Interpretation Bias in the Context of Depressed Mood: Assessment Strategies and the Role of Self-Generation in Cognitive Bias Modification

Von der Fakultät für Human- und Sozialwissenschaften der Technischen Universität Chemnitz genehmigte Dissertation zur Erlangung des akademischen Grades doctor rerum naturalium (Dr. rer. nat.)

von Heike Rohrbacher, geboren am 30.06.1977 in München

Gutachter: Prof. Dr. Stephan Mühlig
Prof. Dr. Mike Rinck

Tag der Einreichung: 30.11.2015
Tag der Verteidigung: 04.07.2016
Für Federico und Laura
## Contents

I. **Introduction** .................................................................................................................. 4  
   Cognitive theories of depression.................................................................................... 6  
   Interpretation bias and depression............................................................................. 11  
   Assessment of interpretation bias.............................................................................. 12  
   Cognitive modification of interpretation bias......................................................... 14  
   Limitations, open questions, and aims of the dissertation.................................... 16

II. **Paper 1** ....................................................................................................................... 20

III. **Paper 2** ...................................................................................................................... 33

IV. **Paper 3** ....................................................................................................................... 47

V. **General discussion** .................................................................................................... 63  
   Summary of empirical findings.................................................................................... 63  
   Theoretical and practical implications....................................................................... 66  
   Limitations and suggestions for future studies....................................................... 69  
   Final conclusions....................................................................................................... 71

VI. **References** ................................................................................................................ 72

VII. **Appendix** .................................................................................................................. 81

### Summary

**Zusammenfassung**

**Danksagung**

**Eidesstattliche Erklärung**

**Lebenslauf**
Introduction

Depression is a highly prevalent mental disorder that is estimated to affect more than 120 million people worldwide (Lépine & Briley, 2011). Projections by the World Health Organisation for the year 2030 regard unipolar depression as the leading cause of burden of all diseases (WHO, 2008). As defined by the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-V) depression comprises symptoms such as low mood, fatigue or loss of energy, loss of interest, withdrawal, feelings of guilt and low self-esteem, diminished ability to think or concentrate and sleeping problems (American Psychiatric Association, 2013). These symptoms can occur at any age and often become chronic (Paykel, Brugha, & Fryers, 2005). Depressive disorders highly impair people’s everyday functioning and are associated with decreased workplace productivity, severe psychosocial problems, and increased mortality risk (see Lépine & Briley, 2011). Despite the availability of a broad range of empirically well-evaluated and well-established therapeutical treatments, relapse rates remain high. Recurrence rates for major depressive disorder after treatment in specialized mental health institutions were reported to range between 60% (after 5 years) and 85% (after 15 years), indicating that current therapies are limited in efficacy (Hardeveld, Spijker, De Graaf, Nolen, & Beekman, 2010). More profound insights into the aetiology and maintenance of the disorder are needed in order to develop novel and more effective treatments.

Over the past three decades, a promising and fast-growing line of research has emphasized the role of cognitive biases as risk factors for mental disorders, including depressive disorders (for reviews see Gotlib & Joormann, 2010; Mathews & MacLeod, 2005). Cognitive biases describe abnormalities in information processing that are not necessarily the product of a deficit in function but can rather be described as a tendency to filter information in a selective way (e.g. Browning, Blackwell, & Holmes, 2013). A review of research in this area indicates that biases in attention, memory and interpretation are common across emotional disorders, although they might vary in form according to the type of disorder (Mathews & McLeod, 2005). In other words, individuals with mental disorders seem to have a different cognitive “software” than individuals without mental diagnoses and process - especially emotionally relevant - information in a rather negative way.

An interpretation bias refers to the tendency to interpret ambiguous stimuli in a threatening or pessimistic way. In everyday life, we encounter numerous examples of information that can be interpreted in more than one way, for instance facial expressions,
feedback from others or physiological cues. How we disambiguate or make sense of these stimuli is important for how we further respond to the world. To interpret one’s dialogue partner’s yawning as a sign of boredom can obviously result in a different, even opposite emotion than attributing it to a simple lack of oxygen in the room. There is empirical consensus that negatively biased interpretation is associated with dysphoria and depression and might play a key role in the development and maintenance of the disorder (e.g. Dearing & Gotlib, 2009; Eley, Gregory, Lau et al., 2008; Mogg, Bradbury, & Bradley, 2006; Rude, Wenzlaff, Gibbs, Vane, & Whitney, 2002).

If cognitive biases are crucial in the aetiology of depression, could a modification of these tendencies be effective in reducing the symptoms? This research hypothesis has led to the development of cognitive bias modification procedures. Cognitive bias modification paradigms (CBM; Grey & Mathews, 2000; Mathews & Macintosh, 2000) are computerized training procedures that encourage individuals to adopt a more positive information processing style. For instance, in interpretation bias training a more positive interpretation style can be trained by repeatedly presenting virtual scenarios that combine initially ambiguous information with a clearly positive outcome. Unlike cognitive therapy (e.g. Beck, 1997), which is complex and comprises a range of different therapeutic strategies, CBM-procedures aim to target as specifically as possible the cognitive process of interest, for instance interpretation of ambiguous material (Browning et al., 2013). Although CBM targeting interpretation (CBM-I) has initially been developed for the training of anxiety-related bias, positive benefits of interpretation modification training could also be reported in the context of dysphoria and depression by a number of studies (e.g. Blackwell & Holmes, 2010; Holmes, Lang, & Shah, 2009; Lothmann, Holmes, Chan, & Lau, 2011). These findings could offer exciting new possibilities for therapy and led to a fast-growing line of research.

Surprisingly, the increasing interest in the modifiability of interpretation bias is contrasted by a severe lack of methodological and measurement development studies. However, reliable and valid measures are required in order to evaluate and further advance these promising procedures. Especially, instruments with high psychometric qualities justifying repeated administration have to be developed to adequately measure the effect of CBM-I procedures.

Further, studies so far yielded inconsistent findings for the existence of depression-related interpretation bias, depending on the (direct or indirect) assessment paradigm. Several accounts have been discussed for these discrepant findings. However, to date, no study has ever
examined the different assessment paradigms in one sample, thus no knowledge about the relatedness of the different direct and indirect assessment strategies exists.

Although studies on CBM-I in the context of depression and dysphoria have already yielded promising results, this branch of research is still relatively young. Before implementing CBM-I procedures into the therapeutic context, more profound knowledge about the underlying work mechanism and optimal ingredients needs to be obtained.

The aim of the present work is to bridge this gap in research by developing, psychometrically evaluating and comparing several direct and indirect assessment paradigms for the measurement of depression-related interpretation bias. A further objective of this dissertation is to increase the knowledge of the underlying work mechanism and optimal components of imagery-based CBM-I in the context of depressed mood, by comparing the efficacy of guided CBM-I to a novel, more active training version that requires participants to generate the positive interpretations themselves.

The first chapter of this dissertation provides an overview of the most influential cognitive frameworks of depression, of the concept of interpretation bias and its relation to depressive symptoms, of diagnostic measures as well as of the modifiability of interpretation bias via CBM-I procedures. At the end of this chapter, gaps and limitations of the current research are illustrated and a description of the precise aims of the dissertation is outlined. In chapters 2 through 4, the empirical research, comprising two methodological studies and a CBM-I study, is presented. Chapter 5 delivers a conclusive review of the work.

**Cognitive theories of depression**

Over the last 40 years, several cognitive frameworks have been advanced to explain the onset and maintenance of depressive disorders. These models emphasize the link between cognitive biases and depressive symptoms and propose that intra-individual differences in information processing (e.g. memory, attention, interpretation) account, among other (e.g. biological) factors, for the vulnerability to develop the disorder. Although virtually all cognitive theories predict that depression is characterized by biased processing of emotional material, they propose distinct causal pathways and differ on how they integrate multiple biased aspects of information processing. The most influential cognitive frameworks are described here below.
I. INTRODUCTION

Beck’s schema model

Beck’s cognitive theory of depression (Beck, 1967; Beck, 1984; Beck, 1997; Beck, Rush, Shaw, & Emery, 1979) has contributed in a major way to the understanding of depression and has served as the foundation for cognitive therapy. Beck argued that depression is, at least in part, the consequence of negative beliefs and biased cognitions. He proposed that early adverse experiences lead to the development of dysfunctional schemata, including themes of loss, failure and rejection. These dysfunctional schemata can be activated by congruent stressors and evoke negative thinking styles. These thinking styles concern negative thoughts about the self, the world and the future, a concept known as Beck’s cognitive triad. Due to these biased cognitions, depression-prone individuals attend selectively to negative information, interpret ambiguous situations in a pessimistic manner and memorize more negative events, increasing depressed mood. These processes can lead into a vicious circle of biased processing, negative automatic thoughts, and depressive symptoms.

Figure I.1. Depiction of interrelations among cognitive biases in Beck’s schema model

I. INTRODUCTION

as Ingram (1984) and Williams, Watts, MacLeod, and Mathews (1988, 1997), assumed a more specific account of cognitive biases in depression.

**Ingram’s information processing analysis**

Following work by Bower (1981), Ingram (1984) formulated an information processing analysis of depression, which emphasized the role of biased elaboration and memory in the development of depression. Bower (1981) postulated in his work on mood and memory the existence of associative networks that lead to biased processing in individuals at risk for depression. In line with this, Ingram (1984) proposed that the initial development of depressive symptoms results from the activation of an affective structure referred to as a depression-emotion node. The activation of this depression-emotion node is determined by the appraisal of a certain life event (e.g. separation from a significant other). Once the unit is activated above threshold, activation spreads through the network, causing its related thoughts, contents and memories to become accessible. Through selective processing styles, for instance negative

**Figure I.2.** Schematic depiction of interrelations among cognitive biases in Ingram’s information processing analysis

I. INTRODUCTION

Appraisal, connected memory networks can be further activated by associative linkages. These connected networks may, for instance, contain memories that are related to past feelings of depression (e.g. failure at a job interview). In other words, the individual might recall more and more negative memories that serve to maintain the depressive symptoms.

**Williams et al.’s cognitive framework.**

In their influential framework on the influences of cognitive biases on depression and anxiety, Williams et al. (1988; revised 1997) argued that the function of the disorder determines the kind of information processing bias. While the function of anxiety is to prevent danger, the function of depression might be to cope with failure or loss. As a result anxiety is assumed to facilitate the perceptual (data-driven) processing of danger-related information, whereas depression is associated with negative biases in conceptual (top-down) processing of internally generated material. In their revised model, Williams et al. (1997) argued that depression is characterized by enhanced or biased elaboration upon negative stimuli. This biased elaboration leads to enhanced or biased memory for similar information. According to Williams et al.,

*Figure I.3. Schematic depiction of interrelations among cognitive biases in Williams et al.’s (revised) cognitive framework*

I. INTRODUCTION

depressed individuals exhibit biases in interpretation and explicit memory, but do not show mood-congruent biases in attention, which contradicts predictions made by Beck’s depression model (Clark, Beck, & Alford, 1999).

Impaired cognitive control account

Joormann, Yoon, and Zetsche (2007) proposed that depressed individuals show deficits in cognitive inhibition, which is a key mechanism of cognitive control. They argued that depressed people have difficulties in the inhibition of negative material. This leads to prolonged processing of depression-related information or rumination, which increases negative affect and therefore depressive symptoms. Depressed individuals are not able to prevent mood-congruent information from entering working memory, which leads – via the process of rumination – to better long-term memory for negative information.

Figure I.4. Schematic depiction of interrelations among cognitive biases in Joorman et al.’s impaired cognitive control account

I. INTRODUCTION

The combined cognitive bias hypothesis

Everaert et al. (2012) applied the combined cognitive bias hypothesis (CCBH) to depressive disorders. The CCBH was originally formulated for anxiety disorders (Hirsch, Clark, & Mathews, 2006) and states that cognitive biases are not isolated processes but influence each other. A number of studies found evidence of the interplay between cognitive biases. In this context associations between depression-related biases in attention and explicit memory (Ellis, Beevers, & Wells, 2011; Koster, De Raedt, Leyman, & De Lissnyder, 2010; Wells, Beevers, Robison, & Ellis, 2010) as well as associations between attention biases and interpretation biases (Hertel & El-Messidi, 2006) were reported in dysphoric individuals. Interrelations were also found between interpretation biases and memory biases (Salemink, Hertel, & Mackintosh, 2010; Tran, Hertel, & Joormann, 2011). Everaert, Duyck, and Koster (2014) attained a good fit for a path analysis model in which selective orienting of attention was associated with interpretation bias, which in turn was associated with a congruent bias in memory.

Interpretation bias and depression

There is a long held theoretical view, that biased interpretation plays an important role in the onset and maintenance of depression (e.g. Beck, 1967). Interpretation refers to the process of making sense of a situation or of a stimulus whose valence is unclear and negatively biased interpretation is defined as the tendency to interpret ambiguous information in a pessimistic manner. In everyday life, we experience numerous examples of information that can be interpreted in more than one way. How we disambiguate or assign meaning to ambiguous stimuli (e.g. facial expressions, feedback from others, physiological cues) is important in determining how we further respond to the world. To interpret the behaviour of a neighbour who does not greet when met on the street as a sign of dislike can obviously lead to a different emotion than attributing his behaviour to a moment of distraction. A number of studies found evidence that biased interpretation is associated with dysphoria and depression (Blackwell & Holmes, 2010; Butler & Mathews, 1983; Eley et al., 2008; Hertel & El-Messidi, 2006; Mathews & Mackintosh, 2000; Mogg et al., 2006; Rude et al., 2002). When presented with ambiguous textual scenarios (e.g. “You wake with a start in the middle of the night thinking you heard a noise, but all is quiet. What do you suppose woke you up?”), individuals with depressed mood
are more likely to select a negative (e.g. “It could be a burglar”) interpretation (Butler & Mathews, 1983). Empirical support was found to suggest that interpretation bias could be a risk factor for the development of depression (e.g. Dearing & Gotlib, 2009; Reinecke, Deiprose, Browning, Williams, & Holmes, submitted). For example, Dearing and Gotlib (2009) reported negative interpretation tendencies in daughters of depressed mothers, whereas daughters of non-disordered mothers did not show these biases. The finding that a population known to be at risk for depression already exhibits biased interpretation can be regarded as a support for cognitive vulnerability models. Further studies indicated that the experimental modification of negative biases towards more optimistic interpretation positively influences other markers of depression (Blackwell & Holmes, 2010; Holmes, Lang, & Shah, 2009; Holmes, Mathews, Dalgleish, & Mackintosh, 2006; Lang, Blackwell, Harmer, Davison, & Holmes, 2011; Reinecke et al., submitted). These findings are in line with Beck’s cognitive theory of depression (Beck, 1967; Beck, 1984; Beck et al., 1979), assuming that biased processing styles, such as negative interpretation, cause and maintain depressive feelings.

Assessment of interpretation bias

Numerous studies found biased interpretation in dysphoric and depressed individuals (e.g. Blackwell & Holmes, 2010; Butler & Mathews, 1983; Eley et al., 2008; Rude et al. 2002). However, empirical findings for the existence of biased interpretation in depression were mixed depending on the respective interpretation bias assessment method. Researchers have proposed a variety of different methods and paradigms to study interpretation styles in depressed and dysphoric individuals. Cognitive interpretation habits are difficult to measure because they are assumed to be largely automatic and beyond the awareness of the individual. Studies so far have therefore tried to assess the interpretation tendencies of individuals by providing them with ambiguous material.

For instance, researchers have presented individuals with a series of homophones (words that have the same sound, but different emotional meanings, e.g. “pane/pain”) and required them to write down the words, to spontaneously form a sentences using them or to report mental images provoked by the homophones (e.g. Eley et al., 2008, Hertel & El-Messidi, 2006, Mogg et al., 2005). In these tasks, negatively biased interpretation was indicated when participants chose the word with the aversive meaning. The homophone paradigm was shown
I. INTRODUCTION

to significantly differentiate between individuals with high and low BDI-II scores (Mogg et al., 2005). Interpretation biases were also detected via the Scrambled Sentences Task (SST; Wenzlaff, 1993). In this test, participants were instructed to unscramble a list of scrambled sentences. Their task was to form sentences using five of six given words (e.g. “good feel very I bad usually”), requiring participants to select either a positively (e.g. “I usually feel very good.”) or a negatively (e.g. “I usually feel very bad.”) valenced sentence. The valence of the selected or unscrambled sentence is regarded to be an indicator of their interpretation style. The SST was shown to significantly differentiate between formerly and non-depressed individuals (Rude, Covich, Jarrold, Hedlund, & Zentner, 2001) as well as to predict subsequent depressive disorders (Rude, Durham-Fowler, Baum, Rooney, & Maestas, 2010). Another paradigm used to measure interpretation bias is the presentation of ambiguous scenarios. In these tasks, ambiguous open-end situations are described (e.g. “You are hosting a dinner party to introduce your friends. At the table you notice how your friends are interacting.”), and subjects’ interpretation is assessed, for instance, via emotional valence ratings (e.g. Berna, Lang, Goodwin, & Holmes, 2011; Holmes et al., 2006). Higher unpleasantness ratings on these measures were found to be significantly associated with depressive symptoms (Berna et al., 2011).

Studies using indirect methods that are supposed to be less prone to response and desirability biases reported mixed findings for the existence of a depression-related interpretation bias. For instance, studies by Lawson and MacLeod (1999) and Bisson and Sears (2007), both using semantic priming paradigms, failed to find evidence for a dysphoria-related interpretive bias. Participants in the first study were required to pronounce a target word following a priming sentence. A subset of the prime sentences were ambiguous (e.g. “The doctor examined little Emily’s growth”), and were followed by a target word, that was either related to the negative or the neutral interpretation (“tumour” or “height” respectively in the former example). Semantic priming effects were measured using the pronunciation latency as dependent variable. The study by Bisson and Sears (2007) used a cross-modal semantic priming paradigm, in which participants had to listen to ambiguous prime sentences and make lexical decisions to target words. Each of the prime sentences (e.g. “Kathy had been committed for some time”) permitted a positive, negative, or neutral interpretation and was either paired with a positively related target (romance), a negatively related target (hospital), a neutrally related word (years), an unrelated word (sauce), or a non-word target. Individuals were required to decide whether the target was a word or a non-word (yes- or no-button). Results did not confirm the expectation that depressed participants would show an increased tendency to focus on the
negative interpretation of the primes and therefore show reaction delays in response to negatively related targets. Both priming studies questioned the existence of a depression-related interpretation bias, but rather suggested that depression and dysphoria could be associated with a negative response bias. However, Sears, Bisson, and Nielson (2011) contradicted this view by finding support for a negative interpretation bias in error rates but not in response latencies in a later priming study. Evidence for the existence of interpretation bias was additionally reported by a study that tested an ambiguous cue-conditioning paradigm (Schick, Wessa, Vollmayr, Kuehner, & Kanske, 2013) as well as a study using eye blink responses to ambiguous words (Lawson, MacLeod, & Hammond, 2002), although the latter finding could not be replicated in a second study using a clinical sample (Käse et al., 2013).

Cognitive modification of interpretation bias

Cognitive bias modification paradigms (Grey & Mathews, 2000; Mathews & Macintosh, 2000; Mathews & MacLeod, 2005) have been developed in order to identify the role of cognitive biases in the development of mental disorders and to reduce symptom severity by changing negative biases towards a more positive direction. They can be described as computerized training procedures that encourage individuals to adopt a more positive (or negative) information processing style. The idea behind CBM procedures is that information processing tendencies are malleable and that individuals can be trained to adapt an alternative processing style. For example, during attentional bias modification (CBM-A), participants are presented with a modified version of the dot-probe test and learn, over many trials, to consistently avoid negative stimuli and instead direct their attention towards positive stimuli (e.g. Browning et al., 2012). In interpretation bias training procedures (CBM-I), a more positive interpretation style can be trained by repeatedly presenting virtual scenarios that combine initially ambiguous information with a clearly positive outcome. Cognitive bias modification targeting interpretation has initially been developed for the training of anxiety-related bias, but positive benefits of interpretation modification training could also be reported in the context of dysphoria and depression by a number of studies (Holmes, Coughtrey, & Connor, 2008; Holmes, Lang, & Shah, 2009; Holmes et al., 2006; Lothmann et al., 2011; Reinecke et al., submitted). An example by Holmes et al. (2006) is as follows: “You have started an evening
I. INTRODUCTION

class which is tough going. You are determined to succeed, and after a while, it becomes *much easier and more enjoyable*” (positive resolution in italics).

Positive bias modification has not only been shown to be effective in modifying the interpretation of fresh ambiguous information (e.g. Reinecke et al., submitted) in the trained direction, but it also ameliorates depressed mood and positively influences resilience to negative mood induction (e.g. Holmes, Lang, & Shah, 2009).

A line of research (e.g. Holmes, Lang, & Shah, 2009) has been particularly interested in identifying the most effective components of these procedures to make them more effective in changing bias. Based on prior research suggesting a special relationship between imagery and emotion (see Holmes & Mathews, 2005), Holmes, Lang, and Shah (2009) found imagery, rather than verbal processing to be essential. In the study, individuals listened to 100 ambiguous scenarios that were consistently resolved in a positive manner. One group was instructed to imagine the events, whereas the other group was instructed to listen to them while thinking about their meaning. The latter condition was not only less effective in ameliorating interpretation bias as well as mood than imagery, but even led to paradoxical, negative mood responses. It was assumed that imagery might directly provoke emotion like a positive “as-if” experience, whereas verbal processing might be perceived as less believable. It might more readily provoke a comparison with one’s actual status quo, which could especially in depressive individuals turn out to be disadvantageous and thus mood deteriorating (Holmes et al, 2006; Holmes, Lang, & Shah, 2009). This finding could be successfully replicated by a second study (Nelis, Vanbrabant, Holmes, & Raes, 2012). Holmes, Coughtrey, and Connor (2008) could further specify that only imagery from a field perspective (“through your own eyes”) improves affect as opposed to imagery from an observer perspective (“looking at you”), that could - like verbal processing - lead to adverse effects. This outcome emphasizes the role of precise task instructions. A study by Clarke et al. (2014) further found the presence of ambiguity to be crucial, indicating that the mere positive valence of training scenarios alone does not lead to a positive bias modification.

A study in the context of anxiety-related interpretation bias found active selection of meaning during the training superior to a mere passive exposure in modifying subsequent emotional responses to new ambiguous stimuli (Hoppitt, Mathews, Yiend, & Mackintosh, 2010). In this study, participants were presented with threat-related ambiguous sentences that were negatively resolved by the final word, e.g. “You have decided to go caving even though you feel nervous about being in such an enclosed space. You get to the caves before anyone else arrived. Going deep inside the first cave you realize you have completely lost your way.”
I. INTRODUCTION

(negative resolution in italics). Individuals in the passive group were presented with the entire passage, whereas individuals in the active group were presented with only a fragment of the final word (“….completely lost your w--.”). They were asked to actively resolve the meaning by themselves (only one possible completion). Hoppitt et al. (2010) suggested that the superiority of the active condition may be due to the induction of an implicit production rule, in which individuals continue to actively generate training-congruent meanings of subsequent ambiguous scenarios after training.

Limitations, open questions, and aims of the dissertation

Lack of methodological and measurement development studies

The modification of depression-related interpretation bias towards a more optimistic interpretation style has become a fast-growing and promising line of research, offering exciting new possibilities for cognitive therapy. Surprisingly, the increasing interest in the modifiability of interpretation bias is in contrast with a severe lack of methodological and measurement development studies. Although a variety of measures and paradigms have been proposed and differences between high and low depressive individuals have been shown on some of these tasks (e.g. Rude et al., 2001), none of them has been fully psychometrically evaluated, especially regarding reliability aspects and factorial structure. Another weakness of earlier research concerns the lack of disorder-content relatedness of the assessment instruments. Some approaches have been used to measure interpretation bias in other disorder groups as well (e.g. Dagleish et al., 2003; Eley et al., 2008), which makes it difficult to attribute the detected bias to one disorder or another. It is therefore important to consider instruments that are based on theory-driven factors (e.g. Beck’s cognitive triad in depression) as well as on core characteristics (e.g. self-devaluation in depression) of the disorder (Beck et al., 1979). Further, the established clinical and research interest in the modifiability of negative interpretation bias calls for interpretation bias measures with high psychometric qualities justifying repeated administration. Instruments that are either very change-sensitive or provide equivalent parallel versions do not exist yet and need to be developed for this purpose.

The study presented in Paper 1 (Chapter II) aimed at filling these methodological gaps by developing and evaluating a depression-related parallel ambiguous scenarios test.
I. INTRODUCTION

Inconsistent findings of biased interpretation depending on the assessment method

As outlined above several measures, such as homophone tasks, ambiguous scenarios tests, scrambled sentences, or semantic priming paradigms have been used to measure depression-related interpretation bias. However, empirical studies investigating biased interpretation in the context of depressed mood have yielded mixed results for the existence of these distortions, depending on the bias assessment method used. Whereas studies using homophone tasks, ambiguous scenarios, and scrambled sentences found evidence for depression-related interpretive biases, studies that used different semantic priming paradigms failed to find support. These inconsistencies clearly demonstrate the need for further methodological studies. Although the growing interest in the study of depression-related interpretation bias has led to the development and implementation of several assessment methods, no consensus has been reached on whether interpretation bias in depression is limited to a rather strategic level of information-processing or whether it generalises into a more automatic bottom-up processing mode. The latter could be rather captured through indirect assessments, such as sequential priming methods. Only few studies have tested the validity of priming strategies as measures for depression-related interpretation bias. These studies have solely focused on semantic priming methods and findings have been mixed. Further, no study, to my knowledge, has yet examined in the same sample the association between direct (e.g. ambiguous scenarios tests) and indirect measures (e.g. priming paradigms). At present, it is therefore not known if the various proposed assessment tools are interrelated or rather comprise different aspects.

With these considerations in mind, the aim of study 2 (presented in Paper 2, Chapter III) is to evaluate an affective priming task using homophones for the indirect assessment of depression-related interpretation bias. As a further goal, the associations between four different assessment paradigms are examined: an ambiguous scenarios task, a scrambled sentences test, a homophone task, and the newly evaluated priming paradigm.

Work mechanisms and optimal ingredients of CBM-I

Cognitive bias modification procedures targeting interpretation have initially been developed for the training of anxiety-related bias. However, promising effects of interpretation modification training could be reported in the context of dysphoria and depression by a number of studies (Holmes, Coughtrey, & Conner, 2008; Holmes, Lang, & Shah, 2009; Holmes et al., 2006; Lothmann et al., 2011; Reinecke et al., submitted). Although research on CBM-I in the context of depression/dysphoria is a fast growing field, this branch is still relatively young.
I. INTRODUCTION

More knowledge about the work mechanism and optimal ingredients is needed before implementing CBM-I procedures into therapeutic treatments. In this context, prior research found imagery, rather than verbal processing to be crucial (Holmes, Lang, & Shah, 2009). Verbal processing of positive training material was not only less effective than imagery, but even led to paradoxical, negative mood responses. However, the mental visualization of pleasant events alone does not guarantee positive emotion but depends on the precise task instruction. Holmes, Coughtrey, and Connor (2008) demonstrated that only imagery from a field perspective (“through your own eyes”) improves affect as opposed to imagery from an observer perspective (“looking at you”), that could - like verbal processing - lead to adverse effects. Clarke et al. (2014) further determined that the presence of ambiguity in the imagery training scenarios is a crucial feature and that the mere positive valence of the trainings scenarios alone does not lead to a positive bias modification.

In line with these studies, the aim of study 3 (presented in Paper 3, Chapter IV) is to further increase the knowledge about the underlying work mechanisms and optimal ingredients, by examining the role of self-generated imagery in imagery-based CBM-I.

The objectives of this work

Taken together, the objectives of this dissertation are to fill the gap regarding measure development studies, to shed light on the inconsistent findings for the existence of depression-related interpretation bias depending on the assessment procedure, as well as to further the knowledge about the mechanisms of action underlying successful CBM-I and to optimize its ingredients. A secondary goal is to develop and adapt assessment as well as CBM-I training materials for the German language context. The precise aims of the dissertation are:

Paper 1

- to develop an easy to use interpretation bias tool for repeated measurement
- to describe the development of the Ambiguous Scenarios Test relevant to Depressed Mood II (AST-D-II) and its two parallel short versions A and B
- to test the factor structure and examine its fit
- to describe the AST-D-II scales and demonstrate its internal consistencies
- to investigate the convergent validity of the scales and its subscales
- to examine the parallelity of version A and version B
I. INTRODUCTION

Paper 2

- to increase the knowledge of the association between direct and indirect assessment strategies for depression-related interpretation bias
- to evaluate an affective priming task using homophones for the measure of depression-related interpretation bias, including convergent validity and internal consistency
- to examine the interrelations of four different paradigms for the measure of depression-related interpretation bias, namely an ambiguous scenarios task, a scrambled sentences test, a homophone task, as well as a newly evaluated priming paradigm
- to examine the associations between the four instruments and depressive symptoms

Paper 3

- to increase the knowledge about the underlying work mechanisms and best ingredients of CBM-I
- to develop a new, more active variant of CBM-I and test its impact on positive mood and interpretation bias in comparison to a control group as well as standard imagery CBM-I
II. Paper 1: Measuring change in depression-related interpretation bias: Development and validation of a parallel ambiguous scenarios test

Abstract

Background: Depressed mood is associated with making negatively biased interpretations of ambiguous everyday events. Experimental modification towards a more optimistic interpretation has become a focus of recent research. However, to date no measures exist that have been tested with respect to their psychometric properties that justify repeated administration to capture change. Aims: We aimed to develop and evaluate a pragmatic assessment instrument, consisting of a 30-item questionnaire (long version) and two 15-item parallel short versions (A and B). Methods: Items were generated as ambiguous sentences, reflecting three relevant content areas based on Beck’s cognitive triad. The sentences were rated for pleasantness, and this emotional appraisal task indicates the emotional valence of the interpretation. Due to the intention to develop a parallel test version, item-twins were generated. Results: All three versions of the instrument were found to be structurally stable, internally consistent and valid. In line with Beck’s cognitive triad in depression, confirmatory factor analyses determined a three factor-solution (self, experiences, future). Significant correlations were found between all scales and depressive mood. The two short versions represent the same underlying constructs, share identical psychometric properties and possess high parallel-test reliability. Discussion: This study is the first to evaluate and confirm the factorial validity as well as the parallel-test reliability of an interpretation bias measure. It is suitable to measure bias modification and has therefore great potential for research and clinical practice.

II. PAPER 1

Introduction

Depressed mood is associated with a bias in interpretation - a tendency to interpret ambiguous stimuli in a relatively negative manner (Blackwell & Holmes, 2010; Butler & Mathews, 1983; Eley et al., 2008; Hertel & El-Messidi, 2006; Mathews & Mackintosh, 2000; Mogg et al., 2006; Rude et al., 2002). For example, when presented with ambiguous textual scenarios (e.g. “You wake with a start in the middle of the night thinking you heard a noise, but all is quiet. What do you suppose woke you up?”), individuals with depressed mood are more likely to select a negative (e.g. “It could be a burglar”) interpretation (Butler & Mathews, 1983). Recent research even suggests that interpretation bias could be a crucial factor in the development of depressive mood (Reinecke et al., submitted), and that the experimental modification of these negative biases towards a more optimistic interpretation tendency positively influences other markers of depression (Blackwell & Holmes, 2010; Holmes, Lang, & Shah, 2009; Holmes et al., 2006; Lang et al., 2012). This research is in line with Beck’s cognitive theory of depression (Beck, 1967; Beck, 1984; Beck et al., 1979), assuming that negative thinking styles provoke and maintain depressive feelings. These thinking styles, supposedly the result of dysfunctional beliefs, concern negative thoughts about the self, the world and the future - a concept known as Beck’s cognitive triad (e.g. Beck, 1997).

A cognitive bias modification paradigm that targets interpretation bias (CBM-I; Mathews & Macintosh, 2000; Mathews & MacLeod, 2005) has been developed as a tool to experimentally investigate the aetiological influence as well as the modifiability of biased interpretative reasoning in mental disorders. CBM-I can be described as a computerized training procedure that encourages individuals to develop a more positive interpretation style by repeatedly presenting virtual scenarios that combine initially ambiguous information with a clearly positive outcome. The modification of these negative biases towards a more optimistic interpretation style can offer exciting new possibilities for cognitive therapy. However, measures with good psychometric properties are required in order to evaluate and further advance these promising procedures. Surprisingly, the growing interest in the modifiability of interpretation bias is contrasted by a severe lack of methodological and measurement development studies.

Cognitive interpretation habits are difficult to measure because they are assumed to be largely automatic and beyond the awareness of the individual. Studies so far have therefore tried to directly assess the interpretation tendencies of participants by providing them with examples of ambiguous stimuli. For instance, individuals were presented with a list of
homophones (words that have the same sound, but different meanings, for instance “die/dye”) and were asked to make up sentences or to report instant mental pictures (e.g. Eley et al., 2008; Hertel & El-Messidi, 2006). If participants chose the word with the aversive meaning, negative interpretation bias was assumed. Alternatively, a Scrambled Sentences Task (SST; Wenzlaff, 1993) and pictures of ambiguous stimuli (Bucks, Garner, Tarrant, Bradley, & Mogg, 2008; Holmes, Mathews, Mackintosh, & Dalgleish, 2008) were used to assess interpretation. Although differences between non-depressed and (formerly) depressed individuals have been shown on some of these tasks, for instance the SST (Rude et al., 2001), none of them has been fully psychometrically evaluated, especially regarding reliability aspects and factorial structure.

Another widely used approach is the presentation of ambiguous scenarios (e.g., “You are hosting a dinner party to introduce your friends. At the table you notice how your friends are interacting.”). Participants are either asked to continue the scenario, to choose or rank different interpretations in terms of likelihood, or to rate the emotional valence of the item (e.g. Eley et al., 2008; Holmes & Mathews, 2005; Holmes et al., 2006; Reinecke et al., submitted; Voncken, Bögels, & De Vries, 2004). In line with this approach a recognition task (Mathews & Macintosh, 2000) has, for example, often been used as a measure for interpretation bias in the context of anxiety research. In this task, participants read several ambiguous social scenarios. Afterwards, participants are provided with four different interpretations to each scenario and are asked to rate the interpretations in terms of similarity to the original scenario. This task was shown to differentiate between high and low levels of neuroticism and to be insensitive to mood state (Salemink & Van Den Hout, 2010). The Ambiguous Scenarios Test relevant to Depressed Mood (AST-D) was developed and validated (Berna et al., 2011) for the measure of depressive interpretation. In this 24-item measure, participants were asked to imagine the scenarios and rate their pleasantness (Holmes, Lang, & Deeprose, 2009). It was found that higher unpleasantness ratings were significantly associated with dysphoric mood, and the instrument was found to be internally consistent. However, the established clinical interest as well as the growing research interest in the modifiability of negative interpretation bias call for interpretation bias measures with high psychometric qualities justifying repeated administration. Instruments that are either very change-sensitive or provide equivalent parallel versions have to be developed for this purpose.

Another weakness of earlier research concerns the lack of disorder-content relatedness of these assessment instruments. Some approaches have been used to measure interpretation bias in other disorder groups as well (e.g. Dagleish et al., 2003; Eley et al., 2008), which makes it difficult to attribute the detected bias to one disorder or another. It is therefore important to
consider instruments that are based on theory driven factors (e.g. Beck’s cognitive triad in depression) as well as core characteristics (e.g. self-devaluation in depression) of the disorder (Beck et al., 1979).

This study aimed at filling these methodological gaps by developing and evaluating a parallel and depression-related ambiguous scenario test based on the AST-D (Berna et al., 2011). The aims of this paper are: (1) to describe the development of the Ambiguous Scenarios Test relevant to Depressed Mood II (AST-D-II) and its two parallel short versions A and B, (2) to test the factor structure and examine its fit, (3) to describe the AST-D-II scales and demonstrate its internal consistencies, (4) to investigate the convergent validity of the scales and its subscales, and (5) to establish the parallelity of version A and version B.

**Method**

**Materials and procedure**

Thirty-eight items partly based on the AST-D (Berna et al., 2011) as well as the contents of interpretation modification materials as described in Holmes et al. (2006) were generated, reflecting three relevant content areas based on Beck’s cognitive triad (Beck et al., 1979): the tendency to make negative interpretations of ambiguous future situations, the tendency to make negative interpretations of ambiguous situations concerning past experiences and the tendency to make negative interpretations of ambiguous situations concerning one’s own skills and performances regarding social and non-social situations (see appendix for final version). For instance, statements such as “You buy a new outfit for a party. Other people’s reactions show how you look.” are ambiguous in that their outcomes could be either positive or negative. Due to the intention to develop a parallel test version, we tried to generate “item-matches” that appear equivalent in their content - the first step of a method described as “item-cloning” (Clause, Mullins, Nee, Pukalos, & Schmitt, 1998). For instance, item 19 (“As you walk into the interview room the panel of interviewers welcomes you and proceeds to ask some tough questions. By the end of the interview you know what the outcome is”) and item 22 (“The probation period at your new job is almost over. You get invited to your boss and receive feedback on how you’ve done so far.”) were generated as “item-twins”. Participants were then presented with these 38 items (e.g. “It is an overcast day and you are sitting on the beach. You look up to notice the weather really beginning to change.”) and instructed to imagine each situation as vividly as possible and then rate the emotional valence of each paragraph, using a 11-point Likert scale ranging from -5 (extremely unpleasant) to +5 (extremely pleasant).
To determine current depressive mood and behaviour, participants were additionally given the German version of the Centre for Epidemiologic Studies Depression Scale (CES-D; Hautzinger & Bailer, 1993; Radloff, 1977). The CES-D is a well validated 20-item questionnaire measure for depressive tendencies in clinical and non-clinical samples and was chosen because of its high capacity to detect differences in depressive severity especially in student samples (Santor, Zuroff, Ramsay, Cervantes, & Palacios, 1995). Scores range from 0 to 60, with higher scores indicating more symptoms of depression. A score of 16 has been proposed as the cut-off point for significant depressive symptoms (Radloff, 1977).

**Participants**

Participants were 176 undergraduate psychology students at the Technische Universität Dresden (Germany). Forty male students and 136 female students anonymously completed the questionnaires during psychology lectures after receiving information about the questionnaire validation study. Completion of the two questionnaires took approximately 10-15 minutes. Participation was voluntary and was not compensated. Participants’ ages ranged from 18 to 52 years, with a mean age of 24 years ($SD = 5.1$). Their CES-D scores ranged between 0 and 48 with a mean score of 14.4 ($SD = 9.1$). Scores of 16 or greater were reported by 36.9% of the student sample, suggesting mood irregularities. No difference was found between the CES-D mean scores of male and female students, $t(174) = 0.08, p = .94$.

**Statistical analyses**

Because the number of factors and the factor assignments were determined a priori according to Beck’s theoretical assumptions of a negative cognitive triad, we conducted confirmatory factor analyses (CFA) to determine the fit of the postulated factor solution as well as factor correlations and item-to-factor loadings (Moosbrugger & Schermelleh-Engel, 2007). We further used confirmatory analysis techniques to examine the parallelity of AST-D-II (A) and AST-D-II (B). For this purpose we assumed the two questionnaires to be parts of one underlying variable, a preferred method to test parallelity suggested by Schermelleh-Engel and Werner (2007). If adequate fit of this model is indicated, the standardized parameter loadings can be regarded as estimates of the parallel-test reliability. CFA was carried out using the maximum likelihood estimation. Missing data (< 1%) were estimated using the Expectation-Maximization (EM) algorithm. As model fit indices we used the root-mean-square-error (RMSEA), the minimum discrepancy divided by its degrees of freedom (CMIN/DF) as well as the parsimonious normed fit index (PNFI); adequate fit is considered to be achieved if the RMSEA is .06 or below (Hu & Bentler, 1999) and the CMIN/DF is 2.00 or below (Byrne, 1989). The PNFI ranges from 0 to 1 with higher values indicating a better fit. Although no cut-
off threshold has been recommended for the PNFI, levels around 0.50 are considered to be typical (Byrne, 1998). To describe the AST-D-II scales and its subscales, we calculated means, standard deviations, consistency coefficients (Cronbach’s alpha) and item-test-correlations. To test the distributional parameters of all administered assessment instruments, we used the Kolmogorov-Smirnov-Test. Because no normal distribution was found for the German CES-D scale ($p = .03$; skewness = 1.07, $SE = .18$; kurtosis = 1.50, $SE = .36$), Spearman rank correlations were carried out to assess the association between the AST-D-II scales and the CES-D. To examine the intercorrelations of the AST-D-II scales, we used Pearson correlations.

**Results**

**Item selection**

A subgroup of the 38 generated items was considered to be weak and thus eliminated using two common selection criteria (e.g. Kelava & Moosbrugger, 2007). First, items were rejected if the range of item scores across participants indicated either a ceiling or floor effect. For example, the average of responses on an 11-point scale (-5 = extremely unpleasant to +5 = extremely pleasant) to the item “You’ve taken an exam some weeks ago and are worried you did not do very well. Your heart is beating when you are approaching the notice board where the grades are displayed” was very low ($M = -2.7$, $SD = 1.9$). This item and other similarly performing items were deleted. Furthermore, items with a low item-total score-correlation were removed. According to these criteria a total of 8 items was considered to be too weak and therefore excluded, resulting in a final selection of 30 items. Seventeen items out of the 30 final items were identical with those of the AST-D (Berna et al., 2011).

**Factor structure**

Results of CFA indicated adequate fit (RMSEA = 0.06; CMIN/DF = 1.56; PNFI = 0.62) for the a priori postulated three factorial structure. Figure II.1 presents the model for the 30-item questionnaire with standardised parameter loadings for item-to-factor relationships and factor correlations. The first factor consists of 12 items (average loading = .50, $p < .05$), assessing the interpretation of the self. The second factor loaded above .65 ($p < .05$) on 8 items measuring negative interpretative bias toward one’s experiences in the past, while the third component consisting of 10 items (average loading = .45, $p < .05$) describes the interpretation tendency towards future situations.
II. PAPER 1

Figure II.1. Factor intercorrelations and standardized loadings of the AST-D-II (N = 176)
Reliability and scale descriptions

Means, standard deviations, internal consistency coefficients and median item-test correlations of the total AST-D-II scale and the three subscales are reported in Table II.1. Female subjects reported significantly higher AST-D-II (LV) “total” scores \((t_{\text{fem-mal}} (174) = 2.25, p = .03, d = 0.34)\) as well as AST-D-II (LV) “experiences” scores \((t_{\text{fem-mal}} (54) = 2.85, p = .01, d = 0.76)\) than male subjects. A normal distribution was found for all AST-D-II scales, except for the subscale “future”. Indices of internal consistency were shown to be good for the AST-D-II \(\alpha = .87\) as well as for its subscale “experiences” \(\alpha = .86\) and acceptable for the subscales “future” \(\alpha = .79\), and “self” \(\alpha = .73\).

Validity

In order to examine the convergent validity, we correlated the AST-D-II and its two parallel forms with the CES-D. As expected, moderate but significant negative correlations could be found between the CES-D and the AST-D-II \(r_s = -.51, p < .001\) as well as between the CES-D and the two parallel forms A \(r_s = -.47, p < .001\) and B \(r_s = -.48, p < .001\). High CES-D scores, indicating high levels of depression are connected with low AST-D-II scores, indicating more negative interpretation styles towards ambiguous situations.

Development of the parallel short versions AST-D-II (A) and AST-D-II (B) and scale descriptions

In order to create two parallel short versions of the AST-D-II we calculated means, standard deviations and item-total-correlations for each of the thirty items and divided them - according to these parameters and within their dimension - into two equal parts as described in Clause et al. (1998). Confirmatory factor analyses indicated adequate fit for the expected three factorial (“self”, “experiences”, “future”) structure for short version A (RMSEA = 0.05; CMIN/DF = 1.36; PNFI = 0.65), as well as for its parallel version B (RMSEA < 0.01; CMIN/DF = 0.95; PNFI = 0.71), see Figure II.2. Regarding version A, the factor “self” loaded above .48 \((p < .05)\) on 6 items measuring negative interpretative bias toward ambiguous past situations. The factor “experiences” consisted of 4 items (average loading = .65, \(p < .05\)), while the third component “future” included 5 items (average loading = .43, \(p < .05\)). A similar pattern could be found for version B with the three factors “self” (6 items), “experiences” (4 items), and “future” (5 items). Figure 2 presents the model for each of the two short versions with standardised parameter loadings for item-to-factor relationships and factor correlations. Table II.1 reports scale descriptions for the two new versions. On both versions females reported significantly higher
II. PAPER 1

**Table II.1.** AST-D-II: Means (M), standard deviations (SD), item-test-correlations (IT) and internal consistencies (α) (N = 176)

<table>
<thead>
<tr>
<th>Scale</th>
<th>M</th>
<th>SD</th>
<th>IT (Median)</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Long Version</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AST-D-II „total“</td>
<td>0.97</td>
<td>1.07</td>
<td>.43</td>
<td>.87</td>
</tr>
<tr>
<td>♂</td>
<td>0.64</td>
<td>1.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>♀</td>
<td>1.07</td>
<td>1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AST-D-II „future“</td>
<td>0.61</td>
<td>1.22</td>
<td>.39</td>
<td>.72</td>
</tr>
<tr>
<td>AST-D-II „self“</td>
<td>0.73</td>
<td>1.26</td>
<td>.47</td>
<td>.80</td>
</tr>
<tr>
<td>AST-D-II „experiences“</td>
<td>1.79</td>
<td>1.69</td>
<td>.60</td>
<td>.85</td>
</tr>
<tr>
<td>♂</td>
<td>1.04</td>
<td>1.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>♀</td>
<td>2.01</td>
<td>1.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Version A</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AST-D-II (A) „total“</td>
<td>0.97</td>
<td>1.12</td>
<td>.39</td>
<td>.77</td>
</tr>
<tr>
<td>♂</td>
<td>0.65</td>
<td>1.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>♀</td>
<td>1.07</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AST-D-II (A) „future“</td>
<td>0.54</td>
<td>1.40</td>
<td>.31</td>
<td>.55</td>
</tr>
<tr>
<td>AST-D-II (A) „self“</td>
<td>0.69</td>
<td>1.28</td>
<td>.38</td>
<td>.62</td>
</tr>
<tr>
<td>AST-D-II (A) „experiences“</td>
<td>1.94</td>
<td>1.79</td>
<td>.50</td>
<td>.74</td>
</tr>
<tr>
<td>♂</td>
<td>1.28</td>
<td>1.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>♀</td>
<td>2.13</td>
<td>1.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Version B</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AST-D-II (B) „total“</td>
<td>0.97</td>
<td>1.15</td>
<td>.38</td>
<td>.78</td>
</tr>
<tr>
<td>♂</td>
<td>0.63</td>
<td>1.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>♀</td>
<td>1.07</td>
<td>1.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AST-D-II (B) „future“</td>
<td>0.68</td>
<td>1.35</td>
<td>.30</td>
<td>.56</td>
</tr>
<tr>
<td>AST-D-II (B) „self“</td>
<td>0.77</td>
<td>1.46</td>
<td>.43</td>
<td>.71</td>
</tr>
<tr>
<td>AST-D-II (B) „experiences“</td>
<td>1.65</td>
<td>1.83</td>
<td>.55</td>
<td>.73</td>
</tr>
<tr>
<td>♂</td>
<td>0.81</td>
<td>2.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>♀</td>
<td>1.89</td>
<td>1.64</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
II. PAPER 1

Figure II.2. Factor intercorrelations and standardized loadings of Form A and B (N = 176)

Form A

- 01 „oral examination“
- 03 „job profile“
- 08 „choosing the right present“
- 10 „going on a blind date“
- 20 „publishing a book“
- 29 „organising a party“

Form B

- 06 „preparing dinner for friends“
- 11 „giving a wedding speech“
- 12 „presenting a work project“
- 23 „entering a choir“
- 24 „language test“
- 30 „reactions to party outfit“

Correlations:

- „self“
- „experiences“
- „future“
scores on the “total”-scale (version A: $t_{\text{fem-mal}} (174) = 2.12, p = .04, d = 0.32$; version B: $t_{\text{fem-mal}} (174) = 2.13, p = .04, d = 0.32$) as well as on the subscale “experiences” (version A: $t_{\text{fem-mal}} (174) = 2.71, p = .01, d = 0.41$; version B: $t_{\text{fem-mal}} (53) = 2.90, p = .01, d = 0.78$). Equally high correlation coefficients between version A and the long version ($r = .95, p < .01$) as well as version B and the long version ($r = .95, p < .01$) underline the parallelity of the two short forms. In addition, the two parallel-versions are highly intercorrelated ($r = .79, p < .01$). No significant difference ($t_{A-B} (175) = 0.54, p = .96, d = 0.08$) could be found between the mean of form A ($M = 0.97, SD = 1.12$) and the mean of form B ($M = 0.97, SD = 1.15$), a further criterion for parallelity. The parallelity of the two questionnaires was additionally supported by results of confirmatory factor analysis. By modelling the two parallel versions as parts of one underlying variable we could prove adequate fit (RMSEA < 0.01; CMIN/DF = 0.24; PNFI = 0.99) as well as a good estimation of the parallel-test reliability ($R^2 = .79$).

**Discussion**

The present study aimed to develop and evaluate a parallel ambiguous scenarios test for the pragmatic assessment of interpretation bias in the context of depressed mood. This goal could be accomplished by providing the AST-D-II, comprising a 30-item long form as well as two 15-item short forms that can be either used as independent short versions or administered alternately to measure change. The parallelity of the two short versions is indicated by the equal means and standard deviations as well as by the equivalent factor structures of form A and B, which is in line with Cronbach’s (1947) criteria for parallel test forms. In addition, the parallelity of the two tests was supported by results from confirmatory analysis techniques. The parallel-form version could be suitable for application in cognitive-clinical research as well as in a therapeutic context, for instance when used to measure improvement in thinking styles after cognitive-behavioural treatment. The fit of the a priori postulated three-factor model, based on Beck’s cognitive triad model, could be supported for all three questionnaire versions. This finding could be highly useful for examining (possible) differential effects of cognitive intervention strategies on these three factors and can further be regarded as an indicator for the disorder-relatedness of our instrument.

Due to its questionnaire form, the AST-D-II can - as opposed to other, more laborious paradigms – be efficiently used in large scale studies in order to examine, for instance, the aetiological influence of interpretation in mood disorders. Because the AST-D-II has a bidirectional scale, it could also be used to assess the potential protective qualities of positive
interpretation bias. After the conduction of a thorough item analysis and selection, indices of internal consistency were found to be good for the total scale of the AST-D-II ($\alpha = .87$) and acceptable for the two short versions A ($\alpha = .77$) and B ($\alpha = .78$). Although some of the subscales only consist of few items, the internal reliability for the subscales can still be regarded as sufficient, with few items exceptions.

Convergent validity was supported by the significant negative correlations between the scales of the AST-D-II and the CES-D. This is consistent with Beck’s cognitive theory of depression and the findings of the above-mentioned studies (e.g. Butler & Mathews, 1983; Hertel & El-Messidi, 2006; Holmes, Lang, & Shah, 2009) that negative interpretation is related to depressed mood. The correlations are further comparable to those that were found between the AST-D and the BDI-II (Berna et al., 2011).

Interestingly, females in this study reported higher scores on the “total”- scale as well as on the subscale “experiences”, which is not in line with findings of the AST-D (Berna et al., 2011). Further research in other samples is needed to examine if this gender difference is systematic or due to a particularity of this sample.

A few limitations of the study need to be highlighted. The ambiguity present in everyday situations is often not recognized as such. By presenting descriptions of ambiguous situations and asking participants to solely rate their emotional valence on a Likert-scale (instead of asking them to choose between ranges of pre-existing options for interpretation), we try to keep participants unaware of their explicit interpretation tendency and therefore assume that the items are less susceptible to self-presentation bias than for instance mood questionnaires. However, one can never exclude these or other effects, for instance depression specific response bias (Bisson & Sears, 2007), in self report questionnaires. Future research should therefore compare the AST-D-II with measures that use different operationalizations for depression-specific interpretation bias. These cross-validations should include truly involuntary measures, as for instance event-related brain potentials (ERP) or eye blink reflex, which have been described as psycho-physiological correlates of interpretation bias in the context of social anxiety (Moser, Hajcak, Simons, Huppert, & Foa, 2008) and depression (Lawson et al., 2002). An interesting aim for follow-up research would be further a comparison of the AST-D-II with others aspects of bias (e.g. memory or attention) in depression (Koster et al., 2010; Lang, Moulds, & Holmes, 2009) and other populations (Lothmann et al., 2011). It was further not possible to examine divergent validity, since no constructs that are known to be theoretically unrelated to interpretation bias, have been assessed in this study.
Another restriction is related to the stability of the instrument as well as the homogeneity of our sample. Interpretive tendencies were examined only once in a sample of mainly young adults. Thus, stability could not be examined and further research in more representative samples is needed to evaluate if the item contents are as equally valid for older individuals. More than one third of the participants in this study showed CES-D scores above the clinical cut-off score. However, studies with samples of patients formally diagnosed with depressive disorders are needed in order to determine whether the findings reported here generalize to clinical populations. In this context it would additionally be worthwhile to explore whether a negative interpretation bias, measured with the AST-D-II, changes during treatment for depression, as has been demonstrated for other bias parameters during antidepressant treatment (Harmer et al., 2009). Although the structure of the questionnaire and the item contents are strictly in line with CBT depression theory, future research is needed to empirically determine its specificity by examining the measure in patients with other disorders. Altogether, despite these limitations, the current investigation revealed that the long version of the AST-D-II and its two parallel forms can be considered as structurally stable, internally consistent and valid instruments. This study is the first to evaluate and confirm the factorial validity as well as the parallel-test reliability, and therefore suitability to measure modification of interpretation bias measure in the context of depressive mood.
III. Paper 2: How to measure depression-related interpretation bias? Affective priming in comparison with an ambiguous scenarios test, a scrambled sentences task and a homophone task

Abstract

Background: Negatively biased interpretation has been hypothesized as an important factor in the aetiology and maintenance of depression. Empirical findings, however, have yielded inconsistent results for the existence of biased interpretation in depression, depending on the interpretation bias assessment method. Methods: The aim of this study was to evaluate an affective priming task using homophones as ambiguous material, to acquire a measure of interpretation bias potentially less prone to response bias. This task was compared to three different established interpretation bias paradigms. Seventy participants were tested. We measured self-reported depression as well as interpretation bias, using 1) an ambiguous scenarios test, 2) a scrambled sentences test under cognitive load, 3) a homophone paradigm, and 4) an affective priming paradigm using homophones. Results: No empirical support for mood-related interpretation bias was found for the newly developed affective priming task as well as the homophone task, which both showed no association with depressive symptoms. The ambiguous scenarios test and the scrambled sentences test under cognitive load were shown to be internally consistent and valid instruments that were also highly intercorrelated. Limitations: A limitation of this study is related to the homogeneity of the sample, which comprises a non-clinical population of mainly young adults. Conclusions: These findings emphasize the challenge to create valid ambiguous test stimuli for the measure of depression-related interpretation bias as well as the differences between assessment paradigms.

---

2 In preparation für submission.
Introduction

Biased interpretation of emotionally ambiguous information has been hypothesized as an important factor in the aetiology and maintenance of depression (e.g. Beck, 1984; Beck, 1997). An interpretation bias in depression refers to the tendency to consistently interpret ambiguous (e.g. the yawning of one’s dialogue partner) information in a pessimistic (e.g. “they are bored”) rather than a neutral or benign (e.g. “they just slept badly” or “there must be a lack of oxygen in the air”) manner, which might in turn contribute to a worsening of the individuals’ emotional state (Beck, 1967). However, empirical studies investigating biased interpretation in depression have yielded mixed results, depending on the bias assessment method used (e.g. Berna et al., 2011; Bisson & Sears, 2007; Lawson & MacLeod, 1999; Mogg et al., 2005; Phillips, Hine, & Thorsteinsson, 2010; Rude et al., 2001). This leaves open whether the various assessment paradigms are related or measure different constructs.

In homophone tasks, for instance, individuals are presented with a series of homophones, words that have the same sound, but different meaning (e.g. “die/dye”). Subjects are asked to write down the word, to spontaneously form a sentences using this word or to report mental images provoked by the word (e.g. Eley et al., 2008; Hertel & El-Messidi, 2006; Mogg et al., 2005). If participants chose the word with the aversive meaning, negative interpretation bias is assumed. The homophone paradigm was shown to significantly differentiate between individuals with high and low BDI-II scores (Mogg et al., 2005), and to have a strong association with depressive symptoms in a meta-analysis (Phillips et al., 2010). In ambiguous scenarios tasks, open-ended situations are described (e.g. “You are hosting a dinner party to introduce your friends. At the table you notice how your friends are interacting.”), and subjects’ interpretation is assessed, for instance, via emotional valence ratings (e.g. Berna et al., 2011; Holmes et al., 2006; Rohrbacher & Reinecke, 2014). It was found that higher unpleasantness ratings on these measures were significantly associated with depressive symptoms (Berna et al., 2011; Rohrbacher & Reinecke, 2014) and prospectively predicted depressive symptoms over a six-month period (Kleim, Thörn, & Ehlert, 2014). Another paradigm to assess interpretation bias was developed by Wenzlaff (1993) with the Scrambled Sentences Task (SST). In this task, participants are instructed to unscramble a list of scrambled sentences. They are asked to form sentences using five of six given words (e.g. “good feel very I bad usually”), requiring participants to select either a positively (e.g. “I usually feel very good.”) or a negatively (e.g. “I usually feel very bad.”) valenced sentence, which is supposed to be an indicator of their interpretation style. The SST was shown to differentiate
between formerly and non-depressed individuals (Rude et al., 2001), to predict subsequent depressive disorders (Rude et al., 2010), and showed a strong association with depressive symptoms in a meta-analysis (Phillips et al., 2010).

However, other studies using priming methods, and therefore more indirect measures that are less prone to response bias, found no evidence for a depression-related interpretation bias (Bisson & Sears, 2007; Lawson & MacLeod, 1999). Lawson and MacLeod (1999), for instance, failed to find evidence for a dysphoria-related interpretive bias using a semantic priming paradigm. Participants in this study were required to say a target word out loud following a priming sentence. A subset of the prime sentences were ambiguous (e.g. “the doctor examined little Emily’s growth”), allowing negative or neutral interpretations. These ambiguous primes were followed by a target word, which was either related to the negative (tumour) or the neutral (height) interpretation. Semantic priming effects were measured using the pronunciation latency as dependent variable. There was also no support for a pessimistic interpretive bias in a series of studies by Bisson and Sears (2007) using a lexical decision task. In this cross-modal semantic priming paradigm participants listened to ambiguous prime sentences and made lexical decisions to target words. Each of the prime sentences (e.g. “Kathy had been committed for some time”) permitted a positive, negative, or neutral interpretation and was either paired with a positively related target (romance), a negatively related target (hospital), a neutrally related word (years), an unrelated word (sauce), or a nonword target. Participants had to indicate whether the target was a word or a nonword (yes- or no-button). Results did not confirm the expectation that depressed participants would show an increased tendency to focus on the negative interpretation of the primes and therefore show reaction delays in response to negatively related targets. The null results of both these priming studies suggest that depression and dysphoria might be associated with a negative response bias, but no actual interpretative bias. However, a later semantic priming study (Sears et al., 2011) contradicted this view by finding evidence for a negative interpretation bias in error rates but not in response latencies. Participants in this study listened to self-referent ambiguous prime sentences (e.g. “my boyfriend said that I am unlike his past girlfriends”) and were presented with target words that were related to a negative (jealous), positive (attractive), or neutrally-related (relationship) interpretation of the sentence. Unrelated words (democracy) were also presented. Individuals had to quickly indicate whether or not the target was related to the prime (yes or no button). Further support for the existence of interpretation bias measured with implicit methods was also found by a study that applied an ambiguous cue-conditioning
These inconsistencies between more implicit and more explicit measures of interpretation bias clearly demonstrate the need for further methodological studies. Although the growing interest in the study of depression-related interpretation bias has led to the development and implementation of several assessment methods, no consensus has been reached on whether interpretation bias in depression is limited to a rather strategic level of information-processing or whether it generalises into a more automatic bottom-up processing mode. Dual process models of depression suggest that cognitive vulnerability to depression reflects an interplay of automatic (implicit) and reflective (explicit) processes, although they differ as to their emphasis on the relative importance of these processes (e.g. Beevers, 2005; Haeffel et al., 2007). Automatic processes could potentially be rather captured through indirect, implicit assessments, such as sequential priming methods. However, only few studies have tested the validity of priming strategies as measures for depression-related interpretation bias. These studies have exclusively focused on semantic priming methods and findings have been mixed. Further, although there is increasing interest in investigating the relationship between different cognitive processes in depression (e.g. Everaert et al., 2014; Vrijsen, Van Oostrom, Isaac, Becker, & Speckens, 2014), no study, to our knowledge, has yet examined the association between direct (e.g. ambiguous scenarios tests) and indirect measures (e.g. priming paradigms) related to the interpretation of ambiguity in one sample. At present, it is therefore not known if the various proposed assessment tools are interrelated or rather comprise different aspects of interpretation bias. In order to advance the aetiological and clinical research of interpretation bias, more methodological knowledge is needed about the validity and interrelatedness of the various assessment tools.

The present study evaluates an implicit priming task for the measure of interpretation bias and examines its relation to three other, already evaluated assessment procedures. Unlike the above mentioned studies (Bisson & Sears, 2007; Lawson & MacLeod, 1999; Sears et al., 2011) using semantic priming methods, this study examines a task that is based on the evaluative (affective) priming paradigm (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). In affective priming tasks, individuals have to evaluate a target as quickly as possible that is preceded by a prime stimulus. The paradigm is based on the idea that target responses following affectively congruent primes are faster than target responses following affectively incongruent
primes. To our knowledge, the evaluative or response priming paradigm has not been tested yet in the context of depression-related interpretation bias assessment. However, the application of homophone words as discriminating material has been shown to be successful in the direct assessment of interpretation bias (Mogg et al., 2005) and will also be used as ambiguous stimuli in the present evaluative priming task.

In this task, participants have to categorize a target word as either positive or negative as quickly as possible. The target is preceded by either a depression-related prime word (e.g. guilty), a happiness-related prime word (e.g. optimistic) or a homophone prime (e.g. “cancer”, which can be interpreted as crab or illness in the German language). We expect depressive individuals to show an increased tendency to focus on the negative interpretation of the homophone word than non-dysphoric individuals and therefore show longer reaction times when homophone primes are paired with positive targets and shorter reaction times when homophone primes are paired with negative targets. As a second goal, we examine the association between the affective priming task and three other, already evaluated interpretation bias measures that rely on different assessment paradigms. These measures comprise a scrambled sentences test, a homophone task, and an ambiguous scenarios test.

**Method**

**Participants**

A total of 70 participants (20 male/50 female) took part in the study, with a mean age of 27.0 years (SD = 5.7). They were recruited via online advertisements on the website of the Technische Universität Dresden. The majority of them were undergraduate students. One third of the participants (31.4%) in this study showed BDI-II-scores above the clinical cut-off score with 18.6% reporting symptoms of mild depression (BDI-II-scores: 14-19), 7.1% reporting symptoms of moderate depression (BDI-II-scores: 20-28) and 5.7% reporting symptoms of severe depression (BDI-II-scores: 29-63). All individuals received a description of the study aims and an informed consent form. Reimbursement for participation consisted of either a small participation fee (5€) or course credits.

**Procedure and materials**

All participants completed the following assessment instruments in one session (completion took 40-50 minutes).
Self-reported depression. The German version of the Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996; Hautzinger, Kühner, & Keller, 2006) was used to measure depressive symptoms. Participants have to respond to 21 depression-related questions concerning their feelings, thoughts and behaviours during the past two weeks. The BDI-II possesses high internal consistency as well as high test-retest reliability (Beck et al., 1996). Participants were assigned to groups based on their BDI-II-scores. Participants in the dysphoric group had scores above the clinical cut-off, that is greater than or equal to 14 (n = 22; 73% female; mean age = 28.6 years; BDI-II-scores: 14-49). Participants in the non-dysphoric group had BDI-II-scores less than 14 (n = 48; mean age = 26.8 years; 71% female; BDI-II-scores: 0-12).

Interpretation bias. Interpretation style of ambiguous material was measured via four different assessment instruments:

(1) Homophone Affective Priming Task (HAP). In this task, volunteers have to categorize a target word as either positive or negative as quickly as possible. The target is preceded by a supraliminally presented positive or negative prime word. Target stimuli were 8 positive (e.g., blossom, rainbow, sunrise) and 8 negative (e.g., assault, disaster, infection) words. Positive and negative target words were matched for word length. Prime stimuli were 8 depression-related (e.g., guilty, disappointing, worthless), 8 happiness-related (e.g., admirable, optimistic, precious) and 8 homophone words (those that were also homographs, as they were read as opposed to listened to, e.g. “cancer”, which can be interpreted as crab or illness in the German language). Words in the three prime groups were matched for word length. All stimuli were presented in uppercase letters at a font size of 24.

Each experimental trial started with the presentation of a fixation cross for 500 ms. Afterwards, a prime word was presented for 200 ms, immediately followed by a target word for up to 5000 ms. Participants were instructed to categorize the target word as either positive or negative as quickly and as accurately as possible, and to merely watch the prime word for a later test. Half of the participants used the left key for a “positive”-response and the right key for a “negative”-response, while the other half received reversed instructions. After receiving written instructions and performing four neutral practice trials, volunteers were presented 256 experimental trials, divided into 12 blocks of 32 each, separated by resting periods. In half of the trials the target required a positive response, in half of the trials a negative answer. Both target types were preceded by a depression-related, happiness-related or homophone prime in half of the cases each. Thus, there were two combinations of compatible trials (depression prime
III. PAPER 2

+ negative target, happiness prime – positive target), two combinations of incompatible trials (depression prime + positive target, happiness prime + negative target) and two combinations of either compatible or incompatible trials depending on the interpretation style of the subject (homophone prime + positive target, homophone prime + negative target). The combination and order of prime and target conditions was pseudo-randomly selected by the experimental script.

Priming effect scores were calculated by subtracting the median RT to positive targets from the median RT to negative targets, separated for each prime condition. Thus, a score close to zero indicates low affective priming effects, a score largely different from 0 indicates reaction time differences between the two target types and, thus, a relevance of the affective association between prime and target. For homophone primes, particularly negative scores are supposed to indicate negative interpretation whereas positive scores are supposed to be an indicator of positive interpretation.

(2) The Ambiguous Scenarios Test relevant to Depressed Mood II (AST-D-II; Rohrbacher & Reinecke, 2014) is a self-report measure to assess interpretation bias in the context of depression, consisting of a 30-item questionnaire (long version) and two 15-item parallel short versions (A and B). The items comprise descriptions of ambiguous everyday life situations reflecting three relevant content areas based on Beck’s cognitive triad (Beck, 1997; Beck et al., 1979): the tendency to make negative interpretations of ambiguous future situations, the tendency to make negative interpretations of ambiguous situations concerning past experiences and the tendency to make negative interpretations of ambiguous current situations concerning one’s own skills and performances regarding social and non-social situations. Subjects are instructed to imagine these descriptions (e.g. “You buy a new outfit for a party. Other people’s reactions show how you look.”) as vividly as possible and rate their emotional valence using an 11-point bipolar scale ranging from “extremely unpleasant” to “extremely pleasant”. Lower scores on the AST-D-II indicate a negative interpretation style, whereas higher scores indicate positive interpretation bias. The long version as well as the two parallel versions have been shown to have good psychometric properties, including good parallel-reliability. In line with Beck’s cognitive triad in depression a three factor-solution (self, experiences, future) could be determined using confirmatory factor analysis. In terms of convergent validity, significant correlations were found between depressive symptoms and all scales of the AST-D-II (Rohrbacher & Reinecke, 2014). 
III. PAPER 2

(3) The Scrambled Sentences Task (SST; Wenzlaff, 1993) is a cognitive processing measure that requires participants to create coherent sentences from scrambled phrases (e.g. “good feel very I bad usually”) under a cognitive load. This requires participants to generate either a positively (e.g. “I feel usually good”) or a negatively (e.g. “I feel usually bad”) valenced sentence. A negative interpretation score is generated by calculating the percentage of negative solutions. Thus, higher scores on this measure indicate a negative interpretation tendency. The SST has been shown to differentiate between formerly and non-depressed individuals (Rude et al., 2001) as well as to predict subsequent depressive disorders (Rude et al., 2010; Rude, Valdez, Odom, & Ebrahimi, 2003; Rude et al., 2002). Participants in this study were asked to resolve 20 mixed scrambled sentences under a cognitive load (holding a six digit number in mind) and constrained time (4 minutes). The sentences used in this study were based on the original material and were translated into German and partly modified to fit the language context. Instead of using six words as in the original version we added a seventh “useless” word in order to conceal the ambiguity of the scrambled sentences.

(4) The Homophone Task (e.g. Eley et al., 2008, Hertel & El-Messidi, 2006). Individuals were acoustically (via headphones) presented with a list of homophones (words that have the same sound, but different meanings, e.g. “die/dye”) and were asked to write down instant associations or synonyms as fast as possible (time was restricted to 10 seconds after the presentation of each word). If participants “chose” the word with the aversive or less positive meaning, negative interpretation bias was assumed. Participants were presented with a list of 30 words, including 22 homophones (see Appendix) and 8 neutral filler words. A negative interpretation score was generated by calculating the total number of negative meanings of the homophone words. Thus, higher scores on the Homophone Task are supposed to indicate a negative interpretation style.

Statistical analyses

To examine the internal reliability for the AST-D-II, the SST, as well as the Homophone Task we calculated consistency coefficients (Cronbach’s alpha). To examine the consistency of the HAP, we divided all items for each condition into two sets (first and second half) and examined the correlation of the two total scores (split-half-reliability). To test the distributional parameters of all administered assessment instruments, we used the Kolmogorov-Smirnov-Test. Because no normal distribution was found for homophone priming effects in the HAP, Spearman rank correlations were carried out to assess the association between this scale and all
other scales. To examine all other correlations, we used Pearson correlations. To test group
differences on the HAP we conducted a 3 x 2 repeated measures ANOVA with the within-factor
Prime stimuli (depression-related, happiness-related, homophone) and the between-factor
Group (non-dysphoric group, dysphoric group). The priming effect scores (calculated by
subtracting the median RT to positive targets from the median RT to negative targets, separated
for each prime condition) as well as the error rates (separated for each prime condition) were
used as the dependent variables. Two-tailed independent-samples t-tests were conducted to test
group differences on all other interpretation bias measures.

Results

Scale descriptions

Means and standard deviations for of all measures are depicted in Table III.1. A normal
distribution was found for all assessment instruments (all, \( p > .06 \)), except for the priming effect
scores “homophone primes” (\( p = .02 \); skewness = 3.49, \( SE = .29 \); kurtosis = 25.23, \( SE = .57 \)).
Indices of internal consistency were shown to be excellent for the AST-D-II (\( \alpha = 0.90 \)), poor
for the SST (\( \alpha = 0.57 \)) and unacceptable for the Homophone Task (\( \alpha = 0.24 \)). The split-half-
reliability for the HAP was high for all three prime stimuli (depression-related: \( rs = .90, p < .001 \); happiness-related: \( rs = .88, p < .001 \); homophone: \( rs = .88, p < .001 \)).

Group differences

Homophone Affective Priming Task. A 3 x 2 repeated measures ANOVA with the
within-factor Prime stimuli (depression-related, happiness-related, homophone) and the
between-factor Group (non-dysphoric group, dysphoric group) was calculated using the
priming effect scores (calculated by subtracting the median RT to positive targets from the
median RT to negative targets, separated for each prime condition) as the dependent variables.
In the Homophone Affective Priming Task (HAP), the Prime (depression-related, happiness-
related, homophone) x Group (non-dysphoric group, dysphoric group) ANOVA yielded a
significant main effect of the factor Prime, \( F(2,67) = 6.16, p < .01, \eta^2_p = .16 \), but no significant
effect of Group, \( F(1,68) = .02, p = .89 \), or interaction, \( F(2,67) = 1.61, p = .21 \). This indicates
that the two groups did not differ on their priming effect scores (Table III.1).

A 3 x 2 repeated measures ANOVA with the within-factor Prime stimuli (depression-
related, happiness-related, homophone) and the between-factor Group (non-dysphoric group,
dysphoric group) was also calculated using the error rates (separated for each prime condition)
III. PAPER 2

as the dependent variables. In the Homophone Affective Priming Task (HAP), the Prime (depression-related, happiness-related, homophone) x Group (non-dysphoric group, dysphoric group) ANOVA yielded a significant main effect of the factor Prime, $F(2,67) = 4.52, p < .05, \eta^2_p = .12$, but no significant effect of Group, $F(1,68) = 2.01, p = .16$, or interaction, $F(2,67) = 0.16, p = .85$. This indicates that the two groups did not differ in their error rates.

**Other interpretation bias measures.** Dysphoric participants reported significantly lower values on the AST-D-II, $t(68) = 5.04, p < .01, d = 1.22$, and significantly higher values on the SST, $t(68) = -3.69, p < .01, d = 0.89$, than non-dysphoric participants. This indicates that dysphoric participants showed a higher tendency to negatively interpret ambiguous stimuli than non-dysphoric participants as measured with the AST-D-II and the SST. No difference between groups was found on the Homophone Task, $t(68) = 0.34, p = .74$ (Table III.1).

**Table III.1.** Means ($M$) and Standard Deviations ($SD$) on the BDI-II and all interpretation bias measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>total sample (n = 70)</th>
<th>non-dysphoric (n = 48)</th>
<th>dysphoric (n = 22)</th>
<th>t-statistics ($df = 68$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>27.3 (5.71)</td>
<td>26.8 (5.37)</td>
<td>28.6 (6.34)</td>
<td></td>
</tr>
<tr>
<td>% females</td>
<td>71.0 (0.46)</td>
<td>71.0 (0.46)</td>
<td>73.0 (0.46)</td>
<td></td>
</tr>
<tr>
<td>BDI-II</td>
<td>11.3 (10.62)</td>
<td>5.77 (3.79)</td>
<td>23.36 (10.37)</td>
<td>-10.11**</td>
</tr>
<tr>
<td>AST-D-II</td>
<td>0.69 (1.33)</td>
<td>1.15 (1.02)</td>
<td>-.32 (1.37)</td>
<td>5.04**</td>
</tr>
<tr>
<td>SST</td>
<td>4.11 (2.62)</td>
<td>3.40 (2.46)</td>
<td>5.68 (2.30)</td>
<td>-3.69**</td>
</tr>
<tr>
<td>Hom. Task</td>
<td>11.00 (2.67)</td>
<td>11.06 (2.53)</td>
<td>10.86 (1.58)</td>
<td>.34</td>
</tr>
<tr>
<td>HAP dep</td>
<td>2.53 (50.13)</td>
<td>6.42 (42.24)</td>
<td>-5.95 (64.46)</td>
<td>.96</td>
</tr>
<tr>
<td>HAP hap</td>
<td>19.65 (51.55)</td>
<td>20.30 (44.11)</td>
<td>18.23 (66.14)</td>
<td>.16</td>
</tr>
<tr>
<td>HAP hom</td>
<td>9.75 (68.87)</td>
<td>6.95 (33.55)</td>
<td>15.86 (114.06)</td>
<td>-.50</td>
</tr>
</tbody>
</table>

Note: BDI-II = Beck’s Depression Inventory II, AST-D-II = Ambiguous Scenarios Test relevant to Depressed Mood II, SST = Scrambled Sentences Test, Hom.task = Homophone task, HAP dep= Homophone affective priming task “depression-related primes”, HAP hap = Homophone affective priming task “happiness-related primes”, HAP hom= Homophone affective priming task “homophone primes”, t-statistics provided are for the comparison of the non-dysphoric vs. the dysphoric group. **$p < .001$ (two tailed)
Convergent validity

To examine the convergent validity, the association between self-reported depression (BDI-II) and all interpretation bias measure results was calculated (Table III.2). In order to simplify the interpretation of the results, the AST-D-II scores were inverted (only for the calculation of the convergent validity) so that positive correlations between the AST-D-II, the SST, and the Homophone Task always suggest relatedness. A large correlation could be found between the BDI-II and the AST-D-II, indicating that higher levels of depression are associated with more negative interpretation styles towards ambiguous situations. A medium correlation could be found between the BDI-II and the SST, indicating that higher levels of depression are associated with more negative interpretation styles as measured on the SST. A low (negative) correlation could be detected between the BDI-II and the priming effect scores “depressive primes”, indicating that higher levels of depression are (weakly) connected to stronger priming effects using “depressive primes”.

To examine the relatedness of the four bias measures all instruments were intercorrelated (Table III.2). A large correlation could be found between the AST-D-II and the SST, which indicates the relatedness of these measures. A low correlation could be found between

<table>
<thead>
<tr>
<th>Measure</th>
<th>AST-D-II</th>
<th>SST</th>
<th>Hom.task</th>
<th>HAP dep</th>
<th>HAP hap</th>
<th>HAP hom</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI-II</td>
<td>0.66**</td>
<td>0.47**</td>
<td>0.00</td>
<td>-0.27*</td>
<td>-0.16</td>
<td>0.10</td>
</tr>
<tr>
<td>AST-D-II</td>
<td>0.51**</td>
<td>0.23*</td>
<td>-0.26**</td>
<td>-0.19</td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>SST</td>
<td></td>
<td>0.16</td>
<td>-0.31**</td>
<td>-0.08</td>
<td>-0.20</td>
<td></td>
</tr>
<tr>
<td>Hom.task</td>
<td></td>
<td></td>
<td>-0.13</td>
<td>-0.05</td>
<td>-0.27*</td>
<td></td>
</tr>
<tr>
<td>HAP dep</td>
<td></td>
<td></td>
<td></td>
<td>0.63**</td>
<td>0.58**</td>
<td></td>
</tr>
<tr>
<td>HAP hap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.43**</td>
<td></td>
</tr>
<tr>
<td>HAP hom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: N = 70. BDI-II = Beck’s Depression Inventory II, AST-D-II = Ambiguous Scenarios Test relevant to Depressed Mood II, SST = Scrambled Sentences Test, Hom.task = Homophone task, HAP dep= Homophone affective priming task “depression-related primes”, HAP hap = Homophone affective priming task “happiness-related primes”, HAP hom= Homophone affective priming task “homophone primes”

* $p < .05$ (two tailed)

**$p < .01$ (two tailed)
the AST-D-II and the Homophone Task, indicating that these two instruments are only weakly related. A low (negative) correlation could be further detected between the AST-D-II and the priming effect scores “depressive primes”, as well as between the SST and the priming effect scores “depressive primes”, indicating that higher levels of negative interpretation bias are (weakly) connected to stronger priming effects using “depressive primes”. A low (negative) correlation was further found between the Homophone Task and the priming effect scores “homophone primes”, indicating that these instruments are only weakly related.

Discussion

In view of prior inconsistent findings for the existence of depression-related interpretation bias depending on the measure used, this study attempted to gain a clearer understanding of the methodological relatedness of different direct and indirect assessment paradigms. For this purpose the present study evaluated a new affective priming task for the measure of interpretation bias and examined its association with three other, already established assessment procedures, an ambiguous scenarios test (Berna et al., 2011; Holmes et al., 2006; Rohrbacher & Reinecke, 2014), a scrambled sentences task (Wenzlaff, 1993) and a homophone paradigm (Eley et al., 2008, Hertel & El-Messidi, 2006).

Our findings show that the evaluative priming task using homophone words as prime stimuli did not differentiate between dysphoric and non-dysphoric participants in terms of response latencies and error rates. It was further not related to the ambiguous scenarios test and the scrambled sentences task. Interestingly, there was also no support for interpretation bias using the Homophone Task, as it did not differentiate between dysphoric and non-dysphoric participants and was only weakly related to the AST-D-II. In this paradigm homophone words are also used as ambiguous stimuli, but the task can – as opposed to the Homophone Affective Priming Task – rather be classified as a direct measure. In contrast, the AST-D-II as well as the SST under cognitive load were shown to significantly differentiate between the dysphoric and the non-dysphoric groups, to be highly associated with the BDI-II and to be significantly intercorrelated.

These findings successfully replicate prior research (e.g. Rohrbacher & Reinecke, 2014; Rude et al., 2003) and further support the validity of both instruments. Given their significant interrelation, the AST-D-II as well as the SST seem to measure related aspects. This is particularly notable given the differences between the paradigms. For example, the AST-D-II involves generation of mental imagery, whereas the SST is verbal or semantic. This has been
What do our findings imply and how can the failure to show differences on the evaluative priming task be explained? There are several differences between the tasks that may be considered. The AST-D-II and SST involve processing the ambiguous information in a self-relevant way, whereas the Homophone Task and HAP do not; depression-relevant interpretation bias may only be present for self-relevant information (e.g., Wisco, 2009). The tasks also differ in the relative complexity of the stimuli and amount of elaborative processing that is required, which is particularly relevant in view of dual-process theories of depression (e.g., Beevers, 2005; Haeffel et al., 2007). It is worth noting that some previous studies using priming paradigms (Bisson & Sears, 2007; Lawson & MacLeod, 1999) have not found evidence for a depression-related interpretation bias. Several possibilities have been discussed in order to explain this. Authors have, for instance, questioned the general suitability of priming methods as an adequate dependent variable, as depressed individuals tend to have slower and more variable response times in speed tasks making it difficult to use them as valid indicator for interpretation (e.g., Lawson et al. 2002; Moretti et al., 1996; Sears et al., 2011). Another explanation for the lack of interpretation bias measured via implicit methods (as opposed to explicit assessments) was that interpretation of ambiguity does not happen automatically in depression, but rather at a later strategic processing level and can therefore not be measured with priming methods, as they only measure the immediate reaction and not the later processing (Sears et al., 2011).

However, an important factor to consider is that we also found no relationship between homophone interpretation and depressive symptoms, and hence it is perhaps not surprising that the HAP showed no homograph priming effect. Given that previous studies have found homophone interpretation to be related to depressive symptoms (e.g., Phillips et al., 2010), why was this not the case in our study? The simplest explanation may be differences in language, i.e. that a less depression-relevant set of homophones was available in German than in previous studies in English. Responses on the Homophone Task correlated with the effect of homophone priming on the HAP, suggesting that to the extent that participants interpreted the homophones
negatively, the HAP provided an implicit (via priming) measure of this. Further, scores on the BDI-II, AST-D-II, and SST correlated with priming effects of depression-relevant words on the HAP, indicating validity of the priming task for detecting individual differences related to depression when relevant primes were used (although this score did not differentiate between dysphoric and non-dysphoric groups). Hence, the pattern of results suggest that the HAP could potentially be able to provide an implicit measure of interpretive bias where there exists a suitable ambiguous set of primes, e.g. homophone sets in English that have been successfully used to detect depression-relevant interpretive bias (e.g. Mogg et al., 2006), or for other disorder populations for whom the (German) homophones in this study would provide a suitable ambiguous stimuli set. The results also illustrate the challenge of developing suitable stimuli sets and the constraints that language differences place on the replication of effects across different countries.

A restriction of this study is related to the homogeneity of the sample, which comprises a non-clinical population of mainly young adults. Studies with samples of patients formally diagnosed with depressive disorders and a wider age range are needed to determine whether the findings reported here generalize to clinical populations and older individuals.

Altogether, the current investigation revealed the AST-D-II and the SST under cognitive load as internally consistent and valid instruments. Significant intercorrelations between these instruments can be regarded as an indicator of a shared underlying construct. No empirical support was found for the newly developed affective priming task using homophones as well as the Homophone Task, which showed low psychometric properties and no association with depression. Findings of this study emphasize the difficulty of finding valid ambiguous test stimuli for the measure of depression-related interpretation bias and clearly entail the need for further methodological evaluation studies.
IV. Paper 3: Optimizing the ingredients for imagery-based interpretation bias modification for depressed mood: Is self-generation more effective than imagination alone?  

Abstract  
Negative interpretation is thought to be crucial in the development and maintenance of depression. Recently developed cognitive bias modification paradigms, intending to change these biases towards a more optimistic interpretation tendency (CBM-I), seem to offer new promising implications for cognitive therapy innovation. This study aimed to increase our knowledge of the underlying mechanisms of action of imagery-based CBM-I in the context of depressed mood. We therefore compared the efficacy of CBM-I requiring participants to imagine standardized positive resolutions to a novel, more active training version that required participants to generate the positive interpretations themselves. Fifty-four participants were randomly allocated to (1) standardized CBM-I, (2) self-generation CBM-I or (3) a control group. Outcome measures included self-report mood measures and a depression-related interpretation bias measure. Both positive training variants significantly increased the tendency to interpret fresh ambiguous material in an optimistic manner. However, only the standardized imagery CBM-I paradigm positively influenced mood.

---

Introduction

In day-to-day life, we encounter numerous examples of information that can be interpreted in more than one way, for instance facial expressions, feedback from others or physiological cues. How we disambiguate or make sense of these stimuli is important for how we further respond to the world. To interpret one's dialogue partner's yawning as a sign of boredom can obviously result in a different, even opposite emotion than attributing it to a simple lack of oxygen in the room. There is empirical consensus that negatively biased interpretation, defined as the tendency to consistently interpret ambiguous stimuli in a negative manner, is associated with dysphoria and depression and might play a key role in the development of disorders (e.g. Butler & Mathews, 1983; Eley et al., 2008; Hertel & El-Messidi, 2006; Lawson et al., 2002; Mathews & Mackintosh, 2000; Mogg et al., 2006; Reinecke et al., submitted; Rude et al., 2002, Wisco & Nolen-Hoeksema, 2010).

Cognitive bias modification paradigms (CBM; Grey & Mathews, 2000; Mathews & Mackintosh, 2000; Mathews & MacLeod, 2002) are computerized training procedures that encourage individuals to adopt a more positive (or negative) information processing style. For instance, in interpretation bias training procedures a more positive interpretation style can be trained by repeatedly presenting virtual scenarios that combine initially ambiguous information with a clearly positive outcome. An example by Holmes et al. (2006) is as follows: “You have started an evening class which is tough going. You are determined to succeed, and after a while, it becomes much easier and more enjoyable” (positive resolution in italics). Such procedures aim to identify the role of cognitive biases in the development of emotional disorders and have the potential to induce reductions in symptom severity by changing bias towards a more positive direction.

Although CBM-I has initially been developed for the training of anxiety-related bias, positive benefits of interpretation modification training could also be reported in the context of dysphoria and depression by a number of studies (Blackwell & Holmes, 2010; Holmes, Coughtrey, & Connor, 2008, Holmes, Lang, & Shah, 2009, Homes et al., 2006; Lang et al., 2012; Lothmann et al., 2011; Reinecke et al., submitted; Williams et al., 2014). Positive bias modification has not only been shown to be effective in modifying the interpretation of fresh ambiguous information (e.g. Reinecke et al., submitted) in the trained direction, but it also ameliorates depressed mood and positively influences resilience to negative mood induction (e.g. Holmes, Lang, & Shah, 2009).
A line of research (e.g. Holmes et al., 2009) has been interested in the most effective ingredients of these procedures. In line with prior research suggesting a special relationship between imagery and emotion (see Holmes & Mathews, 2005), Holmes, Lang, and Shah (2009) found imagery, rather than verbal processing to be crucial. Participants in the study were presented with 100 auditory scenarios that were initially ambiguous but consistently resolved in a positive manner. They were instructed to either imagine the events or listen to them while thinking about their meaning. Verbal processing of positive training material was not only less effective than imagery, but even led to paradoxical, negative mood responses. In other words, vividly imagining positive virtual events can improve your mood as well as how you interpret events, whereas only verbally thinking about the same positive contents can make you feel worse and negatively influence how you resolve ambiguities. Holmes et al. (Holmes, Lang, & Shah, 2009; Holmes et al., 2006) assumed that this surprising discrepancy could have been caused by different underlying mechanisms. While imagery might directly provoke emotion like a positive “as-if” experience, verbal processing might be perceived as less believable. It might more readily provoke a comparison with one's actual status quo, that could – especially in depressive individuals – turn out to be disadvantageous and thus mood deteriorating. However, the mental visualization of pleasant events alone does not guarantee positive emotion but depends on the precise task instruction. Holmes, Coughtrey, and Connor (2008) demonstrated that only imagery from a field perspective (“through your own eyes”) improves affect as opposed to imagery from an observer perspective (“looking at you”), that could – like verbal processing – lead to adverse effects. Nelis et al. (2012) replicated the impact of imagery vs. verbal processing for positive CBM-I, but not of field vs. observer perspective, suggesting that this aspect requires further exploration. Based on these findings, effective interpretation bias modification can be obtained through a guided imagery training that provides participants with acoustically or visually presented descriptions of ambiguous virtual situations that are consistently combined with positive resolutions that need to be imagined from a field perspective (Holmes, Coughtrey, & Connor, 2008; Pictet, Coughtrey, Matthews, & Holmes, 2011).

Research in the context of anxiety-related interpretation bias has found active selection of meaning during the training to be critical for modifying subsequent emotional responses to new ambiguous stimuli (Hoppitt et al., 2010). In this study, participants were presented with threat-related ambiguous sentences that were negatively resolved by the final word, e.g. “You have decided to go caving even though you feel nervous about being in such an enclosed space.
You get to the caves before anyone else arrived. Going deep inside the first cave you realize you have completely lost your way.” (negative resolution in italics). While participants in the passive group were presented with the entire passage, individuals in the active group were presented with only a fragment (one or more letters missing) of the final word (“….completely lost your w--.”) and therefore had to actively resolve the meaning by themselves (only one possible completion). Active selection of meaning was shown to be superior in modifying later emotional responses in a training-congruent way to images of new emotionally ambiguous descriptions presented after training than mere passive exposure. Hoppitt et al. (2010) suggested that this differential effect may be due to the induction of an implicit production rule, in which participants in the active condition continue to actively generate training-congruent meanings of subsequent ambiguous scenarios.

The self-perception theory (Bem, 1972) suggests that people infer (and modify) their attitudes, cognitions and emotions by observing their overt behaviours. According to this theory a person becomes, for instance, more committed to a certain attitude or general cognition if they have to argue on behalf of it (even when this position contradicts a previous attitude), in other words “as I hear myself talk, I learn what I believe” (Bem, 1972; Laird & Bresler, 1992; Miller & Rollnick, 2002). This principle has been supported by numerous studies. For example, manipulated facial expressions can trigger changes in emotion (e.g. Laird & Bresler, 1992) as well as changes in attitudes (racial bias) as assessed by the Implicit Association Test (Ito, Chiao, devine, Lorig, & Cacioppo, 2006). Sharot, Velasquez, and Dolan (2010) studied participants who rated different vacation destinations both before and after making a blind choice that could not be guided by pre-existing preferences. Their results demonstrated that choices not only reveal preferences, but also shape them even when decisions were made randomly. Interestingly, change in preferences was observed only when participants believed they had been instrumental in making a decision, and not when the decision was made by a computer. We hypothesised that the self-perception principle could also be relevant for the modification of cognitive interpretation bias and be integrated in CBM-I procedures and potentially enhance its effect by instructing participants to not only imagine but to positively complete the initially ambiguous scenarios by themselves.

The aim of this study was to develop a new, more active variant of CBM-I and test its impact on positive mood and interpretation bias in comparison to a control group as well as imagery CBM-I. The purpose of this comparison was to further our knowledge about the mechanisms of action underlying successful CBM-I and to optimize its ingredients to enhance
its future therapeutic potential. The positive imagery CBM-I can be described as a standardized guided imagery training that provides participants with auditorily presented descriptions of ambiguous virtual everyday life situations that are consistently combined with positive resolutions (e.g., Holmes et al., 2006). The new training variant, however, instructs participants to not only imagine but to positively complete the initially ambiguous scenarios themselves by speaking one or two phrases into a microphone. As participants in the new training variant have to invent positive resolutions themselves, we expected that they would perceive the scenarios as more authentic than participants in the standardized guided imagery training. Further, the self-generation of positive resolutions (and its subsequent vocalisation) can be seen as a more active process than guided imagery and in terms of the self-perception theory be regarded as “overt behaviour” that could lead to a (greater) modification of cognitions (interpretation bias) and internal states (mood). We therefore hypothesised that the self-generation variant would be more effective in changing interpretation bias and mood towards a more positive direction, based on the principles of self-perception as well as prior findings by Hoppitt et al. (2010).

Method

Participants

54 participants (13 male, 41 female) took part in the study (age in years: \( M = 22.0, SD = 2.9 \)). They were recruited via online advertisements at the website of the Technische Universität Dresden. The majority of them were undergraduate students. One third of the participants (29.6%) in this study showed BDI-scores above the clinical cut-off score with 22.2% reporting symptoms of mild depression (BDI-scores: 14–19) and 7.5% reporting symptoms of moderate depression (BDI-scores: 20–28). Reimbursement for participation consisted of either a small participation fee (5€) or course credits.

Materials

Questionnaire measures. The German version of the Beck Depression Inventory-II (BDI-II; Beck et al., 1996; Hautzinger et al., 2006) was used to measure current levels of depressive symptoms. The trait version of the German version of the Spielberger State-Trait Anxiety Inventory (STAI-T; Laux, Glanzmann, Schaffner, & Spielberger, 1981; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) was given to assess trait anxiety. State positive and negative affect were measured using the German version of the Positive and Negative Affect Schedule (PANAS; Krohne, Eglof, Kohlmann, & Tausch, 1996; Watson, Clark, & Tellegen,
IV. PAPER 3

1988). The 20-item scale was administered with the short-term time instruction (“indicate to what extent you feel this way now”) and rated from 1 (not at all) to 5 (extremely).

**Interpretation bias measure.** Interpretation bias was measured using the Ambiguous Scenarios Test relevant to Depressed Mood II (AST-D-II; Rohrbacher & Reinecke, 2014). This was developed from the Ambiguous Scenarios Test (AST; Berna et al., 2011; Holmes, Lang, & Shah, 2009; Mathews & Mackintosh, 2000), where ambiguous situational descriptions are presented that allow either a positive or a negative outcome interpretation (e.g. It’s the morning of your birthday. The postman comes down the street with their bag). The outcome measure is the emotional valence rating of each description made by participants on a 9-point scale, ranging from extremely unpleasant to extremely pleasant. The AST-D-II consists of two 15-item parallel questionnaires in German, one of which was presented before and after CBM-I, respectively, with the order being pseudo-randomised and counterbalanced between groups. Both versions have been shown to possess high psychometric properties, including high-parallel-reliability (Rohrbacher & Reinecke, 2014).

**Interpretation bias modification training.** Participants were randomly assigned to 1) standardized CBM-I-, 2) Self-generation CBM-I-, or 3) a control group (CG). Using the CBM-I paradigm described by Holmes et al. (2006), participants in all three groups were presented with 50 auditory scenarios. A definition of mental imagery was given. Then participants were instructed to vividly imagine these scenarios from a field-perspective, and to try not to engage in verbal thoughts about the scenarios, such as making comparisons between the scenario and themselves (for details on instructions with practice examples, see Holmes, Coughtrey, & Connor, 2008; Holmes, Lang, & Shah, 2009). Across the three training conditions, the first part of a specific scenario was identical (e.g. “You are on your way to the photographer to pick up your new portrait pictures. As you look at the photos you are quite surprised.”). This part was also ambiguous in that it allowed either a positive or a negative outcome. Each statement lasted about 5-10 seconds. Scenarios in all groups were recorded using the same female voice and presented stereo-phonically through headphones.

The second part of the scenarios differed between training conditions. In the standardized CBM-I condition, the auditory description continued directly on with the positive resolution (e.g. “The pictures are much better than you expected.”). Participants had to vividly imagine the complete scenario including its positive resolution (no time limit was given) and then had to press a “continue”- button to start the next scenario. Participants in the condition self-generation CBM-I were presented the initial part of the scenario. They were instructed (before
training) to vividly imagine the auditorily presented scenario and to resolve its ambiguity by imagining a positive completion by themselves. After the imagery part (no time limit was given) participants had to press a “start”-button and were asked to describe their actively generated positive resolution by speaking a few words into a microphone (all completions were recorded in order to check for compliance). Individuals had to press a “continue”-button to start the next scenario. Participants in the control group were merely presented the first part of the scenario with no resolution. They had to vividly imagine each scenario and had to press a “continue”-button to start the next scenario.

**Filler.** Following previous studies (e.g. Holmes et al., 2006), in order to equalize mood levels after training between the three groups (before the application of the interpretation bias test), participants were presented with a series of classical music extracts for a 10-minutes interval after the interpretation bias training. Participants were asked to rate the extracts with respect to their pleasantness.

**Manipulation checks.** During the training, participants gave ratings at the end of each 10th scenario regarding the previously presented scenario. Participants were asked to rate a) to what degree they had been able to vividly imagine the scenario (vividness), b) as how realistic they had experienced the presented scenario (authenticity), and c) as how pleasant they had perceived the presented scenario (pleasantness). Responses were rated on a 9-point scale, ranging from 1 (not at all) to 9 (extremely).

**Memory questions/expectancy beliefs.** Furthermore, participants were given five memory questions at the end of the session (e.g. “What present did you buy for your partner (in one of the presented scenarios)?”). These questions were part of the cover story and additionally served as a filter to exclude participants with insufficient ability to concentrate on the training. Since imagery has been associated with enhanced learning and memory effects (e.g. Schwartz & Heiser, 2005), we expected participants to recollect the majority of the memory questions. As previous pilot-testing revealed that most pilot-participants were able to correctly answer 4 to 5 of the memory questions, an exclusion cut-off of less than three correct answers was determined a priori. None of the participants fell below that cut-off. In addition, participants were asked to guess the purpose of the study (“In your opinion, what is the purpose of this study?”) after completion of the experiment in order to estimate possible demand or expectancy effects.

**Procedure**

In order to decrease expectancy as well as demand effects, participants were provided with a cover story. They were told that the purpose of the study was to examine the association
between memory effects and spatial representations, and that they would have to answer several questions concerning their mental images after an imagery-task (*experimental phase*). After having given informed consent and being randomly assigned to one of the three training-conditions, participants completed BDI-II, STAI-T, PANAS, as well as the pre-test of the AST-D-II. After the *experimental phase* (CBM-I), they again completed the PANAS. This state questionnaire was given for a third time after the 10 minutes *filler task*. Participants then completed the post-test AST-D-II. They completed the *manipulation check ratings* (vividness, authenticity, pleasantness) during the training. At the end of the study, they answered five *memory questions* and were asked about their *expectancy beliefs* regarding the aim of the study.

**Results**

**Comparison of participants in the three training conditions at baseline**

Questionnaire scores of all baseline measures are depicted in Table IV.1. The three CBM groups were not different in terms of gender ratio, $X^2(2) = 0.20, p = .90$, mean age ($F(2,53) = 0.20, p = .82$), depressive symptoms (BDI-II; $F(2,53) = 2.20, p = .12$), trait anxiety (STAI-T; $F(2,53) = 1.30, p = .28$), mood (PANAS-positive; $F(2,53) = 0.45, p = .64$; PANAS-negative; $F(2,53) = 0.94, p = .40$), nor interpretation tendencies (pre-training AST-D-II; $F(2,53) = 0.40, p = .67$).

**Effects over training – changes in mood measured with the PANAS**

2 x 3 repeated measures ANOVAs with the within-factor time (pre- vs. post training) and the between-factor group (CG, standardized CBM-I, self-generation CBM-I) were calculated using the pre- and post-test scores of the PANAS as the dependent variables, separated for the positive and the negative PANAS subscale (Table IV.1, Figure IV.1). *Negative subscale*: There was no significant main effect of the factors time, $F(1,51) = 0.09, p = .76$, or group, $F(2,51) = 1.15, p = .32$, nor a significant interaction, all $F(2,51) = 0.54, p = .58$, indicating that none of the trainings lead to a change in negative mood. *Positive subscale*: There was no main effect of group, $F(2,51) = 1.27, p = .29$, but a significant effect of time, $F(1,51) = 10.56, p < .01, \eta^2_p = .17$, as well as a significant interaction between time and group, $F(2,51) = 4.05, p = .02, \eta^2_p = .14$, indicating differential changes in positive mood in the three groups during training. Additional paired-samples t-tests comparing pre- and post-CBM mood scores, separated for the three CBM groups, indicated a significant increase of positive mood during standardized CBM-I, $t(17) = 4.32, p < .01$, whereas no change was found in Self-generation CBM-I, $t(17) = 1.17, p = .26$, and the control group, $t(17) = 0.35, p = .73$. There was a trend level difference between
### Table IV.1. Characteristics of participants and effects over training

<table>
<thead>
<tr>
<th></th>
<th>Control group (n = 18)</th>
<th>Standardized CBM-I (n = 18)</th>
<th>Self-generation CBM-I (n = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Age</td>
<td>22.2</td>
<td>2.7</td>
<td>21.6</td>
</tr>
<tr>
<td>% female</td>
<td>78.0</td>
<td>0.4</td>
<td>72.0</td>
</tr>
<tr>
<td>STAI Trait</td>
<td>45.8</td>
<td>5.1</td>
<td>43.9</td>
</tr>
<tr>
<td>BDI-II</td>
<td>8.7</td>
<td>5.6</td>
<td>10.3</td>
</tr>
<tr>
<td>PANAS-positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>31.94</td>
<td>5.88</td>
<td>32.77</td>
</tr>
<tr>
<td>Post-training</td>
<td>32.17</td>
<td>6.25</td>
<td>35.21</td>
</tr>
<tr>
<td>After filler</td>
<td>33.00</td>
<td>6.05</td>
<td>35.26</td>
</tr>
<tr>
<td>PANAS-negative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>12.89</td>
<td>3.45</td>
<td>11.73</td>
</tr>
<tr>
<td>Post-training</td>
<td>13.06</td>
<td>5.03</td>
<td>11.34</td>
</tr>
<tr>
<td>After filler</td>
<td>12.83</td>
<td>3.78</td>
<td>11.29</td>
</tr>
<tr>
<td>AST-D-II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training</td>
<td>0.83</td>
<td>0.74</td>
<td>0.55</td>
</tr>
<tr>
<td>Post-training</td>
<td>0.81</td>
<td>0.87</td>
<td>1.48</td>
</tr>
<tr>
<td>Ratings during training</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vividness</td>
<td>1.98</td>
<td>1.55</td>
<td>2.78</td>
</tr>
<tr>
<td>Authenticity</td>
<td>1.24</td>
<td>1.37</td>
<td>1.77</td>
</tr>
<tr>
<td>Pleasantness</td>
<td>-0.70</td>
<td>0.98</td>
<td>2.73</td>
</tr>
<tr>
<td>Memory questions after training</td>
<td>4.39</td>
<td>0.50</td>
<td>4.22</td>
</tr>
</tbody>
</table>

**Note**: STAI = State-Trait Anxiety Inventory; BDI-II = Beck Depression Inventory II; PANAS = Positive and Negative Affect Schedule; AST-D-II = Ambiguous Scenarios Test relevant to Depressed Mood II.
the three groups’ levels of positive affect post-CBM, $F(2,51) = 2.43, p = .09$. In summary, our analysis revealed no changes in negative mood in any of the training conditions, but an increase of positive mood in the standardized CBM-I used in earlier studies.

**Mood change during the filler task**

To assess whether the mood improvement during training in the standardized CBM-I group survives a neutral filler task or is rather transient, we ran an additional 2 x 3 repeated measures ANOVA with the within-factor time (pre- vs. post filler) and the between-factor group, using the pre- and post-filler task scores of the positive subscale of the PANAS as the dependent variable. The was no significant main effect of group, $F(2,51) = 2.02, p = .14$, but a significant effect of time, $F(1,51) = 57.36, p < .01$, $\eta^2_p = .53$, and a significant interaction between time and group, $F(2,51) = 12.16, p < .01$, $\eta^2_p = .32$, indicating differential changes in mood in the three groups during the filler task. Follow-up paired-samples t-tests, comparing pre- and post-filler positive mood, indicated a significant increase of positive mood in the control group, $t(17) = 9.22, p < .01$ and self-generation CBM-I, $t(17) = 3.69, p < .01$, whereas no change was found in standardized CBM-I, $t(17) = 1.00, p = .33$. In summary, our analysis revealed no mood transition in standardized CBM-I, but mood changes in self-generation CBM-I as well as in the control group.

**Mood after the filler task**

Two one-way ANOVAs, separated for the negative and the positive subscale of the PANAS, were run for the scores recorded after the filler task to compare mood across the three CBM groups. No significant differences between the three training groups were found, neither for negative nor positive affect scores, both $F(2,53) < 1.65$, both $p > .20$. This indicates that the groups were comparable prior to the administration of the bias tasks.
IV. PAPER 3

Figure IV.1. Effects over training – changes in mood

Note: PANAS = Positive and Negative Affect Schedule. Mood change scores are calculated as the difference between the PANAS post-training score and the PANAS pre-training score; positive scores reflect an increase on the measure, negative scores a decrease. Error bars show the standard deviation of the mean.

Effects over training - interpretation bias measured via AST-D-II

A 2 x 3 repeated measures ANOVA with the within-factor time (pre- vs. post training) and the between factor group (CG, standardized CBM-I, self-generation CBM-I) with the pre- and post-training scores of the AST-D-II as the dependent variables was run (Table IV.1, Figure IV.2). The analysis yielded no main effect of group, $F(2,51) = 0.30, p = .74$, but a significant effect of time, $F(1,51) = 16.68, p < .01, \eta^2_p = .25$, and a significant interaction between time and group, $F(2,51) = 5.23, p < .01, \eta^2_p = .17$, indicating differential changes in interpretation bias in the three groups during training. Additional paired samples t-tests separately conducted for each of the three groups indicated a significant increase of positive interpretation bias during training in the standardized CBM-I group, $t(17) = 3.69, p < .01$, and the self-generation CBM-I group, $t(17) = 3.33, p < .01$, whereas no change was found in the CG, $t(17) = 0.13, p = .90$. Independent-samples t-tests suggested no difference between the training effect scores of standardized CBM-I and self-generation CBM-I, $t(34) = 1.08, p = .29$, indicating these two trainings as equally effective in changing interpretation style towards a more positive direction. There was no significant difference in AST-D-II scores post-training between the groups, $F(2,51) = 1.98, p = .15$. In summary, participants in the positive training conditions, but not in the control group, showed a similar increase of positive interpretation as measured with the AST-D-II.
IV. PAPER 3

Figure IV.2. Effects over training – changes in interpretation bias

Note: AST-D-II = Ambiguous Scenarios Test relevant to Depressed Mood II. Bias change scores are calculated as the difference between the AST-D-II post-training score and the AST-D-II pre-training score; positive scores reflect an increase in a positive direction. Error bars show the standard deviation of the mean.

Manipulation checks/memory questions/expectancy beliefs

Using one-way ANOVAs, the groups did not differ regarding their ratings of the vividness and the authenticity of the situation samples, nor their imagery-related memory performance, all $F(2,53) < 1.82$, $p > .17$. However, a significant difference was found regarding their pleasantness ratings, $F(2,53) = 1.82$, $p < .01$, $\eta^2_p = .54$. Subsequent post-hoc pair wise comparisons (Bonferroni) revealed that the control group rated the scenarios as significantly less pleasant than did the two positive training conditions (both $p < .01$). None of the participants correctly guessed the purpose of the study when asked at the end.

Discussion

The goal of this study was to design and evaluate a new self-generation variant of imagery CBM-I and test its effectiveness in comparison to a control group as well as standard or guided imagery CBM-I. The purpose of this comparison was to increase the knowledge about the mechanisms of action behind this cognitive bias modification procedure. Our self-generation version of CBM-I increased positive interpretation, whereas no such change was seen in the control condition. We had assumed that self-generation CBM-I would be more effective than the standardized variant, based on the principles of self-perception as well as prior findings by Hoppitt et al. (2010), but in fact our findings did not support this prediction. Contrary to hypothesis, the effect of self-generation CBM-I on interpretation bias was not
superior compared to the standardized CBM-I. Moreover, the new CBM-I variant had no effect on either positive or negative mood. In contrast, the type of instruction used in earlier studies led to a significant increase in positive mood. These results do not confirm the expected superiority of a more active CBM-I version involving self-generation of positive resolutions in enhancing mood and interpretation, but rather advocate the use of standardized scripts for imagery-based CBM-I to target depressive mood.

How can these findings be explained? We had suggested that the self-generation of positive resolutions (and its subsequent vocalisation) as a more active and overt behaviour would lead – according to the self-perception theory – to greater changes in interpretation bias as well as mood. We had expected that observing themselves generating the (positive) resolutions for the scenarios would facilitate participants internalising these resolutions as their own expectations for the outcomes of the imagined situations. From a self-perception theory perspective, the lack of additional effect on modification of interpretation bias of this self-generation could perhaps be explained by the idea that mental imagery from a field perspective is like an “as-if” experience (cf. Holmes, Lang, & Shah, 2009). That is, even if the imagery is guided, when participants imagine themselves engaging in and experiencing the positive outcomes, they modify their cognitions in line with what they experience themselves thinking, feeling, and doing—albeit in their minds' eye. It may be that the process of creating a scenario resolution to imagine and vocalising it verbally adds nothing on top of this. Consistent with this, analysis of ratings during the training revealed that participants in self-generation CBM-I did not perceive the training situations as more authentic than those in standardized CBM-I. That is, while we had expected that inventing the positive resolutions for themselves might lead to participants experiencing the scenarios as more authentic (i.e. realistic) this was not in fact the case. We suggest that this finding might be because imagery in field perspective – whether guided or self-generated – precludes the kind of evaluative thought processes that might lead to finding the scenarios unrealistic (cf. Holmes, Lang, & Shah, 2009). Therefore, attempting to increase the believability by other means (e.g. self-generation of the resolution) may not add anything above encouraging vivid field-perspective imagery.

Interestingly, the finding of similar effects on interpretation after self-generated and standardized CBM-I could be seen as also contrasting with previous research, showing stronger effects of “active” compared to “passive” procedures in modifying later responses to new emotionally ambiguous descriptions (Hoppitt et al., 2010). However, importantly, substantial differences in study designs might account for this apparent contradiction. First, the present
study explored the potential of two depression-specific positive CBM-I versions in altering interpretation, while the previous study investigated anxiety-specific negative CBM-I. Furthermore, while in our more “active” training variant (self-generated resolutions) participants had to freely invent positive resolutions to then incorporate these into mental imagery, Hoppitt et al. (2011) asked their participants to resolve a fragmented word at the end of a written scenario, with only one solution being possible. Finally, in contrast to the “passive” condition used by Hoppitt et al. (2011), the CBM-I using standardized scripts does require participants to actively engage in the emotional valence of the training material, by generating a field-perspective mental image of the situation described. Such differences in target disorder, CBM-I complexity and demand, and the use of mental imagery might be relevant for the interpretation of results.

Why was there an unexpected lack of mood improvement during our self-generation CBM-I training as opposed to the significant increase of positive mood in standardized CBM-I, that even remained stable over a filler task? Perhaps the generation of positive outcomes may (at least in this study) place a higher demand on concentration and performance, reducing any potential increase in positive affect. The requirement to concentrate on self-generating an ending may have also disrupted participants’ absorption in the positive imagery, reducing its emotional impact, or resulted in more evaluative post-processing of the image via the requirement to monitor the valence of the ending generated. That is, having to “think about” the positive endings may have distracted from the “as if” processing that characterizes mental imagery, and promoted more comparative, evaluative processing (cf. Holmes, Lang, & Shah, 2009). The two training groups did not differ regarding their pleasantness-ratings, we therefore assume that the self-generated resolutions were perceived as equally positive as the standardized resolutions. However, the pleasantness-ratings do not allow conclusions about the objective positivity. This could be an interesting aim for a follow-up study in which the self-generated resolutions are used as guided solutions in the standardized condition to control for objective positivity. It is also worth noting that the explicit information in the self-generation condition that participants would have to generate positive resolutions to all of the training scenarios may have reduced the initial ambiguity of the stems, as participants would be aware from the start of each scenario that it would end positively (cf. Clarke et al., 2014).

What might be the significance of the greater increase in positive mood seen in the standardized imagery CBM-I? Although the main aim of the CBM-I paradigm may be to train more positive interpretation, if the training sessions also boost positive mood then this may
have significant implications for future clinical application in depression. Positive affect may be an important determinant of clinical outcome in depression, with lack of positive affect predicting poorer prognosis (e.g. Morris, Bylsma, & Rottenberg, 2009), and early improvement in positive affect predicting response to antidepressant treatment (Geschwind et al., 2011). Furthermore, there is some preliminary evidence that repeated practice in upregulating the neural areas involved in generating positive emotions may provide a promising route for novel treatment development (Linden et al., 2012). Thus, although the paradigm's primary aim may not be to boost positive affect, this aspect of it may help increase its potential clinical impact. On a very practical level, if a depressed individual was completing the CBM-I in their home environment, then even a transient increase in positive mood following completion of a training session could potentially have a positive impact on subsequent activities. Increasing our understanding of the regulation of positive affect in depression may help to capitalise on such transient increases and increase possible therapeutic benefits (Raes, Smets, Nelis, & Schoofs, 2012; Werner-Seidler et al., 2013). Finally, even if the increase in positive affect had no significant clinical impact, the expectation of a boost in positive mood may enhance motivation to engage with a CBM-I treatment schedule. Our finding that the mood improvement in standardized CBM-I was maintained over a filler task seems very encouraging in terms of stability. However, follow-up studies are needed to clarify prolonged mood effects or their sustained effect after a negative mood induction to verify clinical utility.

On the basis of prior research (e.g. Holmes, Lang, & Shah, 2009; Holmes et al., 2006) and supported by our findings we conclude that imagery CBM-I using standardized scripts, as in previous studies, is the most useful form to pursue. This is perhaps a promising outcome for clinical translation, as people with depressed mood may find it particularly difficult to generate positive interpretations (cf. Wisco & Nolen-Hoeksema, 2010), and thus a standardized imagery CBM-I paradigm may be easier to engage in successfully. Imagery techniques may work relatively well at an emotional level (Holmes & Mathews, 2005), as opposed to cognitive therapy that focuses on rational and verbal techniques such as identifying distorted thinking and challenging dysfunctional beliefs. Further, imagery has been associated with enhanced learning and memory effects (e.g. Schwartz & Heiser, 2005) and may have potential longer term effects. It may be that efforts to increase the effectiveness of imagery-based CBM-I should focus on enriching the use of imagery in the training, or on refining the composition of the standardized training scenarios, rather than by incorporating non-imagery processing requirements into the tasks.
A few limitations of the study need to be highlighted. First, our assessments relied on self-report, involving the risk of response bias. However, as none of the individuals in our sample correctly guessed the study purpose, the influence of demand or expectancy effects on results is rather unlikely. Furthermore, preliminary studies using more involuntary, psychophysiological measures of interpretation provide further evidence for the association between interpretation bias and depression (e.g. Moser, Huppert, Foa, & Simons, 2012). Nevertheless, future studies into the effect of CBM on mood and interpretation bias are encouraged to incorporate more objective measures. Second, although participants rated vividness of imagery equally in all three conditions, we did not measure the participants’ baseline imagery ability and so do not know whether these ratings were equally accurate in all conditions. Follow-up studies should therefore measure trait imagery and also ask participants to provide further ratings of their imagery (e.g. “positiveness”) in order to obtain a fuller picture of their experience of it. A third issue that needs to be considered refers to the generalizability of the present results to patients with depressive disorders and to the interpretation of real-life situations. Given the fact that one third of the participants (29.6%) in this study showed BDI-scores above the clinical cut-off score (BDI-scores>14) makes transfer to clinical samples very promising. Future studies should investigate the longer term impact of the immediate effects of the CBM-I on bias and mood, and their importance in both near and far transfer of the training (Hertel & Mathews, 2011).

In summary, the current study adds further support to the potential therapeutic significance of guided CBM-I-based imagery techniques and evaluated a new, self-generation variant of CBM-I.
General discussion

In this chapter, the empirical findings of the three conducted studies are summarized and their implications are discussed. Further, limitations of this work and suggestions for future research are outlined and a final conclusion is given.

Summary of empirical findings

The study of depression-related interpretation bias and its modification has become a matter of huge interest for cognitive and clinical research over the last years. However, a review of current research revealed 1) a severe lack of psychometrically evaluated measures for depression-related interpretation bias and its modification, 2) inconsistent findings for the existence of depression-related interpretation bias depending on the (direct or indirect) assessment paradigm, 3) a demand for profound knowledge about the underlying work mechanisms and best ingredients for CBM-I procedures, and 4) a lack of measures and CBM-I training materials in the German language.

With these considerations in mind, the objectives of this dissertation were 1) to develop and evaluate an internally consistent and valid measure for the assessment of depression-related interpretation bias and its modification, 2) to validate an indirect priming task to assess interpretation bias and to further examine the interrelations of four different direct and indirect assessment paradigms, 3) to evaluate a new and more active CBM-I variant and test its effectiveness in comparison with guided CBM-I and a control group, and 4) to provide valid test and CBM-I training materials in German.

In Paper 1 (Rohrbacher & Reinecke, 2014), the first objective could be fully accomplished with the development and evaluation of the AST-D-II. In this pragmatic test, individuals are presented with a number of ambiguous sentences, reflecting three relevant content areas based on Beck’s cognitive triad. Participants are asked to imagine the scenarios including its unknown outcome and to rate the emotional valence on a bidirectional Likert scale. The emotional appraisal of the situations indicates the negative (or positive) interpretation tendency. The AST-D-II consists of a 30-item long form as well as two 15-item short forms that can be either used as independent short versions or administered alternately to measure change. Using confirmatory analysis, the fit of the a priori postulated three-factor model, based on Beck’s cognitive triad model, could be supported for the long form as well as the two parallel versions. Internal reliability scores were found to be good for the total scale and acceptable for
the two short versions A and B. Convergent validity was indicated by significant negative correlations between the scales of the AST-D-II and depressive symptoms. The finding is in line with Beck’s theoretical assumptions (Beck, 1997) as well as with outcomes of other studies that found evidence of an association between interpretation bias and depressed mood (e.g. Berna et al., 2011; Butler & Mathews, 1983; Hertel & El-Messidi, 2006; Holmes, Lang, & Shah, 2009). The parallelity of the two short versions was supported by results from confirmatory analysis techniques and was also indicated by equal means, standard deviations and equivalent factor structures of the two forms. Altogether, the current investigation revealed that the long version of the AST-D-II and its two parallel forms can be considered as structurally stable, internally consistent and valid instruments. This study was the first to evaluate and confirm the factorial validity as well as the parallel-test reliability, and therefore suitability to measure modification of interpretation bias measure in the context of depressive mood.

Self-report measures, such as the AST-D-II are generally at risk of being biased through response or social desirability tendencies. A number of studies (Bisson & Sears, 2007; Lawson & MacLeod, 1999; Sears et al., 2011) tried to examine the existence of depression-related interpretation bias on indirect measures, but yielded inconsistent findings. Basically, depression-related interpretation biases were mostly detected via direct measures (for instance ambiguous scenarios tests), whereas findings on indirect measures were mixed. There was also no knowledge about the associations between direct and indirect measures regarding depression-related interpretive bias, since no study yet has examined the proposed direct and indirect measures in one sample. The aim of Paper 2 (Rohrbacher & Reinecke, in preparation) was therefore first, the validation of an affective priming paradigm that has not been tested before for the measurement of interpretation bias and second, the comparison of the indirect priming task with three other, already evaluated, assessment paradigms. Unlike prior studies using semantic priming methods (Bisson & Sears, 2007; Lawson & MacLeod, 1999; Sears et al., 2011), the study in Paper 2 examined an evaluative (affective) priming task based on the paradigm by Fazio et al. (1986) that had not been tested before in the context of depression-related interpretation bias. In this task homophone words were used as ambiguous prime stimuli. Participants had to categorize a target word as either positive or negative as quickly as possible. The target was preceded by either a depression-related prime word, a happiness-related prime word or an ambiguous homophone prime. The expectation was that depressive individuals would show an increased tendency to focus on the negative interpretation of the homophone word than non-dysphoric individuals and therefore show longer reaction times.
V. GENERAL DISCUSSION

when homophone primes were paired with positive-related targets and, conversely shorter reaction times when homophone primes were paired with negative-related targets.

As a second goal of Paper 2, the interrelations between the affective priming task using homophones and three other, already evaluated interpretation bias measures were examined. These measures comprised a scrambled sentences test, a homophone task, and an ambiguous scenarios test. Results did not support the validity of the homophone affective priming task, as no differences between dysphoric and non-dysphoric participants were detected. It was further not related to the ambiguous scenarios test and the scrambled sentences task. There was also no empirical support for the direct homophone task, as it did not differentiate between dysphoric and non-dysphoric individuals and was only weakly related to the AST-D-II. In contrast, the AST-D-II as well as the scrambled sentences task significantly differentiated between dysphoric and non-dysphoric individuals. Both instruments were also highly associated with depressive symptoms and significantly intercorrelated.

In short, the study in Paper 2 found no support for the Homophone Affective Priming Task and the Homophone Task, but outlined further evidence of the validity of the AST-D-II and the SST as well as their interrelatedness.

Based on the methodological findings of Paper 1 and Paper 2, the AST-D-II was regarded to be the most appropriate instrument to measure change after interpretation bias modification and was therefore chosen as outcome measure for the CBM-I study in Paper 3 (Rohrbacher et al., 2014). The aim of this study was to evaluate a new and more active CBM-I variant and test its effectiveness in comparison with guided CBM-I and a control group. Since this study was conducted in Germany, the original CBM-I training materials (Holmes et al., 2006) were translated and adapted to fit the German language context. Participants in study 3 were randomly assigned to standardized CBM-I, self-generation CBM-I, or a control group. All three groups were presented with 50 ambiguous auditory scenarios and instructed to vividly imagine these scenarios from a field perspective. While the first part of a scenario was identical (e.g. “You are on your way to the photographer to pick up your new portrait pictures. As you look at the photos you are quite surprised.”), the second part of the scenario differed between training conditions. Whereas the participants in standardized CBM-I were automatically presented with a positive resolution (e.g. “The pictures are much better than you expected.”), participants in self-generation CBM-I were only presented the initial part of the scenario and had to resolve its ambiguity by imagining a positive completion by themselves. Participants in the control group were merely presented the first part of the scenario with no resolution. Based on prior research in the context of anxiety-related CBM as well as the self-perception principle,
V. GENERAL DISCUSSION

It was hypothesized that the self-generation variant would be more effective than standardized CBM-I in improving interpretation tendencies as well as mood. It was found that the self-generation version as well as the standardized version of CBM-I significantly increased positive interpretation, whereas no such change was seen in the control condition. Interestingly, only the standardized version positively influenced mood, whereas self-generation CBM-I had no effect on either positive or negative mood. These results indicate the effectiveness of both training variants in ameliorating fresh ambiguous stimuli, but do not confirm the expected superiority of a more active CBM-I version involving self-generation.

Theoretical and practical implications

The validity of the newly developed ambiguous scenarios test could be supported in all three studies. Paper 1 revealed that the long version of the AST-D-II and its two parallel forms could be considered as structurally stable, internally consistent and valid instruments. High psychometric were also found for the measure in a new sample in study 2 (Paper 2), which can be regarded as a successful replication of the findings from Paper 1. Further, significant correlations between the AST-D-II and the Scrambled Sentences Task could be detected, which can be regarded as further support for its validity. In Paper 3, the two parallel forms of the AST-D-II were used as outcome measures. Successful bias modification could be measured on these instruments after positive bias training, which further supported the appropriateness of the two parallel forms for the measure of change. AST-D-II is the first instrument, which has been proven to have adequate psychometric properties for the measure of change after any sort of intervention. It could therefore be highly profitable for application in cognitive-clinical research as well as in the therapeutic context, for instance when used to measure improvement in thinking styles after cognitive-behavioural treatment. Due to its questionnaire form and its ease of application, the AST-D-II can - as opposed to other more laborious paradigms – be efficiently used in large scale studies in order to examine, for instance, the aetiological influence of interpretation in mood disorders. Because the AST-D-II has a bidirectional scale, it measures negative as well as positive interpretation tendencies and could also be used to assess the potential protective qualities of positive interpretation bias.

Beck’s assumption of a cognitive triad was mirrored in the three factor solution of the questionnaire. This finding can be regarded as support for the disorder-relatedness of the instrument and could be highly useful in examining (possible) differential effects of cognitive intervention strategies on these three factors. Further, because the AST-D-II is the first German ambiguous scenarios test, the cross-cultural validity of the paradigm as well as the underlying
V. GENERAL DISCUSSION

construct could be demonstrated. Altogether, the current work revealed the AST-D-II and its two parallel short versions as structurally stable, internally consistent and valid instruments.

Findings of Paper 2 showed, that the Homophone Affective Priming Task did not differentiate between dysphoric and non-dysphoric participants and was not related to the AST-D-II and the SST. There was also no support for the Homophone Task, in which homophone words were (explicitly) used as ambiguous stimuli. In contrast, the AST-D-II as well as the scrambled sentences paradigm under cognitive load were shown to significantly differentiate between the dysphoric and the non-dysphoric groups, to be highly associated with the BDI-II and to be significantly intercorrelated. This can be regarded as a successful replication of prior research (e.g. Rohrbacher & Reinecke, 2014; Rude et al., 2003). Further, the AST-D-II as well as the SST - though based on different paradigms - seem to measure related aspects given their high intercorrelation.

What do these findings imply and how can the failure to show differences on the affective priming task as well as the direct homophone paradigm be explained? The lack of evidence for a depression-related interpretation bias on the priming task is in line with prior priming studies (Bisson & Sears, 2007; Lawson & MacLeod, 1999). It has been discussed that depression-related interpretation biases are not measurable through indirect methods for three possible reasons. First, priming methods could be generally unsuitable, as depressed individuals tend to have slower and more variable response times in speed tasks making it difficult to use them as valid indicators for interpretation (e.g., Lawson et al. 2002; Moretti et al., 1996; Sears et al., 2011). Second, the phenomenon of negative interpretation bias does not exist but is rather explained through a negative response bias (Bisson & Sears, 2007; Lawson & MacLeod, 1999). Third, the interpretation of ambiguity does not occur automatically, but rather happens at a later strategic processing level and can therefore not be captured with priming methods (Sears et al., 2011).

Interestingly, the results of Paper 2 do not accentuate the discrepancy between direct and indirect measures, as empirical support was not only found for the ambiguous scenarios test but also for the scrambled sentences paradigm under cognitive load, which has been classified as an indirect measure (Phillips et al., 2010). The findings rather suggest that homophones, at least in the German language, are invalid ambiguous stimuli for the measure of depression-related interpretation bias, since no evidence was found either on the indirect homophone task or on the direct homophone task. Although these findings do not allow implications about the general validity of indirect priming strategies as interpretation bias measures, they clearly accentuate the challenge of developing valid ambiguous test material.
V. GENERAL DISCUSSION

Developing indirect measures that are not at risk of being biased through response tendencies or social desirability is undoubtedly important. Notwithstanding, the most objective assessment is unfeasible without proper ambiguous test stimuli. Since researchers are prone to biased interpretation styles themselves, it seems essential to previously evaluate the validity of ambiguous stimuli in a sample of high and low depressive individuals, rather than to merely rely on its face validity.

Due to its good psychometric properties, the two parallel forms of the AST-D-II were used as outcome measure for the CBM-I study in Paper 3. As summarized above, the new self-generation CBM-I version was found to be successful in ameliorating interpretation bias as compared to a control group. In contrast to the expectation, self-generation CBM-I did not appear to be superior to guided CBM-I and had no effect on positive mood, as opposed to guided CBM-I. Further, participants’ ratings during the training revealed that self-generated scenarios were not perceived as more authentic than standardized scenarios.

How can these unexpected findings be explained? From a self-perception theory perspective, the working mechanism of guided CBM-I might be explained by the idea that mental imagery from a field perspective can be regarded as an active process that is perceived like an “as-if” experience (cf. Holmes, Lang, & Shah, 2009). Apparently, the process of creating a scenario resolution in order to imagine and vocalising it verbally adds nothing on top of this, regarding the perceived authenticity and effectiveness. Surprisingly, these findings could be also seen as contrasting with previous research in the context of anxiety (Hoppitt et al., 2010). However, as outlined in Chapter IV, substantial differences in study designs might account for these inconsistencies.

What could be the reason for the unexpected lack of mood improvement during self-generation CBM-I as opposed to standardized CBM-I, that even remained stable over a filler task? The self-generation of positive outcomes may require a higher demand on concentration and performance and may disrupt participants’ absorption in the positive imagery, reducing any potential increase in positive affect. Further, the information in the self-generation condition that participants would have to generate positive resolutions may reduce the initial ambiguity, which has been reported to be an essential ingredient in CBM-I (cf. Clarke et al., 2014). Although the principal aim of imagery-based CBM-I may not be to boost positive affect, this side effect could be relevant in increasing its clinical impact. Even a transient positive mood shift after CBM-I training might positively influence subsequent activities and enhance the motivation to continue with CBM-I sessions. Positive affect is supposed to be an important determinant of clinical outcome in depression, with lack of positive affect predicting poorer
V. GENERAL DISCUSSION

prognosis (e.g. Morris et al., 2009), and early improvement in positive affect predicting response to antidepressant treatment (Geschwind et al., 2011). It has further been suggested that repeated practice in upregulating the neural areas involved in generating positive emotions may provide a promising route for novel treatment development (Linden et al., 2012). On the basis of prior research (e.g. Holmes, Lang, & Shah, 2009; Holmes et al., 2006) and supported by findings of this work, it can be concluded that imagery CBM-I using standardized scripts, is the most useful form to apply. People with depressed mood may find generating positive resolutions particularly difficult (cf. Wisco & Nolen-Hoeksema, 2010), thus a standardized imagery CBM-I paradigm may be easier to engage in successfully.

Limitations and suggestions for future studies

A few limitations of this work need to be highlighted. The first issue that needs to be considered refers to the generalizability of the present results to patients with depressive disorders. In all three studies, research questions were tested in non-clinical samples. However, given the fact that one third of the participants in all three study samples showed depression-scores above the clinical cut-off score makes transfer to clinical samples very promising. Notwithstanding, future studies with samples of patients formally diagnosed with depressive disorders are needed in order to determine whether the findings reported here generalize to clinical populations. Another restriction is related to the stability of the assessment instruments under research as well as to the homogeneity of the sample. Interpretive tendencies in Paper 1 and Paper 2 were examined only once in a sample of mainly young adults. Thus, stability of the assessment instruments could not be examined and further research in more representative samples is needed to evaluate whether the item contents are equally valid for older individuals. Additionally, it would be worthwhile to explore whether a negative interpretation bias, measured with the AST-D-II (and/or the SST), changes during treatment for depression, as has been demonstrated for other bias parameters during antidepressant treatment (Harmer et al., 2009). An interesting goal for follow-up research would be a comparison of the AST-D-II and the SST with others aspects of bias (e.g. memory or attention) in depression (Koster et al., 2010; Lang et al., 2009) and other populations (Lothmann et al., 2011).

Although the structure of the AST-D-II and the item contents are in line with cognitive theories of depression (e.g. Beck et al., 1979), future research is also needed to empirically determine its specificity by examining the measure in patients with other disorders.

Findings in Paper 2 emphasize the challenge of developing valid ambiguous test material and clearly entail the need to evaluate the validity of ambiguous stimuli in a sample of
high and low depressive individuals before applying these stimuli in indirect tests. Ambiguous test material for the measure of depression-related interpretation bias needs to differentiate between high and low depressive individuals and should therefore show a continuous or normal distribution. More research should therefore be spent to assess the suitability of the proposed ambiguous stimuli and future research should test the validity of indirect methods with ambiguous material that has already been found to be valid in direct assessments. For example, the items of the AST-D-II were found to be valid and could be evaluated using indirect paradigms, such as priming tasks, or event-related brain potentials (ERP) and eye blink reflex, which have been described as psycho-physiological correlates of interpretation bias in the context of social anxiety (Moser et al., 2008) and depression (Lawson et al., 2002).

A restriction of the CBM-I study in Paper 3 concerns the use of self-report assessments as outcome measures. The finding that none of the individuals in the study sample correctly guessed the study purpose makes the influence of demand or expectancy effects rather unlikely. However, the risk of response bias cannot be excluded and future studies into the effect of CBM on mood and interpretation bias are encouraged to incorporate more objective measures.

Future studies should also investigate the longer term impact of the immediate effects of CBM-I on bias and mood, and their importance in both near and far transfer of the training (Hertel & Mathews, 2011). In this context, follow-up studies could, for instance, measure sustained effects after a negative mood induction to verify clinical utility.

Further, although participants rated vividness of imagery equally in all three conditions, the participants’ baseline imagery ability was not assessed. Follow-up studies should measure trait imagery in order to know if dysphoric and non-dysphoric groups differ on this trait.

The two training groups in the CBM-I study did not differ in their pleasantness-ratings. Notwithstanding, these subjective ratings do not allow conclusions about the objective positivity. This could be an interesting aim for a follow-up study in which self-generated resolutions are used as guided solutions in the standardized condition to control for objective positivity.

Due to its positive effect on mood, guided CBM-I seems to have a greater clinical impact than self-generation CBM-I. The more active self-generation version of CBM-I has been also found to be effective in ameliorating interpretation bias and could eventually be of additional benefit in combination with guided CBM-I, for instance, as a sequel at later training stages. This, however, needs to be examined by future research.
V. GENERAL DISCUSSION

Final conclusions

Altogether, this dissertation significantly contributed to the research field of depression-related interpretation bias by 1) filling a gap in measure development studies, 2) increasing the knowledge of the relatedness of different assessment paradigms, 3) increasing the knowledge about the working mechanisms and optimal ingredients of successful CBM-I, and 4) evaluating and providing assessment as well as CBM-I training materials for the German language context. Study 1 of this work revealed the long version of the AST-D-II and its two parallel forms as structurally stable, internally consistent and valid instruments. The study was the first to evaluate and confirm the factorial validity as well as the parallel-test reliability, and therefore suitability to measure modification of interpretation bias in the context of depressed mood. Study 2 was the first study to examine the relatedness of four different direct and indirect assessment paradigms for the measure of depression-related interpretation bias in one sample. Findings revealed the AST-D-II and the SST under cognitive load as internally consistent and valid instruments. Significant intercorrelations between these instruments can be regarded as an indicator of a shared underlying construct. No empirical support was found for the newly evaluated affective priming task using homophones as well as the direct homophone paradigm, which showed low psychometric properties and did not differentiate between dysphoric and non-dysphoric individuals. Findings of study 2 emphasize the threat of using invalid ambiguous test stimuli for the measure of depression-related interpretation bias and clearly entail the need for further methodological evaluation studies. The CBM-I study (study 3) successfully evaluated a new, self-generation variant of CBM-I and added further support to the potential therapeutic significance of imagery-based CBM-I.
References


REFERENCES

Research and Therapy, 5, 51-62.


REFERENCES


REFERENCES


Käse, M., Dresler, T., Andreatta, M., Ehlis, A. C., Wolff, B., Kittel-Schneider, S,...


Linden, D. E. J., Habes, I., Johnston, S. J., Linden, S., Tatineni, R., Subramanian, L., …
REFERENCES


REFERENCES


Rohrbacher, H., & Reinecke, A. (in preparation). Direct and indirect assessment of depression-related interpretation bias: Validation of an evaluative priming task in comparison with a scrambled sentences test, an ambiguous scenarios test and a homophone task.


REFERENCES


REFERENCES


Appendix

AST-D-II (German version)

Instruktion: Versuchen Sie bitte, sich die im Folgenden beschriebenen Szenen so bildlich wie möglich vorzustellen. Bewerten Sie auf einer Skala von "sehr unangenehm" (-5) bis "sehr angenehm" (+5) welche Empfindung sich spontan bei Ihnen bemerkbar macht. Denken Sie nicht zu viel über die Szenen nach, sondern achten Sie auf Ihr erstes Gefühl.

2. Die Firma, für die Sie arbeiten muss große Einsparungen vornehmen. Eines Tages werden Sie zu einem Gespräch bei Ihrem Chef vor geladen.
3. Sie interessieren sich für ein Stellenangebot. Da Sie sich nicht sicher sind, ob Ihre Qualifikationen ausreichend sind, erkundigen Sie sich telefonisch beim Ansprechpartner.
5. Sie gehen zum Bahnhof um eine Freundin zu treffen, die Sie seit Jahren nicht mehr gesehen haben. Sie fragen sich, inwieweit sie sich im Gegensatz zu Ihnen verändert hat.
7. An einem bewölkten Tag sind Sie am Strand. Sie schauen nach oben und bemerken, dass das Wetter sich ändert.
8. Ihr Partner/Partnerin bittet Sie, ein Geschenk für seine/ihre Schwester aus zu suchen, da er/sie keine Zeit hat. Als die Schwester das Geschenk auspackt, erkennen Sie an ihrem Gesichtsausdruck, ob sie richtig gewählt haben.
9. An einem verregneten Sonntag lassen Sie Ihren Gedanken freien Lauf. Viele Erinnerungen werden wach…
10. Ihr bester Freund überredet Sie, zu einem Blind-date zu gehen. An der Bar warten Sie auf die Verabredung und denken darüber nach, was Sie erwartet.
11. Sie halten eine Rede auf der Hochzeit eines befreundeten Paares. Schon nach einigen
APPENDIX

Minuten ist Ihnen klar, wie Ihre Worte aufgenommen werden.

12. Einige "hohe Tiere" besuchen die Abteilung, in der Sie arbeiten. Sie werden spontan gebeten, Ihr Projekt vorzustellen. Am nächsten Tag gibt Ihr Vorgesetzter Ihnen eine Rückmeldung.


16. Sie beginnen demnächst eine neue Stelle, um die sie sich sehr bemüht haben. Sie denken darüber nach, wie es dort wohl sein wird.

17. Ein runder Geburtstag nähert sich. Sie denken darüber nach, wie Ihr Leben bis dato verlaufen ist.

18. Ihr Freund ist ein begeisterter Schlittschuhläufer und überredet Sie, es doch auch einmal zu probieren. Sie ziehen sich Schlittschuhe an und betreten die Eisfläche. Sie gleiten vorwärts, erst langsam und dann immer schneller.


20. Sie sind leidenschaftlicher Hobbyfotograf und überlegen, ob Sie einen Fotoband herausgeben könnten. Eine Freundin, die in einem Verlag arbeitet, sagt Ihnen, was sie von dieser Idee hält.


26. Sie bekommen eine Einladung zum Klassentreffen. Dies lässt Sie zurück an Ihre Schulzeit
APPENDIX

denken…
27. Sie finden beim Aufräumen auf dem Dachboden Fotoalben von sich, die Sie schon lange nicht mehr angeschaut haben. Sie fangen an zu blättern…
29. Sie organisieren die alljährliche Betriebsfeier, wofür Ihnen nur wenig Geld zur Verfügung steht. Während der Feier lassen Sie Ihren Blick schweifen und erkennen schnell, wie Ihre Bemühungen ankommen.
30. Für eine Party haben Sie sich ein neues Outfit gekauft. An den Reaktionen der anderen Leute können Sie erkennen, ob Sie die richtige Wahl getroffen haben.

AST-D-II (English translation)

Instruction: Please imagine the following scenarios as vividly as possible. Rate how the imagined scenario felt on a scale ranging from extremely unpleasant (-5) to extremely pleasant (+5). Don’t think too much, and follow the first idea that comes up.

1. As you enter the room, the commission welcomes you and begins with the oral examination. After just a few minutes you know intuitively how the exam will go.
2. At the company you are working for there have been big cut-backs. One day you are called in to see your boss. When you enter the room, the boss's face is tired.
3. You are interested in a job, but think you might be under-qualified and so ask for details. When you speak to the people, you realize what your chances are to get the job.
4. You are camping in a forest and are very cold. You decide to light a fire. The flames grow in intensity much faster than you imagined.
5. You are going to see a very good friend at the station. You haven't seen them for years. You feel emotional, thinking about how much they might have changed.
6. You are hosting a dinner party for 10 people and got pretty stressed out while preparing the food. You can tell from the initial reaction of the guests how they like the food.
7. It is an overcast day and you are sitting on the beach. You look up to notice the weather really beginning to change.
8. Your partner asks you to buy a present for their sister's birthday, as they are busy. When the sister opens it, her face shows you how she feels.

9. On a rainy Sunday you let your thoughts wander freely. Many memories come back…

10. Your best friend convinces you to go on a blind date and as you sit in the bar waiting to meet your date, you think about how it will go.

11. You give a speech at your friend's wedding. When you have finished, you observe the audience's reaction.

12. Some important people are visiting the office and you are asked at the last minute to present a project to them. Afterwards, you get feedback on your performance.

13. You are in a reflective mood and think back at past achievements and disappointments that you have experienced during your life. Overall, your main feelings about your life so far emerge.

14. You are going to see your sister in her school play. You've left it to the last minute to get there. As you drive up to the school and see the parking bays you anticipate the time it will take you to arrive.

15. You go to a wedding where you know very few other guests. After the party, you reflect on how the other guests behaved.

16. You are starting a new job that you very much want. You think about what it will be like.

17. Your next birthday is approaching soon. You reflect about your life so far.

18. Your friend is very keen on skating and persuades you to try it out. At the rink you put on the skates and step on the ice. You glide forward, slowly at first, then faster.

19. As you walk into the interview room the panel of interviewers welcomes you and proceeds to ask some tough questions. By the end of the interview you know what the outcome is.

20. You are a passionate hobby photographer and wonder, if you could publish a photo book. A friend of yours, who works for a publishing company, tells you what she thinks about this idea.

21. You go to a place you visited as a child. Walking around makes you feel emotional.

22. The probation period at your new job is almost over. You get invited to your boss and receive feedback on how you’ve done so far.

23. You’d love to join a choir and go to an audition. The next day the director of the choir calls you on the phone to tell you if you can join the choir.

24. You want to refresh your Italian language skills und enrol for an advanced-level language course. The teacher however would like to give a placement test first.

25. It is the end of December. You reflect upon the year behind of you.
APPENDIX

26. You got invited to a class reunion. That makes you remember your school days…
27. When you clean up the attic, you find some of your old photo albums you have not looked at in a while. You begin to browse…
28. Your colleague just came back from the holidays and tells you enthusiastically about her experiences. While you listen to her, you think of your last vacation.
29. You are organizing the annual office party on a small budget. On the night of the party, you look around to see if people are enjoying themselves.
30. You buy a new outfit for a party. You can tell if you made the right choice by the reaction of the other people.

Note: Participants are asked to indicate their answers using an 11-point scale anchored from (-5) very unpleasant to (+5) very pleasant.

AST-D-II (LV) comprises the items 1, 3, 6, 8, 10, 11, 12, 20, 23, 24, 29, 30 (factor “self”), 9, 13, 17, 21, 25, 26, 27, 28 (factor “experiences”), 2, 4, 5, 7, 14, 15, 16, 18, 19, 22 (factor “future”).

AST-D-II (A) comprises the items 1, 3, 8, 10, 20, 29 (factor “self”), 9, 17, 27, 28 (factor “experiences”), 4, 5, 14, 15, 22 (factor “future”).

AST-D-II (B) comprises the items 6, 11, 12, 23, 24, 30 (factor “self”), 13, 21, 25, 26 (factor “experiences”), 2, 7, 16, 18, 19 (factor “future”).

List of homophones (English translation in parentheses)

HAP:
KREBS (Cancer/crab), FEST (celebration/fixed) STEUER (tax/wheel) DRACHEN (dragon/kite) ARM (arm/poor) GEFAHREN (danger/drove) RATEN (installment/to advise) HAUT (skin/to beat)

Homophone Task:
Biss/bis (bite/till), trennt/Trend (to separate/trend), verwaist/verweist (orphaned/to refer to sth.), reisst/reist (to crack/to travel), laichen/Leichen (fish spawn/dead body), Rute/Route (rod/route), schaal/Schal (mouldy/scarf), Verband (bandage/union), Mine (mine/face), Leere/Lehre (emptiness/teaching), fiel/viel (to fall/plenty), Steuer (tax/wheel), Fest/fest (celebration/fixed), Drachen (dragon/kite), Krebs (cancer/crab), Schloss (lock/castle), Raten (installment/to
APPENDIX

advise), Gefahren/gefahre (danger/drove), Haut/haut (skin/to beat), arm/Arm (arm/poor),
Kriege/kriege (war/to get), Genossen/genossen (associate/to enjoy)
Summary

Negatively biased interpretation has been hypothesized as an important factor in the aetiology and maintenance of depression. Recently developed cognitive bias modification paradigms, intending to change these biases towards a more optimistic interpretation tendency (CBM-I), seem to offer new promising implications for cognitive therapy innovation. Surprisingly, the increasing interest in the modifiability of interpretation bias is contrasted by a severe lack of methodological and measurement development studies. A review of current research revealed 1) a severe lack of psychometrically evaluated measures for depression-related interpretation bias and its modification, 2) inconsistent findings for the existence of depression-related interpretation bias depending on the (direct or indirect) assessment paradigm, 3) a demand for profound knowledge about the underlying work mechanisms and best ingredients for CBM-I procedures, and 4) a lack of measures and CBM-I training materials in the German language.

With these considerations in mind, the objectives of this dissertation were 1) to develop and evaluate an internally consistent and valid measure for the assessment of depression-related interpretation bias and its modification, 2) to validate an indirect priming task to assess interpretation bias and to further examine the interrelations of four different direct and indirect assessment paradigms, 3) to evaluate a new and more active CBM-I variant and test its effectiveness in comparison with guided CBM-I and a control group, and 4) to provide valid test and CBM-I training materials in German.

The aim of study 1 was to develop and evaluate a pragmatic assessment instrument, consisting of a 30-item questionnaire (long version) and two 15-item parallel short versions (A and B). Items were generated as ambiguous sentences, reflecting three relevant content areas based on Beck’s cognitive triad (Beck et al., 1979). All three versions of the instrument were found to be structurally stable, internally consistent and valid. In line with Beck’s cognitive triad (Beck et al., 1979) in depression, confirmatory factor analyses determined a three factor-solution (self, experiences, future). Significant correlations were found between all scales and depressive mood. The two short versions represent the same underlying constructs, share identical psychometric properties and possess high parallel-test reliability. The study was the first to evaluate and confirm the factorial validity as well as the parallel-test reliability, and therefore suitability to measure modification of interpretation bias measure in the context of depressed mood.
The aim of study 2 was to evaluate an indirect affective priming task using homophones as ambiguous material and to compare this task to three different already established interpretation bias paradigms, namely an ambiguous scenarios test, a scrambled sentences test under cognitive load, as well as a homophone paradigm. No empirical support for mood-related interpretation bias was found for the newly developed affective priming task as well as the homophone task, which both did not differentiate between dysphoric and non-dysphoric individuals. The ambiguous scenarios test and the scrambled sentences test under cognitive load were shown to be internally consistent and valid instruments that were also highly intercorrelated. These findings emphasize the challenge to create valid ambiguous test stimuli for the measure of depression-related interpretation bias as well as the inequality of different assessment paradigms. Study 2 was the first study to examine the relatedness of four different direct and indirect assessment paradigms for the measure of depression-related interpretation bias in one sample.

The aim of study 3 was to compare the efficacy of CBM-I requiring participants to imagine standardized positive resolutions to a novel, more active training version that required participants to generate the positive interpretations themselves. Participants were randomly allocated to (1) standardized CBM-I, (2) self-generation CBM-I or (3) a control group. Outcome measures included self-report mood measures and a depression-related interpretation bias measure. Both positive training variants significantly increased the tendency to interpret fresh ambiguous material in an optimistic manner. However, only the standardized imagery CBM-I paradigm positively influenced mood. Study 3 successfully evaluated a new CBM-I variant, gained insights about the role of self-generation and added further support to the potential therapeutic significance of imagery-based cognitive bias modification strategies.

Altogether, this dissertation significantly contributed to the research field of depression-related interpretation bias by 1) filling a gap in measure development studies, 2) increasing the knowledge of the relatedness of different assessment paradigms, 3) increasing the knowledge about the working mechanisms and optimal ingredients of successful CBM-I, and 4) evaluating and providing assessment as well as CBM-I training materials for the German language context.
Ein negativer Interpretationsbias bezieht sich auf die Tendenz, mehrdruitge Reize (z.B. Situationen, Gesichtsausdrücke, Rückmeldung von anderen Personen) in einer pessimistischen Art und Weise zu interpretieren. Diese kognitive Verzerrung im Rahmen der Informationsverarbeitung wird als wichtiger Faktor für die Entstehung und Aufrechterhaltung von depressiven Störungen diskutiert. Insbesondere die Modifikation dieser Tendenz zu einem positiveren Interpretationsstil (Cognitive bias modification targeting interpretation; CBM-I) erweckte in den vergangenen Jahren großes Forschungsinteresse, da diese innovativen, computergestützten Techniken wirkungsvolle Behandlungsansätze für die Therapie von Depressionen bedeuten könnten.

Nach aktuellem Forschungsstand besteht jedoch 1.) ein Mangel an hinreichend evaluierten Instrumenten für die Erhebung von depressionsbedingten Interpretationsverzerrungen und insbesondere deren Veränderbarkeit, 2.) eine inkonsistente Befundlage für die Existenz von depressionsbedingten Interpretationsverzerrungen in Abhängigkeit von der (direkten oder indirekten) Erhebungsmethode, 3.) ein Bedarf an genauerer Erkenntnissen über die zugrunde liegenden Wirkmechanismen und effektivsten Komponenten von CBM-I und 4.) ein Fehlen von CBM-I Materialien und Erhebungsmethoden in deutscher Sprache. Vor diesem Hintergrund wurden die Ziele dieser Dissertation formuliert, die neben einer kurzen Einführung in das Themengebiet und einer abschließenden Diskussion aus drei Studien besteht.


Danksagung

Mein ganz besonderer Dank gilt Dr. Andrea Reinecke, ohne die ich mit diesem spannenden und innovativen Forschungsfeld nicht in Berührung gekommen wäre. Vielen, vielen Dank für Deine großartige fachliche und menschliche Unterstützung, die vielen hilfreichen Anmerkungen und genialen Anregungen, die Unterstützung bei der Publikation der Fachartikel und die zahlreichen motivierenden Worte.

Mein herzlicher Dank geht an Prof. Dr. Stephan Mühlig, der mir die Unterstützung und den nötigen Freiraum gab, diese Arbeit zu verwirklichen und mich sehr ermutigte, die Dissertation zum Abschluss zu bringen.

Des weiteren möchte ich mich ganz herzlich bei Prof. Dr. Mike Rinck für seine Bereitschaft bedanken, diese Arbeit zu begutachten.

Für die Bereitstellung der englischen Erhebungsinstrumente und originalen CBM-I Trainingsmaterialen sowie der tollen fachlichen Unterstützung möchte ich mich außerdem ganz herzlich bei Prof. Dr. Emily Holmes und Dr. Simon Blackwell bedanken.

Ein großes Dankeschön geht an Prof. Dr. Thomas Goschke und Frau Wobst für die logistische Unterstützung bei der Durchführung der Studien. Großer Dank gilt auch allen Probandinnen und Probanden, die an den drei Studien teilgenommen haben. Herzlichen Dank an Dr. Natalie Vannini, die mir stets mit gutem Rat sowie fachlicher und emotionaler Unterstützung zur Seite stand und somit sehr zum Gelingen dieser Arbeit beigetragen hat. Vielen lieben Dank auch an Dr. Eva Hoch für ihre wertvollen Ratschläge und die vielen motivierenden Gespräche.

Eidesstattliche Erklärung

Hiermit erkläre ich, Heike Rohrbacher, geboren am 30.06.1977 in München, dass ich die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe.

Heike Rohrbacher
München, den 25.11.2015
Curriculum Vita

Persönliche Angaben
Name: Heike Rohrbacher
Geburtsdatum: 30.06.1977
Geburtsort: München

Studium
1998-2000
   Studium der Psychologie, Grundstudium (Universität Amsterdam, Niederlande)
2000-2004
   Studium der Psychologie, Hauptstudium (Technische Universität Dresden)

Fort- und Weiterbildungen
2005-2010
   Aufbaustudiengang Psychologische Psychotherapie (Institutsambulanz und Tagesklinik der TU Dresden GmbH; Approbation 2010)

Berufliche Tätigkeiten
Seit 2011
   Praxis f. Verhaltenstherapie, München
   Psychologische Psychotherapeutin
2009-2010
   Tagklinik Westend/Psychosomatische Klinik Windach, München
   Therapeutische Tätigkeit
2005-2008
   Institutsambulanz und Tagesklinik der TU Dresden GmbH
   Therapeutische Tätigkeit
2005-2007
   Institut f. Klinische Psychologie und Psychotherapie, TU Dresden
   Wissenschaftliche Mitarbeiterin
Lehrtätigkeiten

- Seminar „Sucht und Abhängigkeit im Kindes- und Jugendalter (2014, Aufbaustudiengang KJP, Ruhr-Institut, Bochum)
- Diverse Workshops und Therapeutenschulungen zum Thema „CANDIS-Modulare Therapie von Cannabissstörungen“, seit 2005

Preise und Förderungen

- 2009: Forschungspreis 2010 der Deutschen Gesellschaft für Verhaltenstherapie (CANDIS-Projektgruppe)
- 2007: Travel-Award NIDA/CPDD (USA)

Publikationen

*Therapiemanual*


*Artikel und Buchbeiträge*

Rohrbacher, H., & Reinecke, A. (in preparation). Direct and indirect assessment of depression-related interpretation bias: Validation of an evaluative priming task in comparison with a scrambled sentences test, an ambiguous scenarios test and a homophone task.


Artikel und Buchbeiträge


