initial insights, which help to differentiate between patients who improve during early treatment and those who do not on a macro-level. However, as our results suggest, longitudinal micro-level analyses can shed light on further between-person differences that go beyond initial indicators.

In closing and despite the mentioned limitations, our findings suggest that the integration of methods allowing the assessment of a dimension of patient characteristics beyond regular intake measures, broadens our knowledge of patient variables, which explain differences in early response rates. Daily life patterns of negative and positive emotions may
help therapists to gain insights into the psychopathology of a particular patient, thus decision-making regarding treatment choices may be improved. In order to ensure the best possible patient-centered treatment, the integration of EMA strategies in the diagnostic process before the onset of treatment seems to be a promising endeavor.
6 GENERAL DISCUSSION

6.1. Summary and discussion of study results

The general discussion of this thesis starts with a summary and global discussion of the three studies and then proceeds with combined conclusions and future prospects. For in-depth discussions of the single studies please refer to the corresponding chapters.

6.1.1. Study I

Study I focused on feasibility analyses of EMA applications in routine care – namely outpatient settings. Different aspects of feasibility were highlighted: First, control group comparisons with regard to sampling revealed that those patients who agreed to participate in the EMA application during the waiting period for outpatient treatment were significantly less impaired, compared to the average patient who applied for outpatient treatment in the clinic. Although previous studies have shown that EMA can also be applied in severely impaired patient subgroups (e.g., Solzbacher, Böttger, Memmesheimer, Mussgay, & Rüddel, 2007), our results suggest that applications in outpatient settings must take the risk of missing out on patients with more severe impairment levels into account or adopt a sampling strategy, which explicitly includes patients with all levels of impairment at intake.

Compared to previous studies (e.g., Peters et al., 2000), results from the analyses regarding technical challenges were satisfying. If missings occurred due to technological malfunctions they did not appear systematically. Nevertheless, results from these analyses must be considered cautiously, as the applied software did not allow a distinct differentiation between technically- and participant-induced missings. This limitation and the fact that data was missed because of technical reasons illustrates that even though the past years have brought about rapid developments in the field of electronic data assessment, technical devices are never entirely error proof. While technological malfunctions can be monitored and quickly
eliminated in EMA applications, where participants are close to the research team (e.g., in inpatient settings), applications in outpatients settings must a) thoroughly test different hard- and software applications before starting pilot projects; b) conduct elaborate pilot studies with different participants allowing every sort of technical challenge to arise c) consider missings due to technical challenges in power and sample size analyses and d) enable contact possibilities between participants and the research team during the EMA period.

In order to keep the burden low for participants, the number of items and their length must be limited (Barta et al., 2012). Analyses regarding the duration participants require to answer a single assessment revealed an average time of approximately 2 minutes, which seems reasonable. This conclusion is supported by the participants’ feedback taken from the EMA evaluation form. Longitudinal analyses from the first to the last day of assessment revealed training effects, indicated by a drop in duration to fill out the questionnaire. However, these drops only occurred until day three and were more pronounced for older participants. These results clearly support a thorough introduction of EMA participants to study protocol. The more training participants have in answering the items, the better they know the items and the faster they can respond to them. Especially for older participants, enough time should be calculated for the training session. To ensure a thorough introduction and training session for participants, enough staff must be budgeted when planning and organizing an EMA application in an outpatient setting.

Similar to previous studies (e.g., Broderick et al., 2003; Jamison et al., 2001; Sokolovsky et al., 2014), we found high compliance rates with an average of almost 93%. In contrast to some previous studies (e.g., Courvoisier et al., 2012; Silvia et al., 2013), we did not detect changes in compliance rates over time. This can have several reasons: First, the telephone contact after two days can have potentially functioned as a reminder for participants to accurately answer the assessments (although the length of this reminder’s impact on
compliance rates is arguable). Second, the study design may have significantly influenced compliance rates in several ways: Compared to the studies by Courvoisier et al. (2012) and Silvia et al. (2013), our EMA applications incorporated fewer daily assessments. While Courvoisier signaled study participants 6 times daily, Silvia even signaled 8 times daily, which is twice as many assessments as in the current study. The higher the burden for participants, the more likely losses in compliance rates become. Third and last, the objectives of the EMA applications influence compliance rates. In our application, study participants were patients waiting to be treated at the outpatient clinic. Thus, the sample consisted of participants who were highly motivated to give information on their everyday emotional, cognitive and behavioral experiences. As they were also informed before the onset of the EMA period that they would have the opportunity to receive an evaluation of the assessment after they had started treatment and that they could share this evaluation with their therapist, it was suggested that they could use the chance to gather important pieces of diagnostic information, which could improve the therapeutic process. Although, to our knowledge, no studies exist, which have systematically analyzed whether the announcement of providing patients with an evaluation of the EMA period influences compliance rates, researchers agree that patients highly appreciate the provision of an evaluation (U. W. Ebner-Priemer; Talk at the University of Trier: Ambulatory Assessment in Clinical Psychology, January 22, 2016) and the dissemination of self-tracking apps, which use course or trend charts support this assumption. Furthermore, the results show that patients significantly differed in their compliance rates; however no significant variables explaining these differences could be identified. Results from previous studies (Silvia et al., 2013), showing that compliance rates were lower for men and more severely impaired patients, could not be replicated. To this end, more research is necessary in order to understand the predictors of missing data and include predictors of non-response, allowing researchers to implement methods to increase
compliance and to handle missing data more effectively. In total, the results regarding compliance rates support good feasibility for applications with four daily assessments over 14 consecutive days in outpatient settings.

When analyzing the feasibility of an EMA application, the objective of its application - independent of the setting in which it is applied - must be defined in advance. Our application was intended to be mere assessment, without any intervention or treatment component. Thus, the objective of its application was to gather (diagnostic) information on patients before treatment onset and specifically not to increase their psychological functioning or well-being. The results of control group comparisons suggest that this objective was successfully met: compared to a matched control group, study participants showed no elevated waiting period effects, i.e. the EMA period had no significant effects on their impairment levels. As illustrated earlier, electronically-mediated interventions that intend to influence patients’ impairment levels either involve interactive assessments (in the form of direct feedback to the patient) or treatment components, which was specifically not the case in our application of EMA during the waiting period for outpatient treatment. Here, the introduction of the study objective as a time period to gather individualized pieces of information from the patient may have added to the non-occurrence of effects on impairment levels. Furthermore, the point in time at which the evaluation of the EMA period is given to the patients may be crucial when measuring effects on impairment levels or well-being. In our application, the evaluation was given to patients after they had received their tenth session of treatment (earlier in case they dropped out of treatment sooner). This had several reasons: First, during the waiting period, our objective was to assess the effects of the EMA, not of its evaluation. Second, as we intended to measure the effects of the EMA on early treatment response, we did not want the results to be distorted by providing feedback at evaluation. Third and last, as, in some cases, the evaluation can be rather confrontational for participants
(e.g., for patients with no or only a small social network who experienced no social support during the entire 14-day period), we wanted to give therapists and patients the chance to establish a therapeutic relationship before discussing the evaluation and its consequences.

The results of the evaluation of participants’ acceptance of the EMA application and their perceived burden were promising. The acceptance was generally good and supports applications with four daily assessments over a two-week period in outpatient settings. In-depth qualitative analyses of the open feedback option revealed that a) patients profit from flexible assessments, which they can easily integrate in their everyday routines and b) more research is needed to be able to answer the question “Who profits from EMA applications in what way?”. Regarding a): 98.28% of the sample (all but one participant) made use of the postponement option to answer a signal. This indicates that much data would have been lost if participants had only had one chance to answer a signal. Providing the option of postponing answers or a time window for response seems crucial, especially in outpatient settings, where patients are involved in their everyday routines (e.g., work) during the assessment period. However, the more we adapt and individualize assessment time points, the less comparable the results between patients become. Therefore, the sampling plan must always be a compromise between individualization and comparability and be primarily guided by mapping requirements of the construct of assessment. Regarding b): our analyses showed that although the majority of participants indicated having profited from the EMA period, more research on patient variables that facilitate or impede successful EMA applications in outpatient settings is needed. While some patients experienced higher burdens by the daily assessments than others, others profited in ways that exceeded the character of mere assessments.
6.1.2. Study II

Study II replicated and extended previous findings on the rumination-affect link. The results suggest that real-time rumination is negatively linked to affect on a within-person level. Furthermore, the conducted analyses allowed a statistical estimation of the proportion of participants in the sample to which the identified association applied. The vast majority of participants (95%) displayed the negative link between state rumination and affect, underlining the strong effects of this dysfunctional emotion regulation strategy.

The Response Styles Theory (Nolen-Hoeksema, 1991) suggests a unidirectional influence of rumination on affect. Nevertheless, studies that have attempted to shed light on the rumination-affect link on a within-person level (e.g., Moberly, & Watkins, 2008) have found evidence for a reciprocal relationship between the two constructs. As is often the case in psychological research, the chicken-and-egg dilemma arises. However, the Study II analyses statistically bypass the uncertainties regarding the direction of the influence. Controlled for concurrent and previous affect levels, rumination still predicts concurrent and subsequent affect levels – supporting a direct influence of rumination on affect.

Furthermore, the current EMA application allowed an estimation of how long the deteriorating effects of rumination on affect lasted within the individuals of the sample. The results revealed that the effects of rumination on affect could still be detected at the subsequent assessments, which were conducted max. 5 hours later. This result highlights the extensive negative consequences and the need to tackle this dysfunctional emotion regulation strategy in order to improve emotional well-being of affected individuals.

Analysing the influences of state and trait rumination simultaneously provided the possibility of identifying person-variables that influence the rumination-affect link. The results replicated earlier findings, which showed that trait rumination predicts mean levels of affect. Nevertheless, the aspects of trait rumination that predicted the intercepts in our models
contradicted earlier findings by Moberly, & Watkins (2008): the results suggest that symptom-focused rumination rather than self-focused rumination predicted mean levels of affect. These contradicting results could have several explanations; one may simply lay in the utilization of different versions of the questionnaires. While the German version of the RSQ-D short-version applied in Study II consists of three subscales (symptom-focused rumination, self-focused rumination and distraction), different English versions exist. Moberly and Watkins, who found that self-focused and not symptom-focused rumination predicted mean affect levels, applied the first version of the RSQ (Nolen-Hoeksema, & Morrow, 1991) with 4 subscales (rumination, distraction, problem-solving, and dangerous situations) and additionally calculated scores on the separate brooding and reflection subscale identified by Treynor, Gonzales, & Nolen-Hoeksema (2003). The application of different questionnaire versions might have led to slightly different results regarding the separate subscales. Results from the models that used aggregated state rumination as a person-specific predictor for intercept and slope variance revealed a dampening effect of average rumination on the rumination-affect link. The more individuals ruminated on average over the two-week period, the lesser the deteriorating effects of state rumination on affect. Individuals who habitually ruminate more than others seem to display decreased effects of state rumination on affect. These individuals have possibly undergone a habituation process comparable to chronic pain patients, who also do not experience a deterioration of affect consequential to every pain symptom during the day. Likewise, these individuals may also be characterized by a generally limited level of emotional responsiveness. More research on person-variables is necessary to shed light on individual differences regarding the rumination-affect link.

Due to the fact that a sample with mixed primary diagnoses was analysed in Study II, the results can be interpreted in favour of the concept that rumination is a relevant factor in many Axis-I disorders – supporting Ehring, & Watkins’ (2008) assumption. Though the
study’s sample size limits its interpretation to an exploratory character, it can at least be precluded that rumination only functions as a relevant factor in depressive disorders. Future studies should systematically include different Axis-I disorders and extend the sample size to include comparisons between different diagnostic subgroups.

The results of Study II support the general notion of an existing discordance between retrospective self-reports and real-time assessments. Conner, & Feldman Barrett (2013) explain this phenomenon by differences in qualitative aspects captured by the two assessment strategies. They argue that while the former assesses aspects of the remembering and believing self, the latter rather assesses aspects of the experiencing self. As pointed out in the theoretical introduction of this thesis, retrospective biases occur when time passes and prediction, intention, choice global memories, attitudes and preferences converge and influence the reconstruction of experiences. This conglomerate of cognitions constitutes the remembering and believing self, while the experiencing self is less prone to be affected by it (although it shall not be denied that attitudes and preferences, especially with regard to perception and interpretation, also affect the experiencing self). The results of Study II cannot confirm this hypothetical explanation, but provide yet another piece of evidence for an existing discordance between the two assessment strategies. None of the two can be regarded as generally superior to the other; only the research question can decide which method should be applied.

6.1.3. Study III

After the feasibility of the EMA application was analysed in Study I and the coherence of the relationship between different constructs within the EMA period was tested in Study II, Study III finally used the information obtained from the EMA period before treatment onset in an application of patient-oriented psychotherapy research. The results indicate that EMA
applications open up the possibility to assess patient characteristics before the onset of treatment that allow the prediction of early treatment response beyond commonly applied intake indicators (e.g., Lambert et al., 2002; Lutz et al., 1999; Rubel et al., 2014). Specifically, the results of Study III show that affective states and dynamics in daily life can shed light on how individual impairment levels change within an early treatment period (sessions 1-5) beyond initial impairment levels.

Importantly, different aspects of affective states and dynamics predict early treatment progress. While the interrelation of PA and NA (specifically the PA/NA ratio) - and not PA or NA levels separately - predicts early response, only fluctuation in NA and not in PA predicts the slope coefficient, i.e. the change in impairment levels from sessions 1-5. Additionally, both indicators explained significant proportions of the slope variance beyond one another. In our sample, the individuals most likely to show improvement during the early treatment period were those with a) high initial impairment levels, b) high PA/NA ratios and c) low fluctuations in NA. This result highlights that PA/NA ratios do not merely reflect initial impairment levels, as the direction of influence on early response is inverse, i.e. individuals with affect ratios in favour of PA over NA respond faster and not individuals with affect ratios in favour of NA over PA, which may correspond to higher impairment levels. The in-depth analyses of the specific values of the predictor and the corresponding rates of change during treatment reveal that patients with average PA/NA levels and PA/NA levels above the average of the sample are more likely to improve during an early treatment period. Furthermore, the results indicate that in a mixed sample of patients seeking outpatient treatment, fluctuation in NA before treatment onset predicts early treatment response. Patients with higher fluctuations in NA before treatment onset are less likely to improve during an early treatment period. The in-depth analyses of the specific values of the predictor and the corresponding rates of change during treatment reveal that patients with NA fluctuation levels
above the sample average do not only not display early response, but seem to deteriorate during an early treatment period. This result is in line with results from Wichers et al. (2010), who showed that negative affect variability predicts relapse in patients with major depression. High fluctuation in NA seems to be a crucial indicator for the clinical course during and after treatment.

Most importantly, Study III explicitly illustrated that EMA applications can improve prediction models in patient-oriented psychotherapy research. Although it is still unclear, whether real-time assessments capture qualitatively different aspects of cognition, behaviour and emotion than retrospective self-reports, they help to reduce unexplained variance in treatment response. Especially the combination of predictors derived from retrospective and real-time self-reports in prediction models for treatment response seems beneficial.

6.2. Conclusions and Future Prospects

The present thesis highlights the challenges and chances associated with the application of EMA in patient-oriented psychotherapy research. It can guide future applications of EMA in outpatient settings. The results of the feasibility analyses support the conclusions that the introduction of study objectives, the setting, in which it is applied, and the provision of an evaluation influence compliance and possibly affect the impact of the EMA application on impairment levels. Additionally, the results show that more research on patient characteristics with regard to acceptance and burden is required, although they generally support good applicability in outpatient settings.

It could be shown that the analysis of EMA patient data allows the identification of interesting findings on how different cognitive and emotional mechanisms are interrelated on a within-person level. We analyzed the relationship between rumination and affect, as it is a well-studied phenomenon with potential relevance for many disorders. In the field of emotion
regulation strategies, there are other potential mechanisms (e.g., distraction or self-aggressive behavior) that future studies could consider, which would have beneficial consequences for understanding the mechanisms that cause and maintain different disorders on a highly individualized level.

Furthermore, the current thesis underlines the possibility of achieving practice-oriented improvements using the methodology in patient-oriented research. Integrating predictors derived from EMA applications can improve prediction models regarding treatment progress and inform treatment decisions. Diary assessment strategies can enrich the repertoire of retrospective self-report assessment tools (Schneider, & Stone, 2016), especially when both strategies are combined. The advantages of applying EMA before treatment onset can be regarded from two different clinically relevant perspectives: From the patient’s perspective, it not only bridges the time between registration for and actual onset of treatment (which is especially relevant for health care systems with long waiting periods for psychotherapeutic treatment), but also provides the chance to provide a detailed and undistorted account of how the symptoms, for which a patient seeks treatment, affect everyday experiences. Thereby applications of EMA before treatment onset potentially have critical implications for the assessment and classification of psychopathology. EMA enables dynamic diagnostic processes, in which psychopathology can be considered in terms of the functional relationship between states, in addition to the correlational relationships that contribute to the syndrome structure. EMA data captures syndrome-specific, as well as person-specific information about the structure of psychopathology. Thus important diagnostic variance that might be lost due to dichotomization of presence versus absence in the existing categorical diagnostic taxonomies is recorded on an individual basis (Fischer, 2015). From the perspective of the clinician or researcher who applies EMA, the advantage consists of allowing predictions of treatment response before the onset of the treatment itself, and by
improving and individualizing the diagnostic process. Real-time assessments in a patient’s natural environment provide insights into the causes and maintaining factors of any disorder, which cannot be assessed by retrospective instruments, because of several biases that distort human memory. Standard diagnostic taxonomies do not consider the unique combinative presentations of symptoms among individual patients or the dynamic relationship therein. Dynamic symptomatology that spans over a defined period of time (as determined in the criteria of different disorders in DSM-5 and ICD-10) should – if possible – be assessed with an ecologically valid method that allows repeated assessments over time. The only reasons that would oppose the application of such methods - if available - are costs that exceed the benefits. Costs can arise for the patient, when the burden of the repeated assessments becomes unbearable and also for the applying researcher or clinician.

Even if the field of electronic devices that can be utilized for EMA continues to develop rapidly, lowering the cost of its application, EMA still appears resource demanding. It is not primarily the financial burden connected to equipment, but to the resources associated with time, staff (both points resulted in the rather small sample size in this thesis project) and know-how, which may present difficulties. Patients must be instructed in handling the electronic device used for EMA and the generation and interpretation of relevant data outputs requires technical and statistical methods, which may not be applicable for every clinician. However, there are promising technical developments integrating software, which generates relevant output that can be used for assessments, monitoring, and also for patient feedback (Wichers et al., 2011). Developments of this kind would enable more clinicians to integrate EMA strategies in their diagnostic routines.

For the routine application of EMA in diagnostics another big (and currently still unresolved) issue arises: reliable and valid short versions of diagnostic instruments are needed that can be used in EMA applications. Currently there is only a small number of validated
questionnaires for EMA applications that mainly focus on the assessment of affect, emotion regulation, self-esteem, social behavior and selected disorder-specific symptoms (e.g. Carlson et al., 2016). This poses difficult questions for researchers on what items to take when planning an EMA study and leads to a diversity of instruments or items which are applied which again leads to decreased comparability between study results. If EMA applications were to find their way into routine practice (and before that into routine research) future research should focus on the development and validation of applicable short-versions of the established (diagnostic) questionnaires.

This thesis focused on EMA applications, which support the diagnostic process in outpatient settings. The results underline Fisher’s (2015) assumption that dynamic, intraindividual assessment could help “to parse the shared versus unique variance across putative syndromes, as well as that occurring across individuals” (p. 10). If psychopathology was classified that way the functional relationship among syndromes of distress within individuals would be emphasized and core transdiagnostic dimensions across individuals could potentially be identified. At the same time individualized diagnostic facilitates personalized interventions. For interventions to be maximally efficient – and thus effective - they should seek to (a) target active psychopathological dimensions within individuals and (b) be capable of delivering therapeutic content in a hierarchical sequence that directly maps onto presenting dynamics (Fisher, 2015). Interventions can only be personalized when data is at hand that encompasses idiographic and dynamic aspects of the interaction of different syndromes within a certain psychopathology of an individual. Thus EMA applications can help to select a specific set of empirically supported treatments for a specific disorder within a specific individual and it can potentially even guide the order of delivery of specific modules within the treatment protocol as a function of symptomatic dynamics within each individual. Future research is needed to test whether the collection of clinically relevant data, over
appropriate time scales with sufficient frequency, can elucidate potent factors for the most
effective personalized treatment (Fisher, 2015).

Clearly, the field of EMA applications in clinical psychology is not limited to the
objectives of supporting diagnostic and informing treatment decisions. Interactive
assessments and electronically-mediated interventions open up a field of application that is
rich in manifold options, which clearly reach beyond the scope of this thesis, but should not
remain unmentioned. Ethical considerations with regard to privacy protection and the need to
promptly intervene, when EMA data discloses risky, self-destructive behavior or destructive
behavior directed at others pose specific challenges for this area of application.

The results from the three studies of this thesis clearly allow the conclusion that the
integration of EMA strategies in patient-oriented research and in evidence-based practice
seems to be a promising and valuable endeavor.
7 REFERENCES


Huffziger, S., Ebner-Priemer, U., Eisenbach, C., Koudela, S., Reinhard, I., Zamoscik, V.,
life: An experimental ambulatory assessment study. *Journal of Behavior Therapy and
Experimental Psychiatry, 44*, 322-328.

trajectories and quantitative electroencephalography (QEEG) biomarkers in major

plan for research on mental illness. *Archives of General Psychiatry, 66*, 128–133.

ecological momentary assessment. In A. A. Stone, S. Shiffman, A. Atienza, & L. Nebeling
(Eds.), *The science of real-time data capture: Self-reports in health research* (pp. 308-337).
New York: Oxford University Press.

*Bundesgesundheitsbl.-Gesundheitsforsch.-Gesundheitsschutz, 47*, 736-744.

momentary assessment: Indices using successive difference and group comparison via

REFERENCES


Mutz, G. iDialogPad (Version 1.922) [iOS-based software]. Unpublished.


Figure 2.1: Recruitment flow chart...........................................................................22
Figure 2.2: Study flow chart with relevant measures or assessed constructs..............25
Figure 3.1: Reductions in the duration to answer the EMA protocol from day 1 to day 14
of the EMA period.................................................................................................50
Figure 3.2: Average number of missings per day from day 1 until day 14 of the
EMA period...........................................................................................................51
Figure 3.3: Flow chart of recruitment and study participants....................................52
Figure 3.4: To what extent did the daily self-observation help you?............................56
Figure 3.5: Did the daily self-observation have a positive effect on your
mood?......................................................................................................................56
Figure 3.6: Did the surveys lead to a more conscious perception of your emotions?........56
Figure 3.7: Did the surveys encourage you to pay more attention to your own interests
and needs?..............................................................................................................56
Figure 3.8: Did you feel stressed by the daily surveys?..............................................57
Figure 3.9: Did the beep sound at inconvenient times?.............................................57
Figure 3.10: Did you worry about missing the beep?...............................................57
Figure 3.11: Did the daily self-observation have a negative effect on your mood?........57
Figure 4.1: Path diagrams of the analyzed Level 1 relationship between state rumination
and positive-negative affect....................................................................................72
Figure 5.1: Study Flow Chart with relevant measures or assessed constructs.............88
Figure 5.2a. and Figure 5.2b: Rates of early response depending on daily life PA/NA ratios (Figure 5.2a, left) and daily life fluctuation in NA measured by MSSD (Figure 5.2b, right)
9 TABLES

Table 3.1: Demographics of EMA sample and matched controls……………………………….39
Table 3.2: Results of multilevel analyses predicting the daily number of signals with an
unconditional random slope model (Model 1).........................................................46
Table 3.3: Results of multilevel analyses predicting duration with an unconditional random
slope model…………………………………………………………………………………48
Table 3.4: Results of multilevel analyses predicting duration to answer EMA assessments
by demographic variables (Model 2).................................................................49
Table 3.5: Results of multilevel analyses predicting missings due to participants with an
unconditional random slope model .................................................................51
Table 4.1: Descriptive statistics..................................................................................73
Table 4.2: Results of multilevel analyses predicting positive-negative affect by state
rumination...........................................................................................................75
Table 5.1: Descriptive statistics..................................................................................91
Table 5.2: Results of multilevel analyses predicting treatment response with an
unconditional random slope model (Model 1).....................................................92
Table 5.3: Results of multilevel analyses predicting treatment response by EMA
affective states and temporal dynamics (Model 2).................................................93
10 APPENDICES

APPENDIX A: EMA-ITEMS AND QUESTIONNAIRES

A1: Brief Symptom Inventory (BSI)
A2: Becks Depression Inventory (BDI-II)
A3: Response Styles Questionnaire (RSQ-D)
A4: Hopkins Symptom Checklist -11 (HSCL-11)
A5: EMA Items
A6: EMA Evaluation Form

APPENDIX B: RESULTS

B1: Pie charts of single item analyzes from the EMA evaluation form

APPENDIX C: STUDY MATERIAL

C1: Informed Consent
C2: Liability Agreement
APPENDIX A1: BRIEF SYMPTOM INVENTORY (BSI)

Sie finden nachstehend eine Liste von Problemen und Beschwerden, die man manchmal hat. Bitte lesen Sie jede Frage sorgfältig durch und entscheiden Sie, wie stark Sie durch diese Beschwerden gestört oder bedrängt worden sind, **und zwar während der vergangenen sieben Tage bis heute.** Überlegen Sie bitte nicht erst, welche Antwort „den besten Eindruck“ machen könnte, sondern antworten Sie so, wie es für Sie persönlich zutrifft. Machen Sie bitte hinter jeder Frage ein Kreuz bei der für Sie am besten zutreffenden Antwort.

Bitte beantworten Sie jede Frage!

<table>
<thead>
<tr>
<th>überhaupt nicht</th>
<th>ein wenig</th>
<th>ziemlich</th>
<th>stark</th>
<th>sehr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Wie sehr litten Sie in den letzten sieben Tagen unter...

1. Nervosität oder innerem Zittern

2. Ohnmachts- oder Schwindelgefühlen

3. die Idee, dass irgend jemand Macht über Ihre Gedanken hat

4. dem Gefühl, dass andere an den meisten Ihrer Schwierigkeiten Schuld sind

5. Gedächtnisschwierigkeiten

6. dem Gefühl, leicht reizbar oder verärgerbar zu sein

7. Herz- oder Brustschmerzen

8. Furcht auf offenen Plätzen oder auf der Straße

9. Gedanken, sich das Leben zu nehmen

10. dem Gefühl, dass man den meisten Menschen nicht trauen kann

11. schlechtem Appetit
<table>
<thead>
<tr>
<th>überhaupt nicht</th>
<th>ein wenig</th>
<th>ziemlich</th>
<th>stark</th>
<th>sehr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Wie sehr litten Sie in den letzten sieben Tagen unter ...

<table>
<thead>
<tr>
<th>12. plötzlichem Erschrecken ohne Grund</th>
</tr>
</thead>
<tbody>
<tr>
<td>13. Gefühlsausbrüchen, denen gegenüber Sie machtlos waren</td>
</tr>
<tr>
<td>14. Einsamkeitsgefühlen, selbst wenn Sie in Gesellschaft sind</td>
</tr>
<tr>
<td>15. dem Gefühl, dass es Ihnen schwer fällt, etwas anzufangen</td>
</tr>
<tr>
<td>16. Einsamkeitsgefühlen</td>
</tr>
<tr>
<td>17. Schwermut</td>
</tr>
<tr>
<td>18. dem Gefühl, sich für nichts zu interessieren</td>
</tr>
<tr>
<td>19. Furchtsamkeit</td>
</tr>
<tr>
<td>20. Verletzlichkeit in Gefühlsdingen</td>
</tr>
<tr>
<td>21. dem Gefühl, dass die Leute unfreundlich sind oder Sie nicht leiden können</td>
</tr>
<tr>
<td>22. Minderwertigkeitsgefühlen gegenüber anderen</td>
</tr>
<tr>
<td>23. Übelkeit oder Magenverstimmung</td>
</tr>
<tr>
<td>24. dem Gefühl, dass andere Sie beobachten oder über Sie reden</td>
</tr>
<tr>
<td>25. Einschlafschwierigkeiten</td>
</tr>
<tr>
<td>26. dem Zwang, wieder und wieder nachzukontrollieren, was Sie tun</td>
</tr>
<tr>
<td>...</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>27</td>
</tr>
<tr>
<td>28</td>
</tr>
<tr>
<td>29</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>31</td>
</tr>
<tr>
<td>32</td>
</tr>
<tr>
<td>33</td>
</tr>
<tr>
<td>34</td>
</tr>
<tr>
<td>35</td>
</tr>
<tr>
<td>36</td>
</tr>
<tr>
<td>37</td>
</tr>
<tr>
<td>38</td>
</tr>
<tr>
<td>39</td>
</tr>
<tr>
<td>40</td>
</tr>
<tr>
<td>41</td>
</tr>
</tbody>
</table>
Wie sehr litten Sie in den letzten sieben Tagen unter ...

<table>
<thead>
<tr>
<th></th>
<th>überhaupt nicht</th>
<th>ein wenig</th>
<th>ziemlich</th>
<th>stark</th>
<th>sehr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

42. starker Befangenheit im Umgang mit anderen

43. Abneigung gegen Menschenmengen, z. B. beim Einkaufen und im Kino

44. dem Eindruck, sich einer anderen Person nie so richtig nahe fühlen zu können

45. Schreck- oder Panikanfälle

46. der Neigung, immer wieder in Erörterungen und Auseinandersetzungen zu geraten

47. Nervosität, wenn Sie alleine gelassen werden

48. mangelnder Anerkennung Ihrer Leistungen durch andere

49. so starker Ruhelosigkeit, dass Sie nicht still sitzen können

50. dem Gefühl, wertlos zu sein

51. dem Gefühl, dass die Leute Sie ausnutzen, wenn Sie dies zulassen würden

52. Schuldgefühlen

53. dem Gedanken, dass irgendetwas mit Ihrem Verstand nicht in Ordnung ist
A2: BECKS DEPRESSION INVENTORY (BDI-II)


1.) Traurigkeit
   0   Ich bin nicht traurig.
   1   Ich bin oft traurig.
   2   Ich bin ständig traurig.
   3   Ich bin so traurig oder unglücklich, dass ich es nicht aushalten kann.

2.) Pessimismus
   0   Ich bin nicht mutlos, was meine Zukunft angeht.
   1   Ich bin mutloser als früher, was meine Zukunft angeht.
   2   Ich glaube nicht, dass sich meine Lage verbessert.
   3   Ich habe das Gefühl, dass es keine Hoffnung gibt für meine Zukunft und es nur noch schlimmer wird.

3.) Frühere Misserfolge
   0   Ich fühle mich nicht als Versager.
   1   Ich habe öfter versagt, als ich sollte.
   2   Wenn ich zurück blicke, sehe ich eine Menge Misserfolge.
   3   Ich fühle mich persönlich als totaler Versager.

4.) Verlust von Freude
   0   Ich habe so viel Freude wie immer an den Dingen, die mir Spaß machen.
   1   Ich habe nicht mehr so viel Spaß an den Dingen wie früher.
   2   Ich habe sehr wenig Freude an den Dingen, die mir früher Spaß gemacht haben.
   3   Ich habe keine Freude an den Dingen, die mir früher Spaß gemacht haben.
5.) **Schuldgefühle**

<table>
<thead>
<tr>
<th></th>
<th>Beschreibung</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Ich habe keine besonderen Schuldgefühle.</td>
</tr>
<tr>
<td>1</td>
<td>Ich habe bei vielen Dingen, die ich getan habe oder hätte tun sollen, Schuldgefühle.</td>
</tr>
<tr>
<td>2</td>
<td>Ich habe die meiste Zeit Schuldgefühle.</td>
</tr>
<tr>
<td>3</td>
<td>Ich habe ständig Schuldgefühle.</td>
</tr>
</tbody>
</table>

6.) **Gefühle, bestraft zu werden**

<table>
<thead>
<tr>
<th></th>
<th>Beschreibung</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Ich habe nicht das Gefühl, für etwas bestraft zu werden.</td>
</tr>
<tr>
<td>1</td>
<td>Ich habe das Gefühl, dass ich vielleicht für etwas bestraft werde.</td>
</tr>
<tr>
<td>2</td>
<td>Ich glaube, dass ich für etwas bestraft werde.</td>
</tr>
<tr>
<td>3</td>
<td>Ich habe das Gefühl für etwas bestraft zu werden.</td>
</tr>
</tbody>
</table>

7.) **Abneigungen gegen sich selbst**

<table>
<thead>
<tr>
<th></th>
<th>Beschreibung</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Meine Gefühle mir gegenüber sind die gleichen geblieben.</td>
</tr>
<tr>
<td>1</td>
<td>Ich habe das Vertrauen in mich verloren.</td>
</tr>
<tr>
<td>2</td>
<td>Ich bin von mir selbst enttäuscht.</td>
</tr>
<tr>
<td>3</td>
<td>Ich mag mich nicht.</td>
</tr>
</tbody>
</table>

8.) **Selbstvorwürfe**

<table>
<thead>
<tr>
<th></th>
<th>Beschreibung</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Ich bin mir selbst gegenüber nicht kritischer als sonst und mache mir nicht mehr Vorwürfe als sonst.</td>
</tr>
<tr>
<td>1</td>
<td>Ich bin mir selbst gegenüber kritischer als früher.</td>
</tr>
<tr>
<td>2</td>
<td>Ich mache mir Vorwürfe für alle meine Fehler.</td>
</tr>
<tr>
<td>3</td>
<td>Ich gebe mir die Schuld, für alles Schlimme, was passiert.</td>
</tr>
</tbody>
</table>

9.) **Selbstmordsgedanken oder -wünsche**

<table>
<thead>
<tr>
<th></th>
<th>Beschreibung</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Ich denke nie daran, mich umzubringen.</td>
</tr>
<tr>
<td>1</td>
<td>Ich habe Selbstmordsgedanken, aber ich würde sie nicht ausführen.</td>
</tr>
<tr>
<td>2</td>
<td>Ich möchte mich umbringen.</td>
</tr>
<tr>
<td>3</td>
<td>Ich würde mich umbringen, wenn ich die Möglichkeit hätte.</td>
</tr>
</tbody>
</table>

10.) **Weinen**

<table>
<thead>
<tr>
<th></th>
<th>Beschreibung</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Ich weine nicht mehr als früher.</td>
</tr>
<tr>
<td>1</td>
<td>Ich weine mehr als früher.</td>
</tr>
<tr>
<td>2</td>
<td>Ich weine wegen jeder Kleinigkeit.</td>
</tr>
<tr>
<td>3</td>
<td>Mir ist nach Weinen zumute, aber ich kann nicht.</td>
</tr>
</tbody>
</table>
11.) **Unruhe**

0  Ich bin nicht unruhiger oder erregter als sonst.
1  Ich bin unruhiger oder erregter als sonst.
2  Ich bin so unruhig oder erregt, dass es schwer ist, mich nicht zu bewegen.
3  Ich bin so unruhig oder erregt, dass ich ständig in Bewegung bleiben oder etwas tun muss.

12.) **Interesselosigkeit**

0  Ich habe das Interesse an anderen Menschen oder an Tätigkeiten nicht verloren.
1  Ich bin weniger an anderen Menschen oder Dingen interessiert als vorher.
2  Ich habe mein Interesse an anderen Menschen oder Dingen zum größten Teil verloren.
3  Es ist schwer für irgendetwas Interesse aufzubringen.

13.) **Entschlussfähigkeit**

0  Ich treffe Entscheidungen etwa so leicht wie immer.
1  Es fällt mir schwerer als sonst, Entscheidungen zu treffen.
2  Ich habe viel größere Schwierigkeiten, Entscheidungen zu treffen, als früher.
3  Ich habe Mühe, überhaupt Entscheidungen zu treffen.

14.) **Wertlosigkeit**

0  Ich fühle mich nicht wertlos.
1  Ich halte mich nicht für so wertvoll und nützlich wie früher.
2  Ich habe das Gefühl, weniger Wert zu sein als andere Menschen.
3  Ich habe das Gefühl, völlig wertlos zu sein.

15.) **Verlust an Energie**

0  Ich habe so viel Energie wie immer.
1  Ich habe weniger Energie als früher.
2  Ich habe nicht genügend Energie, sehr viel zu tun.
3  Ich habe nicht genügend Energie, irgend etwas zu tun.

16.) **Veränderung der Schlafgewohnheiten**

0  Meine Schlafgewohnheiten haben sich nicht geändert.
1a  Ich schlafe etwas mehr als sonst.
1b  Ich schlafe etwas weniger als sonst.
2a  Ich schlaf viel mehr als sonst.
2b  Ich schlafe viel weniger als sonst.
3a  Ich schlafe die meiste Zeit des Tages.
3b Ich wache 1-2 Stunden zu früh auf und kann dann nicht mehr einschlafen.

---------------------------------------

17.) Reizbarkeit
0 Ich bin nicht reizbarer als sonst.
1 Ich bin reizbarer als sonst.
2 Ich bin viel reizbarer als sonst.
3 Ich bin ständig reizbar.

---------------------------------------

18.) Veränderung des Appetits
0 Mein Appetit hat sich nicht verändert.
1a Mein Appetit ist etwas kleiner als sonst.
1b Mein Appetit ist etwas größer als sonst.
2a Mein Appetit ist viel kleiner als vorher.
2b Mein Appetit ist viel größer als vorher.
3a Ich habe überhaupt keinen Appetit.
3b Ich habe ständig großen Hunger.

---------------------------------------

19.) Konzentrationsschwierigkeiten
0 Ich kann mich so gut konzentrieren wie immer.
1 Ich kann mich nicht so gut konzentrieren wie sonst.
2 Es fällt mir schwer, mich sehr lange auf etwas zu konzentrieren.
3 Ich kann mich auf gar nichts konzentrieren.

---------------------------------------

20.) Müdigkeit
0 Ich bin nicht müder als sonst.
1 Ich werde schneller müde als sonst.
2 Ich bin für viele Dinge, die ich früher gern getan habe, zu müde.
3 Ich bin für die meisten Dinge, die ich früher gern getan habe, zu müde.

---------------------------------------

21.) Verlust des Interesses am Sex
0 Ich habe in letzter Zeit keine Veränderung meines Interesses am Sex bemerkt.
1 Ich habe weniger Interesse am Sex als früher.
2 Ich habe jetzt viel weniger Interesse am Sex.
3 Ich habe das Interesse am Sex völlig verloren.

---------------------------------------
**APPENDIX A3: RESPONSE STYLES QUESTIONNAIRE (RSQ-D)**

Menschen denken und verhalten sich ganz unterschiedlich, wenn sie sich traurig oder niedergeschlagen fühlen.

Bitte kreuzen Sie bei allen nachfolgenden Aussagen an, ob Sie dies "fast nie", "manchmal", "oft" oder "fast immer" denken oder tun, wenn Sie sich traurig, niedergeschlagen oder deprimiert fühlen.

Bitte geben Sie jeweils an, was Sie üblicherweise tun, wenn Sie sich traurig oder niedergeschlagen fühlen, *nicht*, was Sie Ihrer Meinung nach tun sollten!

<table>
<thead>
<tr>
<th>Wenn ich mich traurig oder niedergeschlagen fühle:</th>
<th>fast nie</th>
<th>manchmal</th>
<th>oft</th>
<th>fast immer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. denke ich daran, wie allein ich mich fühle</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. denke ich, „ich werde nicht fähig sein, meine Arbeit zu tun, weil ich mich so schlecht fühle“</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. denke ich daran, wie erschöpft ich mich fühle</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. denke ich, wie schwer es ist, mich zu konzentrieren</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. versuche ich etwas Positives an der Situation zu finden oder etwas, was ich dabei gelernt habe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. denke ich, „ich werde jetzt etwas tun, um mich besser zu fühlen“</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. helfe ich jemand anderem bei irgendetwas, um mich abzulenken</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. denke ich daran, wie passiv und unmotiviert ich bin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. sage ich mir, dass diese Gefühle nicht anhalten werden</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. denke ich daran, dass ich überhaupt nichts mehr zu fühlen scheine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. denke ich, „warum komme ich nicht in Schwung“</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. denke ich, „warum reagiere ich immer so“</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. gehe ich an einen Lieblingsort, um mich von meinen Gefühlen abzulenken</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. tue ich etwas, das mich in der Vergangenheit hat besser fühlen lassen</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wenn ich mich traurig oder niedergeschlagen fühle:</td>
<td>fast nie</td>
<td>manchmal</td>
<td>oft</td>
<td>fast immer</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------------------------------------</td>
<td>----------</td>
<td>----------</td>
<td>-----</td>
<td>------------</td>
</tr>
<tr>
<td>16. denke ich über eine zurückliegende Situation nach und wünsche, dass diese besser gelaufen wäre</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. denke ich, „ich werde jetzt ausgehen und etwas Spaß haben“</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. denke ich daran, dass ich mich nicht stark genug fühle um irgendetwas zu tun</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. denke ich über meine Persönlichkeit nach und versuche zu verstehen, weshalb ich depressiv bin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. gehe ich irgendwohin, wo ich alleine bin, um über meine Gefühle nachzudenken</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. höre ich traurige Musik</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. ziehe ich mich zurück und denke über die Gründe nach, weshalb ich mich so traurig fühle</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. versuche ich, mich selbst zu verstehen, indem ich mich auf meine depressiven Gefühle konzentriere</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX A4: HOPKINS SYMPTOM CHECKLIST -11 (HSCL-11)

Im Folgenden sind belastende Probleme und Beschwerden aufgelistet, die man manchmal hat. Bitte lesen Sie jede Aussage sorgfältig durch und entscheiden Sie, wie sehr Sie jedes einzelne Problem in der letzten Woche gestört oder geplagt hat.

Bitte beantworten Sie jede Frage!

<table>
<thead>
<tr>
<th>überhaupt</th>
<th>ein</th>
<th>ziemlich</th>
<th>sehr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Wie sehr litten Sie in den letzten sieben Tagen unter...

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Furchtsamkeit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Nervosität oder innerem Zittern</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. dem Gefühl, gespannt oder aufgeregt zu sein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Schreck- oder Panikanfällen</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Einschlafschwierigkeiten</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. einem Gefühl der Hoffnungslosigkeit angesichts der Zukunft</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Schwermut</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Einsamkeitsgefühlen</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. dem Gefühl, sich für nichts zu interessieren</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Gedanken an den Tod und das Sterben</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. dem Gefühl, wertlos zu sein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX A5: EMA ITEMS

I. Compass Item

Wie gut sind Sie in den letzten 3-4 Stunden zurecht gekommen?

II. PANAS Items

Negative Affect:
1. Momentan fühlen Sie sich niedergeschlagen?
2. Momentan fühlen Sie sich beschämt?
3. Momentan fühlen Sie sich ängstlich?
4. Momentan fühlen Sie sich nervös?

Positive Affect:
1. Momentan fühlen Sie sich freudig erregt?
2. Momentan fühlen Sie sich entschlossen?
3. Momentan fühlen Sie sich wach?
4. Momentan fühlen Sie sich aktiv?

III. Rumination und einhergehende sensorische Empfindungen

1. Während der letzten drei Stunden musste ich ständig über eine Situation nachdenken und wünschte mir, sie wäre anders verlaufen.

5. Waren Ihre Gedanken mit auditiven Eindrücken - z.B. dem Hören eines Inneren Kritikers - verbunden?
6. Waren Ihre Gedanken mit visuellen Eindrücken - z.B. dem Wiedererleben vor dem Inneren Auge - verbunden?
7. Waren Ihre Gedanken mit körperlichen, taktilen oder olfaktorischen Eindrücken verbunden?
IV. Soziale Unterstützung und Selbstwirksamkeitserleben

1. Während der letzten drei Stunden hat mir jemand Vertrauen entgegen gebracht.
2. Während der letzten drei Stunden hat mir jemand konkrete Hilfe angeboten.
5. Während der letzten drei Stunden habe ich meine eigenen Interessen und Bedürfnisse verfolgt.

V. Life Events

1. Ist Ihnen während der letzten drei Stunden etwas Ärgerliches widerfahren?
2. Ist Ihnen während der letzten drei Stunden etwas Angenehmes widerfahren?
Sie haben an einer Studie der Forschungsambulanz der Universität Trier teilgenommen, in deren Rahmen Sie mehrfach pro Tag einen kurzen Fragebogen auf dem iPod ausgefüllt haben. Im Folgenden möchten wir gerne mehr über Ihre Erfahrungen und Einschätzungen bezüglich des zweiwöchigen Erhebungszeitraumes erfahren.

<table>
<thead>
<tr>
<th></th>
<th>gar nicht</th>
<th>Leicht</th>
<th>mäßig</th>
<th>stark</th>
<th>sehr stark</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Inwieweit hat Ihnen die tägliche Selbstbeobachtung geholfen?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Hat sich die tägliche Selbstbeobachtung positiv auf Ihre Stimmung ausgewirkt?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Hat sich die tägliche Selbstbeobachtung negativ auf Ihre Stimmung ausgewirkt?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Haben Sie die täglichen Erhebungen als Belastung empfunden?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Haben die Menschen in Ihrer Umgebung positiv auf Ihre Studienteilnahme reagiert?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Haben die Menschen in Ihrer Umgebung negativ auf Ihre Studienteilnahme reagiert?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Kam das Tonsignal zu ungünstigen Zeiten?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Hatte Sie Erklärungsnot anderen gegenüber?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frage</td>
<td>gar nicht</td>
<td>Leicht</td>
<td>Mäßig</td>
<td>stark</td>
<td>sehr stark</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-----------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>------------</td>
</tr>
<tr>
<td>9. Hatten Sie Sorge, den iPod zu vergessen/haben ihn tatsächlich vergessen?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Hatten Sie Sorge, den iPod falsch zu bedienen / ihn zu beschädigen / vergessen zu laden?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Hatten Sie Sorge, dass Tonsignal zu verpassen?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Dauerte Ihnen das Ausfüllen des Fragebogens zu lang?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Hatten Sie Schwierigkeiten, den iPod zu bedienen?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Haben die Erhebungen bei Ihnen ein bewussteres Wahrnehmen Ihrer Gefühle gefördert?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Haben die Erhebungen Sie dazu angeregt, mehr Ihren eigenen Interessen und Bedürfnissen nachzugehen?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Empfinden Sie die Aufwandsentschädigung als angemessen?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Ist Ihnen während des Erhebungszeitraumes etwas Unangenehmes widerfahren?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Gibt es ansonsten etwas, was Sie uns bezüglich des 2-wöchigen Erhebungszeitraumes rückmelden möchten?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Vielen Dank für Ihre Teilnahme!
APPENDIX B1: PIE CHARTS OF SINGLE ITEM ANALYSES FROM THE EMA EVALUATION FORM

Figure B1.1: Did the people in your surroundings react positively to your study participation?

Figure B1.2: Did the people in your surroundings react negatively to your study participation?

Figure B1.3: Did you have difficulty giving others an explanation?

Figure B1.4: Did you have difficulties using the iPod?
Figure B1.5: Did you worry about forgetting the iPod/Did you actually forget it?

Figure B1.6: Did you feel like it took too long to fill out the questionnaire?

Figure B1.7: Did you worry about handling the iPod incorrectly/Did you damage it/forget to charge it?

Figure B1.8: Do you feel the compensation for expenses was appropriate?
APPENDIX C1: INFORMED CONSENT

INFORMATIONEN ZUR STUDIENTEILNAHME

1. INHALT

2. RISIKO UND NEBENWIRKUNGEN


3. DATENSCHUTZ

Bei dieser Studie werden personenbezogene und gesundheitsbezogene Daten erhoben, gespeichert und ausgewertet. Die erhobenen Daten werden an einen Code gekoppelt und so anonymisiert. Daten der Probanden/innen, die nicht zur Teilnahme zugelassen werden können (nach dem telefonischen Interview) oder die ihre Teilnahme zu einem beliebigen Zeitpunkt einfach abbrechen möchten, werden gelöscht.

4. BEDINGUNGEN

Der Proband / die Probandin erhält für die Studienteilnahme eine Aufwandsentschädigung von 80,00€. Die Teilnahme ist freiwillig und die Wartezeit wird dadurch nicht beeinflusst. Die Teilnahme kann jederzeit von dem Probanden/ der Probandin ohne Angabe von Gründen und ohne dass ihm/ihr daraus Nachteile entstehen widerrufen werden.
ERKLÄRUNG

Name: 

Vorname: 

Straße: 

Wohnort: 


☐ Ich habe alle Informationen über die Studie verstanden. Ich bestätige zudem, dass ich aus freiem Willen an dieser Studie teilnehme.

☐ Mir ist auch bekannt, dass die von mir erhobenen Daten anonymisiert werden und dass ich jederzeit das Recht habe, die Löschung dieser Daten zu verlangen.

☐ Auch bestätige ich, dass ich mir im Fall eines Notfalles (z. B. akute Suizidalität) sofort Hilfe holen werde.

☐ Hiermit bestätige ich, dass ich über eventuelle Risiken und Nebenwirkungen, die mit der Teilnahme an dieser Studie verbunden sind, informiert wurde. Ich bin mir dieser Risiken bewusst und wurde informiert, dass die Studienleiter nicht für eventuelle Schäden haftbar gemacht werden können.

Datum

Unterschrift Proband/in
APPENDIX C2: LIABILITY AGREEMENT

ERKLÄRUNG

Name: Vorname:

Straße: Wohnort:

NUTZUNGSREGELN

Ich bin ausführlich und verständlich in die Nutzung des iPods eingewiesen worden. Ich werde den iPod nur für den Zweck dieser Untersuchung gebrauchen.

HAFTUNG

Ich verpflichte mich, den iPod sachgerecht und pfleglich zu behandeln und für eine sichere Aufbewahrung zu sorgen. Für Schäden und Verluste, die durch mein Verschulden entstanden sind, übernehme ich die Haftung. In diesem Fall leiste ich der Universität Trier den vollen Ersatzwert.

LEIHZET

Die Leihe wird vom ______________ bis zum ______________ vereinbart.

Ich verpflichte mich, die mir entliehenen Gegenstände (Apple iPod touch 4G MP3-Player, 32 GB, sowie Kabel und Netzteilstecker) nach Ablauf der Laufzeit ohne Aufforderung zurückzugeben.

Datum Datum

Unterschrift Verleiher Unterschrift Entleiher
DECLARATION OF AUTHORSHIP

I hereby declare that the thesis submitted is my own unaided work. All direct or indirect sources used are acknowledged as references.

This dissertation was not previously presented to another examination board in order to obtain an academic title.

EIDESSTATTLICHE ERKLÄRUNG

Hiermit erkläre ich, dass die vorliegende Dissertationsschrift von mir selbständig angefertigt wurde und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet wurden. Zudem wurde die Arbeit an keiner anderen Universität zur Erlangung eines akademischen Grades eingereicht.

Trier, den ______________

________________________
Kristin Husen