

Affect dynamics : a focus on genes, stress, and an opportunity for change

Citation for published version (APA):

Menne-Lothmann, C. (2015). *Affect dynamics : a focus on genes, stress, and an opportunity for change*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20150129cm>

Document status and date:

Published: 01/01/2015

DOI:

[10.26481/dis.20150129cm](https://doi.org/10.26481/dis.20150129cm)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

AFFECT DYNAMICS

A focus on genes, stress, and an opportunity for change

Claudia Menne-Lothmann

AFFECT DYNAMICS

A focus on genes, stress, and an opportunity for change

PROEFSCHRIFT

Ter verkrijging van de graad van doctor aan de
Universiteit Maastricht,
op gezag van de
Rector Magnificus,

prof. Dr. L.L.G. Soete,

volgens het besluit van het College van Decanen,

in het openbaar te verdedigen op
donderdag 29 januari 2015 om 14:00

door

Claudia Menne-Lothmann

Geboren op 22 September 1983 te Düren, Duitsland

Lay out and printing: Ridderprint, Ridderkerk, The Netherlands

Cover: Claudia Menne-Lothmann, Ridderprint BV

Cover Foto: Martin Klimas, 'Paul Hindemith – Ludus Tonalis' from series
'Sonic Sculptures'

ISBN: 978-90-5335-820-7

© Claudia Menne-Lothmann | Maastricht | The Netherlands

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, mechanically, electronically, by photocopy, by recording, or otherwise, without prior written permission from the proprietor.

Promotor

Prof. dr. Jim van Os

Co-promotors

Dr. M. Wichers

Dr. N. Jacobs

Beoordelingscommissie

Prof. dr. Inez Myin-Germeys (voorzitter)

Prof. dr. M. deVries

Prof. dr. M. Peters

Dr. P. Kuppens

Dr. E. Salemink

für Daniel

für Nora und den Zwerg in meinem Bauch

für meine Eltern

The research presented in this thesis was performed at the School for Mental Health and Neuroscience, Department of Psychiatry & Psychology, Maastricht University, Maastricht, the Netherlands and at the Department of Experimental Psychology, University of Oxford, United Kingdom.

Paranimfen

Martine van Nierop

Jeroen Decoster

Table of contents

Chapter 1	General introduction	9
Chapter 2	The dynamic interplay between negative and positive emotions in daily life predicts response to treatment	27
Chapter 3	Genetic and environmental causes of individual differences in daily life positive affect and reward experience and its overlap with stress-sensitivity	51
Chapter 4	A modest impact of genes on real-world reward experience: a special role for the OPRM1 gene	71
Chapter 5	Psychological and biological validation of a novel digital social peer evaluation experiment (digi-SPEE)	93
Chapter 6	Cognitive bias modification training in adolescents: Effects on interpretation biases and mood	119
Chapter 7	How to boost positive interpretations? A meta-analysis of the effectiveness of cognitive bias modification for interpretation	141
Chapter 8	General Discussion	205
	Summary	225
	Samenvatting	231
	Valorization	237
	Dankwoord	245
	Curriculum Vitae	251
	List of Publications	255

CHAPTER

1

General Introduction



AFFECT DYNAMICS

Positive as well as negative affective states are continuously present throughout our everyday lives. Most importantly, affective states are *momentary* experiences: they are dynamic and constantly subject to change. These changes can be associated with one another, with particular situations a person is exposed to, or with the way a person appraises these situations. These processes are inherently subjective and may differ between individuals. In sum, ‘affect dynamics’ thus refer to the subjective experience of affective states as they occur in the moment, as they change from one moment to the next, and as they change in response to particular situations or appraisals of these situations¹.

Understanding affect dynamics bears great importance as they have been proposed to constitute a basic mechanism of mental health (e.g., see Wichers, 2013). The current thesis aims to increase the understanding about selective affect dynamic processes that remain ill-understood so far. More specifically, the current thesis aims to (i) examine the interaction between positive and negative affect dynamics as they occur in the flow of daily life, (ii) determine the influence of genetic variation on positive affect dynamics in the flow of daily life, (iii) investigate the impact of daily social stressors on momentary affective states, and finally (iv) investigate whether and how the experience of momentary stress can be reduced through computerized trainings targeting one specific affect dynamic process. In order to gain a better understanding about the implications of affect dynamics for mental health, an overarching aim is to link each of these affect dynamics with symptoms of and risk factors for psychopathology, most prominently depression.

In this chapter I will introduce the aims of this thesis in more detail by differentiating between positive and negative affect dynamics. I will first explain how *positive* affect dynamics are related to mental wellbeing and how *negative* affect dynamics are associated with psychopathology. I will furthermore elaborate on the interaction of positive and negative affect dynamics and emphasize the influence of genes and stress on affect dynamics. Throughout the introduction I will highlight affect dynamic processes that are in need for further empirical investigation and describe how these will be addressed in the studies included in this thesis. As affect dynamics occur in the moment they are difficult to capture. I will therefore explicate two suitable methods to validly assess these momentary processes: experience sampling and experimental designs. I will conclude the introduction with an outline of the current thesis.

¹ In this thesis, the terms ‘affective states’, ‘affect’, ‘emotions’, and ‘mood (-states)’ are employed to refer to in-the-moment experiences of emotions and are used interchangeably.

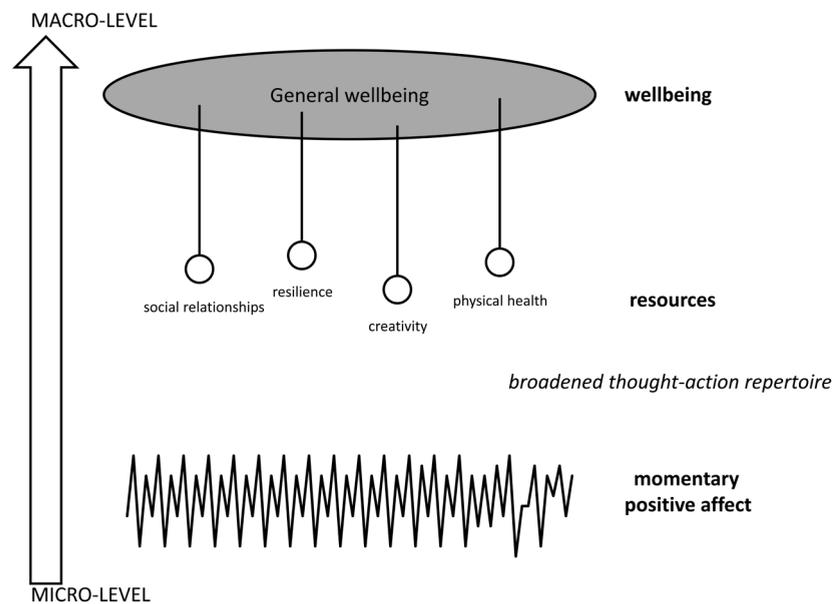
Affect dynamics and mental health

Positive affect dynamics and wellbeing

For a long time the experience of happiness or positive affect has been regarded as an end-state of favourable events and life-circumstances (Lyubomirsky, King, & Diener, 2005) signalling goal achievement (King, Hicks, Krull, & Del Gaiso, 2006). While intuitively quite reasonable, the view that positive affect solely indicates satisfactory progress in one’s life has been seriously challenged in the last decades. It has been shown that positive affect itself is not only associated with favourable life circumstances (e.g., see Berry & Hansen, 1996; Cohen & Pressman, 2006), but *predicts* successful outcomes in all areas of life as well, including social relationships, employment status, and physical health (Cohn, Fredrickson, Brown, Mikels, & Conway, 2009; King et al., 2006; Lyubomirsky et al., 2005). It may therefore not be surprising that the experience of positive affect has been called “the hallmark of [happiness]” (Lyubomirsky et al., 2005, p.840).

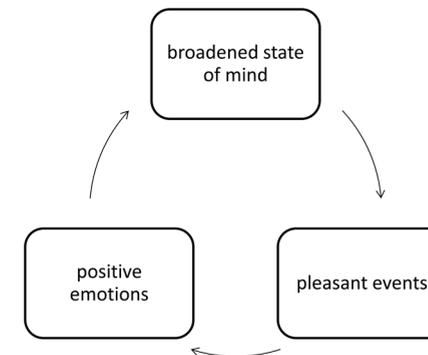
Positive affect is usually thought to be experienced at the micro-level, i.e., as a momentary fleeting state, while its effects on general wellbeing become apparent at the macro-level, i.e., as durable and stable outcomes. This reflects the two parts of hedonic wellbeing: the affective and the cognitive part. While the affective part encompasses affective states, the cognitive one includes evaluations about life satisfaction (Kuppens, Realo, & Diener, 2008; Nes, 2010; Shmotkin, 2005). Both components are distinct but strongly related (Diener, 1994). Barbara Frederickson’s ‘broaden-and-build theory’ of positive emotions offers an explanation for *how* these two parts may be connected. Momentary positive affect is thought to broaden our ‘thought-action-repertoire’ making our mind open to a variety of opportunities. This open-mindedness will cause us to build several enduring social, physical, and intellectual resources we can rely upon in times of stress (Fredrickson, 1998, 2001). Early evidence underpins this theory. For example, Cohn and colleagues showed that a prospective association of daily positive affect and general life satisfaction was partially mediated by building resilience (Cohn et al., 2009). Figure 1 depicts the ‘broaden-and-build theory’ along the axis of the micro – to macro-level of experience paralleling the theory on the dynamic nature of depression by Wichers (2013).

Figure 1. The ‘broaden-and-build theory’ of positive emotions along the axis of micro-macro-level of experience (adapted from Fredrickson, 1998, 2001; Wichers, 2013)



What does positive affect set into motion at the micro-level to build resources and enduring wellbeing? To fully grasp the working mechanisms of positive affect Garland has elegantly zoomed in to the processes of *momentary* positive affect in daily life (Garland et al., 2010). In line with the ‘broaden-and-build theory’ he proposes that the experience of positive affect broadens the momentary state of mind resulting in a widened focus and increased chances for encountering pleasant events, which then further increase positive affect (Garland et al., 2010) (Figure 2).

Figure 2. The self-perpetuating action of positive affect at the micro-level of experience. The experience of positive emotions results in a broadened state of mind² which then fosters the experience of pleasant events. This again leads to increases in positive affect (Garland et al., 2010).



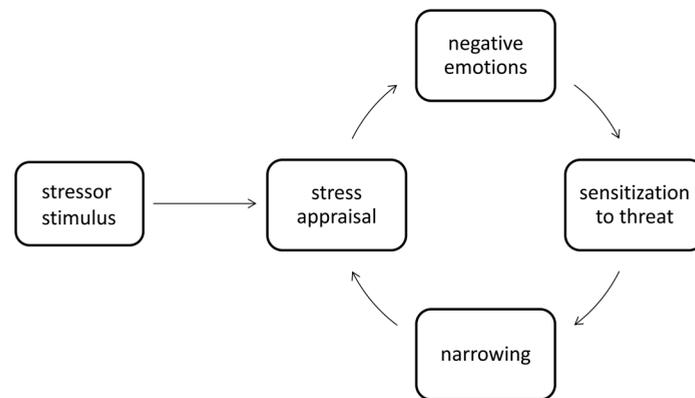
This self-perpetuating action of positive emotions becomes particularly relevant in the face of stress. When encountering one of the inevitable stressors of daily life, people will feel capable of effectively dealing with this stressor because their broadened mind-set allows them to positively re-appraise the stressor which then results in increased positive affect and reduced stress-appraisal. Garland has called these processes the ‘upward spiral of positive emotions’ leading to more general flourishing and wellbeing (Garland et al., 2010; Garland, Gaylord, & Fredrickson, 2011). In support for this, observational studies at the macro-level of wellbeing suggest that the system of wellbeing is able to sustain considerable homeostasis even when stress increases (Cummins, 2010; Tomy & Cummins, 2011), presumably by keeping stress appraisals at a low level (Garland, Gaylord, & Park, 2009).

Negative affect dynamics and depression

However, once the stressors are becoming too strong or are experienced chronically, this homeostasis will be lost (Cummins, 2010) and individuals might enter a self-perpetuating ‘downward spiral of negative emotions’ (Figure 3) (Garland et al., 2010). In this downward spiral stress appraisals result in negative emotions facilitating the detection of threat (i.e. sensitization to stress) and a narrowed focus preparing the individual for fight-flight responses further facilitating subsequent stress appraisals (Fredrickson, 2001; Garland et al., 2010).

² Note. Garland elaborates on this upward spiral of positive emotions in the context of mindfulness. For the purpose of the current thesis we here only include ‘broadened state of mind’ to summarize the potential cascade of mental operations (i.e., decentering – state of mindfulness – broadening), because the current focus does not lie on mechanisms of mindfulness.

Figure 3. The downward spiral of negative emotions adapted from Garland et al (2010; 2011). The appraisal of stress is supposed to increase the experience of negative emotions, which cause a sensitization to threat. This results in a narrowed focus on stressors, which again increases the tendency to appraise stress.



In situations of acute threat negative emotions have a great survival value by enabling the organism to detect and encounter threat. However, when stressors are experienced as overwhelming or are experienced chronically, individuals might get caught in this downward spiral. It is easy to imagine then that wellbeing will decrement and mood-symptoms will occur. Indeed, the processes of the downward spiral have been found typical for depression (Cummins, 2010; Garland et al., 2010). Furthermore, the two main symptoms of depression include increased negative mood (fueling the downward spiral of negative emotions) and decreased ability to experience reward and positive affect in daily life (preventing entry into the upward spiral of positive emotions) (APA, 2000). It thus seems that individuals suffering from mood-symptoms are caught in the downward spiral and have little access to counteracting positive affect.

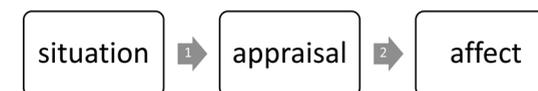
Interaction between positive and negative affect dynamics

The processes of the upward and downward spirals are incompatible at every single moment as “the affective system cannot simultaneously be both narrowed and broadened” (Garland et al., 2010, p. 855). In support for this, positive emotions are proposed to counteract downward spiral processes by reducing stress-appraisals (Garland et al., 2010). There is now emerging evidence that the experience of momentary positive affect indeed reduces downward spiral processes. For example, it has been demonstrated that the long-term effects of momentary positive affect include resilience against future depressive symptoms (Geschwind et al., 2010; Wichers et al., 2009). Research focusing on the micro-level processes has shown that momentary positive affect attenuates increases in negative affect in response to daily life stress-appraisals (Wichers et al., 2007a). Other research has shown that positive affect

can linger immediate cardiovascular effects of negative affect (undoing hypothesis; Fredrickson & Levenson, 1998). However, one of the most fundamental questions about the interaction of positive and negative affect dynamics remained unaddressed so far. It is still unknown whether the interaction between positive and negative affect in the flow of daily life is associated with current and future psychopathology. Therefore, the current thesis aims to examine negative affect in response to changes in positive affect and vice versa and to investigate whether these processes are associated with symptoms of depression (**Chapter 2**).

Next to being influenced by one another, affective states are also supposed to be strongly influenced by appraisals. Appraisals include evaluations of the situation with respect to “own concerns, goals, and competencies” (Kuppens & Tong, 2010, p. 1139). The appraisal theory of affective experience claims that the experience of affective states is not so much contingent on outside stimuli per se but rather on the perceived implications of this outside stimuli for personal wellbeing. Thus, *how* a situation is appraised impacts on affective experience (e.g., see Roseman & Edvokas, 2004; Smith & Kirby, 2009). Figure 4 depicts the role of appraisal in affective experience.

Figure 4. Appraisal theory for affective experience (adapted from Kuppens & Tong, 2010).



The spiral processes emphasize the role of stress-appraisals which are (through a series of mental processes) reciprocally linked with both positive and negative affective experiences (Garland et al., 2010). However, the role of pleasant appraisals as a source of positive affect has received less attention. Based on the appraisal theory pleasant appraisals should relate to increases in positive affect, which has been demonstrated to be indeed the case (Geschwind et al., 2010; Wichers et al., 2009). Regarding positive and negative affect dynamics and their influences on one another, this leads to the question whether increases in positive affect in response to pleasant appraisals, called reward sensitivity, is a process independent of the increase in negative affect in response to unpleasant appraisals, called stress sensitivity (also see figure 3). This question is of great interest for the treatment of mood disorders. If stress sensitivity and reward sensitivity are two ends of one continuum (e.g., ‘sensitivity’) it might be sufficient to reduce stress sensitivity in order to escape the downward spiral of negative emotions and to stimulate the upward spiral of positive emotions. However, if stress sensitivity and reward sensitivity are different entities, this would suggest two separate targets for intervention (i.e., increasing reward sensitivity and decreasing

stress sensitivity). In order to address this issue, the current thesis aims to assess the association between stress sensitivity and reward sensitivity in the flow of daily life (**Chapter 3**).

Influences on affect dynamics

Various influential factors have been determined for mental health. Especially, genes and stress are consistently reported as being important risk factors for the development of psychopathology (e.g., Booij, Bouma, de Jonge, Ormel, & Oldehinkel, 2012; Cuellar & Johnson, 2009; Hammen, 2005; Peerbooms et al., 2011; Rao, 2010; Rietschel et al., 2010; Risch et al., 2009). One way of gaining a deeper understanding about the role of affect dynamic processes in mental health is to comprehend how these risk-factors impact on affect dynamics.

Genes

Individuals differ in their affective experience. According to the appraisal model of affective experience there are two different sources for these differences between individuals (Kuppens & Tong, 2010). First, individuals may differ in how they appraise the same situation, which concerns the situation – appraisal association (arrow ‘1’ in Figure 4). Second, individuals may differ in how the same appraisal is associated with momentary affective experience, which concerns the appraisal – affect association (arrow ‘2’ in Figure 4) (Kuppens & Tong, 2010). Genetic variation may cause individual differences in the situation-appraisal association as well as in the appraisal-affect association. For example, genetic variation may make some individuals more prone to appraise situations in a certain way or to show particularly pronounced affective reactions to certain appraisals than others. This would indicate that some individuals are – due to their genetic makeup – more likely to enter downward or upward spiral affective processes. Supporting that genetic variation is associated with negative affect dynamics, it has been shown that individual differences in the situation – negative appraisal link (arrow ‘1’ in Figure 4) is influenced by genetic variation which is partly shared with mood disorders (Eley et al., 2008; Fox & Standage, 2012; Lau & Eley, 2008; Lau, Rijdsdijk, & Eley, 2006). It has also been demonstrated that the experience of momentary negative affect as it is (Jacobs et al., 2012), and in response to stress appraisals (i.e. stress sensitivity; arrow ‘2’ in Figure 4) (Jacobs, Rijdsdijk, et al., 2005; Wichers et al., 2007b) is under the influence of genetic variation. It therefore seems that some individuals have a higher propensity of entering the downward spiral than others because they are genetically more vulnerable to experience stressful appraisals, negative affect, and stress sensitivity. While the cognitive component of wellbeing (i.e. evaluations about life satisfaction) has been shown to differ as a function of genetic

variation (Nes, 2010; Nes, Czajkowski, & Tambs, 2010), the findings for positive affect dynamics such as positive affect and reward sensitivity are less conclusive. Positive affect as measured by asking individuals whether they felt several positive affective states within the past few weeks did not differ between individuals as a source of genetic variation (Baker, Cesa, Gatz, & Mellins, 1992). However, assessing affect retrospectively may result in a flawed picture of *momentary* affective experience as it partly taps into the concept of cognitive wellbeing (Diener, 1994) and may be subject to recall biases (Shiffman, Stone, & Hufford, 2008). Additionally, to date but one study suggests a role of one single genetic variant in momentary reward sensitivity (Wichers et al., 2008). This latter study is in urgent need of replication. Furthermore, the *degree* to which reward sensitivity is influenced by genetic variation and whether additional genetic variants play a role needs yet to be determined. The current thesis aims to address these gaps by studying the influence of genetic variation on prospectively assessed *momentary* positive affect and reward sensitivity (**Chapter 3, Chapter 4**).

Stress

Next to genes, the experience of stress is another important risk factor for the development of psychopathology (e.g., Booij et al., 2012; Cuellar & Johnson, 2009; Hammen, 2005; Rao, 2010). One conceptualization of stress is the affective reaction to a stressful situation (situation – affect association) (e.g., Cuellar & Johnson, 2009). Today, empirical research has shown that psychopathology is associated with aberrant affective reactivity to stressful situations (e.g., Strout, Salovey, & Epel, 2002; Williams, Cheing, & Choi, 2000). Particularly social stress has been identified as a potent stressor (Williams et al., 2000). The pure possibility of being evaluated has already been shown to be experienced as stressful (Dickerson & Kemeny, 2004). The odds for being exposed to social evaluation have recently reached a new dimension. Today social interactions occur increasingly in online social networks, particularly among adolescents (Donnerstein, 2012; Pempek, Yermolayeva, & Calvert, 2009), who experience increased vigilance to peer relationships (Sommerville, 2013) and are at increased risk for mood-disorders (Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993). Thus, chances for being evaluated by peers become virtually unlimited in a vulnerable period of life. It therefore bears great importance to understand the momentary affective reactions to peer-evaluations in online interactions among adolescents. In order to achieve this, tools enabling the assessment of sensitivity to online social evaluations are required. To this end the current thesis aims to validate a new computerized task mimicking online social interactions between adolescent peers as stressful (**Chapter 5**).

According to the appraisal model of affective experience, the affective reaction to a stressful situation can be split into a situation – appraisal (arrow ‘1’ in figure 4) and an appraisal – affect association (arrow ‘2’ in figure 4), both of which have been associated with psychopathology (Beck, 1970; Beck, Emery, & Greenberg, 1985; Koster, Fox, & MacLeod, 2009; MacLeod & Cohen, 1993; Myin-Germeys et al., 2003; Wichers et al., 2007b). In the current thesis I want to focus on the possibility of altering the situation – negative appraisal association as one way of reducing the level of experienced stress. It has long been recognized that mood-disorders are strongly associated with a tendency to appraise situations that are inherently neutral or ambiguous as stressful, which has been called a negative *interpretation bias* (Beck, 1970; Beck et al., 1985; Koster et al., 2009; MacLeod & Cohen, 1993). This aberrant processing of neutral or ambiguous situations is proposed to be a consequence of the threat-sensitization in the downward spiral (Garland et al., 2010) (also see figure 3). That is, the increased awareness for stressful stimuli expands to ambiguous stimuli, which are then automatically interpreted as stressful (Garland et al., 2010; Koster et al., 2009; MacLeod & Cohen, 1993) further facilitating downward affect dynamics. Being able to put a hold on such negative interpretation biases may therefore be a promising way of stopping downward spiral processes. Trainings have been developed to modify maladaptive interpretation biases, called cognitive bias modification for interpretation (CBM-I). It has been demonstrated that these trainings are effective in adult samples (e.g., see Mathews & Mackintosh, 2000). However, CBM-I’s effectiveness in younger samples and factors influencing CBM-I’s overall effectiveness across studies remain yet to be determined. The current thesis therefore aims at investigating whether CBM-I’s are effective in adolescent samples and to what degree and how negative interpretations (i.e. appraisals) of external stimuli can most effectively be manipulated by CBM-I (**Chapter 6, Chapter 7**).

Methods to capture affect dynamics

One major challenge in validly investigating affect dynamics is a methodological one because the experience of momentary affect is fleeting, varying, and readily influenced by immediate environmental factors (e.g., see Csikszentmihalyi & Hunter, 2003). In order to measure momentary and varying affective states as they occur *in the flow of daily life*, the experience sampling method (ESM) is a perfectly suited option. In order to measure momentary affect and appraisals in response to *controlled environmental factors*, experimental designs are indicated. In the remainder of this section I will shortly describe both methods and relate them to the applicable research questions addressed in the current thesis.

Experience sampling

The experience sampling method (ESM) has been developed to validly capture the experience of momentary affect in the flow of daily life. ESM prompts individuals to simply state their momentary level of several affective states (e.g., happy, sad, interested, irritated) and circumstances (e.g., being alone or in company, being at work) at several unpredictable moments during a day for a couple of days (Csikszentmihalyi & Hunter, 2003; Nes, 2010). ESM has repeatedly been shown to be a valid and reliable assessment tool for momentary experiences (Delespaul, 1995; Jacobs, Nicolson, et al., 2005). Because of the high frequency with which this affect is prospectively measured this methodology is able to zoom in to the micro-level of experience and to unravel dynamic relations between appraisals of context and affect from one moment to the next.

The first part of the current thesis describes research exploring the interaction between the positive and negative affective dynamics and the role of genes in the experience of positive affect dynamics in the flow of daily life. This part will report on experience sampling studies. More specifically, the association between positive affect and negative affect and the association between reward sensitivity and stress sensitivity are investigated (**Chapter 2, Chapter 3**). Furthermore, genetic variation is examined as a source of individual differences in momentary positive affect and reward sensitivity by means of a behaviour genetic twin design (**Chapter 3**) and molecular genetic data (**Chapter 4**).

Experimental designs

While ESM enables the ecologically valid assessment of momentary affect and appraisals as they occur in the flow in daily life, it does not allow for investigating the association between outside stimuli and both appraisal and affect. This is because ESM always assesses the subjective *appraisal* of outside stimuli and cannot assess or control for these outside stimuli per se. In order to understand affect dynamics in relation to such stimuli, experimental designs are implicated. Experimental designs allow that the outside stimulus is controlled and that the effect of this very stimulus on both appraisals and affective states can be examined systematically.

The second part of the current thesis will focus on the role of stress in affect dynamics. This part will report on experimental studies addressing the affective reaction to online social evaluative stress as a risk factor for mood symptoms and the possibility of reducing the negative appraisal of ambiguous situations as an intervention tool for mood symptoms. More specifically, one study reports about the effects of a

controlled online social evaluation stressor on the momentary affective states in adolescents (**Chapter 5**). Furthermore one experimental study aiming at manipulating the interpretation of ambiguous situations in adolescents (**Chapter 6**) and one meta-analytic study on the effectiveness of such manipulation trainings in adults (**Chapter 7**) are reported.

Outline of the current thesis

Part I – Affect dynamics in daily life: interaction and genes

Chapter 2 examines the effects of changes in momentary positive affect on subsequent negative affect as well as the effects of changes in momentary negative affect on subsequent positive affect in the flow of daily life. More specifically, differences in these dynamics between healthy individuals and individuals suffering from depression as well as the predictive value of these affect dynamics for future mood-symptoms are investigated.

Chapter 3 investigates with a behavioural genetic twin design to what degree daily life positive affect and reward sensitivity are influenced by genes, shared, and unshared environmental factors, and explores whether daily life reward sensitivity and stress sensitivity are associated. To the extent these concepts are associated the source of this association is examined.

Chapter 4 investigates whether several candidate genetic variations involved in the brain's reward circuitry play a role in daily life positive affect and reward sensitivity. This is investigated in a general population sample and a sample of individuals suffering from residual depressive symptoms.

Part II – Affect dynamics in response to controlled stressors: social evaluation and manipulation of appraisals

Chapter 5 describes the development and validation of a computerized social interaction paradigm mimicking daily life social interactions between peers in online social networks. Individuals introduced themselves with a short profile and video including information about their interests, hobbies, relationship status, and stance in life. They were also asked to evaluate five peers on their looks, congeniality, and intelligence. Participants were then exposed to manipulated slight negative social evaluation from their potential peers. It is assessed to what degree individuals believed in the tasks authenticity, and whether positive and negative affect, as well as self-esteem and hormonal stress-levels (cortisol) are affected. It is furthermore examined

whether vulnerable groups, specifically females and individuals with current mood symptoms, show a particularly pronounced reaction.

Chapter 6 and 7 explore the possibility of training individuals to appraise their environments as less stressful. **Chapter 6** reports on the first interpretation bias training paradigm for adolescents as both the importance for peer relationships (Sommerville, 2013) and the prevalence for mood-symptoms (Lewinsohn et al., 1993) are at a peak during this period of life. The effectiveness of this training paradigm in manipulating interpretation style (i.e., stress appraisal) on positive and negative affect is determined. **Chapter 7** addresses the clinical potential of benign interpretation bias trainings in affecting stress appraisals, affect, and stress-reactivity and pin down the factors that increase these trainings' effectiveness. To this end a meta-analysis combining all available experimental data on interpretation bias modification trainings is conducted. Several training factors that significantly increase trainings' effectiveness are identified and the question of whether vulnerable groups can particularly benefit from these trainings is addressed.

References

- APA. (2000). *Quick reference to the diagnostic criteria from DSM-IV-TR*. Arlington, VA: American Psychiatric Association.
- Baker, L. A., Cesa, I. L., Gatz, M., & Mellins, C. (1992). Genetic and environmental influences on positive and negative affect: support for a two-factor theory. *Psychology and Ageing, 7*(1), 158-163.
- Beck, A. T. (1970). *Depression. Causes and Treatment*. Philadelphia: University of Pennsylvania Press.
- Beck, A. T., Emery, G., & Greenberg, R. (1985). *Anxiety Disorders and Phobias. A Cognitive Perspective*. New York: Basic Books.
- Berry, D. S., & Hansen, J. S. (1996). Positive affect, negative affect, and social interaction. *Journal of Personality and Social Psychology, 71*(4), 796-809.
- Booij, S. H., Bouma, E. M. C., de Jonge, P., Ormel, J., & Oldehinkel, A. J. (2012). Chronicity of depressive problems and the cortisol response to psychosocial stress in adolescents: The TRAILS study. *Psychoneuroendocrinology*. doi: DOI:10.1016/j.psyneuen.2012.08.004
- Cohen, S., & Pressman, S. D. (2006). Positive affect and health. *Current directions in psychological science, 15*(3), 122-125.
- Cohn, M. A., Fredrickson, B. L., Brown, S. L., Mikels, J. A., & Conway, A. M. (2009). Happiness unpacked: positive emotions increase life satisfaction by building resilience. *Emotion, 9*(3), 361-368.
- Csikszentmihalyi, M., & Hunter, J. (2003). Happiness in everyday life: the use of experience sampling. *Journal of Happiness Studies, 4*, 185-199.
- Cuellar, A. K., & Johnson, S. L. (2009). Depressive symptoms and affective reactivity to maternal praise and criticism. *Journal of Social and Clinical Psychology, 28*(9), 1173-1194.
- Cummins, R. A. (2010). Subjective wellbeing, homeostatically protected mood and depression: a synthesis. *Journal of Happiness Studies, 11*, 1-17.
- Delespaul, P. (1995). *Assessing schizophrenia in daily life. The experience sampling method*. Maastricht: IPSE Foundation.
- Dickerson, S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*(3), 355-391.
- Diener, E. (1994). Assessing subjective well-being: progress and opportunities. *Social Indicators Research, 31*, 103-157.
- Donnerstein, E. (2012). Internet bullying. *Pediatric Clinics of North America, 59*(3), 623-633.
- Eley, T. C., Gregory, A. M., Lau, J. Y. F., McGuffin, P., Napolitano, M., Rijdsdijk, F., & Clark, D. M. (2008). In the face of uncertainty: a twin study of ambiguous information, anxiety and depression in children. *Journal of Abnormal Child Psychology, 36*, 55-65.
- Fox, E., & Standage, H. (2012). Variation on the serotonin transporter gene and bias in the interpretation of ambiguity. *Journal of Cognitive Psychology, 24*(1), 106-114.
- Fredrickson, B. L. (1998). What good are positive emotions? *Review of General Psychology, 2*(3).
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*(3), 218-226.
- Fredrickson, B. L., & Levenson, R. W. (1998). Positive emotions speed recovery from the cardiovascular sequelae of negative emotions. *Cognition and Emotion, 12*(2), 191-220.
- Garland, E. L., Frederickson, B. L., Kring, A. M., Johnson, D. P., Meyer, P. S., & Penn, D. L. (2010). Upward spirals of positive emotions counter downward spirals of negativity: insights from the broaden/and/build theory and affective neuroscience on the treatment of emotion dysfunctions and deficits in psychopathology. *Clinical Psychology Review, 30*, 849-864.
- Garland, E. L., Gaylord, S. A., & Frederickson, B. L. (2011). Positive reappraisal mediates the stress-reductive effects of mindfulness: an upward spiral process. *Mindfulness, 2*(59-67).
- Garland, E. L., Gaylord, S. A., & Park, J. (2009). The role of mindfulness in positive reappraisal. *Explore, 5*(1), 37-44.
- Geschwind, N., Peeters, F., Jacobs, N., Delespaul, P., Derom, C., Thiery, E., . . . Wichers, M. (2010). Meeting risk with resilience: high daily life reward experience preserves mental health. *Acta Psychiatrica Scandinavica, 122*(129-138).
- Hammen, C. (2005). Stress and depression. *Annual Review of Clinical Psychology, 1*, 293-319.
- Jacobs, N., Menne-Lothmann, C., Derom, C., Thiery, E., van Os, J., & Wichers, M. (2012). Deconstructing the familiarity of variability in momentary negative and positive affect. *Acta Psychiatrica Scandinavica, 1-10*.
- Jacobs, N., Nicolson, N. A., Derom, C., Delespaul, P., van Os, J., & Myin-Germeys, I. (2005). Electronic monitoring of salivary cortisol sampling compliance in daily life. *Life Sciences, 76*, 2431-2443.
- Jacobs, N., Rijdsdijk, F., Derom, C., Vlietinck, R., Delespaul, P., van Os, J., & Myin-Germeys, I. (2005). Genes making one feel blue in the flow of daily life: a momentary assessment study of gene-stress interaction. *Psychosomatic Medicine, 68*, 201-206.
- King, L. A., Hicks, J. A., Krull, J. L., & Del Gaiso, A. K. (2006). Positive affect and the experience of meaning in life. *Journal of Personality and Social Psychology, 90*(1), 179-196.
- Koster, E.H.W., Fox, E., & MacLeod, C. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *Journal of Abnormal Psychology, 118*(1), 1-4.
- Kuppens, P., Realo, A., & Diener, E. (2008). The role of positive and negative emotions in life satisfaction judgment across nations. *Journal of Personality and Social Psychology, 95*(1), 66-75.
- Kuppens, P., & Tong, E. M. W. (2010). An appraisal account of individual differences in emotional experience. *Social and Personality Compass, 4*(1138-1150).
- Lau, J. Y. F., & Eley, T. C. (2008). Attributional style as a risk marker of genetic effects for adolescent depressive symptoms. *Journal of Abnormal Psychology, 117*(4), 849-859.
- Lau, J. Y. F., Rijdsdijk, F., & Eley, T. C. (2006). I think, therefore I am: a twins study of attributional style in adolescents. *Journal of Child Psychology and Psychiatry, 47*(7), 696-703.
- Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & Andrews, J. A. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology, 102*(133-144).
- Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: does happiness lead to success? *Psychological Medicine, 131*(6), 803-855.
- MacLeod, C., & Cohen, L. (1993). Anxiety and the interpretation of ambiguity: a text comprehension study. *Journal of Abnormal Psychology, 102*, 238-247.
- Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology, 109*, 62-615.
- Myin-Germeys, I., Peeters, F., Havermans, R., Nicolson, N. A., De Vries, M., Delespaul, P., & van Os, J. (2003). Emotional reactivity to daily life stress in psychosis and affective disorder: an experience sampling study. *Acta Psychiatrica Scandinavica, 107*, 124-131.
- Nes, B. R. (2010). Happiness in behaviour genetics: findings and implications. *Journal of Happiness Studies, 11*, 369-381.
- Nes, B. R., Czajkowski, N., & Tambs, K. (2010). Family matters: happiness in nuclear families and twins. *Behavioral Genetics, 40*, 577-590.

- Peerbooms, O., van Os, J., Drukker, M., Kenis, G., Hoogveld, L., Group, M. i. P., . . . Rutten, B. (2011). Meta-analysis on MTFHR gene variants in schizophrenia, bipolar disorder and unipolar depressive disorder: evidence for a common genetic vulnerability. *Brain, Behavior, and Immunity, 25*, 1530-1543.
- Pempek, T. A., Yermolayeva, Y. A., & Calvert, S. L. (2009). College students' social networking experiences. *Journal of Applied Developmental Psychology, 30*(227-238).
- Rao, U. (2010). Comorbidity between depressive and addictive disorders in adolescents: Role of stress and hpa activity. *US Psyc, 1*(3), 39-43.
- Rietschel, M., Mattheisen, M., Frank, J., Treutlein, J., Degenhardt, F., Breuer, R., . . . Cichon, S. (2010). Genome-wide association-, replication-, and neuroimaging study implicates HOMER1 in the etiology of major depression. *Biological Psychiatry, 68*, 578-585.
- Risch, N., Herrell, R., Lehner, T., KLiang, K., Eaves, L., Hoh, J., . . . Merikangas, K. R. (2009). Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression *Journal of the American Medical Association, 301*(23), 2462-2471.
- Roseman, I., & Edvokas, A. (2004). Appraisals cause experienced emotions: experimental evidence. *Cognition and Emotion, 18*(1), 1-28.
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*(1), 1-32.
- Shmotkin, D. (2005). Happiness in the face of adversity: reformulating the dynamic and modular bases of subjective well-being. *Review of General Psychology, 9*(4), 291-325.
- Smith, A. C., & Kirby, L. D. (2009). Putting appraisals in context: toward a relational model of appraisal and emotion. *Cognition and Emotion, 23*(7), 1352-1372.
- Sommerville, L. H. (2013). The teenage brain: sensitivity to social evaluation. *Psychological Science, 22*(2), 121-127.
- Strout, L. R., Salovey, P., & Epel, E. S. (2002). Sex differences in stress responses: social rejection versus achievement stress. *Biological Psychiatry, 52*, 318-327.
- Tomyn, A. J., & Cummins, R. A. (2011). Subjective wellbeing and homeostatically protected mood: theory validation with adolescents. *Journal of Happiness Studies, 12*, 897-914.
- Wichers, M. (2013). The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine, 1*-12.
- Wichers, M., Aguilera, M., Kenis, G., Krabbendam, L., Myin-Germeys, I., Jacobs, N., . . . van Os, J. (2008). The catechol-O-methyl transferase val158met polymorphism and experience of reward in the flow of daily life. *Neuropsychopharmacology, 33*, 3030-3036.
- Wichers, M., Barge-Schaapveld, D. Q. C. M., Nicolson, N. A., Peeters, F., de Vries, M., Mengelers, R., & van Os, J. (2009). Reduced stress-sensitivity or increased reward experience: the psychological mechanism of response to antidepressant medication. *Neuropsychopharmacology, 34*, 923-931.
- Wichers, M., Myin-Germeys, I., Jacobs, N., Peeters, F., Kenis, G., Derom, C., . . . van Os, J. (2007a). Evidence that moment-to-moment variation in positive emotions buffer genetic risk for depression: a momentary assessment twin study. *Acta Psychiatrica Scandinavica, 115*, 451-457.
- Wichers, M., Myin-Germeys, I., Jacobs, N., Peeters, F., Kenis, G., Derom, C., . . . van Os, J. (2007b). Genetic risk for depression and stress-induced negative affect in daily life. *British Journal of Psychiatry, 191*, 218-223.
- Williams, K., Cheing, C., & Choi, W. (2000). Cyberostracism: effects of being ignored over the internet. *Journal of Personality and Social Psychology, 79*, 748-762.

The dynamic interplay between negative and positive emotions in daily life predicts response to treatment in depression: a momentary assessment study

Wichers, M.¹; Lothmann, C.¹; Simons, C.J.P.^{1,2}; Nicolson, N.A.¹; and Peeters, F.¹

¹ Department of Psychiatry and Psychology, South Limburg Mental Health Research and Teaching Network, European Graduate School of Neurosciences (EURON), Maastricht University, The Netherlands

² GGzE, Mental Health Care Institution, The Netherlands

British Journal of Clinical Psychology 2011, 51 (2), 206-222



Abstract

Objectives: Although treatment of depressive illness aims to restore the imbalance between an excess of negative affect (NA) and a shortage of positive affect (PA), no study has examined how NA and PA may influence each other in depression. This study examines how NA and PA dynamically influence each other in depression and how this may impact on treatment response.

Design: Depressed help-seeking individuals participated in the Experience Sampling Method (ESM), which enables visualisation of subtle dynamic alterations of momentary affective states over time. Thereafter participants received a combination of antidepressant treatment and psychotherapy, and were followed-up each month.

Methods: NA and PA were assessed during ESM at 10 random moments per day for six days. Depressive symptoms were assessed at baseline and at monthly intervals during treatment.

Results: Future response to treatment was associated with altered baseline NA-PA dynamics in individuals with previous depressive episodes. Their daily life boosts of PA were followed by a stronger suppression of NA over subsequent hours than in other depressed groups or controls.

Conclusions: Subtle individual differences in daily life emotional dynamics predict future treatment outcome in depression.

Depressive illness is characterized by sizeable and unexplained variability in treatment response (Fournier, et al., 2009; Kemp, Gordon, Rush, & Williams, 2008). Little is known about individual differences in emotional resilience that may promote recovery in the face of mental illness. Because depression is a complex behavioural phenotype, the processes underlying its development and course can be unravelled and understood only when studies are sensitive to the subtle and dynamic nature of these processes (Wichers, Geschwind, van Os, & Peeters, 2010). There is accumulating evidence that ongoing subtle differences in the way people deal with everyday life experiences may impact on both vulnerability and resilience with respect to future course of symptoms. Intensity and variability or instability in NA and PA (Kovacs, Rottenberg, & George, 2009; Olino, et al., 2010; Thompson, Berenbaum, & Bredemeier, 2010; Wichers, Peeters, et al., 2010) may provide information on risk for future depressive symptomatology. Also, subtle increases in negative affect (NA), in response to minor daily life stressors, reflect genetic vulnerability and predict future depression (Wichers, et al., 2009; Wichers, Myin-Germeys, et al., 2007), while the ability to generate higher levels of positive affect (PA) in response to minor daily life rewards may prevent future symptoms in the face of adversity (Geschwind, et al., 2010). These subtle dynamic changes in response to changing environments are relevant as these effects impact cumulatively on overall experience of NA or PA in daily life. Because of the frequent and repetitive nature of these experiences, small individual differences may nevertheless have a sizeable impact on the way the future course of symptoms unfolds.

Watson and Clark previously identified a two-factor structure underlying the experience of affective states. A factor analysis indicated that although PA and NA are moderately correlated (around -0.30) PA and NA in fact reflect separate dimensions of affect and do not represent opposite ends of a single continuum (Watson & Clark, 1997). A more recent study replicated the model with NA and PA as separate, but correlated, factors in a large representative population sample (Crawford & Henry, 2004). It has been suggested that PA and NA represent the subjective components of underlying biobehavioural systems regulating approach and withdrawal (Watson, Wiese, Vaidya, & Tellegen, 1999), and thus have different biological underpinnings and function. Clinical depression is characterised by a strong imbalance in emotional experience, with high levels of NA and low levels of PA (Watson, et al., 1995), compared to healthy individuals. Although an increasing number of studies suggest that NA and PA are closely linked to mechanisms of vulnerability and resilience (Trute, Benzie, Worthington, Reddon, & Moore, 2010; Tugade & Fredrickson, 2004; Wichers, Peeters, et al., 2010; Wichers, et al., 2009; Zautra, Affleck, Tennen, Reich, & Davis, 2005), very little is known about the way in which these emotions dynamically influence each other over time. Since alterations in PA and NA in major depressive disorder go hand in hand and since it is

known that sudden changes in NA will influence the course of PA to a degree, as well as the other way around, there is a need to study the dynamic relationship between NA and PA. Individual differences within the NA-PA dynamic interplay - such as stronger reductions in NA following boosts in PA, or the preservation of PA experience in the face of increased NA - may be especially relevant and predictive for the process of recovery. To date, studies examining response to treatment in depression have not had access to research designs that are able to shed light on the ongoing dynamic and subtle emotional processes and their associations with recovery. Moreover, to our knowledge, only one study (Finan, Zautra, & Davis, 2009) has prospectively examined the dynamic interplay between NA and PA over time. This was a study in fibromyalgia and osteoarthritis patients. Although prospective, measurements were performed only once daily (at the end of the day) and therefore were not sensitive to immediate subtle fluctuations in affect embedded in daily life NA-PA dynamics. The current study, therefore, examined prospectively the emotional dynamics between NA and PA in depressive illness, and how individual differences therein may predict response to treatment.

Momentary assessment technology, such as the Experience Sampling Method (ESM) (Csikszentmihalyi & Larson, 1987), allows for prospective, repeated sampling of experiences in the flow of daily life. Unlike cross-sectional questionnaire data, ESM enables the visualisation of subtle dynamic alterations in momentary affective states (Myin-Germeys, et al., 2009; Trull & Ebner-Priemer, 2009). In the current study, ESM was used to examine the course of PA before and following changes in NA and, vice versa, the course of NA before and following changes in PA, in relation to depression and recovery of depression over time. In a prospective comparison of future non-responders and future responders to treatment, as well as a reference group of healthy controls, it was hypothesized that more favourable emotional dynamics, i.e. an increased tendency to keep PA high and NA low in response to positive or negative affective changes, would be associated with a more favourable future course of depressive symptomatology.

Method

Subjects

Forty-seven consecutive participants diagnosed with Major Depressive Disorder (MDD) and meeting inclusion criteria were recruited among individuals seeking treatment at a university-affiliated outpatient treatment center in the city of Maastricht, The Netherlands. Inclusion criteria were (i) between 18-65 years, (ii) having 17-item Hamilton Depression Rating Scale (HDRS) scores of ≥ 18 (indicating moderate illness)

and (iii) no antidepressant treatment at study entry. Thirty-nine healthy participants, frequency-matched to the patient sample on gender and age, were also recruited as reference group. For more details, see Peeters et al (Peeters, Nicolson, Berkhof, Delespaul, & deVries, 2003). The study was approved by the Medical Ethics Committee and written informed consent was obtained.

Study design

An observational longitudinal study was conducted. In the patient group, ESM and HDRS were administered at baseline; controls also underwent ESM. After baseline, patients immediately received a combination of treatment with an antidepressant and supportive psychotherapy (Peeters, Berkhof, Rottenberg, & Nicolson, 2010). In case of non-response, subsequent antidepressant medications were prescribed following national and international guidelines (APA, 2000). Patients were assessed with HDRS again in telephone interviews each month for six months following the start of treatment.

Experience Sampling Method

ESM is a structured diary technique to assess subjects' experiences in their daily living environment (Csikszentmihalyi & Larson, 1987; Delespaul, 1995; DeVries, 1992; Myin-Germeys, et al., 2003; Peeters, et al., 2003). Subjects received a digital wristwatch and a set of ESM self-assessment forms collated in a booklet for each day. The wristwatch was programmed to emit a signal ("beep") at an unpredictable moment in each of ten 90-minute time blocks between 7:30 and 22:30, on six consecutive days. After each beep, subjects were asked to fill out the ESM self-assessment forms to record current mood state, context (activity, persons present, and location) and appraisals of the current situation. All self-assessments were rated on 7-point Likert scales. Trained research assistants with ample experience in momentary assessment technology explained the ESM procedure to the participants during an initial briefing session and a practice form was completed to confirm that subjects were able to understand the 7-point Likert scale. Subjects could call a telephone number in case they had questions or problems during the ESM sampling period. Subjects were instructed to complete their reports immediately after the beep, thus minimizing memory distortion, and to record the time at which they completed the form. Previous research has shown that measures of individuals with less than 30% completed reports are less reliable (Delespaul, 1995) and that reports become less reliable when the retrospective window becomes too large (Delespaul, 1995). Therefore, the criterion for inclusion of data in the analyses was a completion of more than 30% of the ESM reports within 25 minutes after the programmed time of the beep (Peeters, et al., 2003). In a previous

study Jacobs and colleagues verified, using an electronic monitoring device of which participants were unaware, that 81% of the ESM reports completed were adherent to the protocol and that inclusion of non-adherent reports did not distort the data (Jacobs, et al., 2005; Peeters, et al., 2003).

Measurements

Momentary affect was assessed with a range of mood adjectives using ESM (see previous paragraph for a description of this procedure) and were rated on 7-point Likert scales. The choice for particular ESM mood adjectives was guided in part by the PANAS questionnaire (Watson, Clark, & Tellegen, 1988) and in part by the results of previous ESM studies (revealing items with high loadings on NA and PA latent factors and substantial within-person variability). Since we hypothesized that NA and PA are separate but correlated latent factors (Crawford & Henry, 2004) we used principal component analysis with oblique rotation. This resulted in two factors (with eigenvalues of 8.5 and 1.9, respectively) explaining 69% of the variance. In order to check whether this structure extrapolated also to the within-subject level, a second factor analysis on within-subject z-scores was performed, which confirmed the existence of a similar factor structure. The ESM mood adjectives ‘cheerful’, ‘satisfied’, ‘enthusiastic’, ‘strong’ and ‘happy’ loaded on the PA factor while ‘gloomy’, ‘tense’, ‘insecure’, ‘lonely’, ‘irritated’, ‘anxious’, ‘guilty’, ‘hurried’ ‘tired’ and ‘edgy’ loaded on the NA factor. Reliability estimates (as indicated by the within-person correlations calculated from the variance components of a null multilevel model) for the NA and PA scales were 0.74 and 0.82, respectively.

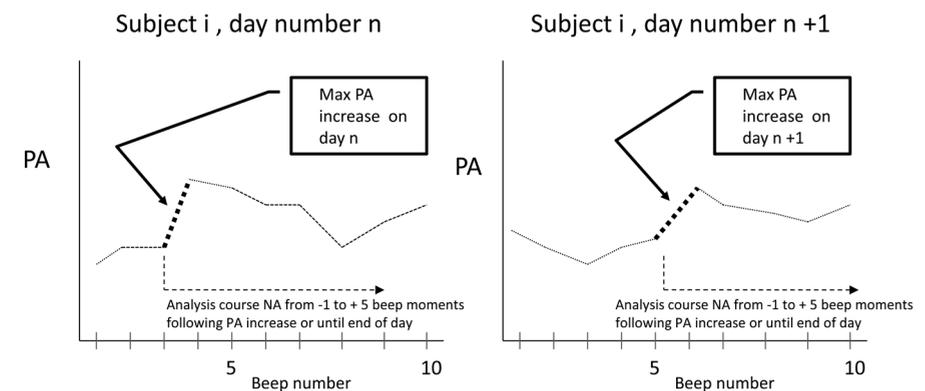
The HDRS is a clinical rating scale for the assessment of the severity of depressive symptoms. This study used the HDRS-17 in which 17 items contribute to the total depression score (Hamilton, 1960). Response to treatment was defined as a 50% reduction in HDRS scores (Kim, et al., 2006). Response was assessed at week 8, the timing conform to other work on response to treatment (Brent, et al., 2008; Choi, Kang, Lim, Oh, & Lee, 2006; Papakostas, et al., 2007). Furthermore, a variable was created representing the amount of residual depressive symptomatology over the six-month follow-up.

Analyses

Since ESM measurements of negative and positive affect were available for six days with a maximum of 10 measurements per day per person with on average 90 minutes in between, the data were ideally suited to explore the dynamic within-subject temporal associations between the two emotional states. To this end, a prospective analysis was

carried out of (i) how negative emotions changed in response to within-subject increases or decreases in positive emotions in daily life, and (ii) how positive emotions changed in response to increases or decreases in negative emotions. Within-subject increase in NA or PA was defined as reporting higher levels of negative or positive emotions, respectively, compared to the beep before. Likewise, within-subject decrease in NA or PA was defined as reporting lower levels of negative or positive emotions, respectively, compared to the beep before. Thus, four types of NA-PA dynamics were examined: (i) NA following increase in PA, (ii) NA following decrease in PA, (iii) PA following increase in NA and (iv) PA following decrease in NA. Below, the procedure of analysis is explained for the NA-PA dynamics as described under (i) above; the other three types were examined in the same way. Thus, for the examination of NA fluctuations following increase in PA, all the beep moments with the maximum within-subject increase in PA of the day were selected. Only the moments of the *maximum* increase of the day were selected in order to prevent overlap between periods examined for subsequent affective changes in NA. Figure 1 shows an example of the selection of moments with the daily maximum increases in PA. Using this method, each subject accumulated, over the course of the 6-day ESM period, 6 beep moments (1 for each day) defined as the daily moment of maximum increase in PA. For these 6 beep moments of largest daily increase in PA, the course of NA was examined starting from one beep before to five beeps following the moment of maximum increase in PA (in the same day). For the purpose of analysis, a variable ‘time’ was created relating to the time period from beep moment -1 to beep moment +5, with the beep moment ‘minus one’ (the beep before the observed increase in PA) as reference category. The same procedure was followed for the other three types of NA-PA dynamics.

Figure 1. Example of selection of ‘maximum increase in PA’ moments on two out of the six days that were available for subject *i*. The selected moments were used in the analysis to examine pre and post PA increase changes in NA. Similar selections were performed for the other three types of NA-PA dynamics.



STATA 11.0 (StataCorp., 2009) was used to conduct multilevel random regression analyses, in which observations were clustered within individuals. PA or NA was regressed on the interactions between group (healthy controls, future treatment non-responders and future treatment responders) x time for the four types of NA-PA dynamics, with time modelled as a continuous variable. Since there is evidence that experience of previous episodes of depression may affect treatment response (Lekman, et al., 2008; Portella, et al., 2009), separate groups were constructed for (non)-responders with a first episode of depression and those with a recurrent episode. These groups hereafter will be referred to as ‘first-episode group’ and ‘recurrent-episode group’. In addition to the group x time interaction based on the outcome of future dichotomous response status, interaction with time was also examined for the outcome of future continuous level of depressive symptoms over the 6 months following start of treatment (follow-up level of depression x time interaction). Significant interactions were followed up with analyses using time as a dummy variable to examine at what specific time points groups significantly differed from each other. The mathematical coupling issue that arises – i.e., that it is not possible to control for unequal levels of NA and PA at the moment before the affective change in a non-randomized setting – was resolved by adding the random effect of time in the regression model and by centring the time variable as described by Blance and colleagues (Blance, Tu, & Gilthorpe, 2005). The statistical model is given below. In this model ‘time’ represents, for example, the distance in time from the moment of the maximum daily increase in PA, and ‘status’ represents responder group status or average level of follow-up HDRS scores. At the deep level the i -th affect score of subject j is modelled as follows: $Affect_{ij} = \beta_0 + \beta_1 status1_j + \beta_2 status2 + \beta_3 status3 + \beta_4 time_{ij} + \beta_5 status1_j * time_{ij} + \beta_6 group\ status2_j * time_{ij} + \beta_7 group\ status3_j * time_{ij} + \zeta_{1j} + \zeta_{2j} * time_{ij} + \epsilon_{ij}$. Here ζ_{1j} represents the subject’s deviation from the overall mean (random intercept) and $\zeta_{2j} * time_{ij}$ represents the subject’s deviation from the overall slope of time (random slope), indicating that affect change over time was allowed to vary randomly between subjects. Note that group status is a dummy variable with status0 as the reference category and time is modeled as a continuous variable.

Results

Compliance was generally good. Only one depressed patient was excluded because of insufficient valid ESM reports within the defined time window. The final sample consisted of 46 patients (20 men and 26 women) and 39 control subjects (16 men and 23 women). Average age in years was 40 (SD=11) and 44 (SD=12) for patients and controls, respectively. Three patients had missing data for response to treatment at week 8. Twenty-five patients had a first-episode depression, of which 12 were classified

as responders to treatment. Eighteen patients suffered a recurrent episode of which 9 patients were classified as responder. Additional demographic information can be found in Peeters et al. (Peeters, Berkhof, Delespaul, Rottenberg, & Nicolson, 2006). Gender did not significantly differ between responder groups ($\chi^2(4)=0.89, p=0.927$). The first-episode responders were significantly younger than the other patient groups and control group ($B=-9.7, p=0.011$). Participants had an average of 50.7 valid responses (85%) (Peeters, et al., 2003). At the level of ESM momentary assessments the overall correlation between NA and PA was -0.27; stratified at the level of responders and non-responders the correlations were -0.33 and -0.28, respectively. At the level of the individual (NA and PA averaged per subject), these correlations were -0.25, -0.37 and -0.27, respectively. Figure 2 shows the HDRS scores per responder group over the six month follow-up period. Baseline pre-treatment HDRS scores did not differ significantly ($F=0.62, df=2, p=0.54$) between the responder groups. The recurrent-episode participants who at follow-up were non-responders had the lowest baseline HDRS score (23.1) and the first-episode participants who at follow-up were responders the highest (25.0). Depressed patients showed significantly higher baseline average NA ($F=22.1, df=3, p<0.001$) and lower PA ($F=69.2, df=3, p<0.001$) than healthy controls; baseline NA and PA did not significantly differ between the different types of responders (responders/non-responders with either a first or a recurrent episode; NA: $F=0.19, df=2, p=0.83$; PA: $F=0.70, df=2, p=0.50$). Table 1a shows the mean baseline HDRS, NA and PA levels and the average daily maximum of emotional change (\uparrow PA, \downarrow PA, \uparrow NA, \downarrow NA) corresponding to the four types of baseline NA-PA dynamics, stratified by follow-up responder groups. Table 1b shows the same information, but now stratified by the level of post-treatment follow-up residual depressive symptoms (table 1b). These descriptive results show that the first-episode future responders and controls exhibited similar average increases in PA for the selected moments of maximum daily increase in PA, whereas the other groups showed lower daily maximum increases in PA (see table 1a for indications of statistical significance). In addition, the recurrent-episode future responder group was the only group showing significantly lower average maximum decreases in PA compared to healthy controls and compared to the first-episode future responders. While striking differences in emotional patterns were evident for the recurrent-episode and first-episode responder groups, recurrent-episode and first-episode non-responder groups showed very similar averages for all indicators of affect (see table 1a). In order to maximize statistical power, both non-responder groups were therefore merged together for the remainder of the analyses. The group with high levels of average residual symptoms over 6 months showed a tendency for lower average of maximum *change* in NA, but higher average *absolute* levels of NA compared to the groups with low and middle levels of residual symptoms

(see table 1b). The number of subsequent measurement occasions -starting from the defined moment of maximum daily change in affect- will decrease when maximum daily changes occur nearer the end of the day. Table 2 shows information on the amount of data included in the analyses separately for each measurement occasion.

Figure 2. HDRS scores over time from baseline to 6 months following the start of treatment. The black lines represent non-responder groups; the grey lines represent the responder groups; the solid lines represent first episode patients and the dotted lines represent patients of whom this episode is not the first one.

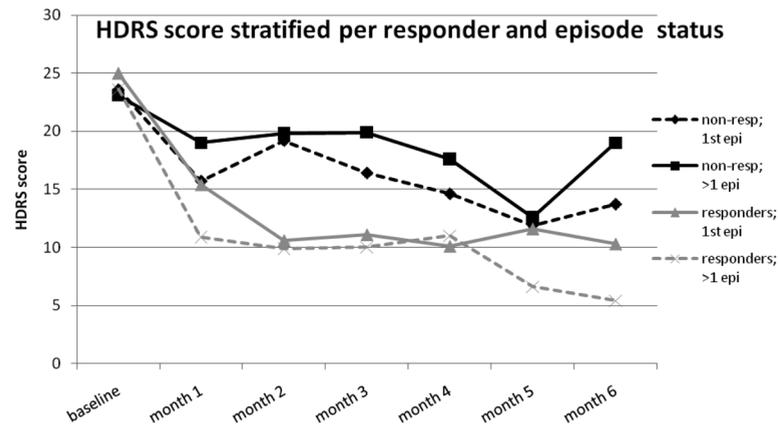


Table 1a. Characteristics of depressive symptoms, average PA and NA, and the four average maximum daily changes in NA or PA corresponding to the four selections (PA increase, PA decrease, NA increase and NA decrease), stratified for responder groups.

	HDRS	Average PA	Average NA	Max PA increase	Max PA decrease	Max NA increase	Max NA decrease
Controls	-	0.78	-0.59	0.56	0.50	0.41	0.43
Future non-responders first-episode	23.6	-0.81*	0.52*	0.41 Φ Π	0.42	0.76*	0.79*
Future non-responders >1 episode	23.1	-0.87*	0.62*	0.38 * Ψ	0.43	0.74*	0.82*
Future responders first-episode	25.0	-0.72*	0.71*	0.60	0.54	0.84*	0.93*
Future responders > 1 episode	23.6	-0.63*	0.48*	0.40 Φ Π	0.30* Ψ	0.56 Π	0.68 Φ

*sign difference with control group with p<0.05
 Ψ sign difference with first-episode future responder group with p<0.05
 Φ sign difference with control group with p<0.1
 Π sign difference with first-episode future responder group with p<0.1

Table 1b. Characteristics of depressive symptoms, average PA and NA, and the four average maximum daily changes in NA or PA corresponding to the four selections (PA increase, PA decrease, NA increase and NA decrease), stratified for levels of residual depressive symptomatology.

	HDRS	Average PA	Average NA	Max PA increase	Max PA decrease	Max NA increase	Max NA decrease
Controls	-	0.78	-0.59	0.56	0.50	0.41	0.43
Low post-baseline 6-month HDRS	23.3	-0.67*	0.54*	0.48	0.42	0.74*	0.85*
Middle post-baseline 6-month HDRS	23.6	-0.81*	0.45*	0.47	0.46	0.88* Ψ	0.95* Π
High post-baseline 6-month HDRS	24.9	-0.81*	0.81*	0.43	0.43	0.60*	0.71*

*sign difference with control group with p<0.05
 Ψ sign difference with high HDRS group with p<0.05
 Φ sign difference with control group with p<0.1
 Π sign difference with high HDRS group with p<0.1

Table 2 shows the percentages of data included in the analyses per measurement occasion from one beep before to five beeps following the moment of maximal daily change in affect. For example, 8.7% of the data on NA in relation to the maximal daily increase in PA stemmed from moments five beeps following this increase.

	Max PA increase in %	Max PA decrease in %	Max NA increase in %	Max NA decrease in %
One beep before*	20.4	21.5	23.1	21.3
Moment of maximal change	20.4	21.5	23.1	21.3
One beep later	15.5	15.9	15.4	14.9
Two beeps later	13.8	14.0	13.1	13.6
Three beeps later	11.4	10.7	10.8	11.5
Four beeps later	9.9	9.1	8.7	9.7
Five beeps later	8.7	7.2	5.8	7.7
Total %	100	100	100	100

*Percentages at the first (one beep before) and second (moment of maximal change) row are equal because all occasions that have data on affect at that first time point have also data at the second moment. For subsequent moments less data is available: the further in time from the moment of change in affect, the closer to the end of the day where data is no longer collected.

NA-PA dynamics and response to treatment

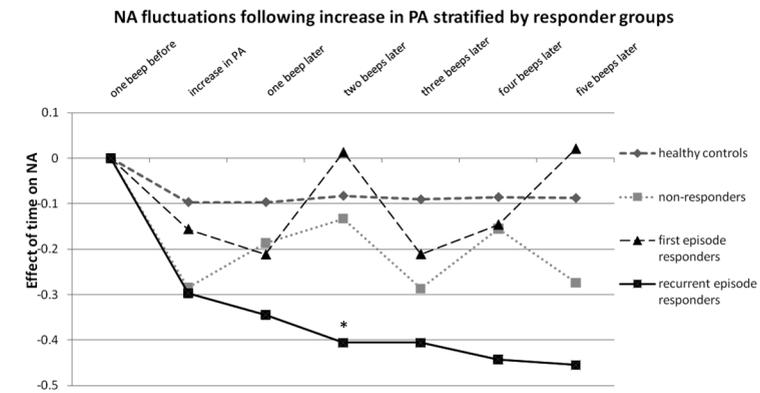
Significant time x group interactions were found with recurrent-episode future responders showing significantly different baseline NA-PA dynamic patterns following increase in PA, compared to the other patient groups. In the recurrent-episode future responder group, daily maximum PA increase resulted in significantly lower levels of subsequent NA over the next few hours than in future non-responders (interaction: $B=-0.046, p=0.045$) and first-episode responders (interaction: $B=-0.067, p=0.009$; see

table 3a). Furthermore, recurrent-episode future responders also showed greater reductions in subsequent NA levels over the few hours following daily maximum PA increase compared to healthy controls (interaction: $B=-0.063, p=0.003$). Figure 3a shows that at the exact moment of PA increase, NA was reduced equally strongly in future non-responders and recurrent-episode future responders. However, in the hours following the maximum daily PA increase, both groups displayed different trajectories of negative emotions over time, the difference between groups becoming significant two beeps (3 hours) following the initial increase in PA ($B=-0.41, p=0.030$). There were no significant differences between the future response status groups for any of the three other types of NA-PA dynamics.

Table 3a. Results of the interaction analyses of time (in beep distance from the moment of the affective change) x responder groups. Significant interaction effects are reported in bold.

NA-PA dynamics	group	Effect size (SE) of interaction group x time (in beep distance)	p-value	Effect size (SE) of interaction group x time (in beep distance)	p-value	Effect size (SE) of interaction group x time (in beep distance)	p-value
		Non-responder group is the reference category		Control group is the reference category		Responder group first episode is the reference category	
NA following PA increase	Controls	0.016 (0.015)	0.267	-	-	-	-
	Responders first episode	0.020 (0.021)	0.319	0.004 (0.019)	0.822	-	-
	Responders >1 episode	-0.046 (0.023)	0.045*	-0.063 (0.021)	0.003*	-0.067 (0.026)	0.009*
NA following PA decrease	Controls	0.006 (0.016)	0.724	-	-	-	-
	Responders first episode	-0.002 (0.021)	0.907	-0.008 (0.019)	0.676	-	-
	Responders >1 episode	-0.032 (0.025)	0.207	-0.037 (0.023)	0.112	-0.029 (0.027)	0.287
PA following NA increase	Controls	0.007 (0.013)	0.601	-	-	-	-
	Responders first episode	0.013 (0.017)	0.442	0.007 (0.015)	0.665	-	-
	Responders >1 episode	0.017 (0.019)	0.373	0.010 (0.017)	0.553	0.004 (0.021)	0.864
PA following NA decrease	Controls	0.015 (0.013)	0.239	-	-	-	-
	Responders first episode	0.006 (0.018)	0.733	-0.009 (0.016)	0.593	-	-
	Responders >1 episode	0.021 (0.019)	0.282	0.006 (0.018)	0.740	0.015 (0.022)	0.504

Figure 3a depicts NA fluctuations following increase in PA, stratified by responder groups with NA levels one beep before the PA increase as reference category. Beeps are approximately 90 minutes apart from each other. * Significant differences between recurrent-episode future responders and non-responders.



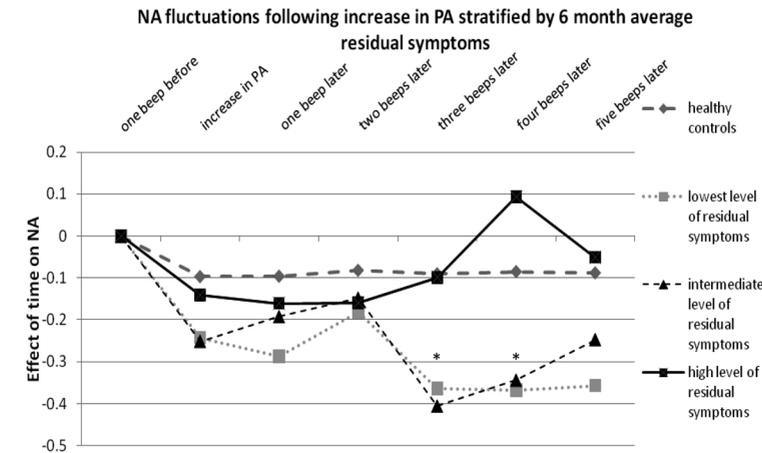
Baseline NA-PA dynamics and follow-up HRDS depressive symptomatology

For the effect of PA increase on subsequent NA levels a significant cross-level interaction was apparent between time (in beep distance from the PA increase) and 6-month follow-up average HDRS scores ($B=0.005, p=0.007$). The analysis was rerun comparing tertile groups of HDRS scores (high, intermediate, and low) and controls. The group with the highest continuous level of depressive symptomatology over the 6 months following start of treatment had a significantly smaller drop in NA following increases in PA than the group with the lowest levels of future symptoms (interaction: $B=0.065, p=0.002$) and the group with intermediate levels of future symptoms (interaction: $B=0.056, p=0.007$; see table 3b and figure 3b). Compared to the group with the lowest average levels of symptoms over the 6-month period following start of treatment, the group with the highest levels started to deviate significantly in their reduction of NA three beeps following the increase in PA ($B=0.26, p=0.05$). Four beeps later the difference remained significant ($B=0.46, p=0.002$) and five beeps later the effect size remained large ($B=0.31, p=0.06$). Compared to the group with intermediate levels of symptoms, similar differences were found (three beeps following increase in PA: $B=0.31, p=0.02$; four beeps later: $B=0.44, p=0.002$). While both in the lowest level ($B=-0.045, p=0.007$) and intermediate level ($B=-0.036, p=0.022$) groups NA was reduced significantly stronger than in healthy controls, this was not the case for the group with the highest level of future depressive symptomatology ($B=0.020, p=0.262$; see table 3b). No significant differences between the groups, including the control group, were found for any other type of NA-PA dynamic associations over time.

Table 3b. Results of the interaction analyses of time (in beep distance from the moment of the affective change) x post-baseline HDRS level groups. Significant interaction effects are reported in bold.

NA-PA dynamics	group	Effect size (SE)	p-value	Effect size (SE)	p-value	Effect size (SE)	p-value
		of interaction group x time (in beep distance)		(SE) of interaction group x time (in beep distance)		of interaction group x time (in beep distance)	
		Low post-baseline 6-month HDRS score group is the reference category		Control group is the reference category		Medium post-baseline 6-month HDRS score group is the reference category	
NA following PA increase	Controls	0.045 (0.017)	0.007*	-	-	-	-
	Medium post-baseline 6-month HDRS score	0.009 (0.020)	0.638	-0.036 (0.016)	0.022*	-	-
	High post-baseline 6-month HDRS score	0.065 (0.022)	0.002*	0.020 (0.018)	0.262	0.056 (0.021)	0.007*
NA following PA decrease	Controls	0.030 (0.018)	0.106	-	-	-	-
	Medium post-baseline 6-month HDRS score	0.014 (0.021)	0.526	-0.016 (0.017)	0.352	-	-
	High post-baseline 6-month HDRS score	0.034 (0.023)	0.141	0.004 (0.019)	0.816	0.020 (0.022)	0.358
PA following NA increase	Controls	-0.009 (0.014)	0.520	-	-	-	-
	Medium post-baseline 6-month HDRS score	-0.005 (0.017)	0.776	0.004 (0.014)	0.748	-	-
	High post-baseline 6-month HDRS score	-0.010 (0.019)	0.603	-0.001 (0.016)	0.917	-0.005 (0.018)	0.787
PA following NA decrease	Controls	0.013 (0.015)	0.391	-	-	-	-
	Medium post-baseline 6-month HDRS score	-0.001 (0.018)	0.971	-0.014 (0.014)	0.335	-	-
	High post-baseline 6-month HDRS score	0.016 (0.019)	0.394	0.003 (0.016)	0.840	0.017 (0.018)	0.355

Figure 3b depicts NA fluctuations following increase in PA, stratified by levels of residual depressive symptomatology with NA levels one beep before the PA increase as reference category. Beeps are approximately 90 minutes apart from each other. * Significant differences between the groups with high and low level of residual symptoms.



Discussion

As far as we are aware, this is the first study in depressed patients examining prospectively the dynamic association between baseline NA and PA in relation to future treatment outcome. Furthermore, NA-PA dynamics were examined using fine-grained moment-to-moment assessments over the day. Results showed that a more favourable future course of depression was associated with stronger reductions of NA following increase in PA over the course of the day. This result is consistent with previous studies reporting that positive experiences (here: moments of increase in PA) – more than absence of negative experiences- are associated with a favourable future course of illness (Geschwind, et al. 2011; Olino, et al., 2010; Wichers, et al., 2009). Responders to treatment with a history of depressive episodes, but not first-episode responders, benefitted more from the natural rewards of daily life in terms of their effect on lowering subsequent levels of NA. Although the initial reduction of NA at the moment of increase in PA was similar for both groups, a few hours later the future responder group still benefitted from the positive emotional change, while the future non-responders had already lost most of the impact of the natural reward encountered. The delayed effect suggests that cognitive processes may be involved in the differentiation between the two groups. One possibility is that negative cognitive schemas in non-responders interfere with the transfer of positive emotions from one moment to the next. Another possibility is that future responders with a history of

depressive episodes have learnt to amplify the effect of their positive emotions in order to compensate their emotional imbalance. This possibility will be further discussed below. The two possible mechanisms imply different strategies to boost recovery, which stresses the urgency to examine these processes in more depth.

The same type of NA-PA dynamics, that is, NA reductions following increases in PA, was associated with a more favourable long-term course of depressive symptomatology (HDRS) measured following start of treatment. People who continued to experience high levels of depressive symptoms during the six-month period following baseline were people in whom PA increases resulted in significantly less reduction of NA than in people belonging to the two other patient groups (intermediate and low residual symptoms). The significant differences found in comparison with control subjects regarding the reduction of NA following increase in PA may be due to a floor effect in healthy controls, who are already at the bottom end of the NA distribution. The only study, to our knowledge, that has examined the dynamic association between NA and PA over time (Finan, et al., 2009) found a similar lower reduction of subsequent NA (average daily level) following previous PA experience (average level previous day) in fibromyalgia patients compared to osteoarthritis patients. This finding may be related to the increased levels of depressive symptoms found in fibromyalgia patients compared to other patient groups with somatic disorders (Thieme, Turk, & Flor, 2004).

Absence of defective NA-PA dynamics

Although most depression responder groups showed a defective ability to generate boosts of PA, it is interesting to note that the dynamic associations between NA and PA over time did not seem to be defective in depressed patients. Depressed patients seemed to be similar to healthy controls in their emotional responses following *both increases and decreases* in NA and *decreases* in PA experienced in daily life. Whether the increased reduction of NA following increase in PA in recurrent-episode future responders is indicative of any particular strength of these patients or defective emotion regulation in the other patients remains undetermined. The superior ability, compared to healthy controls, in the recurrent-episode future responder group and the intermediate and low depressive symptom level groups, likely reflects the absence of negative emotions to suppress in the control group. Since we do not know how well controls would be able to suppress *high levels* of NA, there is no way to establish whether some patient groups have defective NA-PA dynamics or not. However, since (i) the recurrent-episode future responder group showed signs of a defect in the ability to generate PA at baseline that were not observed in the first-episode future responder group, (ii) the first-episode future responders were similar to healthy controls in their

ability to generate PA and (iii) the first-episode future responder group did not show the enhanced ability to suppress NA with previously acquired PA, the superior ability of the recurrent-episode future responders may more likely reflect a compensatory mechanism to improve the NA-PA imbalance rather than a specific defect in the other patient groups.

Two types of responders

The finding that the first-episode responders did not benefit more from daily life natural rewards than the non-responder groups suggests an apparent contradiction. However, combining the findings from both the descriptive results on NA and PA separately (table 1) with the multilevel analyses regarding the dynamic association between NA and PA may clarify the processes underlying these findings. While the other three patient groups, when their six maximum increases were averaged, clearly showed smaller sized increases in PA compared to healthy controls, the first-episode responder group, surprisingly, showed maximum PA increases that were equal to those observed in the control subjects. First-episode responders thus appear to have a “healthy” tendency to experience boosts of PA in daily life while the other depressed patient groups seem to display alterations in the tendency to experience PA boosts. However, the first-episode responders also experienced levels of NA that were higher than those observed for any other group, while the recurrent-episode responders displayed the lowest NA levels. The hypothesis that depression arises due to an imbalance between negative and positive emotions indicates that both *too much NA* as well as *too little PA* may cause depression (Wichers, Jacobs, Derom, Thiery, & Van Os, 2007). According to the current findings, the emotional imbalance in first-episode responders may be shifted towards strongly increased levels of NA, overwhelming the levels of positive emotions, resulting in clinical depression. Since in first-episode responders the system to generate PA still remains intact, the depression may be hypothesized to resolve as soon as increased levels of NA drop, allowing the healthy system of generating PA to kick in and compensate for any remaining levels of NA with newly acquired positive emotions. The recurrent-episode future responder group, however, was similar to both non-responder groups in terms of defective generation of strong boosts of PA. Nevertheless, the results in table 1a show that the recurrent-episode future responder group exhibited lower average decreases in PA compared to both controls and first-episode responders, which suggests that in this group, positive emotions were not lost as easily as in the other groups. Furthermore, the analyses showed that the recurrent-episode responder group profited more from the acquired PA, as evidenced by more effective post-PA reduction of negative emotions over the course of the day. Therefore, although a defect in the ability to generate PA may have

triggered their depressive episode, the recurrent-episode future responder group also appears to be able to compensate for that defect, unlike the non-responder groups, through the enhanced effect of acquired PA to reduce subsequent NA. In contrast, the first-episode future responders had no need to compensate because their ability to generate PA was not defective in the first place.

The above mechanisms remain speculative and should be regarded with caution until replicated by new studies on NA-PA dynamics in depression in relation to future treatment response.

Clinical implications

HDRS score did not differ between patient groups and thus did not have predictive value regarding future outcome. In contrast, the subtle daily life patterns of negative and positive emotions and their dynamic associations may be much more informative. Knowledge of these dynamic emotional characteristics may help attending therapists to get insight into the psychopathology of a particular patient, thus improving decision making with regard to treatment choice. To this end, a new electronic device with touch screen has been developed (“PsyMate”) in order to conduct automatic experience sampling without the need for pencil and paper and research assistance, opening the possibility for routine monitoring of daily life contexts and emotions in treatment settings (Wichers, et al. 2011). The device automatically stores all data collected by the patient and comes with prefabricated routines to generate relevant output. This procedure will allow therapists to have easy access to the daily life emotional patterns of their patients. In addition to assessment of psychopathology, the data output of the PsyMate device may also help to give patients insight into their emotional patterns and to reshape their daily life routines in a way that will be optimal for their level of positive emotions. Since the current results and previous literature (Nutt, et al., 2007) shows that many depressed patients are likely defective in their ability to generate PA, there is a clinical need to help patients compensate, by facilitating strategies to build contexts and situations that are conducive to the experience of high levels of positive emotions.

Limitations

Because in ESM affective states are followed over the day for as long as seven consecutive beeps, affective states may be confounded by a circadian rhythm of affect. However, when analyses were conducted with time of the day as an additional confounder or an additional interaction term (daytime x time), results remained similar. A second limitation of the current study is the small patient sample size, stratified by responder

status and level of future depressive symptomatology. Low power to perform between-group analyses may have caused type II error, given that relatively large effect sizes yielded small or non-significant effects. Also, the number of subsequent measurement occasions was lower when maximum daily changes occurred nearer the end of the day, which results in somewhat reduced power to examine differences between groups in the more persistent effects of up to 5 beeps following the maximum affect change. Therefore, we may have missed some relevant differences between the groups. Future research is necessary to replicate the current findings and shed more light on the subtle dynamic characteristics of emotions in relation to treatment response.

Funding

M.C. Wichers was supported by the Dutch Medical Council (VENI grant nr 916.76.147).

References

- APA (2000). Practice guideline for the treatment of patients with major depressive disorder (revision). *American Journal of Psychiatry*, 157(4 suppl), 1-45.
- Blance, A., Tu, Y. K., & Gilthorpe, M. S. (2005). A multilevel modelling solution to mathematical coupling. *Statistical Methods in Medical Research*, 14(6), 553-565.
- Brent, D., Emslie, G., Clarke, G., Wagner, K. D., Asarnow, J. R., Keller, M., et al. (2008). Switching to another SSRI or to venlafaxine with or without cognitive behavioral therapy for adolescents with SSRI-resistant depression: the TORDIA randomized controlled trial. *JAMA*, 299(8), 901-913.
- Choi, M. J., Kang, R. H., Lim, S. W., Oh, K. S., & Lee, M. S. (2006). Brain-derived neurotrophic factor gene polymorphism (Val66Met) and citalopram response in major depressive disorder. *Brain Research*, 1118(1), 176-182.
- Crawford, J. R., & Henry, J. D. (2004). The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 43(Pt 3), 245-265.
- Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *J Nervous and Mental Disease*, 175(9), 526-536.
- Delespaul, P. (1995). *Assessing schizophrenia in daily life: The experience sampling method*. Maastricht: University of Limburg.
- DeVries, M. W. (Ed.). (1992). *The experience of psychopathology: investigating mental disorders in their natural settings*. Cambridge: Cambridge university press.
- Finan, P. H., Zautra, A. J., & Davis, M. C. (2009). Daily affect relations in fibromyalgia patients reveal positive affective disturbance. *Psychosomatic Medicine*, 71(4), 474-482.
- Fournier, J. C., DeRubeis, R. J., Hollon, S. D., Dimidjian, S., Amsterdam, J. D., Shelton, R. C., et al. (2010). Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA*, 303(1), 47-53 doi: 10.1001/jama.2009.1943.
- Fournier, J. C., DeRubeis, R. J., Shelton, R. C., Hollon, S. D., Amsterdam, J. D., & Gallop, R. (2009). Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *Journal of Consulting and Clinical Psychology*, 77(4), 775-787.
- Geschwind, N., Nicolson, N. A., Peeters, F., van Os, J., Barge-Schaapveld, D., & Wichers, M. (2011). Early improvement in positive rather than negative emotion predicts remission from depression after pharmacotherapy. *European Neuropsychopharmacology*, 21(3), 241-7.
- Geschwind, N., Peeters, F., Jacobs, N., Delespaul, P., Derom, C., Thiery, E., et al. (2010). Meeting risk with resilience: high daily life reward experience preserves mental health. *Acta Psychiatrica Scandinavica*, 122(2), 129-138.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology Neurosurgery and Psychiatry*, 23, 56-62.
- Jacobs, N., Nicolson, N. A., Derom, C., Delespaul, P., van Os, J., & Myin-Germeys, I. (2005). Electronic monitoring of salivary cortisol sampling compliance in daily life. *Life Sciences*, 76(21), 2431-2443.
- Kemp, A. H., Gordon, E., Rush, A. J., & Williams, L. M. (2008). Improving the prediction of treatment response in depression: integration of clinical, cognitive, psychophysiological, neuroimaging, and genetic measures. *CNS Spectrums*, 13(12), 1066-1086; quiz 1087-1068.
- Kim, H., Lim, S. W., Kim, S., Kim, J. W., Chang, Y. H., Carroll, B. J., et al. (2006). Monoamine transporter gene polymorphisms and antidepressant response in Koreans with late-life depression. *JAMA*, 296(13), 1609-1618.

- Kovacs, M., Rottenberg, J., & George, C. (2009). Maladaptive mood repair responses distinguish young adults with early-onset depressive disorders and predict future depression outcomes. *Psychological Medicine, 39*(11), 1841-1854.
- Lekman, M., Laje, G., Charney, D., Rush, A. J., Wilson, A. F., Sorant, A. J., et al. (2008). The FKBP5-gene in depression and treatment response--an association study in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Cohort. *Biological Psychiatry, 63*(12), 1103-1110.
- Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: opening the black box of daily life. *Psychological Medicine, 39*(9), 1533-1547.
- Myin-Germeys, I., Peeters, F., Havermans, R., Nicolson, N. A., DeVries, M. W., Delespaul, P., et al. (2003). Emotional reactivity to daily life stress in psychosis and affective disorder: an experience sampling study. *Acta Psychiatrica Scandinavica, 107*(2), 124-131.
- Nutt, D., Demyttenaere, K., Janka, Z., Aarre, T., Bourin, M., Canonico, P. L., et al. (2007). The other face of depression, reduced positive affect: the role of catecholamines in causation and cure. *Journal of Psychopharmacology, 21*(5), 461-471.
- Olino, T. M., Lopez-Duran, N. L., Kovacs, M., George, C. J., Gentzler, A. L., & Shaw, D. S. (2010). Developmental trajectories of positive and negative affect in children at high and low familial risk for depressive disorder. *Journal of Child Psychology and Psychiatry*. doi: 10.1111/j.1469-7610.2010.02331.x
- Papakostas, G. I., Montgomery, S. A., Thase, M. E., Katz, J. R., Krishen, A., & Tucker, V. L. (2007). Comparing the rapidity of response during treatment of major depressive disorder with bupropion and the SSRIs: a pooled survival analysis of 7 double-blind, randomized clinical trials. *J Clinical Psychiatry, 68*(12), 1907-1912.
- Peeters, F., Berkhof, J., Delespaul, P., Rottenberg, J., & Nicolson, N. A. (2006). Diurnal mood variation in major depressive disorder. *Emotion, 6*(3), 383-391.
- Peeters, F., Berkhof, J., Rottenberg, J., & Nicolson, N. A. Ambulatory emotional reactivity to negative daily life events predicts remission from major depressive disorder. *Behaviour Research and Therapy, 48*(8), 754-760.
- Peeters, F., Nicolson, N. A., Berkhof, J., Delespaul, P., & deVries, M. (2003). Effects of daily events on mood states in major depressive disorder. *Journal of Abnormal Psychology, 112*(2), 203-211.
- Portella, M. J., de Diego-Adelino, J., Puigdemont, D., Perez-Egea, R., Alvarez, E., Artigas, F., et al. (2009). Pindolol augmentation enhances response outcomes in first depressive episodes. *European Neuropsychopharmacology, 19*(7), 516-519.
- StataCorp. (2009). Stata Statistical Software: College Station, TX.
- Thieme, K., Turk, D. C., & Flor, H. (2004). Comorbid depression and anxiety in fibromyalgia syndrome: relationship to somatic and psychosocial variables. *Psychosomatic Medicine, 66*(6), 837-844.
- Thompson, R. J., Berenbaum, H., & Bredemeier, K. (2010). Cross-sectional and longitudinal relations between affective instability and depression. *Journal of Affective Disorders*. doi: 10.1016/j.jad.2010.09.021
- Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section. *Psychological Assessment, 21*(4), 457-462.
- Trute, B., Benzies, K. M., Worthington, C., Reddon, J. R., & Moore, M. (2010). Accentuate the positive to mitigate the negative: Mother psychological coping resources and family adjustment in childhood disability. *Journal of Intellectual and Developmental Disability, 35*(1), 36-43.
- Tugade, M. M., & Fredrickson, B. L. (2004). Resilient individuals use positive emotions to bounce back from negative emotional experiences. *Journal of Personality and Social Psychology, 86*(2), 320-333.
- Watson, D., & Clark, L. A. (1997). Measurement and mismeasurement of mood: recurrent and emergent issues. *Journal of Personality Assessment, 68*(2), 267-296.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.
- Watson, D., Clark, L. A., Weber, K., Assenheimer, J. S., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: II. Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. *Journal of Abnormal Psychology, 104*(1), 15-25.
- Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: structural findings, evolutionary considerations and psychobiological evidence. *Journal of Personality and Social Psychology, 76*, 820-838.
- Wichers, M., Geschwind, N., Jacobs, N., Kenis, G., Peeters, F., Derom, C., et al. (2009). Transition from stress sensitivity to a depressive state: longitudinal twin study. *British Journal of Psychiatry, 195*(6), 498-503.
- Wichers, M., Geschwind, N., van Os, J., & Peeters, F. (2010). Scars in depression: is a conceptual shift necessary to solve the puzzle? *Psychological Medicine, 40*(3), 359-365.
- Wichers, M., Hartmann, J. A., Kramer, I. M., Lothmann, C., Peeters, F., van Bemmelen, L., et al. (2011). Translating assessments of the film of daily life into person-tailored feedback interventions in depression. *Acta Psychiatrica Scandinavica*. doi: 10.1111/j.1600-0447.2011.01684.x
- Wichers, M., Jacobs, N., Derom, C., Thiery, E., & Van Os, J. (2007). Depression: too much negative affect or too little positive affect? *Twin Research and human genetics, 10 Supplement*, 19-20.
- Wichers, M., Myin-Germeys, I., Jacobs, N., Peeters, F., Kenis, G., Derom, C., et al. (2007). Genetic risk of depression and stress-induced negative affect in daily life. *British Journal of Psychiatry, 191*, 218-223.
- Wichers, M., Peeters, F., Geschwind, N., Jacobs, N., Simons, C. J., Derom, C., et al. (2010). Unveiling patterns of affective responses in daily life may improve outcome prediction in depression: a momentary assessment study. *Journal of Affective Disorders, 124*(1-2), 191-195.
- Wichers, M. C., Barge-Schaapveld, D. Q., Nicolson, N. A., Peeters, F., de Vries, M., Mengelers, R., et al. (2009). Reduced stress-sensitivity or increased reward experience: the psychological mechanism of response to antidepressant medication. *Neuropsychopharmacology, 34*(4), 923-931.
- Zautra, A. J., Affleck, G. G., Tennen, H., Reich, J. W., & Davis, M. C. (2005). Dynamic approaches to emotions and stress in everyday life: Bolger and Zuckerman reloaded with positive as well as negative affects. *Journal of Personality, 73*(6), 1511-1538.

**Genetic and environmental causes
of individual differences in daily
life positive affect and reward
experience and its overlap with
stress-sensitivity**

Menne-Lothmann, C.¹; Jacobs, N.^{1,2}; Derom, C.³; Evert Thiery⁴; van Os, J.^{1,5}; and Wichers, M.¹

¹Department of Psychiatry and Psychology, South Limburg Mental Health Research and Teaching Network, EURON, Maastricht University, The Netherlands

² Faculty of Psychology, Open University of the Netherlands, Heerlen, The Netherlands

³ Department of Human Genetics, University Hospital Gasthuisberg, Catholic University of Leuven, Leuven, Belgium

⁴ Association for Scientific Research in Multiple Births, Gent, Belgium

⁵ King's College London, King's Health Partners, Department of Psychosis Studies, Institute of Psychiatry, London, UK



Abstract

Momentary positive affect (PA) and reward experience may underlie subjective wellbeing, and index mental health resilience. This study examines their underlying sources of variation and the covariation with stress-sensitivity.

The experience sampling method was used to collect multiple appraisals of mood and daily life events in 520 female twins. Structural equation model fitting was employed to determine sources of variation of PA, reward experience, and the association between reward experience and stress-sensitivity.

PA was best explained by shared and non-shared environmental factors, and reward experience by non-shared environmental factors only, although the evidence was also suggestive of a small genetic contribution. Reward experience and stress-sensitivity showed no association.

PA was not heritable. Most - if not all - variance of reward experience was explained by environmental influences. Stress-sensitivity, indexing depression vulnerability, and reward experience were non-overlapping, suggesting that resilience traits are independent from stress-sensitivity levels in a general population sample.

Keywords positive affect; reward; psychological resilience; experience sampling method; twin study

People scoring high on subjective wellbeing (SWB) have better immune systems, higher incomes, happier marriages and a better mental health (Lyubomirsky, King, & Diener, 2005; Lyubomirsky, Sheldon, & Schkade, 2005). In short: they are more successful in life. Although SWB appears to be moderately determined by genes (Bartels & Boomsma, 2009; Lyubomirsky, Sheldon, et al., 2005; Nes, Czajkowski, & Tambs, 2010), it is unlikely that people are born with a certain level of SWB. Instead, SWB may be the product of various developmental mechanisms. According to Fredrickson's broaden-and-build theory and early supporting evidence, it is *momentary* positive affect (*hereafter*: PA) that causes people to build up personal, intellectual, physical, and social resources that – in the long run – serve as the person's support system, allowing them to generate SWB (Cohn, Fredrickson, Brown, Mikels, & Conway, 2009; Fredrickson, 2001). To further understand the determinants and mechanisms underlying SWB it is important to investigate to what degree *momentary* PA is genetically and/or environmentally influenced. To achieve this, momentary PA needs to be assessed with an ecologically valid assessment tool, which is able to “detect variations in emotional states over time” (p. 186, Csikszentmihalyi & Hunter, 2003). The experience sampling method (ESM) is best suited for this task because it prompts individuals to state their feelings at various unpredictable moments over an extended period of time (Csikszentmihalyi & Hunter, 2003; Nes, 2010). An important advantage of ESM is that subjects are assessed *prospectively* and *repeatedly* for momentary emotions and events, thus reducing distortion by recall biases and noise. Furthermore, traits involving person-context interplay can be measured by prospectively assessing affective responses to minor daily life situations, because ESM enables repeated assessment of the immediate environment. It has been suggested that momentary happiness is more easily influenced by proximal environmental factors (Csikszentmihalyi & Hunter, 2003) than overall SWB. The tendency to experience PA in the context of positive environmental stimuli has been called ‘reward experience’. That is, the more a person is able to translate the pleasantness of events into daily life PA, the more reward experience he or she has. Reward experience has particularly been linked to wellbeing and is usually assessed behaviorally or with neuroimaging techniques. Reward experience as assessed in these studies appears to be genetically influenced (Blum, et al., 2010; Bogdan & Pizzagalli, 2009; Dreher, Kohn, Kolachana, Weinberger, & Berman, 2009; Forbes, Shaw, & Dahl, 2007; Haas, et al., 2009; Schmack, et al., 2008; Watson, Ghodasra, & Platt, 2009; Yacubian, et al., 2007). Furthermore, reward experience appears to make people resilient against depression liability (Forbes, 2009). This seems particularly true for daily life momentary reward experience. Although most studies on reward experience lack the ecological validity that is required to draw conclusions about the phenomenological experience of reward (Forbes, 2009), recent studies employing ESM showed that daily

life reward experience predicts positive future response to treatment in individuals diagnosed with depression (Wichers, Barge-Schaapveld, et al., 2009; Wichers, et al., 2010), positive emotional outcome in individuals exposed to environmental risks for depression (Geschwind, et al., 2010), and is significantly influenced by the COMT^{Val158Met} polymorphism, a genetic mutation with functional impact on dopamine signaling (Wichers, Kenis, et al., 2008). Daily life reward experience therefore appears to be a genetically influenced resilience phenotype; however, this finding awaits further exploration as it is not yet clear to *what degree* genes and environment influence this phenotype.

Although reward experience is a potential resilience phenotype protecting against depression (Forbes, 2009), *reduced* reward experience has also been shown to characterize depression directly (APA, 2000; Caldú & Dreher, 2007; Forbes, et al., 2007; Gotlib, et al., 2010; Henriques & Davidson, 2000; Henriques, Glowacki, & Davidson, 1994). Conversely, stress-sensitivity (i.e. tendency to experience negative affect (NA) in response to negative environmental stimuli) is a well-accepted risk factor for depression (Mezulis, Funasaki, Charbonneau, & Hyde, 2010; Morris, Ciesla, & Garber, 2010; Siegrist, 2008). More specifically, *daily life* stress-sensitivity index genetic liability to depression and predicts the transition to a clinical depression outcome (Jacobs, et al., 2006; Wichers, Geschwind, et al., 2009; Wichers, et al., 2007b). It is well known that more prolonged stress-exposure reduces the ability to experience reward in animals and humans (Berenbaum & Connelly, 1993; Pizzagalli, Bogdan, Ratner, & Jahn, 2007; Willner, 2005). While it is important to investigate how *accumulative momentary* stress can influence subsequent reward responses, it is also relevant to understand how the general abilities to experience PA and NA in response to positive, respective negative, environmental stimuli relate to each other *directly* under *normal* circumstances in the general population. This will shed more light on the potential of reward experience in preventing mood disturbances. The most parsimonious relationship between reward experience and stress-sensitivity is one where the two simply represent two ends of a single phenotype of emotional reactivity to environmental influences. The question therefore rises to what degree reward experience reflects a resilience factor that is entirely independent of stress-sensitivity or rather reflects one extreme of a single continuum with reward experience and stress-sensitivity at opposite ends. If the latter were true, we would expect reward experience to be significantly associated with stress-sensitivity because then the one would simply be a slightly different expression of the other. There is preliminary evidence that this is the case (Bogdan & Pizzagalli, 2009). However, the study by Bogdan and Pizzagalli was based on a small sample, retrospective stress and behavioral reward assessments. It is therefore still unclear

how these two dynamic phenotypes relate to each other during everyday life in the general population.

To the best of our knowledge, the genetic and environmental influences on the variation of *daily life* positive affect, reward experience, and the association between reward experience and stress-sensitivity have not been examined yet. The current study aimed to address the following open questions with an ecologically valid assessment tool (ESM) and a genetically informative twin design: (1) Is momentary positive affect heritable?, (2) To what degree is momentary reward experience heritable?, and finally (3) Are daily life reward experience and stress-sensitivity associated and if so, what is the source of their association?

Methods

Subjects

Five-hundred and twenty female twins who took part in a longitudinal general population twin survey (East Flanders Prospective Twin Survey, EFPTS) participated in the current study, including 158 MZ and 102 DZ twin pairs. The EFPTS has been recording multiple births from 1964 (Derom, et al., 2006). Zygosity was determined from several analyses based on sex, fetal membranes, blood groups and DNA fingerprints. Approval from the local Ethics committee was obtained. The sample was exclusively female because of well known sex differences in aetiology and expression of mood disorders (Fanous, Gardner, Prescott, Cancro, & Kendler, 2002). The response rate was 29.1%. After nature and procedure of the study was fully explained to participants, they gave written informed consent before study-onset. Two subjects with outlying values on the reward-variable [-75 (10.17SDs from the mean) and -135 (13.47SDs from the mean) with an overall mean of 111 (SD=18.26)] were excluded from analyses to increase data normality. Demographics, means of phenotypes, and *p*-values of the comparisons between MZ and DZ twins are listed in table 1.

Table 1. Demographics and phenotypes per zygosity.

	MZ (N=316)	DZ (N=204)	P
Marital status			
Married	98	78	0.06
Divorced	6	4	0.06
Employment status			
Fulltime	115	81	0.32
Parttime	61	38	0.32
Looking for work	5	5	0.32
Level of education			
University degree	193	131	0.27
Secondary education	116	65	0.36
Age	26.79 (7.34)	27.85 (7.67)	0.12
Phenotypes			
PA	3.54 (0.68)	3.54 (0.73)	0.99
Reward experience	111.07 (18.15)	109.44 (18.34)	0.33
Stress-sensitivity	103.32 (6.01)	102.99 (5.67)	0.67

Note. MZ = monozygotic twins, DZ = dizygotic twins. All variables for marital status, employment status and level of education are given in counts. The variables for phenotypes and age are given in means and standard deviations in brackets. The p-value concerns the difference between monozygotic and dizygotic twins for each of the variables.

Experience Sampling Method (ESM)

The experience sampling method (ESM) is a structured diary technique to assess subjects in their daily living environment, and has been exhaustively validated for the use of assessing the immediate effects of daily life events on mood (Csikszentmihalyi & Larson, 1987; Delespaul, 1995; Jacobs, et al., 2005). Subjects received a digital wristwatch and a set of ESM self-assessment forms collated in a booklet for each day. The wristwatch was programmed to emit a signal ('beep') at an unpredictable moment in each of ten 90-min time blocks between 7:30am and 10:30pm, on five consecutive days. After each beep, subjects were asked to fill in the ESM self-assessment forms about thoughts, activity, company, location, appraisals of current situation, and mood. All these assessments were rated on seven-point Likert scales (1 = absolutely not true, 7 = very much true). Furthermore, participants filled in the most important event since the last signal and rated its pleasantness (-3 very pleasant, 0 neutral, 3 very unpleasant). Trained research assistants explained the ESM procedure to participants during an initial briefing session and a practice form was completed to confirm that subjects were able to understand the scales. Subjects could call a telephone number in case they had questions or problems during the ESM sampling period. Subjects were instructed to complete their reports immediately after the beep, thus minimizing memory distortion, and to record the time at which they completed the form. To

verify whether subjects had completed their record within 15 minutes of the beep, the time participants indicated in the booklet was compared with the actual time of the beep. All reports not filled in within this time window were excluded from analysis, as previous work (Delespaul, 1995) has demonstrated that these reports are less reliable and therefore less valid. For the same reason, subjects with less than 17 valid reports (out of 50) were excluded from the analysis (Delespaul, 1995).

In a random subsample of above sample (59 individuals), compliance to the sampling procedure was assessed with electronic monitoring devices adjusted to 100cc bottles. These bottles contained saliva swabs that participants were required to collect at every 'beep'. Participants were unaware that the timing was monitored. Compliance, thus assessed, was high (>90%) (Jacobs, et al., 2005).

Positive affect (PA) was created by aggregating responses on the items 'cheerful', 'content', 'energetic', and 'enthusiastic' per beep and person to one factor PA (Cronbach's alpha=.86 over participant mean). Items were weighted according to their individual factor loadings.

Reward experience was defined as the degree to which PA increased in response to event pleasantness. For example, if a person reported more PA after a very pleasant event than after a pleasant event, this person showed more reward experience than a person reporting the same level of PA after both events. Reward experience was thus conceptualized by regressing PA on the pleasantness of all pleasant events for each individual with at least three reported pleasant events. Consistent with previous work, only events with a rating from 0 (neutral) to -3 (very pleasant) were considered because it is highly unlikely that people generate PA from unpleasant events and a PA-decrease in response to unpleasant events does not fit the concept of reward experience (Geschwind, et al., 2010; Wichers, Schrijvers, et al., 2009). The event-variable was recoded (0-neutral to 3-very pleasant); higher values indicating higher event pleasantness. As the regression was based on all events for each individual, repeated measurements were taken into account, as a consequence of which noise in this phenotype could be minimized. A steeper slope of this regression equation indicated a higher level of reward experience.

Likewise, stress-sensitivity was defined as the degree to which negative affect (NA) increased in response to event unpleasantness. NA was regressed on unpleasant events for each individual with at least three reported negative events. NA was created by aggregating the items 'insecure', 'lonely', 'anxious', 'irritated', 'low', 'guilty', and 'suspicious' per beep and person to one factor NA (Cronbach's alpha=.76 over the

participant mean). Each item was weighted according to its factor loading. All events with a rating from 0 (neutral) to 3 (very unpleasant) were considered. The slope of the regression was considered to be stress-sensitivity, a steeper slope indicating a higher level of stress-sensitivity.

Analyses

Correlations

Cross-twin, within-trait correlations of PA, reward experience, and stress-sensitivity were calculated for monozygotic (MZ) and dizygotic twins (DZ) separately. Within-twin, cross-trait correlations between reward experience and stress-sensitivity estimated the overall association between both phenotypes. Cross-twin, cross-trait correlations allowed for an estimation of the genetic influence on the association between reward experience and stress-sensitivity. The correlation analyses were used to inform the structural equation model fitting.

Structural equation model fitting

Structural equation models were implemented with Mx (32) (Neale, Boker, Xie, & Maes, 2006) to specify the underlying sources of the phenotype-variations. For PA, reward experience, and stress-sensitivity separate univariate structural equation models were fitted to the data. Univariate structural equation models decompose the overall phenotypic variance into three different sources: additive genetic (A) (variation due to shared genes), shared environmental (C) (variation due to shared environment), and non-shared environmental (E) effects (variation due to individual-specific environmental influences and measurement error). The reward- and stress-variables were linearly transformed for convenience of interpretation and better model convergence. Several models were fitted to the data. The best fitting model was chosen based on fit and parsimony, which are both considered by the AIC-parameter; the smaller the AIC-parameter, the better the model fits. Nested models were compared with the difference in χ^2 - and p -values. Additionally, results from the correlational analyses were employed to guide and support model choice. As stress-sensitivity was defined in a slightly different fashion compared to earlier studies in the same sample (see Jacobs, et al., 2006; Wichers, et al., 2007b) the model for stress-sensitivity was rerun to best inform the following bivariate analysis.

Depending on the outcome of the cross-twin cross-trait correlation and the univariate structural equation modeling, a bivariate structural equation model was planned to estimate the genetic and environmental influences on the covariance between reward experience and stress-sensitivity and the genetic and environmental correlations.

Results

Correlations

Table 2 depicts cross-twin within-trait correlations for PA, reward experience and stress-sensitivity separately for MZ and DZ twins and their differences. The cross-trait cross-twin correlation for reward experience and stress-sensitivity was small ($r=-0.05$, ns), and did not differ significantly between MZ and DZ twins.

Table 2. Correlations between proband twin and co-twin on PA, reward experience, and stress-sensitivity.

Proband twin phenotypes		correlation		p
		MZ	DZ	
positive affect	co-twin positive affect	0.37 ^a	0.34 ^a	0.79
reward experience	co-twin reward experience	0.20 ^b	-0.10	0.02
stress-sensitivity	co-twin stress-sensitivity	0.08	0.04	0.76
reward experience	co-twin stress-sensitivity	0.06	-0.08	0.28

^a $p < 0.001$; ^b $p < 0.05$; MZ = monozygotic twins; DZ = dizygotic twins, p = p -values for the difference between MZ and DZ-correlations were calculated with Fisher's r -to- z transformation ($zr=1/2 \log(r+1)/(r-1)$)

Structural equation model fitting

Table 3 depicts all univariate models with estimated parameters, degrees of freedom, fit-statistics (AIC), and p -values for the difference with the most saturated model.

The CE-model for PA had the best fit as indicated by the AIC-value and did not significantly differ in explanatory power from the full ACE-model but was significantly better than E-model ($p<.001$). Thirty-four percent of the variance was explained by shared environmental influences.

The AE-model for reward experience showed the best AIC-value. However, when comparing the χ^2 -values, this model was not significantly better than the more parsimonious E-model ($p=.07$), which in turn did not fit worse than the full ACE-model.

Stress-sensitivity was best explained by the E-model. This model was not significantly worse than any of the other models (all p 's > 0.6).

The bivariate modeling was not executed because the cross-trait cross-twin correlation between reward experience and stress-sensitivity was weak and not significant.

Table 3. Univariate structural equation models for PA, reward experience, and stress-sensitivity.

phenotype	model	-2LL	df	AIC	A	C	E	p
positive affect	ACE	1068.02	514	40.02	0.19 (0; 0.50)	0.20 (0; 0.43)	0.62 (0.49; 0.76)	-
	AE	1069.15	515	39.15	0.41 (0.27; 0.52)	-	0.59 (0.48; 0.73)	0.29
	CE	1068.78	515	38.78	-	0.34 (0.23; 0.44)	0.66 (0.56; 0.77)	0.38
	E	1100.13	516	68.13	-	-	1 (1; 1)	0.00
reward experience	ACE	4249.68	488	3273.68	0.15 (0; 0.30)	0 (0; 0.19)	0.85 (0.69; 1)	-
	AE	4249.68	489	3271.68	0.15 (0; 0.30)	-	0.85 (0.69; 1)	nc
	CE	4251.38	489	3273.38	-	0.09 (0; 0.22)	0.91 (0.78; 1)	0.19
	E	4252.93	490	3272.93	-	-	1 (1; 1)	0.20
stress-sensitivity	ACE	3644.70	417	2810.70	0.11 (0; 0.33)	0 (0; 0.25)	0.89 (0.67; 1)	-
	AE	3644.70	418	2808.70	0.11 (0; 0.33)	-	0.89 (0.67; 1)	nc
	CE	3644.70	418	2808.83	-	0.08 (0; 0.26)	0.92 (0.74; 1)	0.72
	E	3645.55	419	2807.55	-	-	1 (1; 1)	0.65

Best fitting models in bold. A = standardized additive genetic effects; C = standardized shared environmental effects; E = standardized individual-specific environmental effects; -2LL = minus twice the log likelihood of the raw data; df = degrees of freedom; AIC = Akaike's Information Criterion; p = p-value of difference in χ^2 -value from most saturated model; nc = not calculable because difference in χ^2 literally zero.

Discussion

The current study investigated whether and to what degree momentary positive affect, reward experience, stress-sensitivity, and the association between reward experience and stress-sensitivity were genetically and environmentally influenced in a general population twin sample.

Findings

The overall level of momentary positive affect (PA) appeared to be determined by environment only, indicated by almost equal correlations between MZ and DZ twins and the best-fitting model only including shared and non-shared environmental influences. The current findings on reward experience, i.e. PA in response to positive daily life events, were less consistent. Although the correlation was significantly larger in MZ than DZ twins, indicating a genetic influence, the best-fitting model included non-shared environmental factors only. Stress-sensitivity, i.e. negative affect (NA) in response to negative daily life events, was influenced by non-shared environment only. Momentary reward experience and stress-sensitivity showed no significant correlation, indicating that the two phenotypes were practically independent from each other.

The major strength of the current study was the conceptualization of the phenotypes. Employing ESM guaranteed *prospective* data collection in people's *natural* living environment, reducing potential data distortion by retrospective report bias and increasing ecological validity. Furthermore, the assessment of dynamic phenotypes like stress-sensitivity and reward experience includes a certain level of noise. Because the current conceptualizations included repeated measurements per individual, noise in these phenotypes could be reduced to a potential minimum. To the best of our knowledge, this is the first study to carry out a behavioral genetic analysis on thus conceptualized daily life PA and reward experience.

Best-fitting E-models for reward experience and stress-sensitivity were not expected based on previous research (e.g. Blum, et al., 2010; Bogdan & Pizzagalli, 2009; Dreher, et al., 2009) and appear rather unique in family-based research. Previous behavioral genetic research was usually conducted on questionnaire-data, whereas ESM-data remain very rare, and to date has not been used for PA and reward experience. The uniqueness of the current findings might partly be due to the uniqueness of the method of assessment. Immediate emotions as assessed with ESM are likely more environmentally influenced than the more general perceptions of one's emotional states in the past or traits as assessed with questionnaires. To shed more light on this issue, it would be important to replicate and extend the current findings by employing questionnaire-, experimental-, and ESM assessments in the same sample. Additionally, assuming small (if any) genetic effects, possibly larger samples are needed to detect those in ESM-studies. We elaborate more on this latter issue below.

Sources of PA

Although the AE-model for PA differed little from the CE-model, the smallest AIC-value of the latter as well as the almost equal and equally significant correlations between MZ and DZ twins strongly suggest that PA is uniquely environmentally influenced and *not* heritable. Earlier research has shown that momentary PA can act as an antidote to depression-risk (Wichers, et al., 2007a). Therefore, we cautiously suggest that momentary PA is a *genetically independent* resilience phenotype for depression.

According to the broaden-and-build theory, momentary PA builds resources that result in higher subjective wellbeing (SWB) and successes in several areas of life (Fredrickson, 2001; Lyubomirsky, King, et al., 2005; Lyubomirsky, Sheldon, et al., 2005). Earlier studies found that SWB, defined as overall happiness and satisfaction with life, is to a moderate degree genetically influenced (Bartels & Boomsma, 2009; Lyubomirsky, Sheldon, et al., 2005; Nes, et al., 2010). Assuming that the current results reflect a true

trend towards a non-genetically influenced momentary PA-phenotype, how can the hypothesized pre-requisite of SWB, momentary PA, not be heritable if SWB is? One possible explanation is that momentary PA and SWB are truly unrelated. We consider this unlikely, because previous studies have demonstrated that the *frequency* of positive emotions (Kuppens, Realo, & Diener, 2008) and even the experience of retrospectively assessed *daily* PA (Cohn, et al., 2009) are in fact associated with resources (ego-resilience) and life satisfaction, a core concept of SWB. Moreover, although momentary PA is hypothesized to precede SWB, it is still a different phenotype. This is the first study to investigate the heritability of *momentary* PA with ESM. This is considerably different from the usual SWB measurement, which specifically requires people to report *once* and *retrospectively* on broad concepts like life satisfaction, reflecting a broad and general experience of wellbeing. It is therefore possible that one phenotype is genetically influenced while the other is not, particularly if the genetically influenced phenotype is more complex (SWB). Presumably, cognitive style is heritable (Zavos, Rijdsdijk, Gregory, & Eley, 2010) and has considerable impact on how people generally view their lives (SWB) (Csikszentmihalyi & Hunter, 2003; Lyubomirsky, Sheldon, et al., 2005) and *retrospectively* interpret the experience of momentary PA (assessment of SWB). Possibly, momentary PA in combination with a genetically influenced cognitive style is associated with SWB. However, this is hypothetical and requires further investigation.

Sources of reward experience

The general experience of momentary PA should be contrasted with the ability to experience PA in response to small positive daily life events (reward experience). Reward experience may reflect a trait-like characteristic, i.e. a particular tendency to respond to positive situations. The current within-pair correlations indicate genetic influences whereas the structural equation models point to non-shared environmental influences only; although the most parsimonious model including genetic influences (AE) fitted the data evenly well as the full ACE-model and almost better than the E-model ($p=.07$). Admittedly, the power for an AE-model to fit *significantly* better than an E-model was low (.44). Assuming small genetic effects, this might explain the inconsistent results. Future research needs to show whether the same analyses in a larger sample will result in an AE-model. We consider this likely, because previous behavioral and neuroimaging studies indicated moderate heritability of reward experience (e.g. Bogdan & Pizzagalli, 2009; Dreher, et al., 2009; Forbes, et al., 2007; Yacubian, et al., 2007). Importantly, one previous ESM-study indicated a dose-response influence of a particular functional polymorphism, COMT^{Val158Met} (a polymorphism influencing reward-related neurotransmitter signaling) on *momentary* reward experience (Wichers, Aguilera, et

al., 2008). This result needs to be viewed with caution given the possibility of Type I error; however, a related neuroimaging study replicated this finding (Dreher, et al., 2009). Although we did not find clear evidence for a genetic influence in reward experience in the current sample, we consider it acceptable that future studies with preferably larger samples will be able to do so.

Momentary reward experience can counteract depression vulnerability (Geschwind, et al., 2010) and increases the likelihood of remission (Wichers, Barge-Schaapveld, et al., 2009). Irrespective of possible genetic influences on its variability, the current findings clearly indicate that this resilience phenotype is environmentally influenced. Increasing momentary reward experience may thus foster wellbeing in general and remission from and prevention of depression in particular. A recent study indeed demonstrated that remitted depressed patients could be trained to increase their momentary reward experience, which was associated with reduction in depressive symptoms (Geschwind, et al., 2011).

The possibility to increase momentary reward experience gains validity in the light of the current finding that momentary reward experience and stress-sensitivity were virtually unrelated in a general population sample implying that the two phenotypes are not mere opposite ends of a single dimension. Rather, this suggests that systems marking and supporting wellbeing and resilience operate independently from systems marking vulnerability to affective dysregulation. Thus, a vulnerable (stress-sensitive) person can still have resilience traits (reward experience) and might therefore be protected against severe mood symptoms. This makes sense from both an evolutionary and neurobiological perspective. First, it is important to react differentially to pleasant and unpleasant events in order to adapt to constantly changing environments. Furthermore, different hormonal mechanisms and circuits have been suggested to impact neurobiological systems associated with reward- and stress-reactions (Dedovic, Duchesne, Andrews, Engert, & Pruessner, 2009; Ikemoto, 2010; Krishnan & Nestler, 2010). This does not imply however, that the systems are generally unable to influence each other. It has repeatedly been shown that cumulative stress-exposure reduces the ability to experience reward (Berenbaum & Connelly, 1993; Pizzagalli, et al., 2007; Willner, 2005). Therefore overstimulation of the stress-system may cause a down-regulation of the reward-system. This association may be characteristic for mood disorders as likewise positive and negative affect are mostly unrelated in healthy individuals (Diener & Emmons, 1985), whereas they are negatively correlated in individuals suffering from depression (Viechtbauer, unpublished communication). Possibly, the balance of the two systems becomes disturbed in depressive states. The current findings merely demonstrate that in a general population sample, the tendency

to experience NA in response to negative events is independent from the tendency to experience PA in response to positive events. Future research needs to investigate under which circumstances the two systems become associated.

Methodological considerations

The current findings need to be viewed in the light of some methodological considerations. First, ESM presents with a potential source for measurement error in the form of the subjects' compliance with the sampling time. A too large discrepancy between the beep and the actual report can make ESM-measurements unreliable (Delespaul, 1995). As it was demonstrated in a subgroup of the current sample that compliance to the timing was high (Jacobs, et al., 2005) and as a general procedure (Delespaul, 1995), all records filled in too late (> 15min after beep) were excluded, we suggest that this source of measurement error was at its potential minimum.

Second, the unknown time-lag between the most important event and the measurement of emotions may have resulted in an underestimation of reward experience and stress-sensitivity. At each beep, participants were asked to indicate the (un)pleasantness of the most important event between the current and the previous beep. If the event was more remote in time, the influence on current emotions might have been smaller. Furthermore, if the most important event was *unpleasant*, other positive but less important events might have been missed, introducing some noise to the measurement. This was not true for PA, as *immediate* PA was assessed at *each single beep*, so PA included less noise than the dynamic phenotypes. This may account for the clearer results from correlational and univariate model fitting-analyses for PA. Therefore, future studies may consider explicitly assessing positive (respectively negative) events at each moment and approximate time-lags to overcome these potential problems.

Third, unlike previous studies (Jacobs, et al., 2006; Wichers, et al., 2007b), the current study did not employ all events (from -3(very pleasant) to 3(very unpleasant)) to define stress-sensitivity, but unpleasant events only. This ensured that a possible association between reward experience and stress-sensitivity was not artificially increased by including the *same* events in both phenotypes. An increase in NA from very pleasant to neutral (-3 – 0) events may reflect a *decrease* in NA in response to pleasant events, occasioning a strong correlation with reward experience. Furthermore, laboratory stress tasks also consider stress as a negative emotional reaction to *unpleasant* events (e.g. Kirschbaum, Pirke, & Hellhammer, 1993).

Fourth, it needs to be acknowledged that a number of models could not well be distinguished from alternative ones, likely due to the rather small sample size for structural equation model fitting, reducing the power to detect small effects. For example, the power to distinguish the AE-model from the E-model for reward experience was not higher than 0.44. Although we argue that the current results point into particular directions, without a doubt do these results need replications in other and larger samples before they can be viewed as conclusive.

Fifth, the current study employed event-related reward because it was considered to be a good proxy for natural reward experiences and behavioral and neuroimaging assessments, where individuals are prompted with pleasant stimuli (e.g. Blum, et al., 2010; Bogdan & Pizzagalli, 2009; Schmack, et al., 2008). However, momentary activity- and company-related reward experiences may at least be equally important (Geschwind, et al., 2011). If and to what degree these different concepts relate equally well to wellbeing and mental health remains to be explored.

Finally, the current general population sample was female only with above average educational level. The findings may therefore not generalize to clinical or male samples, or samples with (below) average educational level.

Besides these methodological issues, the current findings suggest that daily life reward experience and positive emotions can thrive in all individuals irrespective of their level of stress-sensitivity and genetic variation. Future research needs to further investigate the associations between momentary positive affect, reward experience, stress-sensitivity, and SWB to elucidate in more detail the factors driving risk and resilience in wellbeing and mental health.

Funding

This research was funded by the Association for Scientific Research in Multiple Births.

Acknowledgements

The authors thank Marjolein Bonte, Veerle van Kerckhove, and Annelies Verlinden for assistance with data collection and all twins who volunteered for participation. This research was supported by the Netherlands Organisation for Scientific Research (VENI grant nr 916.76.147 to Dr Wichers) and by the European Community's Seventh Framework Program under grant agreement No. HEALTH-F2-2009-241909 (Project EU-GEI).

Conflict of interests

There are no conflicts of interest. The authors have full control of the primary data and agree to allow the journal to review these data if requested.

References

- APA (2000). *Quick reference to the diagnostic criteria from DSM-IV-TR*. Arlington, VA: American Psychiatric Association.
- Bartels, M., & Boomsma, D. I. (2009). Born to be happy? The etiology of subjective well-being. *Behavioral Genetics, 39*, 605-615.
- Berenbaum, H., & Connelly, J. (1993). The effects of stress on hedonic capacity. *Journal of Abnormal Psychology, 102*(3), 474-481.
- Blum, K., Chen, T. J. H., Chen, A. L. H., Madigan, M., Downs, B. W., Waite, R. L., et al. (2010). Do dopamine gene polymorphisms affect mesolimbic reward activation of music listening response? Therapeutic impact deficiency syndrome (RDS). *Medical Hypotheses, 74*, 513-520.
- Bogdan, R., & Pizzagalli, D. A. (2009). The heritability of hedonic capacity and perceived stress: a twin study evaluation of candidate depressive phenotypes. *Psychological Medicine, 39*, 211-218.
- Caldú, X., & Dreher, J.-C. (2007). Hormonal and genetic influences on processing reward and social information. *Annals of the New York Academy of Sciences, 1118*, 43-73.
- Cohn, M. A., Fredrickson, B. L., Brown, S. L., Mikels, J. A., & Conway, A. M. (2009). Happiness unpacked: positive emotions increase life satisfaction by building resilience. *Emotion, 9*(3), 361-368.
- Csikszentmihalyi, M., & Hunter, J. (2003). Happiness in everyday life: the use of experience sampling. *Journal of Happiness Studies, 4*, 185-199.
- Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *J Nerv Ment Dis, 175*, 526-536.
- Dedovic, K., Duchesne, A., Andrews, J., Engert, V., & Pruessner, J. C. (2009). The brain and the stress-axis: the neural correlated of cortisol regulation in response to stress. *NeuroImage, 47*(3), 864-871.
- Delespaul, P. (1995). *Assessing schizophrenia in daily life. The experience sampling method*. Maastricht: IPSE Foundation.
- Derom, C. A., Vlietinck, R. F., Thiery, E. W., Leroy, F. O., Fryns, J. P., & Derom, R. M. (2006). The east Flanders prospective twin survey (EFPTS). *Twin Res Human Genetics, 9*(6), 733-738.
- Diener, E., & Emmons, R. A. (1985). The independence of positive and negative affect. *Journal of Personality and Social Psychology, 47*(5), 1105-1117.
- Dreher, J.-C., Kohn, P., Kolachana, B., Weinberger, D. R., & Berman, K. F. (2009). Variation in dopamine genes influences responsivity of the human reward system. *PNAS, 106*(2), 617-622.
- Fanous, A., Gardner, C. O., Prescott, C. A., Cancro, R., & Kendler, K. S. (2002). Neuroticism, major depression and gender: a population-based twin study. *Psychological Medicine, 32*, 719-728.
- Forbes, E. E. (2009). Where's the fun in that? Broadening the focus on reward function in depression. *Biological Psychiatry, 66*, 199-200.
- Forbes, E. E., Shaw, D. S., & Dahl, R. E. (2007). Alterations in reward-related decision making in boys with recent and future depression. *Biological Psychiatry, 61*, 633-639.
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology. The broaden-and-build theory of positive emotions. *American Psychologist, 56*(3), 218-226.
- Geschwind, N., Nicolson, N. A., Peeters, F., van Os, J., Barge-Schaapveld, D. Q. C. M., & Wichers, M. (2011). Mindfulness training increases momentary positive emotions and reward experience in adults vulnerable to depression. A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 79*(5), 618-628.
- Geschwind, N., Peeters, F., Jacobs, N., Delespaul, P., Derom, C., Thiery, E., et al. (2010). Meeting risk with resilience: high daily life reward experience preserves mental health. *Acta Psychiatrica Scandinavica, 122*(129-138).

- Gotlib, I. H., Hamilton, J. P., Cooney, R. E., Singh, M. K., Henry, M. L., & Joormann, J. (2010). Neural processing of reward and loss in girls at risk for major depression. *Arch Gen Psychiatry*, *67*(4), 380-387.
- Haas, B. W., Mills, D., Yam, A., Hoeft, F., Bellugi, U., & Reiss, A. (2009). Genetic influences on sociability: heightened amygdala reactivity and event-related responses to positive stimuli in Williams Syndrome. *The Journal of Neuroscience*, *29*(4), 1132-1139.
- Henriques, J. B., & Davidson, R. J. (2000). Decreased responsiveness to reward in depression. *Cognition and Emotion*, *14*(5), 711-724.
- Henriques, J. B., Glowacki, J. M., & Davidson, R. J. (1994). Reward fails to alter response bias in depression. *Journal of Abnormal Psychology*, *103*(3), 460-466.
- Ikemoto, S. (2010). Brain reward circuitry beyond the mesolimbic dopamine system: A neurobiological theory. *Neuroscience and Biobehavioral Reviews*, *35*, 129-150.
- Jacobs, N., Nicolson, N. A., Derom, C., Delespaul, P., van Os, J., & Myin-Germeys, I. (2005). Electronic monitoring of salivary cortisol sampling compliance in daily life. *Life Sciences*, *76*, 2431-2443.
- Jacobs, N., Rijdsdijk, F., Derom, C., Vlietinck, R., Delespaul, P., van Os, J., et al. (2006). Genes making one feel blue in the flow of daily life: a momentary assessment study of gene-stress interaction. *Psychosomatic Medicine*, *68*, 201-206.
- Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The "Trier Social Stress Test" - A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*, 76-81.
- Krishnan, V., & Nestler, E. J. (2010). Linking molecules to mood: new insight into the biology of depression. *Am J Psychiatry*, *167*, 1306-1320.
- Kuppens, P., Realo, A., & Diener, E. (2008). The role of positive and negative emotions in life satisfaction judgment across nations. *Journal of Personality and Social Psychology*, *95*(1), 66-75.
- Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: does happiness lead to success? *Psychological Medicine*, *131*(6), 803-855.
- Lyubomirsky, S., Sheldon, K. M., & Schkade, D. (2005). Pursuing happiness: the architecture of sustainable change. *Review of General Psychology*, *9*(2), 111-131.
- Mezulis, A. H., Funasaki, K. S., Charbonneau, A. M., & Hyde, J. S. (2010). Gender differences in the cognitive-vulnerability stress model of depression in the transition to adolescence. *Cognitive Therapy and Research*, *34*, 501-513.
- Morris, M. C., Ciesla, J. A., & Garber, J. (2010). A prospective study of stress autonomy versus stress sensitization in adolescents at varied risk for depression. *Journal of Abnormal Psychology*, *119*(2), 341-354.
- Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2006). *Mx: Statistical Modelling* (7 ed.). VCU Box 900126, Richmond, VA 23298: Department of Psychiatry.
- Nes, B. R. (2010). Happiness in behaviour genetics: findings and implications. *Journal of Happiness Studies*, *11*, 369-381.
- Nes, B. R., Czajkowski, N., & Tambs, K. (2010). Family matters: happiness in nuclear families and twins. *Behavioral Genetics*, *40*, 577-590.
- Pizzagalli, D. A., Bogdan, R., Ratner, K. G., & Jahn, A. L. (2007). Increased perceived stress is associated with blunted hedonic capacity: potential implications for depression research. *Behavior Research and Therapy*, *45*, 2742-2753.
- Schmack, K., Schlagenhauf, F., Sterzer, P., Wrase, J., Beck, A., Dembler, T., et al. (2008). Catechol-O-methyltransferase val158met genotype influences neural processing of reward anticipation. *NeuroImage*, *42*, 1631-1638.
- Siegrist, J. (2008). Chronic psychosocial stress at work and risk of depression: evidence from prospective studies. *European Archives of Psychiatry & Clinical Neuroscience*, *258*, 115-119.
- Viechtbauer, W. (unpublished communication). Factor structure of positive and negative affect in healthy and pathological samples.
- Watson, K. K., Ghodasra, J. H., & Platt, M. L. (2009). Serotonin transporter genotype modulates social reward and punishment in rhesus macaques *PLoS ONE*, *10*(1), e4156.
- Wichers, M., Aguilera, M., Kenis, G., Krabbendam, L., Myin-Germeys, I., Jacobs, N., et al. (2008). The catechol-O-methyl transferase val158met polymorphism and experience of reward in the flow of daily life. *Neuropsychopharmacology*, *33*, 3030-3036.
- Wichers, M., Barge-Schaapveld, D. Q. C. M., Nicolson, N. A., Peeters, F., de Vries, M., Mengelers, R., et al. (2009). Reduced stress-sensitivity or increased reward experience: the psychological mechanism of response to antidepressant medication. *Neuropsychopharmacology*, *34*, 923-931.
- Wichers, M., Geschwind, N., Jacobs, N., Kenis, G., Peeters, F., Derom, C., et al. (2009). Transition from stress-sensitivity to a depressive state: longitudinal twin study. *The British Journal of Psychiatry*, *195*, 498-503.
- Wichers, M., Kenis, G., Jacobs, N., Myin-Germeys, I., Derom, C., & Vlietinck, R. (2008). The psychology of psychiatric genetics: evidence that positive emotions in females moderate genetic sensitivity to social stress associated with the BDNF Val(66)Met polymorphism. *Journal of Abnormal Psychology*, *117*(699-706).
- Wichers, M., Myin-Germeys, I., Jacobs, N., Peeters, F., Kenis, G., Derom, C., et al. (2007a). Evidence that moment-to-moment variation in positive emotions buffer genetic risk for depression: a momentary assessment twin study. *Acta Psychiatrica Scandinavica*, *115*, 451-457.
- Wichers, M., Myin-Germeys, I., Jacobs, N., Peeters, F., Kenis, G., Derom, C., et al. (2007b). Genetic risk for depression and stress-induced negative affect in daily life. *British Journal of Psychiatry*, *191*, 218-223.
- Wichers, M., Peeters, F., Geschwind, N., Jacobs, N., Simons, C. J. P., Derom, C., et al. (2010). Unveiling pattern of affective responses in daily life may improve outcome prediction in depression: A momentary assessment study. *Journal of Affective Disorders*, *124*, 191-195.
- Wichers, M., Schrijvers, D., Geschwind, N., Jacobs, N., Myin-Germeys, I., Thiery, E., et al. (2009). Mechanisms of gene-environment interactions in depression: evidence that genes potentiate multiple sources of adversity. *Psychological Medicine*, *39*, 1077-1086.
- Willner, P. (2005). Chronic mild stress (CMS) revisited: consistency and behavioral-neurobiological concordance in the effects of CMS. *Neuropsychobiology*, *52*, 90-110.
- Yacubian, J., Sommer, T., Schroeder, K., Gläscher, J., Kalisch, R., Leuenberger, B., et al. (2007). Gene-gene interaction associated with neural reward sensitivity. *PNAS*, *104*(19), 8125-8130.
- Zavos, H. M. S., Rijdsdijk, F. V., Gregory, A. M., & Eley, T. C. (2010). Genetic influences on the cognitive biases associated with anxiety and depression symptoms in adolescents. *Journal of Affective Disorders*, *124*(1-2), 45-53.

A modest impact of genes on real-world reward experience: a special role for the OPRM1 gene

Menne-Lothmann, C.¹; Viechtbauer, W.¹; Bakker, J.¹; Kenis, G.¹; Jacobs, N.^{1,2}; Geschwind, N.³; Collip, D.¹; Pishva, E.¹; Derom, C.^{4,5}; Thiery, E.⁶; van Os, J.^{1,7}; and Wichers, M.^{1,8}

¹Department of Psychiatry and Psychology, South Limburg Mental Health Research and Teaching Network, EURON, Maastricht University, The Netherlands

² Faculty of Psychology, Open University of the Netherlands, Heerlen, The Netherlands

³Clinical Psychological Science, Maastricht University, The Netherlands

⁴Centre of Human Genetics, University Hospitals Leuven, Belgium

⁵Department of Human Genetics, KU Leuven, Belgium

⁶Department of Neurology, Ghent University Hospital, Ghent University, Ghent, Belgium

⁷King's College London, King's Health Partners, Department of Psychosis Studies, Institute of Psychiatry, London, United Kingdom

⁸ Department of Psychiatry, University Medical Centre of Groningen, Groningen University Groningen, The Netherlands

submitted in adapted form



Abstract

Background Reward experience plays an important role in mood disorders. A limited number of studies suggest a weak influence of particularly two single nucleotide polymorphisms (SNPs) involved in the brain's reward circuits on experimental reward. It remains open whether and how these findings translate to real-world reward when multiple potential reward SNPs are considered.

Objective To investigate the role of multiple genetic dopamine and opioid variations in real-world reward experience in two samples differing in level of mood alterations: a general population sample and a sample with residual depressive symptoms.

Methods 435 individuals from the general population and 126 individuals with residual depressive symptoms provided material for genotyping and participated in ≥ 5 days of experience sampling assessing pleasant events and positive affect up to ten times a day.

Results After correction for multiple testing, no influence of genes was found in the general population sample. However, one variation (OPRM1 rs1799971) reached significance in the sample with residual depressive symptoms. Individuals with two A-alleles showed significantly more positive affect in the context of pleasant events than individuals with at least one G-allele.

Conclusions The overall evidence for genetic influences on daily life reward experience was weak. The effect of the OPRM1 rs1799971 SNP may be enhanced in individuals with residual depressive symptoms. Individuals with depressive symptoms, compared to individuals from the general population, may be less capable to compensate genetic risk for deficits in reward experience.

Keywords reward, major depression, genetic variation, resilience, dopamine, opioids

Reward experience, defined as a positive emotional response to pleasant stimuli, has lately received much attention in the area of psychopathology. Depression is alarmingly common worldwide (WHO, 2013) and one of its core symptoms constitutes the inability to experience reward, or anhedonia (APA, 2000). Additionally, a deficit in reward experience has repeatedly been linked to addiction problems (Kraschewski et al., 2009; Shabalina et al., 2009; Wang et al., 2004). Conversely, it has recently been proposed that reward can also foster resilience against depression (Forbes, 2009; Geschwind et al., 2010; Lambert, 2006). Therefore, understanding the mechanisms of reward experience will help understanding development and prevention of mood disorders.

One way of elucidating such mechanisms is by investigating the role of genes in reward experience (Rutter, Moffitt, & Caspi, 2006). Reward experiences have been found to be reliant on dopamine action in the brain. The mesocorticolimbic dopamine system, originating in the ventral tegmental area in the midbrain and projecting to subcortical regions (striatum) (mesolimbic pathway) and prefrontal, cingulate and perirhinal areas (mesocortical pathway), has been associated with emotion-based behavior including goal-directed motivation like reward (Arias-Carrión, Stamelou, Murillo-Rodríguez, Menéndez-González, & Pöppel, 2010). This dopaminergic reward circuitry also innervates the opioid system (Chen et al., 2009). Opioids have strong addictive properties (van Ree, Gerrits, & Vanderschuren, 1999) and directly influence the 'liking' of particular stimuli (e.g., drugs) (Berridge & Robinson, 1998; Berridge, Robinson, & Aldridge, 2009). However, opioids also stimulate the mesolimbic dopamine system (Dilts & Kalivas, 1989; van Ree et al., 1999; Wise & Rompre, 1989). Therefore, opioids are of particular interest in reward research as they can act both directly and indirectly (by influencing dopamine) on the experience of reward. Genetic variations influencing dopamine and opioid regulation may therefore be hypothesized to account for individual differences in reward experience.

To date, research on genetic influences on reward-functioning is scarce. Recent pioneering work has started to address this gap. In a behavioral genetic study, Bogdan and Pizzagalli (2009) were able to illustrate that reward experience as assessed with an experimental learning task showed latent genetic involvement. It needs to be noted though that this study consisted of a small sample ($n=35$ twin pairs). Other studies focused on the role of mainly two genetic variations (the Catechol-O-Methyltransferase (COMT) Val¹⁵⁸Met coding gene and the dopamine transporter (DAT1 or SLC6A3) gene) in brain activity in response to experimental reward tasks (e.g., Camara et al., 2010; Dreher, Kohn, Kolachana, Weinberger, & Berman, 2009; Haas et al., 2009; Hahn et al.,

2011; Schmack et al., 2008; Yacubian et al., 2007). For example, Dreher et al. (2009) showed in a group of 22 subjects that individuals with a particular variation of the COMT Val¹⁵⁸Met gene showed higher activation of reward-related brain areas when reward was anticipated. COMT is responsible for breaking down dopamine in the synaptic cleft. In another group of 26 subjects, those with a specific variation of the dopamine transporter gene SLC6A3 showed increased activity in reward-areas when reward was perceived (Dreher et al., 2009). In a larger ($N=105$) male sample, brain activation in response to a reward guessing task was associated with both the COMT and SLC6A3 genotypes (Yacubian et al., 2007). The influence of the COMT Val¹⁵⁸Met genotype on reward processing in the brain was also found in at least two other experimental fMRI studies (Camara et al., 2010; Schmack et al., 2008) and the influence of the SLC6A3 genotype in at least one more study (Hahn et al., 2011). These initial findings suggest that genetic variations impacting on the neural reward system may contribute to individual differences in the experience of anhedonia or resilience against depression.

However, these studies are in urgent need of replication and expansion. First, earlier work reported two main genetic associations; the COMT Val¹⁵⁸Met and the SLC6A3. However, if genes influence complex phenotypes such as reward experience, many genetic variants will be involved. In support for this, various other genetic variations have previously been suggested to play a role in opioid- and dopamine signaling (e.g. see Kraschewski et al., 2009; Shabalina et al., 2009) as well as in disorders or behaviors related to reward-functioning (e.g. see Licinio, Dong, & Wong, 2009; Wang et al., 2004; Zhang et al., 2006; Zion et al., 2006). Furthermore, most of the above findings were relatively weak and often inconsistent with behavioral reward data that did not demonstrate significant genetic effects (e.g. see Hahn et al., 2011; Schmack et al., 2008). Finally, studies were rather homogenous by conceptualizing reward experience as behavioral or neural reactions to experimental reward tasks. However, anhedonia is defined as the lack of reward as experienced in everyday situations (APA, 2000). The *phenomenological* experience of reward in the real world is what matters clinically, and recent studies found such real-world reward experience to have resilient properties (Forbes, 2009; Geschwind et al., 2010; Wichers et al., 2009). To date, only two studies investigated the role of genes in daily life reward experience. One twin study found weak evidence for a genetic component in daily life reward experience (Menne-Lothmann et al., 2012), however, this study may have been underpowered to detect a small genetic effect. Another study suggested a role of the COMT Val¹⁵⁸Met genotype on daily life reward experience (Wichers, Aguilera, et al., 2008).

In sum, most existing genetic studies on reward experience are of an experimental nature, all studies so far –both experimental and real-world studies– were limited to one or two genetic variants at the time, and combined they did not convincingly demonstrate robust genetic effects. The current study therefore aimed at expanding previous research by combining an ecologically valid real-world measure of reward experience with examination of the influence of multiple genetic variations associated with dopamine and opioid signaling in the brain. Since we sought to explore the genetic mechanisms of daily life reward experience, a system which is highly relevant to depression, we addressed the role of depressive psychopathology by including two different samples: a general population sample and a sample of individuals with depressive symptoms in the context of a recent episode of Major Depression (MDD).

Methods

Participants

Study participants were drawn from two independent samples. One general population sample and one sample of individuals with residual depressive symptoms. The general population sample consisted of 621 female twins from the East Flanders Population Twin Registry (Derom et al., 2013) (for more information about the sample, see Jacobs et al., 2006; Wichers, Aguilera, et al., 2008; Wichers, Kenis, Jacobs, Myin-Germeys, et al., 2008). The sample of individuals with residual depressive symptoms consisted of 130 participants (for more detailed information about this sample, see Geschwind et al., 2011).

Experience sampling method

The experience sampling method (ESM) is a structured diary technique to assess subjects in their daily living environment (Csikszentmihalyi & Larson, 1987; Delespaul, 1995; Jacobs et al., 2005). Individuals received a digital wristwatch that emitted a signal at ten unpredictable moments between 7:30am and 10:30pm for six (individuals with residual symptoms) or five (general population sample) consecutive days. After each signal, individuals were asked to fill in self-assessment forms about mood on 7-point Likert scales (1 = absolutely not true, 7 = very much true) and rate the pleasantness of the most important event since the last signal on a bipolar scale (-3 very pleasant, 0 neutral, 3 very unpleasant). (for more information, see Delespaul, 1995; Jacobs et al., 2005).

Reward experience

Positive affect (PA) was defined as the mean score over four available positive affect items (enthusiastic, relaxed, content, cheerful) per beep and person. These four items were selected because (1) they were available in both samples, (2) each has previously been shown to have sufficient within-person variation and load on a general PA-factor (Geschwind et al., 2010; Menne-Lothmann et al., 2012; Wichers, Kenis, Jacobs, Mengelers, et al., 2008; Wichers, Kenis, Jacobs, Myin-Germeys, et al., 2008), and (3) in the current sample the Cronbach's alpha for these items was acceptable in both samples, $\alpha = .85$ for the general population sample and $\alpha = .87$ for the sample with residual depressive symptoms (based on within-person residuals).

Event pleasantness was assessed by asking individuals to appraise the most important event since the last ESM assessment on a bipolar scale (-3 very unpleasant, 0 neutral, 3 very pleasant). Only events rated neutral to very pleasant were included in the analysis, as it is counterintuitive to assume that PA reactivity to unpleasant events involves reward experience (Geschwind et al., 2010; Menne-Lothmann et al., 2012; Wichers, Aguilera, et al., 2008; Wichers et al., 2009).

Reward experience was defined conform previous studies (Geschwind et al., 2011; Geschwind et al., 2010; Menne-Lothmann et al., 2012; Wichers, Aguilera, et al., 2008; Wichers et al., 2009) as the impact of event pleasantness appraisal on the level of positive affect (PA). Thus, if an individual reports more PA than another individual in the context of a similar rating for event pleasantness, this first individual shows a higher level of reward experience than the second.

Genotyping and gene-selection

For the twins from the general population, three sample types were available for genotyping: placental tissue, blood, and buccal cells. The buccal cells were collected with specifically designed sterile swabs (Omni Swabs; Whatman plc, Brentford, England). Genomic DNA was isolated using the QUIAamp DNA Mini Kit (Qiagen, Westburg, Leusden, the Netherlands) according to the appropriate protocol for each sample type (Jacobs et al., 2006).

For the sample of individuals with residual depressive symptoms, genomic DNA was collected in Oragene-DNA Self Collection Kits (DNA Genotek, Ottawa, Canada), and DNA was isolated using the AutoGenFlex DNA isolation system (Autgen, Hillison, MA, USA) according manufacturer's instructions. Single Nucleotide Polymorphisms (SNPs) were determined by Sequenom (Hamburg, Germany) using the Sequenom MassARRAY iPLEX platform at the facilities of the manufacturer (Bakker et al., 2014).

In 2009, a selection of single nucleotide polymorphisms (SNP) was made based on the following criteria: (i) the SNP is directly (by acting on dopamine function) or indirectly (by acting on related neurotransmitter systems) involved in dopamine functioning in the brain, or (ii) the SNP has previously been associated with disorders characterized by reward deficiency or anhedonia (Azzato et al., 2009; Balciuniene, Emilson, Oreland, Petterson, & Jazin, 2002; Kraschewski et al., 2009; Licinio et al., 2009; Nackley et al., 2009; Shabalina et al., 2009; Wang et al., 2004; Zhang et al., 2006; Zion et al., 2006).

Of the 38 SNPs that were selected on these grounds:

- 3 SNPs were excluded due to design failures, that is, these SNPs could not be included in any of the multiplex assays due to neighboring SNPs or overlapping sequences (BDNF rs56820186, BDNF rs11030103, DRD2 rs1799732) (also see Bakker et al., 2014);
- 3 SNPs were excluded due to genotyping failure, that is, these SNPs had a call rate $<.90$ (COMT rs2097603, COMT rs6267, DRD4 rs747302);
- 1 SNP was excluded because according to information from the Database of Single Nucleotide Polymorphisms (dbSNP; <http://www.ncbi.nlm.nih.gov/SNP/>) its variation was suspected to be a false positive due to artifacts of the presence of a paralogous sequence in the genome, or because evidence suggested sequencing error or computation artifacts to be a false positive (BDNF rs28722151) (also see Bakker et al., 2014);
- 1 SNP was excluded because of a software error in meaningfully determine the genotypes. The gene in question was located on the X-chromosome resulting in a total of five possible genotypes (A and G in men and AA, AG, and GG in women). The employed software was unable to analyze this (MAO-B rs1799836);
- 6 SNPs had to be excluded due to no or hardly any variation, that is, these SNPs either presented with only two genotypes in both samples and one of the remaining two genotypes was represented with < 10 in both samples (BDNF rs57083135, BDNF rs12273539, COMT rs769224, COMT rs740602, COMT rs8192488) or presented with only one genotype in the sample of individuals with residual depressive symptoms (COMT rs165599).

This resulted in the inclusion of 24 SNPs. All of these SNPs were in Hardy-Weinberg equilibrium as assessed with the GENHWI command in STATA 12; $p < .01$.

Analyses

First, main effects were analyzed, that is i) the association between event pleasantness and PA (testing the main effect of daily life pleasant events on positive affect), ii) the association between each of the 24 SNPs and PA (testing the main effect of genetic variation on positive affect), and iii) the association between each of the 24 SNPs and event pleasantness (testing gene-environment correlations). Finally, interaction effects were analyzed to investigate whether the effect of event pleasantness on PA was moderated by any of the 24 SNPs. These analyses were performed separately in each of the two samples.

For the genotypic effects we adopted a co-dominant model which does not assume superior function of one allele over the other or a linear effect of number of alleles (Peerbooms et al., 2011). For this purpose all SNPs were coded in 0, 1, 2 format (0 = most frequent homozygous genotype, 1 = heterozygous genotype, 2 = least frequent homozygous genotype) and entered as two dummy variables in the statistical models. Tests for SNPs were therefore based on χ^2 -tests with two degrees of freedom (for full statistical model please see below). The analyses were carried out for each SNP separately and therefore were corrected for multiple testing. However, most SNPs within one gene were significantly correlated, (e.g., indicating that the SNPs are not independently inherited), and can therefore not be regarded as independent predictors in statistical analyses (**Table 1**). To avoid overcorrection for multiple testing, the number of effective tests was estimated to be equal to 18 according to the methods by Li and Ji (2005) and Galway (2009). Correcting for 18 tests resulted in a critical alpha of .00277.

ESM data have a hierarchical structure. The repeated observations (level 1) are nested within individuals (level 2) and for the general population sample these are again nested within a twin pair (level 3). Multilevel regression analyses take this nesting into account. For the current analyses, the command XT MIXED of Stata 12.1 was employed to model both the main effects (i) through (iii) and the interaction effects. The latter was followed by the TEST command (Wald test) to assess whether there was any significant difference between the three possible genotypes. A significant result was followed up by interpreting the interaction effects per genotype by primary stratification. All analyses were controlled for age and sex and allowed for a random intercept at the subject level (level 2), accounting for overall differences in the outcome (PA or event pleasantness) between individuals. For models with event pleasantness as predictor, we also allowed the slope of the association with PA to vary randomly at the subject level (level 2), taking into account that PA may increase differentially when the event becomes more pleasant (even in individuals with the same genotype). The covariance

Table 1. Correlations between all tested SNPs

	BDNF	COMT	CHRM2	DRD2	DRD4	OPRM1	SLC6A3
BDNF	rs11030102	rs11030101	rs11030101	rs11030101	rs11030101	rs11030101	rs11030101
	1						
	rs11030101						
		1					
	rs6265						
			1				
COMT	rs6269						
				1			
	rs4633						
					1		
	rs4680						
						1	
	rs737865						
							1
	rs4818						
CHRM2	rs1824024						
							1
	rs2061174						
	rs324650						
DRD2	rs6276						
							1
	rs6277						
DRD4	rs936461						
	rs1800955						
							1
OPRM1	rs1799971						
	rs3823010						
	rs495491						
	rs609148						
	rs648893						
	rs563649						
SLC6A3	rs6350						
	rs6413429						
	rs6347						

Note. In bold correlations significant at $p < .00277$ (multiple testing corrected alpha)

structure was set to be unstructured, allowing the random intercept and random slope to correlate. For the general population twin sample a random intercept was also added at the twin-level (level 3); adding a random slope at the twin level did not add to the proportion of accounted variance and was therefore not included.

Results

Participants

Of the 621 twins, 451 provided both genotyping material and ESM measures. Previous research has shown that ESM measures are only valid when they are answered within 10 minutes after the signal and can only be meaningfully analyzed if at least 30% of all reports per person are validly answered. Of the original 17,370 reports, 1,016 had to be removed because they were not answered within 10 minutes after the signal, and a total of 16 individuals had less than 30% valid reports. Of the remaining valid reports 5,171 included reports over unpleasant events and could therefore not be included in the current analyses on reward experience. The final general population twin sample therefore consisted of 435 individuals who provided a total of 11,068 valid ESM reports; that is, each participant provided a mean of 25 reports over 5 days of sampling (**Table 2**).

Of the 130 participants with residual depressive symptoms, 126 provided genotyping material and ESM measures. Of the original 6,283 valid ESM reports, 239 had to be removed because they were not answered within 10 minutes after the signal. All participants had sufficient valid ESM reports. Of the remaining valid reports 1,128 were not included in the current analyses on reward experience as they included reports on unpleasant events. The final sample of individuals with residual depressive symptoms therefore consisted of 126 individuals who provided a total of 4,916 valid ESM reports; that is each participant provided a mean of 39 reports over 6 days of sampling (**Table 2**). Please note that these data were drawn from a larger study on the effects of mindfulness based stress-reduction and consisted only of the baseline ESM-assessment (for more detail please see Geschwind et al., 2011).

Table 2. Descriptives

	General population	Sample with residual depressive symptoms
N	435	126
ESM variables		
Valid ESM reports (valid reports without unpleasant events)	16.239 (11.068)	6.044 (4.916)
PA (SD, range)	4.6 (1.22; 1-7)	3.93 (1.35; 1-7)
Event appraisal (SD, range)	1.5 (1.16; 0-3)	1.9 (1; 0-3)
Age		
Age (SD, range)	27.67 (7.94; 18-58)	43.67 (9.66; 19-65)
% female	100	76.98
Mean SCL90 (SD)	32.11 ¹ (28.87)	71.61 ² (41.13)
Civil status		
% Married	36.87	46.15
% Divorced	2.07	13.85
% Other	61.06	40.00
Medication³		
% Yes	0	95.00
% No	100	5.00
Work situation		
% Full time	39.63	28.46
% Part time	20.28	36.92
% other ⁴	40.09	34.58

Note. ¹ Means based on n=399, missing data for rest; ² means based on n=121, missing data for rest, ³ this includes antidepressants for the general population sample and antidepressants, antipsychotics, and anxiolytics for the sample with residual depressive symptoms, ⁴ the remaining include volunteering work, unemployment and 'other' for the general population sample and household, structured activities, education, and non-structured activities for the sample with residual depressive symptoms

Main effects

In both samples, event pleasantness was significantly associated with momentary experienced PA (general population: $\beta = .17, p < .001$; patients with residual depression: $\beta = .34, p < .001$). The more pleasant the event was rated the more PA was experienced.

Three individual SNPs were associated with momentary PA in the general population sample (COMT rs6269, rs4818, OPRM1 rs4954912) and two in the sample of individuals with residual depressive symptoms (DRD4 rs1800955, SCL3A6 rs6350) when a critical alpha of .05 was employed. However, when correcting for multiple testing with the effective number of tests (alpha = .00277), none of these remained significant.

In the general population sample, no SNP was associated with event pleasantness. Two SNPs showed associations with event pleasantness in the sample of individuals with residual depressed symptoms (OPRM1 rs609148, rs648893) employing an alpha of .05, but again, these did not survive the multiple testing correction.

Interaction effects

Table 3 depicts the results of the interaction analyses. In the general population sample, none of the tested SNPs interacted with event pleasantness in predicting PA. BDNF rs11030102 reached a significance of $p=.02$, however, this result did not survive when the critical alpha of .00277 was employed. In the sample of individuals with residual depressive symptoms, a single SNP (OPRM1 rs1799971) reached significance surviving multiple testing correction ($\chi^2(2) = 13.66, p<.001$). When following up this general effect, it became apparent that the individuals heterozygous on this SNP (A/G) ($n=26, n_{\text{observations}}=1.041$) experienced significantly less PA in the context of pleasant events ($\beta=.23, p<.001$) than individuals homozygous for the A-allele ($n=98, n_{\text{observations}}=3789$) ($\beta=.39, p<.001$) with a difference of $\beta=-.16$ ($p<.001$) (see **Figure 1**). The difference between the most frequent homozygous genotype (A/A) and the least frequent homozygous genotype (G/G) ($n=2$) did not reach significance ($\beta=-.25, p=.07$), however; as only two individuals presented with this latter genotype, the result cannot be interpreted.

Figure 1. Difference in reward experience in sample with residual depressive symptoms depending on OPRM1 rs1799971 genotype. PA= positive affect. Genotype G/G was only present in two individuals and is therefore omitted.

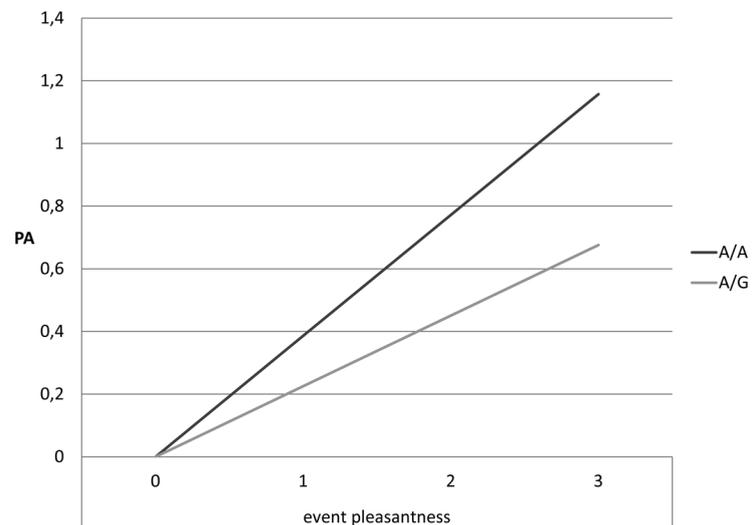


Table 3. Results of SNPs on reward experience

	SNP	General population			Individuals with residual depressive symptoms		
		N	χ^2	p	N	χ^2	P
BDNF	rs11030102	428	7.48	.02*	126	2.40	.30
	rs11030101	428	1.93	.38	126	3.85	.15
	rs6265	426	0.94	.63	126	0.36	.84
COMT	rs6269	428	5.77	.06	126	1.69	.43
	rs4633	427	5.29	.07	126	0.48	.79
	rs4680	427	5.03	.08	126	0.48	.79
	rs737865	428	3.86	.15	126	1.05	.59
	rs4818	409	5.77	.06	126	3.47	.18
CHRM2	rs1824024	428	0.05	.97	126	2.54	.28
	rs2061174	427	0.86	.65	126	0.18	.91
	rs324650	428	0.30	.86	126	0.19	.91
DRD2	rs6276	428	1.65	.44	126	0.30	.86
	rs6277	428	0.20	.91	126	1.63	.44
DRD4	rs936461	427	1.05	.59	126	2.01	.37
	rs1800955	407	0.25	.88	116	1.17	.56
OPRM1	rs1799971	428	1.24	.54	126	13.66	.001
	rs3823010	428	0.32	.85	126	1.82	.40
	rs495491	428	0.48	.79	126	0.85	.65
	rs609148	427	0.54	.77	126	0.21	.90
	rs648893	428	0.61	.74	126	0.21	.90
	rs563649	428	0.32	.85	126	0.85	.65
	rs6350	428	3.36	.19	126	2.58	.28
SLC6A3	rs6413429	428	3.59	.17	126	2.09	.35
	rs6347	428	1.36	.51	126	1.46	.48

Note. In bold results that are significant after multiple testing correction ($p<.00277$). * $p<.05$.

Discussion

The ability to experience reward appears to play a key role in mood disorders and resilience against these (Cohn, Fredrickson, Brown, Mikels, & Conway, 2009; Forbes, 2009; Forbes, Shaw, & Dahl, 2007; Henriques & Davidson, 2000; Pizzagalli, Bogdan, Ratner, & Jahn, 2007; Wichers et al., 2009). To date, research on the genetic effects in reward experience is limited and to the best of our knowledge, this is the first study to look at real-world reward experience in relation to multiple genetic variations of the dopamine and opioid system. Essentially, we found weak evidence for genetic involvement in daily life reward experience. In the general population sample, no SNP impacted on daily life reward experience; whereas one SNP acting on the μ -opioid receptor (OPRM1 rs1799971) influenced the effect of real-world reward in the sample with residual depressive symptoms.

Genetic effects in daily life reward experience

Overall, these results are in line with an earlier *behavioral* genetic study indicating weak – if any – genetic influence on daily life reward experience (Menne-Lothmann et al., 2012). The limited number of studies on experimental reward to date have shown that variations in the COMT Val¹⁵⁸Met and the SLC6A3 genotype may influence brain activity in response to experimental reward (e.g. see Dreher et al., 2009; Haas et al., 2009; Yacubian et al., 2007). Besides the fact that these genetic effects could seldom be demonstrated in the behavioral reward-data (Hahn et al., 2011; Schmack et al., 2008), most importantly they frequently would not have survived the multiple testing correction employed in the current study. This indicates that specific genetic effects may only show when investigated in isolation and become negligible when considered in the larger constellation of many potential genetic variations and multiple testing. This consideration becomes intriguingly obvious with regard to the findings of the COMT Val¹⁵⁸Met genotype. Not only experimental fMRI studies (Dreher et al., 2009; Schmack et al., 2008; Yacubian et al., 2007) but also one real-world study (Wichers, Aguilera, et al., 2008) demonstrated an effect of this genotype on different measures of reward experience when only one or two genetic variations were considered at the same time. In the current analyses, this SNP also demonstrated significant differences between the most common homozygous genotype and both the heterozygous ($\beta=.05$, $p=.049$) and the least common homozygous genotype ($\beta=.06$, $p=.047$) on the experience of positive affect in response to a pleasant daily life experience; however, neither of these effects survived multiple testing correction.

The current results not only add to the existing literature by showing that the effects of single genetic variations may not become obvious when considered in the larger constellation of genetic variation, but also by suggesting a potential role of the OPRM1 gene in daily life reward experience.

The role of OPRM1 rs1799971 in reward experience

In the current study, only one SNP acting on the μ -opioid receptor influenced real-world reward in the sample with residual depressive symptoms and survived multiple testing correction: the OPRM1 rs1799971 SNP. The OPRM1 gene codes for the μ -opioid receptor. The rs1799971 SNP is a functional adenine to guanine substitution, which results in an asparagine (Asn) to aspartic (Asp) amino acid substitution. The effects include an increased binding to beta-endorphins, a reduction in messenger RNA, a drastic reduction in OPRM1 protein level (Garriock et al., 2010) and a downregulation of opioid release (Ray et al., 2011). Consistent with this, in the current sample with residual depressive symptoms, those individuals with at least one G-allele showed significantly less reward experience than individuals with two A-alleles. Previous research has distinguished between two main aspects of reward: the motivational aspect or ‘wanting’ and the consumption aspect of reward or ‘liking’. Whereas the ‘wanting’ aspect is usually ascribed to dopamine action, ‘liking’ has been more associated with opioid functioning (Berridge & Robinson, 1998; Berridge et al., 2009). It may be the case that the current concept of real-world reward experience is closer to the experience of ‘liking’ than ‘wanting’, because it measured mainly the enjoyment of pleasant daily life events without addressing the ‘wanting’ of these events. This may explain why we find a functional SNP involved in opioid neurotransmission being associated with this measure of reward experience. However, as opioids stimulate dopamine in mesolimbic areas (e.g. Diltz & Kalivas, 1989; van Ree et al., 1999; Wise & Rompre, 1989), the less active G- (or Asn) genotype furthermore results in decreased dopamine levels (Chen et al., 2009). The current results could therefore likewise reflect reduced dopamine action. Future studies need to consider how the ‘wanting’ aspect can be captured in daily life assessments of reward and how both the ‘wanting’ and ‘liking’ aspects in daily life can be clearly distinguished from each other.

The current results are in good agreement with previous research, which was mainly conducted in the area of pain and addiction research. A meta-analysis suggests that the G-allele increases the risk for opioid dependence (Haerian & Haerian, 2013). Other research not only found that the presence of a G-allele increased cue-induced craving for alcohol in male heavy drinkers (van den Wildenberg et al., 2007) but also increased the likelihood for alcohol dependency (Bart et al., 2005). Although not all

studies support this association (e.g. Garriock et al., 2010), the findings may indicate that the G-genotype of the rs1799971 OPRM1 SNP reduces the capability to naturally experience (enough) reward, which may result in compensation with drugs that stimulate the opioid system and dopamine release. Crucially, the current study is the first to demonstrate that this effect also translates to real-world experience of natural rewards and survives multiple testing corrections for other potentially important reward SNPs. Furthermore, a second study of our group showed additional evidence for the role of the OPRM1 gene in real-world positive affective experiences using a within-person design. The SNPs (rs495491, rs609148, rs648893, rs3823010) were shown to influence the increase in positive affect that individuals could gain from mindfulness based cognitive therapies (Bakker et al., 2014).

Although results match with theoretical notions and previous findings, the effect of the SNP in question was not consistent across study populations. Thus, OPRM1 rs1799971 did not affect reward experience in the general population sample. Although this fact weakens the robustness of the finding, it is not *necessarily* a sign of non-replication. It may also indicate that the expression of this genetic effect is conditional on other environmental or genetic factors. The impact of a single genetic variation is likely small. It is possible that in the general population, the expression of such genetic effects is obscured by other factors influencing positive affective experience. It may be speculated that individuals with depressive symptoms have less ability to compensate for a (genetically) reduced reward experience in their daily lives. Not being able to compensate a somewhat lower than average reward response may relate to underlying environmental or genetic vulnerability for depressive symptoms. Other sources of positive emotions can be used by vulnerable people to compensate, such as holding on to positive emotions over time (Höhn et al., 2013), learning to obtain more positive emotions from pleasant situations (Geschwind et al., 2011) or finding environments or new behavioral patterns that benefit positive emotional experience. It has previously been shown that even small environmental influences like behavioral interventions can indeed prevent the expression of the genetic effect of OPRM1 rs1799971 on medical treatment response (Anton et al., 2008). We therefore do not necessarily expect to find the same effects in both samples.

Limitations

The current findings need to be viewed in light of several methodological considerations. First, the number of participants was moderate ($N_{\text{general population}}=435$, $N_{\text{residual depressive symptoms}}=126$), but not very large in the context of genetic studies. This fact may have resulted in low power to detect relevant effects. However, instead of one

phenotype measure, we had on average 37 measurements per person in the general population sample and 48 in the sample with residual depressive symptoms, yielding a total of 16.239 and 6.044 observations, respectively. Still, theoretically, it remains possible that the sample size to detect a small effect of OPRM1 rs1799971 in the general population sample was not large enough. We consider this unlikely, as it was the larger of the two samples that did not show an association. Additionally, not only the significance but more importantly the effect sizes differed between the samples (general population sample: $\beta_{A/A \text{ vs } A/G} = .01$, $\beta_{A/A \text{ vs } G/G} = .07$; sample with residual depressive symptoms: $\beta_{A/A \text{ vs } A/G} = -.16$, $\beta_{A/A \text{ vs } G/G} = -.25$).

Another related consideration is that effects may not have been detected due to another reason. Likely, many factors influence reward experience in a complex way, including many gene-gene and gene-environmental interactions. As explained in the previous section, environmental influences may be able to obscure genetic expression. Therefore, it is possible that true genetic effects were hidden not only in one but in both samples examined. Likewise, these findings do not exclude the possibility that some genes come to expression only within certain environmental contexts or within a certain genetic background.

Third, the role of sex remains unclear in the current study. This is because only females were assessed in the general population sample, whereas the sample with residual depressive symptoms was mixed. Although the analyses in the latter were controlled for sex, future research may consider taking a mixed sample to replicate associations.

References

- Anton, R., Orozi, G., O'Malley, S., Couper, D., Swift, R., Pettinati, H., & Goldman, D. (2008). An evaluation of mu-opioid receptor (OPRM1) as a predictor of naltrexone response in the treatment of alcohol dependence: results from the combined pharmacotherapies and behavioral interventions for alcohol dependence (COMBINE) study. *Arch Gen Psychiatry*, *65*(2), 135-144.
- APA. (2000). *Quick reference to the diagnostic criteria from DSM-IV-TR*. Arlington, VA: American Psychiatric Association.
- Arias-Carrión, O., Stamelou, M., Murillo-Rodríguez, E., Menéndez-González, M., & Pöppel, E. (2010). Dopaminergic reward system: a short integrative review. *International Archives of Medicine*, *3*(24), 1-6.
- Azzato, E. M., Morton, L. M., Bergen, A. W., Wang, S. S., Chatterjee, N., Kvale, P., . . . Caporaso, N. E. (2009). SLC6A3 and body mass index in the prostate, lung, colorectal and ovarian cancer screening trial. *BMC Medical Genetics*, *10*(9), 1-9.
- Bakker, J. M., Lieveise, R., Menne-Lothmann, C., Kenis, G., Geschwind, N., Peeters, F., . . . Wichers, M. (2014). Therapygenetics in mindfulness-based cognitive therapy: do genes impact on therapy-induced change in real-life reward experience? *Translational Psychiatry*, *4*:e384.
- Balciuniene, J., Emilson, L., Orelund, L., Petterson, U., & Jazin, E. E. (2002). Investigation of the functional effect of monoamine oxidase polymorphisms in human brain. *Human Genetics*, *110*, 1-7.
- Bart, G., Kreek, M. J., Ott, J., LaForge, K., Proudnikov, D., Pollak, L., & Heilig, M. (2005). Increased attributable risk related to a functional mu-opioid receptor gene polymorphism in association with alcohol dependence in central Sweden. *Neuropsychopharmacology*, *30*, 417-422.
- Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, *28*, 309-369.
- Berridge, K. E., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: 'linking', 'wanting', and 'learning'. *Current opinion in Pharmacology*, *9*(65-73).
- Bogdan, R., & Pizzagalli, D. A. (2009). The heritability of hedonic capacity and perceived stress: a twin study evaluation of candidate depressive phenotypes. *Psychological Medicine*, *39*, 211-218.
- Camara, E., Krämer, U. M., Cunillera, T., Marco-Pallarés, J., Cucurell, D., Nager, W., . . . Münte, T. F. (2010). The effects of COMT (Val108/158) and DRD4 (SNP - 521) dopamine genotypes on brain activations related to valence and magnitude of rewards. *Cerebral Cortex*, *20*, 1985-1996.
- Chen, A. L.-C., Chen, T. J. H., Waite, R. L., Reinking, J., tung, H. L., Rhoades, P., . . . Blum, K. (2009). Hypothesizing that brain reward circuitry genes are genetic antecedents of pain sensitivity and critical diagnostic and pharmacogenomic treatment targets for chronic pain conditions. *Medical Hypotheses*, *72*, 14-22.
- Cohn, M. A., Fredrickson, B. L., Brown, S. L., Mikels, J. A., & Conway, A. M. (2009). Happiness unpacked: positive emotions increase life satisfaction by building resilience. *Emotion*, *9*(3), 361-368.
- Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *J Nerv Ment Dis*, *175*, 526-536.
- Delespaul, P. (1995). *Assessing schizophrenia in daily life. The experience sampling method*. Maastricht: IPSE Foundation.
- Derom, C., Thiery, E., Peeters, H., Vlietinck, R., Defoort, P., & Fijns, J. P. (2013). The East Flanders prospective twin survey (EFPTS). An actual perception. *Twin Research and Human Genetics*, *16*(1), 58-63.
- Dilts, R. P., & Kalivas, P. W. (1989). Autoradiographic localization of mu-opioid and neurotensin receptors within the mesolimbic dopamine system. *Brain Research*, *488*, 311-327.
- Dreher, J.-C., Kohn, P., Kolachana, B., Weinberger, D. R., & Berman, K. F. (2009). Variation in dopamine genes influences responsivity of the human reward system. *PNAS*, *106*(2), 617-622.
- Forbes, E. E. (2009). Where's the fun in that? Broadening the focus on reward function in depression. *Biological Psychiatry*, *66*, 199-200.
- Forbes, E. E., Shaw, D. S., & Dahl, R. E. (2007). Alterations in reward-related decision making in boys with recent and future depression. *Biological Psychiatry*, *61*, 633-639.
- Galwey, N. W. (2009). A new measure of the effective number of tests, a practical tool for comparing families of non-independent significance tests. *Genetic Epidemiology*, *33*(7), 559-568.
- Garriock, H. A., Tanowitz, M., Kraft, J., Dang, V., Peters, E., Jenkins, G., . . . Hamilton, S. (2010). Association of Mu-opioid receptor variants and response to citalopram treatment in major depressive disorder. *Am J Psychiatry*, *167*(5), 565-573.
- Geschwind, N., Nicolson, N. A., Peeters, F., van Os, J., Barge-Schaapveld, D. Q. C. M., & Wichers, M. (2011). Mindfulness training increases momentary positive emotions and reward experience in adults vulnerable to depression. A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *79*(5), 618-628.
- Geschwind, N., Peeters, F., Jacobs, N., Delespaul, P., Derom, C., Thiery, E., . . . Wichers, M. (2010). Meeting risk with resilience: high daily life reward experience preserves mental health. *Acta Psychiatrica Scandinavica*, *122*(129-138).
- Haas, B. W., Mills, D., Yam, A., Hoeft, F., Bellugi, U., & Reiss, A. (2009). Genetic influences on sociability: heightened amygdala reactivity and event-related responses to positive stimuli in Williams Syndrome. *The Journal of Neuroscience*, *29*(4), 1132-1139.
- Haerian, B. S., & Haerian, M. S. (2013). OPRM1 rs1799971 polymorphism and opioid dependence: evidence from a meta-analysis. *Pharmacogenomics*, *14*(7), 813-824.
- Hahn, T., Heinzel, S., Dresler, T., Plichta, M. M., Renner, T. J., Markulin, F., . . . Fallgatter, A. J. (2011). Association between reward-related activation in the ventral striatum and trait reward sensitivity is moderated by dopamine transporter gene. *Human Brain Mapping*, *32*(1557-1563).
- Henriques, J. B., & Davidson, R. J. (2000). Decreased responsiveness to reward in depression. *Cognition and Emotion*, *14*(5), 711-724.
- Höhn, P., Menne-Lothmann, C., Peeters, F., Nicolson, N. A., Jacobs, N., Derom, C., . . . Wichers, M. (2013). Moment-to-moment transfer of positive emotions in daily life predicts future course of depression in both general population and patient samples. *PLoS ONE*, *8*(9), e75655.
- Jacobs, N., Nicolson, N. A., Derom, C., Delespaul, P., van Os, J., & Myin-Germeys, I. (2005). Electronic monitoring of salivary cortisol sampling compliance in daily life. *Life Sciences*, *76*, 2431-2443.
- Jacobs, N., Rijdsdijk, F., Derom, C., Vlietinck, R., Delespaul, P., van Os, J., & Myin-Germeys, I. (2006). Genes making one feel blue in the flow of daily life: a momentary assessment study of gene-stress interaction. *Psychosomatic Medicine*, *68*, 201-206.
- Kraschewski, A., Reese, J., Angheliescu, I., Winterer, G., Schmidt, L. G., Gallinat, J., . . . Wernicke, C. (2009). Association of the dopamine D2 receptor gene with alcohol dependence: haplotypes and subgroups of alcoholics as key factors for understanding receptor function. *Pharmacogenetics and Genomics*, *19*, 513-527.
- Lambert, K. (2006). Rising rates of depression in today's society: Consideration of the roles of effort-based rewards and enhanced resilience in day-to-day functioning. *Neuroscience & Behavioral Reviews*, *30*(4), 497-510.
- Li, J., & Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, *95*(3), 221-227.

- Licinio, J., Dong, C., & Wong, M. (2009). Novel sequence variations in the brain-derived-neurotrophic factor gene and association with major depression and antidepressant treatment response. *Archives of General Psychiatry*, *66*(5), 488-497.
- Menne-Lothmann, C., Jacobs, N., Derom, C., Thiery, E., van Os, J., & Wichers, M. (2012). Genetic and environmental causes of individual differences in daily life positive affect and reward experience and its overlap with stress-sensitivity. *Behavior Genetics*, *42*, 778-786.
- Nackley, A. G., Shabalina, S. A., Lambert, J. E., Conrad, M. S., Gibson, D. G., Spiridonov, A. N., . . . Diatchenko, L. (2009). Low enzymatic activity haplotypes of the human catechol-o-methyltransferase gene: enrichment for marker SNPs. *PLoS ONE*, *4*(4), e7237.
- Peerbooms, O., van Os, J., Drukker, M., Kenis, G., Hoogveld, L., Group, M. i. P., . . . Rutten, B. (2011). Meta-analysis on MTFHR gene variants in schizophrenia, bipolar disorder and unipolar depressive disorder: evidence for a common genetic vulnerability. *Brain, Behavior, and Immunity*, *25*, 1530-1543.
- Pizzagalli, D. A., Bogdan, R., Ratner, K. G., & Jahn, A. L. (2007). Increased perceived stress is associated with blunted hedonic capacity: potential implications for depression research. *Behavior Research and Therapy*, *45*, 2742-2753.
- Ray, R., Ruparel, K., Newberg, A., Wileyto, E. P., Loughhead, J. W., Divgi, C., . . . Lerman, C. (2011). Human Mu Opioid (OPRM1 A118G) polymorphism is associated with brain mu-opioid receptor binding potential in smokers. *PNAS*, *108*(22), 9268-9273.
- Rutter, M., Moffitt, T. E., & Caspi, A. (2006). Gene-environment interplay and psychopathology: multiple varieties but real effects. *Journal of Child Psychopathology and Psychiatry*, *47*(3/4), 226-261.
- Schmack, K., Schlagenhauf, F., Sterzer, P., Wrase, J., Beck, A., Dembler, T., . . . Gallinat, J. (2008). Catechol-O-methyltransferase val158met genotype influences neural processing of reward anticipation. *NeuroImage*, *42*, 1631-1638.
- Shabalina, S. A., Zaykin, D. V., Gris, P., Ogurtsov, A. Y., Gauthier, J., Shibita, K., . . . Diatchenko, L. (2009). Expansion of the human m-opioid receptor gene architecture: novel functional variants. *Human Molecular Genetics*, *18*(6), 1037-1051.
- van den Wildenberg, E., Wiers, R. W., Dessers, J., Janssen, R., Lambrichs, H., Smeets, H., & van Breukelen, G. (2007). A functional polymorphism of the mu-opioid receptor gene (OPRM1) influences cue-induced craving for alcohol in male heavy drinkers. *Alcoholism: Clinical and Experimental Research*, *31*(1), 1-10.
- van Ree, J. M., Gerrits, M. A. F. M., & Vanderschuren, L. J. M. J. (1999). Opioids, reward and addiction: an encounter of biology, psychology, and medicine. *Pharmacological Reviews*, *51*(2), 342-396.
- Wang, J. S., Hinrichs, A. L., Stock, H., Budde, J., Allen, R., Bertelsen, S., . . . Bierut, L. J. (2004). Evidence of common and specific genetic effects: association of the muscarinic acetylcholine receptor M2 (CHRM2) gene with alcohol dependence and major depressive syndrome. *Human Molecular Genetics*, *13*(17), 1903-1911.
- WHO. (2013). Depression. Key facts. Retrieved 24.07.2013 from <http://www.who.int/mediacentre/factsheets/fs369/en/2013>
- Wichers, M., Aguilera, M., Kenis, G., Krabbendam, L., Myin-Germeys, I., Jacobs, N., . . . van Os, J. (2008). The catechol-O-methyl transferase val158met polymorphism and experience of reward in the flow of daily life. *Neuropsychopharmacology*, *33*, 3030-3036.
- Wichers, M., Barge-Schaapveld, D. Q. C. M., Nicolson, N. A., Peeters, F., de Vries, M., Mengelers, R., & van Os, J. (2009). Reduced stress-sensitivity or increased reward experience: the psychological mechanism of response to antidepressant medication. *Neuropsychopharmacology*, *34*, 923-931.
- Wichers, M., Kenis, G., Jacobs, N., Mengelers, R., Derom, C., Vlietinck, R., & van Os, J. (2008). The BDNF Val66Met x 5-HTTLPR x child adversity interaction and depressive symptoms: an attempt at replication. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, *147B*, 120-123.
- Wichers, M., Kenis, G., Jacobs, N., Myin-Germeys, I., Derom, C., & Vlietinck, R. (2008). The psychology of psychiatric genetics: evidence that positive emotions in females moderate genetic sensitivity to social stress associated with the BDNF Val(66)Met polymorphism. *Journal of Abnormal Psychology*, *117*(699-706).
- Wise, R. A., & Rompre, P.-P. (1989). Brain dopamine and reward. *Annual Review of Psychology*, *40*, 191-225.
- Yacubian, J., Sommer, T., Schroeder, K., Gläscher, J., Kalisch, R., Leuenberger, B., . . . Büchel, C. (2007). Gene-gene interaction associated with neural reward sensitivity. *PNAS*, *104*(19), 8125-8130.
- Zhang, H., Luo, X., Kranzler, H. R., Lappalainen, J., Yang, B., Krupitsky, E., . . . Gelernter, J. (2006). Association between two m-opioid receptor gene (OPRM1) haplotype blocks and drug or alcohol dependence. *Human Molecular Genetics*, *15*(6), 807-819.
- Zion, I. Z., Tessler, R., Cohen, L., Lerer, E., Raz, Y., Bachner-Melman, R., . . . Ebstein, R. P. (2006). Polymorphisms in the dopamine D4 receptor gene (DRD4) contribute to individual differences in human sexual behavior: desire, arousal and sexual function. *Molecular Psychiatry*, *11*, 782-786.

Psychological and biological validation of a novel digital social peer evaluation experiment (digi-SPEE)

Menne-Lothmann, C.*¹; Decoster, J.^{1,2}; van Winkel, R.^{1,2}; Collip, D.^{#1}; van Os, J.^{1,3}; Rutten, B.P.F.¹;
Delespaul, P.¹; De Hert, M.²; Derom, C.^{4,5}; Thiery, E.⁶; Jacobs, N.^{1,7}; and Wichers, M.^{1,8}

¹ Department of Psychiatry and Psychology, South Limburg Mental Health Research and Teaching Network, EURON,
Maastricht University, The Netherlands

² University Psychiatric Centre KU Leuven, Belgium

³ King's College London, King's Health Partners, Department of Psychosis Studies, Institute of Psychiatry, London, United
Kingdom

⁴ Centre of Human Genetics, University Hospitals Leuven, Belgium;

⁵ Department of Human Genetics, KU Leuven, Belgium

⁶ Department of Neurology, Ghent University Hospital, Ghent University, Ghent, Belgium

⁷ Faculty of Psychology, Open University of the Netherlands, Heerlen, The Netherlands

⁸ Department of Psychiatry, University Medical Centre of Groningen, Groningen University Groningen, The Netherlands

* contributed equally; # contributed equally

submitted for publication



Abstract

Negative social evaluation has been associated with distress and psychopathological symptoms. Given frequent evaluation through increasingly prevalent virtual social networks, a better understanding of the effects of such social evaluation is urgently required.

A new digital social peer evaluation experiment (*digi-SPEE*) was developed mimicking everyday life online social interactions between peers. Participants received mild negative feedback on their appearance, intelligence, and congeniality. The current aim was to validate whether this task would act as a social evaluation stressor in terms of biological, implicit, and explicit psychological changes relevant to mechanisms of psychopathology.

Two-hundred-and-forty-one young participants (58.9% female, aged 18.9 years (15-34)), from an ongoing novel general population twin study, participated. Positive affect (PA), negative affect (NA), implicit self-esteem, and cortisol were assessed before and after exposure to the social evaluation experiment.

The social evaluation experiment decreased PA ($B=-5.25$, $p<.001$) and implicit self-esteem ($B=-.19$; $p<.001$), whereas it increased NA ($B=5.99$; $p<.001$) and cortisol levels ($B=.07$; $p<.001$). Females ($PA: B=-7.62$; $p<.001$; $NA: B=8.28$; $p<.001$) and participants with higher levels of general psychological distress ($PA: B=-.04$, $p=.035$; $NA: B=.06$; $p=.028$) showed stronger affective responses. Stressor-induced cortisol increase was stronger in adolescents under the age of 18 as compared to those 18 years and older ($B=-.06$, $p=.002$).

The current social evaluation experiment mimicking peer-evaluations occurring in online social networks induced significant stress on psychological (affect, implicit self-esteem) and biological (cortisol) stress measures. The experiment is sensitive to expected differences in specific populations regarding sensitivity to negative social evaluation. Longitudinal designs are required to demonstrate predictive value.

Keywords social evaluation; psychological stress; psychopathology; experimental design; cortisol; risk; resilience

Belongingness and being accepted by peers are basic human needs (Williams, Cheing, & Choi, 2000). Humans are vigilant to social evaluation (Dickerson & Kemeny, 2004). Today social interactions occur increasingly in online social networks – particularly among young people (e.g., see Donnerstein, 2012; Pempek, Yermolayeva, & Calvert, 2009; Sommerville, 2013). These networks offer the possibility of disclosing personal information and evaluating others, “including likes/dislikes, hobbies, and personal musings via ‘wall posts’, and ‘status updates’” (Gonzales & Hancock, 2011, p. 79). The need to belong and self-presentation are the two main reasons for participating in online social networks (Nadkarni & Hoffmann, 2012) and research indicates that these desires are partially met. For example, time spent in online social networks can enhance explicit self-esteem when watching one’s own profile (Gonzales & Hancock, 2011), may foster healthy identity development in adolescents (Reid & Boyer, 2013), and supports individuals with low self-esteem to build social capital (Steinfeld, Ellison, & Lampe, 2008). It may then not be surprising that individuals disclose a considerable amount of sensitive personal information in their social network profiles (Nosko, Wood, & Molema, 2010). However, this may also put people at risk for negative social evaluation (Nosko et al., 2010). For example, there is emerging evidence that bullying over the internet has even more disastrous consequences than its face-to-face counterpart, possibly because it occurs at a greater psychological distance and is more anonymous (Donnerstein, 2012). Therefore, a better understanding of the effects of the most typical feature of online social networks, namely purely being evaluated by peers on personal characteristics, is urgently required.

Laboratory studies have reliably demonstrated that social rejection/exclusion stressors, which usually include that the participant is rejected from joining any kind of group, induce significant stress and yield affective, self-esteem, and cortisol reactions (Blackhart, Eckel, & Tice, 2007; Blackhart, Nelson, Knowles, & Baumeister, 2009; Buckley, Winkel, & Leary, 2004; Eisenberger, Jarcho, Lieberman, & Nalibhoff, 2006; Hsu et al., 2013; Maner, Miller, Schmidt, & Eckel, 2010; Moor, Crone, & van der Molen, 2010; Nesdale & Lambert, 2007; Sandstorm, Cillessen, & Eisenhower, 2003; Williams et al., 2000). Other laboratory studies demonstrated that performance evaluation stressors, which usually include that the participant is evaluated on performance by a panel or peers, increases the stress-hormone cortisol (Booij, Bouma, de Jonge, Ormel, & Oldehinkel, 2012; Dickerson & Kemeny, 2004; Dickerson, Mycek, & Zaldivar, 2008; Elzinga et al., 2008; Kirschbaum, Pirke, & Hellhammer, 1993; van Eck, Nicolson, Berkhof, & Sulon, 1996). Research has indicated that stress-responses differ between social rejection/exclusion and performance evaluation stressors (Stroud et al., 2009).

As online social interactions include selective elements of both kinds of stressors, i.e. feedback by peers as in social rejection/exclusion stressors, and social evaluation as in performance evaluation stressors, there is good reason to assume that slightly negative evaluation of personal characteristics by peers imposes significant stress.

Moreover, exposure to social stressors has been associated with psychopathological symptoms (e.g., see Björkqvist, 2001; Selten & Cantor-Graae, 2005; Strout, Salovey, & Epel, 2002; Williams et al., 2000) and sensitivity to such stressors may be enhanced in individuals at risk for psychopathological symptoms (e.g., see Maner et al., 2010; Sommerville, 2013). The group that most frequently interacts in online social networks are adolescents (Donnerstein, 2012; Pempek et al., 2009; Sommerville, 2013), who show increased importance for peer-relationships (Sommerville, 2013) and are particularly vulnerable to develop mood-symptoms (Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993). This adds urgency to tools that can assess sensitivity to digital social negative evaluation of personal characteristics by peers in an accurate and sensitive way.

Therefore, we developed a novel digital *social peer evaluation experiment (digi-SPEE)*. To realistically mimic online social network interactions, it is important to “deliver self-relevant, salient, and believable social feedback” (Sommerville, 2013, p.123). We therefore took into account that social evaluation as experienced in online social networks includes a greater degree of psychological distance than face-to-face contact, is generally based on personal traits as plotted in the individual’s online profile, and includes feedback by peers. The experiment exposes participants to subtle negative digital social peer-evaluation of fundamental personal characteristics (appearance, intelligence, and stance in life).

The aim of the current study was to validate this experiment as a negative social evaluation stressor in three ways. First, the experiment was hypothesized to elicit psychological and biological responses following exposure to subtle negative social evaluation. As psychological changes are not always directly observable or measurable when asked explicitly, both explicit and implicit measures are required. As an explicit measure of sensitivity to negative social evaluation we used subjective ratings of affect (Blackhart et al., 2007; Buckley et al., 2004; Nesdale & Lambert, 2007; Strout et al., 2002; Williams et al., 2000). In order to obtain an implicit response we measured self-esteem implicitly using an implicit association task (Bosson, Swann, & Pennebaker, 2000). And cortisol was included as a biological response measure (Booij et al., 2012; Brooks & Robles, 2009; Dickerson & Kemeny, 2004; Jacobs et al., 2007).

Secondly, we aimed at demonstrating that the current experiment is sensitive to individual differences to social stress sensitivity. Previous research has indicated that women are more sensitive to peer-induced stress than men (Strout et al., 2002). Also young people, i.e. adolescents, are assumed to be particularly sensitive to social stress situations (Collins & Thomas, 1972; Sommerville, 2013). Finally, individuals with psychopathological problems have been indicated to show increased reactions to social stressors (Booij et al., 2012; Cuellar & Johnson, 2009; Hankin, Badanes, Abela, & Watamura, 2010). Therefore, this study examines whether the current experiment is sensitive to these demographic differences.

Finally, we aimed at showing that responses to the current experiment are specifically due to the social evaluation exposure (i.e. the idea of being evaluated by others) and not to negative stimuli *per se*, as employed in the experiment.

Methods

Participants

Participants were recruited in the context of a novel general adolescent/young adult sampling frame from a population twin register (East Flanders Prospective Twin Survey, EFPTS), which prospectively registers multiple births from 1964 onwards (Derom et al., 2013). In order to oversample adolescent participants, twins between 15 and 18 years of age were sent letters inviting them to participate in an ongoing longitudinal research project called ‘TwinssCan’, which started in April 2010. Additionally, all twins and their sibs between 15 and 34 years were eligible to participate and could register via the twin register newsletter. Two hundred and forty-one subjects (58.9% female; 18.8 years) participated (see **Table 1** for sample demographics).

Table 1. Demographics of both samples and difference between the samples

	Deceived sample		Undeceived sample		P
	N	Mean (SD)	N	Mean (SD)	
Demographics					
% female	241	58.92	25	76.00	.10 ¹
Age (years) ²	241	18.76 (4.54)	25	20.52 (4.77)	.07 ¹
SCL-90-R	239	41.47 (35.25)	24	35.83 (28.52)	.45 ¹
% white ethnicity	239	99.17	25	100	.64 ¹
Work situation					.24 ³
% household	241	0	25	1	
% studies	241	80.91	25	68.00	
% regular work	241	1.66	25	8.00	
Highest level of completed education					.24 ³
% lower education	241	0	25	0	
% secondary education	241	55.19	25	44.00	
% bachelor	241	24.90	25	32.00	
% master	241	18.26	25	16.00	

Note. SCL-90-R = Symptom Checklist 90 – Revised (higher scores = more psychological distress during the last week), ¹based on independent sample t-tests, ²median = 18, range 15-34 years. ³ based on Chi-Square tests.

Instruments for assessing cortisol and self-esteem (SC-IAT, see below) were only available starting from August and September 2010, respectively, resulting in 202 participants (59.4% female; 18.7 years) for analyses on cortisol and 187 participants (53.5% female; 18.2 years) for analyses on self-esteem. Additionally, one participant was excluded due to very high (and physiologically implausible) cortisol values (>40 nmol/L) (Jacobs et al., 2007; van Eck et al., 1996). For another individual, post-social evaluation data were missing for all indices due to technical recording failure (see **Table 2** for N, means, SD for all outcome measures). All siblings of twins were invited to participate. The undeceived subsample consisted of a convenient sample of those 25 siblings of twins (76% female; 20.5 years) that agreed to participate (**Table 1**). Approval from the local Ethics Committee was obtained (Commissie Medische Ethiek van de Universitaire ziekenhuisen KULeuven, Nr. B32220107766). Participants provided written informed consent before study inclusion. In case participants were younger than 18 years, parents provided additional informed consent.

Table 2. N, Means, SD, of outcome measures for both samples and p-value of difference between samples

	Deceived sample		Undeceived sample		P
	N	Means (SD)	N	Means (SD)	
Outcome measures					
PA _{pre}	241	50.82 (13.75)	25	52.96 (13.71)	.46
PA _{post}	240	45.60 (16.67)	25	50.45 (12.92)	.093
NA _{pre}	241	10.85 (9.72)	25	9.85 (8.82)	.60
NA _{post}	240	16.76 (15.33)	25	12.53 (13.69)	.16
Self-esteem _{pre}	187	0.36 (0.35)	25	0.39 (0.35)	.69
Self-esteem _{post}	186	0.17 (0.28)	23	0.28 (.27)	.076
Cortisol t ₁	202	1.78 (0.61)	24	1.86 (0.58)	.50
Cortisol t ₂	200	1.75 (0.62)	24	1.74 (0.57)	.95
Cortisol t ₃	201	1.89 (0.67)	23	1.72 (0.56)	.24
Cortisol t ₄	201	1.90 (0.68)	23	1.78 (0.61)	.35

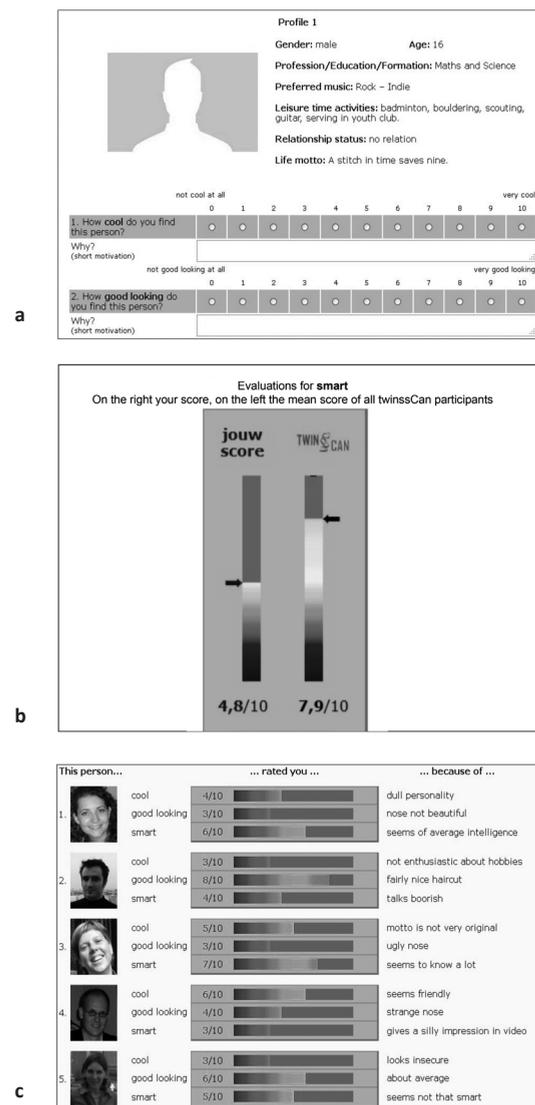
Note. Cortisol is depicted in log-transformed values. pre = before social evaluation, post = after social evaluation, PA = positive affect (higher scores = more positive affect), NA = negative affect (higher scores = more negative affect); t₁ = before social evaluation, t₂ = 10min, t₃ = 20min, and t₄ = 30min after social evaluation.

The digital social peer evaluation experiment (digi-SPEE)

Participants were told that the general aim of the experiment was to investigate why people like or dislike each other based on short profiles and videos. To this end, participants were asked to provide a short written profile and to record a short video (≤ 1 min) introducing themselves, including information about age, profession or education, preferred music, leisure time activities, relationship status, and stance in life. For the same purpose, participants were asked to rate, on a 7-point rating scale (higher scores more positive), the profiles and videos of five other study participants regarding appearance, intelligence and congeniality, and to give a short reason for their rating (**Fig. 1a**). Although participants were told that they would see videos of other study participants, in reality they were presented with prepared videos from five volunteering peers from the same population (female and male individuals between 15 and 34 years from East Flanders, all white). Depending on the age and gender of participants, they were confronted with different sham-peers. These peers were always age (+/- 1 year)-, and gender (three of the same gender and two of the opposite gender as participant)- matched to each single participant. These peers did not in any way participate in the study themselves; their videos were recorded before study onset. Also, in reality the participant's profile and video were never shown to anyone else, so privacy was kept intact throughout the procedure. Twin couples were usually

tested simultaneously but separately by two trained research assistants to prevent them talking about the experiment before social evaluation had taken place. If sibs volunteered as well, they were usually tested immediately after the twins, again, to avoid them talking to one another. The experiment consisted of two sessions which both took place in participants' homes and were conducted by the same experimenter.

Figure 1. Profiles and evaluations of online social peer interaction. Profiles about peers that are rated by the participant (a). Overall comparison between the participant and all other study participants (b). Detailed evaluations per rater about the participant (c).



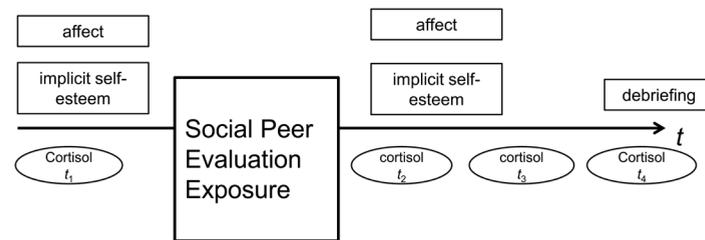
Session 1

During the first session, participants entered their profile into a pre-programmed computerized form and recorded their video with a webcam. Participants furthermore evaluated the profiles and videos of three of the five sham participants (**Fig. 1a**). Participants were told that their profile and video would be shown to the same five participants for evaluation between the sessions, and that they could learn how they had been evaluated by their peers during the second session.

Session 2

The second session took also place at participants' homes several days later, started approximately an hour before the social evaluation exposure took place, and was conducted by the same researcher individuals had seen a week before, to prevent elevated cortisol-levels at the start of the experiment due to experienced novelty. Participants were first asked to rate the remaining two volunteers to increase the impression of a real ongoing interaction. Next, participants again saw their own video, in order to increase self-awareness. Subsequently, the actual social evaluation experiment took place. Participants learned how they had been rated by their sham peers. In reality, all participants received exactly the same evaluation, which was previously generated by the experimenter. First, participants would see two vertical bars, one with the heading "your evaluation" and one with the heading "average evaluation for all individuals within the study". Both bars would start filling up simultaneously – the bar with the highest fill indicating a more positive evaluation and therefore less negative evaluation. The bar of "your evaluation" would not fill further than approximately halfway, whereas the average bar filled up to approximately 80% (**Fig. 1b**). This result was very similar for the three rated characteristics (intelligence, appearance, and congeniality). Following this, participants were presented with the individual evaluations of the five sham peers, accompanied by a short written reason for their evaluation (**Fig. 1c**). Of the fifteen evaluations for appearance, intelligence, and congeniality, seven were neutral or positive, for example "seems friendly" (congeniality) or "seems to know a lot" (intelligence), and eight were mildly negative, for example "strange nose" (appearance). Reasons for the evaluations were provided in a very general manner (i.e., "funny nose") so that they would fit every single individual. To prevent participants from avoiding to look at their evaluations in the first place, they had to indicate their level of agreement with the evaluation of each of the five sham peers by sorting them from 1 (agree most) to 5 (agree least). This evaluation constituted the social evaluation exposure. Participants were debriefed on the true nature of the procedure immediately after finishing the experiment with the accompanying affective and implicit self-esteem measurements. Thus, participants were under the impression that other participants truly evaluated them for not more than 30 minutes (**Fig. 2**).

Figure 2. Design of social evaluation exposure at session 2. t_1 = before stressor, t_2 = 10 (+/- 2)min after stressor, t_3 = 20 (+/- 2)min after stressor, t_4 = 30 (+/- 2)min after stressor



Design

A within-subject (pre-post stressor) design was employed to assess the effects of the social evaluation experiment. Positive affect (PA), negative affect (NA), implicit self-esteem, and cortisol were assessed both before and (for cortisol: multiple times) after exposure to the social evaluation exposure (Fig. 2). In addition, the subsample of 25 sibs underwent the same procedure, but was fully debriefed and therefore undeceived about the nature of the evaluation *before* it was presented. They were thus exposed to the same stimuli but they knew that evaluations had been invented by the experimenter and that no other peers had actually seen their video or evaluated their profile.

Measures

Subjective report of positive and negative affect

The **positive and negative affect scales schedule (PANAS)** (Watson, Clark, & Tellegen, 1988) was assessed using visual analogue scales (VAS). The PANAS consists of 10 momentary positive and 10 momentary negative affect items. Per affect characteristic (i.e. irritable, distressed, enthusiastic, etc.) participants were presented with a horizontal line of 105mm length with the label 'not' on its left and 'very much' on its right end on which they had to indicate by mouse click to what degree they were experiencing this affective state at this very moment. The more this position was on the right hand side of the line, the stronger the affective state in question was experienced. The standard procedure for analyzing the PANAS was followed (Watson et al., 1988). That is, the ratings for positive and negative affect items were averaged per person and per assessment to form measurements of positive (PA) and negative affect (NA), respectively. As affect was assessed on visual analogue scales, values were expressed in mm, higher scores indicating higher level of PA or NA, respectively.

Implicit measurement of self-esteem

In the **single category implicit association task (SC-IAT)** for self-esteem (Greenwald & Farnham, 2000; Karpinsky & Steinman, 2006), individuals were asked to categorize 12 personalized self-words (e.g. their first name, last name) either with positive or with negative words in a reaction time task. The faster people are able to categorize self-words with positive words relatively to negative words, the higher their implicit self-esteem is (i.e., $RT_{\text{self+neg words}} - RT_{\text{self+pos words}}$ in ms). In total there are 2 blocks: one block where self-words have to be sorted in the same category as positive words and one block where self-words have to be sorted together with negative words. Each block contains a total of 81 stimuli (36 self-items, 30 words in the opposite direction (i.e. negative in the positive block and positive in the negative block), and 21 in the direction of the block). Data were prepared according to recommendations for analyses from previous literature on the SC-IAT (for more detail see Greenwald & Farnham, 2000). As the SC-IAT is a reaction time task, the values are expressed in ms, with higher scores indicating higher implicit self-esteem.

Cortisol measurements

Cortisol was measured in saliva samples which were collected immediately before the computer paradigm started at session 2 and after (at 10 (+/- 2)min, 20 (+/- 2)min, and 30 (+/- 2)min) the social evaluation exposure (Fig. 2). These time points were based on suggestions in the literature and several reported experiments on cortisol assessment (Brooks & Robles, 2009; Nicolson, 2007; Young & Nolen-Hoeksema, 2001). Participants were restrained from eating and smoking for at least one hour and from drinking water at least 40 minutes before the first cortisol measure. Subjects collected saliva samples with a cotton swab (Salivette; Sarstedt), and placed the swab in the salivette tube. The interviewer recorded the exact collection time. Uncentrifuged samples were kept at -20°C until analysis. Saliva samples were frozen and stored at -20°C until analysis. After thawing, salivettes were centrifuged at 3,000 rpm for 5 min, which resulted in a clear supernatant of low viscosity. Salivary concentrations were measured using commercially available chemiluminescence-immunoassay with high sensitivity (IBL, 2009). The intra and interassay coefficients for cortisol were below 8%.

Cortisol values, in nmol/L, were then log transformed to reduce skewness of the distribution. We aimed to assess individuals in the afternoon to avoid influence of time of day on cortisol measurements. It was however not feasible to avoid variation in timing. The average time of baseline cortisol assessment was 3:21pm (median 3:45pm; standard deviation 2.75 hours; range 10:45am – 9:50pm). The time between the baseline measurement and the final cortisol measurement was on average 36 minutes

(range 23-46 min). Sampling times of measurements were added as a covariate to control for diurnal changes in cortisol. For this purpose, both the variables *time* and *time*² were included in the model, as described previously in similar work (Collip et al., 2011). Addition of higher order polynomial terms did not improve model fit.

General psychological distress

Current level of general psychological distress was measured using the symptom checklist-revised (SCL-90-R), a 90-item self-assessment questionnaire about the presence of several psychopathological symptoms within the last week including several subscales (Arrindell & Ettema, 1986). It has previously been suggested that the overall scale is a good indicator for general psychological distress (Bufka & Camp, 2010). For the current sample the scale had a high internal consistency ($\alpha = .97$). Higher total scores indicate greater level of general psychological distress.

Deception

After being debriefed all individuals filled in three questions about how much they believed the evaluations were real. These scores were combined to one deception score, with lower scores indicating stronger belief in the evaluation's authenticity.

Analyses

Multilevel regression models were employed to investigate whether affect (PA, NA), implicit self-esteem, and salivary cortisol levels changed from before to after the social rejection stressor. The current data had a hierarchical structure because multiple observations (assessment of affect, cortisol and implicit self-esteem each pre- and post social evaluation exposure) were clustered within subjects (level 1), who were related as they were part of twin pairs or sibs of twins (level 2).

Impact of the digi-SPEE on affect, implicit self-esteem, and cortisol

We first investigated whether the time of assessment (pre-or post social evaluation exposure) predicted the level of PA, NA, or implicit self-esteem with multilevel linear regression, using the XT MIXED command in STATA (StataCorp, 2009; College Station, TX, USA). PA, NA, and implicit self-esteem each were regressed on time of assessment. For cortisol levels, the same analyses were conducted, however, this time with four repeated measures per individual and controlled for contraception and alcohol use as described previously in similar work (Collip et al., 2011) (S1). Cortisol increase thus was expressed as the regression slope (i.e. change) of cortisol concentration in nmol/L over time.

Differential impact of the digi-SPEE depending on age, gender, or level of general psychological distress

Age, gender, and level of general psychological distress were added to the multilevel linear regression models, each in a separate analysis, as moderators of the time effects. Interaction effects between moderator and time (pre- and post social evaluation exposure) on the outcome measures (PA, NA, implicit self-esteem, cortisol) were investigated and followed up by simple effects if applicable.

Manipulation of perceived social evaluation

Manipulation of perceived social evaluation was assessed through two strategies. Deception scores were explored by calculating the percentage of individuals that had no, some, or many doubts or knew that the experiment was not real. Associations between deception scores on the one hand and SCL-90-R score, age, and gender on the other hand were assessed. The second strategy was the inclusion of an undeceived subgroup. It was examined whether a similar impact of the experiment on affect, implicit self-esteem, and cortisol was present for undeceived as for deceived participants. For this purpose unpaired sample t-tests for the pre-post evaluation change in PA, NA, and self-esteem or interactions with above time-effects (only for cortisol) were employed.

Results

Impact of the digi-SPEE on affect, implicit self-esteem, and cortisol

Exposure to the digital social evaluation resulted in a significant decrease in PA ($N=241$, $B=-5.25$, $p<.001$) and a significant increase in NA ($N=241$, $B=5.99$, $p<.001$) (**Fig. 3a**). Mean PA after the stressor was approximately 5mm (on the VAS) lower and mean NA approximately 6mm higher than before. Second, implicit self-esteem decreased significantly from before to after the social evaluation exposure ($N=187$, $B=-.19$; $p<.001$) indicating that the difference between positive and negative self-evaluation as assessed in the SC-IAT decreased by 0.19ms (**Fig. 3b**). Finally, cortisol levels increased significantly from before to after the exposure ($N=143$, $B=.07$; $p<.001$) (**Fig. 3c**). The observation that cortisol concentrations did not differ between T1 and T2 ($b=-.01$, $p=.55$) suggests that cortisol levels were not increased at baseline. The changes in PA and NA were negatively ($r=-.51$, $p<.001$), the changes in PA and self-esteem positively ($r=.15$, $p=.003$), and the changes in NA and self-esteem negatively ($r=-.15$, $p=.003$) correlated with each other. Change in cortisol was also significantly associated in the expected directions with change in psychological measures (PA: $B=-.002$, $p=.008$; NA: $B=.001$, $p=.015$; implicit self-esteem: $B=-.04$, $p=.015$).

Figure 3a. Change in affect. PANAS = positive and negative affect scales. VAS = visual analogue scales. PA = positive affect. NA = negative affect. Pre = before social evaluation. Post = after social evaluation. * $p < .001$.

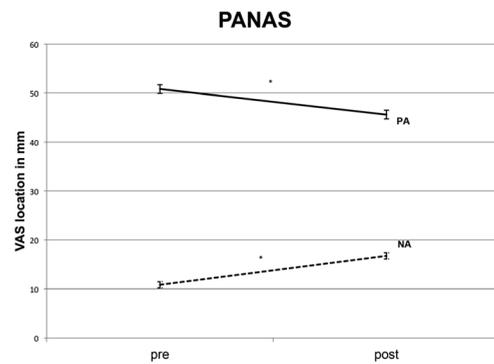


Figure 3b. Change in implicit self-esteem. Pre = before social evaluation. Post = after social evaluation. * $p < .001$.

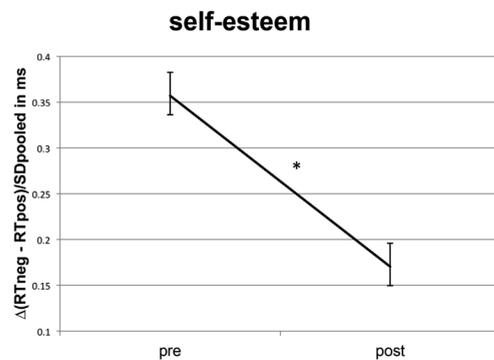
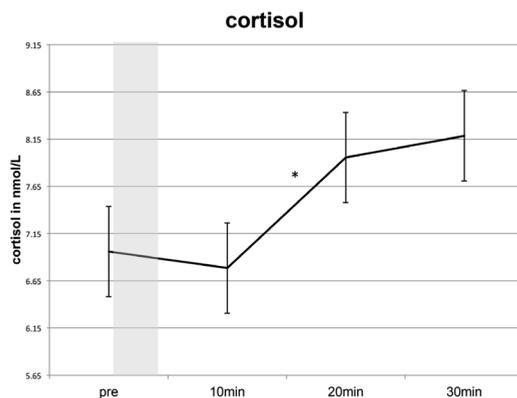


Figure 3c. Change in cortisol. Grey bar indicates the social evaluation part. Pre = before social evaluation. Post = after social evaluation. * $p < .001$



Moderation of the digi-SPEE-impact by age, gender, or current general psychological distress

Changes in affect were significantly moderated by gender ($N=241$) and current general psychological distress ($N=239$) (Table 3). Females and participants with higher levels of psychological distress showed larger decreases in PA and larger increases in NA (simple effects for moderation by gender: PA: males: $B=-2.62$, $p=.012$; females: $B=-7.09$, $p<.001$; NA: males: $B=2.73$, $p=.10$; females: $B=8.28$, $p<.001$). Cortisol responses were significantly moderated by age ($N=186$), indicating that younger participants showed stronger increases in cortisol following the social evaluation exposure (Table 3). Further following up this age-effect we dichotomized age in post-hoc analyses, and found that when participants were 18 years or above ($N=137$), the increase in cortisol was significantly decreased as compared to those younger than 18 years ($N=104$) ($B=-.06$, $p=.002$).

Table 3. Regression betas for the interaction between assessment time (pre and post social evaluation) and potential moderators.

	Outcome			
	PA	NA	Self-esteem	Cortisol
time x gender ¹	- 4.47** ($p=.002$)	5.55** ($p=.001$)	-0.07 ($p=.28$)	-0.03 ($p=.16$)
time x age	0.02 ($p=.90$)	0.00 ($p=.99$)	0.00 ($p=.61$)	-0.01** ($p<.001$)
time x SCL-90-R total score	-0.04* ($p=.035$)	0.06* ($p=.018$)	0.00 ($p=.37$)	0.00 ($p=.38$)

Note. PA = positive affect; NA = negative affect; time = pre and post social evaluation for PA, NA, and self-esteem and pre and 10min, 20min, and 30min post evaluation for cortisol; SCL-90-R = symptom checklist-90 revised. ¹male=0, female=1. * $p < .01$, * $p < .05$

Manipulation of perceived social evaluation

The majority of individuals believed that the experiment was real (67.52% had no doubts at all about the experiment's authenticity, 27.35% had little doubts and 5.13% had some doubts, 0% had serious doubts or knew that the experiment was not real). Furthermore, the level of deception was not significantly associated with general psychological distress (SCL-90-R; $\beta=-.09$, $p=.19$) or age ($-.06$, $p=.35$) but it was with gender (mean_{males} = 2.89, mean_{females} = 2.29, difference = .60, $p < .001$). Post hoc correction for deception in the analyses on the moderating effect of gender on PA- and NA-measures did not change results.

Second, in the undeceived subgroup neither affect (PA: $B=-2.52$, $p=.14$; NA: $B=2.68$, $p=.19$) nor implicit self-esteem ($B=-.10$, $p=.21$) nor cortisol ($B=-.02$, $p=.38$) significantly changed. None of these changes were significantly different when compared to the deceived sample, although effects were in the expected direction, i.e. smaller than in the deceived sample (also see **Table 2** for differences at separate assessment times). Controlling these analyses for age and gender (as this subgroup was slightly older and consisted of more females than the deceived sample) did not change results.

Discussion

Validation of the digital Social Peer Evaluation Experiment (digi-SPEE)

The current study introduced a new digital social peer evaluation experiment, the *digi-SPEE*, that can easily be administered in individuals' homes. The current results demonstrate that the *digi-SPEE* elicits stress-responses on measures of subjective affect, implicit self-esteem, and cortisol, which were particularly pronounced in groups expected to be more sensitive to social stress. The fully-computerized real life interaction format opens the door for innovative research in the area of social peer evaluation experiences as experienced in online social networks and emerging psychopathological symptoms.

More specifically, the results showed that (i) the experiment induced all the expected changes in affect, implicit self-esteem, and cortisol. Furthermore, (ii) groups hypothesized to be more sensitive to social stress – women (Strout et al., 2002), young adolescents (Collins & Thomas, 1972; Sommerville, 2013) and individuals with higher current level of psychological distress (Booij et al., 2012; Cuellar & Johnson, 2009; Hankin et al., 2010) - showed stronger responses to the current social peer evaluation exposure. This suggests that the experiment is sensitive enough to pick up differences between specific populations. Finally, (iii) the results cautiously suggest that not the mere exposure to unspecific experiment characteristics but the experience of being digitally evaluated by peers accounted for experiment-induced stress responses. These findings support that the *digi-SPEE* may prove a valid tool to induce the experience of digital social peer evaluation.

Group differences in peer evaluation-induced stress-responses

In accordance with previous studies (Björkqvist, 2001; Selten & Cantor-Graae, 2005; Strout et al., 2002) we found here that individuals that were female, of adolescent age or with more current psychological distress showed a particularly pronounced effect. Results indicated that females demonstrate an increased *affective* sensitivity

to the digital social peer evaluation exposure. These findings are in line with previous research reporting that females experience unpleasant stimuli as more unpleasant than men (Bianchin & Angrilli, 2012; Codispoti, Surcinelli, & Baldaro, 2008). However, other research showed that differences between women and men were absent for mood ratings in response to a social rejection stressor (Strout et al., 2002). Alternatively, the gender effect may be interpreted as a cultural report bias (i.e. in Western cultures it is well accepted for women to express emotions, whereas this is less accepted for men). Also, that symptoms of mental distress were associated with a more pronounced affective reaction to the current experiment concurs with previous findings. For example, it has previously been reported that individuals with mood symptoms respond with a particular pronounced increase in negative affect to situations of (slight) disapproval (Cuellar & Johnson, 2009). Furthermore, the current results showed that participants younger than 18 years were particularly biologically reactive to the experiment as indicated by amplified cortisol-responses. Adolescence is marked by an increased attunement to social evaluative information and therefore includes increased vigilance to only slightly evaluative contexts (Sommerville, 2013), such as daily online social interactions. Additionally, the hypothalamic-pituitary-adrenal (HPA) axis is particularly sensitive during adolescence (Hankin et al., 2010). As the HPA-axis is responsible for translating social and other stressful experiences to cortisol release (Booij et al., 2012), the current findings were to be expected. Even though we cannot conclude from the current findings whether it is the attunement to social evaluations or the increased HPA-axis reactivity or both that can be made responsible for this effect, it is valid to conclude that younger individuals react with more release of stress-hormones to social evaluative situations. We also expected to find that cortisol levels were moderated by levels of psychological distress; however, this did not emerge from the current data. This could be due to heterogeneity in cortisol reactions, as it is known that increased levels of mood symptoms are associated with both *increased* and *decreased* cortisol-responses. The direction of this effect has been suggested to depend on trait differences (Tops, Riese, Oldenhinkel, Rijdsdijk, & Ormel, 2008) and chronicity of depressive symptoms (Booij et al., 2012). Thus, overall, the different outcome measures were not equally affected by gender, age, and level of psychological distress. Nevertheless, the current study could demonstrate that the reactivity to the *digi-SPEE* could differentiate between individuals at varying levels of expected sensitivity to social evaluation.

Importantly, in the current sample general psychological distress and age did not concur with higher levels of deception, and controlling for deception did not change gender effects, suggesting that none of the moderator-effects are attributable to an increased perception of the experiment's authenticity.

The digi-SPEE and future psychopathology

Based on the current cross-sectional findings (i.e. increased reactivity in females, adolescents, and individuals with current psychological distress) we cautiously propose that the current experiment will demonstrate predictive validity for the development of psychopathological symptoms and may contribute to elucidating genetic and environmental causes in reactivity to peer evaluation as important risk factors for psychopathology. This becomes particularly important given the increased chances of exposure to peer evaluation through online social network interactions (Pempek et al., 2009). Being able to identify individuals at increased risk to develop future symptoms of psychopathology through their sensitivity to the current experiment will ultimately enable research into resilience factors. For example, potential resilience factors that may modify the impact of social peer evaluation sensitivity on long-term outcome are warm parenting (Bowes, Maughan, Caspi, Moffit, & Arseneault, 2010) and interpretation biases in ambiguous situations (MacLeod & Cohen, 1993). Understanding the risk and resilience mechanisms operating on the association between experience of social peer evaluation and incidence of psychopathology particularly in adolescence will enable development of target-oriented prevention.

Methodological considerations

The current findings need to be viewed in the light of certain methodological considerations. First, as this study particularly aimed to include participants between 15 and 18 years of age relative to older participants, findings may not generalize to older samples. Second, the undeceived subgroup was considerably smaller than the deceived sample. This was because reactivity to the social evaluation experiment was the variable of interest and the number of undeceived individuals was kept to a reasonable minimum. However, the non-significant result within the undeceived sample is unlikely the result of low power as also effect sizes were much smaller in this sample. Third, negative social evaluative situations may give rise to a mixture of experiences such as feeling evaluated but also feeling rejected or excluded. The current study is not able to disentangle these separate effects. Future studies with this specific aim can use the current experiment for this purpose combined with response questionnaires on evaluation, rejection, and exclusion. Finally, the current analyses regarding cortisol were controlled for age, gender, alcohol use, and use of contraception. There are, however, some other factors that might have influenced cortisol levels, such as phase of the menstrual cycle, body mass index, socio-economic status or smoking behaviour. These potential influences were not incorporated in the current analyses. To the extent that the above factors influence the within-person *change* in cortisol response from pre to post exposure this may have impacted on the

results, likely by inducing noise. Additionally, it needs to be acknowledged that there was variation in cortisol sampling time on the day. However, time in between baseline and final sampling time was short (average 36 minutes) and time of day was included as a covariate in the analyses.

Acknowledgements

The authors thank Jill Ielegems, Katrien Lyssens, Davinia Verhoeven, and Debora op't Eijnde for data-collection and all twins who volunteered for participation.

References

- Arrindell, W. A., & Ettema, J., M. (1986). *SCL-90. Handleiding bij een multidimensionele psychopathologie-indicator*. Swets & Zeitlinger.
- Bianchin, M., & Angrilli, A. (2012). Gender differences in emotional responses: a psychophysiological study. *Physiology & Behavior, 105*, 925-932.
- Björkqvist, K. (2001). Social defeat as a stressor in humans. *Physiology & Behavior, 73*, 435-442.
- Blackhart, G., Eckel, L. A., & Tice, D. M. (2007). Salivary cortisol in response to acute social rejection and acceptance by peers. *Biological Psychology, 75*, 267-276.
- Blackhart, G., Nelson, B. C., Knowles, M. L., & Baumeister, R. F. (2009). Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: a meta-analytic review of 192 studies on social exclusion. *Personality and Social Psychology Review, 13*(4), 269-309.
- Booij, S. H., Bouma, E. M. C., de Jonge, P., Ormel, J., & Oldehinkel, A. J. (2012). Chronicity of depressive problems and the cortisol response to psychosocial stress in adolescents: The TRAILS study. *Psychoneuroendocrinology*. doi: DOI:10.1016/j.psyneuen.2012.08.004
- Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79*(4), 631-343.
- Bowes, L., Maughan, B., Caspi, A., Moffit, T. E., & Arseneault, L. (2010). Families promote emotional and behavioral resilience to bullying: evidence of an environmental effect. *The Journal of Child Psychology and Psychiatry, 51*(7), 809-8017.
- Brooks, K. P., & Robles, T. F. (2009). Recent depressive and anxious symptoms predict cortisol responses to stress in men. *Psychoneuroendocrinology, 34*, 1041-1049.
- Buckley, K. E., Winkel, R. E., & Leary, M. R. (2004). Reactions to acceptance and rejection: effects of level and sequence of relational evaluation. *Journal of Experimental Social Psychology, 40*, 14-28.
- Bufka, L. F., & Camp, N. (2010). Brief measures for screening and measuring mental health outcome. In M. M. Antony & D. H. Barlow (Eds.), *Handbook of Assessment and Treatment Planning for Planning for Psychological Disorders* (Vol. 2, pp. 62-94). New York, London: The Guilford Press.
- Codispoti, M., Surcinelli, P., & Baldaro, B. (2008). Watching emotional movies: affective reactions and gender differences. *International Journal of Psychophysiology, 69*, 90-95.
- Collins, J. K., & Thomas, N. T. (1972). Age and susceptibility to same sex peer pressure. *British Journal of Educational Psychology, 42*(1), 83-85.
- Collip, D., Nicolson, N. A., Lardinois, M., Lataster, T., van Os, J., & Myin-Germeys, I. (2011). Daily cortisol, stress reactivity and psychotic experiences in individuals at above average genetic risk for psychosis. *Psychological Medicine, 41*(11), 2305-2315.
- Cuellar, A. K., & Johnson, S. L. (2009). Depressive symptoms and affective reactivity to maternal praise and criticism. *Journal of Social and Clinical Psychology, 28*(9), 1173-1194.
- Derom, C., Thiery, E., Peeters, H., Vlietinck, R., Defoort, P., & Fijns, J. P. (2013). The East Flanders prospective twin survey (EFPTS). An actual perception. *Twin Research and Human Genetics, 16*(1), 58-63.
- Dickerson, S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*(3), 355-391.
- Dickerson, S. S., Mycek, P. J., & Zaldivar, F. (2008). Negative social evaluation, but not mere social presence, elicits cortisol responses to a laboratory stressor task. *Health Psychology, 27*(1), 116-121.
- Donnerstein, E. (2012). Internet bullying. *Pediatric Clinics of North America, 59*(3), 623-633.

- Eisenberger, N. I., Jarcho, J. M., Lieberman, M. D., & Nalibhoff, B. D. (2006). An experimental study of shared sensitivity to physical pain and social rejection. *Pain, 126*, 132-138.
- Elzinga, B. M., Roelofs, K., Tollenaar, M. S., Bakvis, P., van Pelt, J., & Spinhoven, P. (2008). Diminished cortisol responses to psychosocial stress associated with lifetime adverse events: a study among healthy young subjects. *Psychoneuroendocrinology, 33*, 227-237.
- Gonzales, A. L., & Hancock, J. T. (2011). Mirror, mirror on my facebook wall: effects of exposure to facebook on self-esteem. *Cyberpsychology, Behavior, and Social Networking, 14*(1-2), 79-83.
- Greenwald, A. G., & Farnham, S. D. (2000). Using the implicit association test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79*(6), 1022-1038.
- Hankin, B. J., Badanes, L. S., Abela, J. R. Z., & Watamura, S. E. (2010). Hypothalamic-pituitary-adrenal axis dysregulation in dysphoric children and adolescents: cortisol reactivity to psychosocial stress from preschool through middle adolescence. *Biological Psychiatry, 68*, 484-490.
- Hsu, D. T., Sanford, B. J., Meyers, K. K., Love, T. M., Hazlett, K. E., Wang, H., . . . Zubieta, J.-K. (2013). Response of the mu-opioid system to social rejection and acceptance. *Molecular Psychiatry, 18*, 1211-1217.
- IBL. (2009). Cortisol Luminescence immunoassay - Luminescence immunoassay for the in-vitro diagnostic quantitative determination of cortisol un human saliva and serum. In I. I. GmbH (Ed.). Hamburg.
- Jacobs, N., Myin-Germeys, I., Derom, C., Delespaul, P., van Os, J., & Nicolson, N. A. (2007). A momentary assessment study of the relationship between affective and adrenocortical stress responses in daily life. *Biological Psychology, 74*, 60-66.
- Karpinsky, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91*(1), 16-32.
- Kirschbaum, C., Pirke, K., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test' - a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*, 76-81.
- Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & Andrews, J. A. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology, 102*(133-144).
- MacLeod, C., & Cohen, L. (1993). Anxiety and the interpretation of ambiguity: A text comprehension study. *Journal of Abnormal Psychology, 102*, 238-222.
- Maner, J. K., Miller, S. L., Schmidt, N. B., & Eckel, L. A. (2010). The endocrinology of exclusion: rejection elicits motivationally tuned changes in progesterone. *Psychological Science, 21*(4), 581-588.
- Moor, B. G., Crone, E. A., & van der Molen, M. W. (2010). The heartbrake of social rejection: heart rate deceleration in response to unexpected peer rejection. *Psychological Science, 21*(9), 1326-1333.
- Nadkarni, A., & Hoffmann, S. (2012). Why do people use facebook? *Personality and Individual Differences, 52*, 243-249.
- Nesdale, D., & Lambert, A. (2007). Effects of experimentally manipulated peer rejection in children's negative affect, self-esteem, and maladaptive social behavior. *International Journal of Behavioral Development, 31*(2), 115-122.
- Nicolson, N. A. (2007). Measurement of cortisol. In L. J. Luecken & L. C. Gallo (Eds.), *Handbook of Physiological Research Methods in Health Psychology*. (pp. 37-74). Thousand Oaks, CA,: Sage Publications.
- Nosko, A., Wood, E., & Molema, S. (2010). All about me: disclosure in online social networking profiles: the case of FACEBOOK. *Computers in Human Behavior, 26*, 406-418.
- Pempek, T. A., Yermolayeva, Y. A., & Calvert, S. L. (2009). College students' social networking experiences. *Journal of Applied Developmental Psychology, 30*(227-238).
- Reid, G., & Boyer, W. (2013). Social network sites and young adolescent identity development. *Childhood Education, 243*-253.
- Sandstorm, M. J., Cillessen, A. H. N., & Eisenhower, A. (2003). Children's appraisal of peer rejection experiences: impact on social emotional adjustment. *Social Development, 12*(4), 530-550.
- Sarstedt. Salivette. Hygienic saliva collection for diagnostics and monitoring.
- Selten, J. P., & Cantor-Graae, E. (2005). Social defeat: risk factor for schizophrenia? *British Journal of Psychiatry, 187*, 101-102.
- Sommerville, L. H. (2013). The teenage brain: sensitivity to social evaluation. *Psychological Science, 22*(2), 121-127.
- Steinfeld, C., Ellison, N. B., & Lampe, C. (2008). Social capital, self-esteem, and use of online social networks: a longitudinal analysis. *Journal of Applied Developmental Psychology, 29*, 434-445.
- Stroud, L. R., Foster, E., Papandonatos, G. D., Handwerker, K., Granger, D. A., Kivlighan, K. T., & Niaura, R. (2009). Stress response in the adolescent transition: Performance versus peer rejection stressors. *Development and Psychopathology, 21*, 47-68.
- Strout, L. R., Salovey, P., & Epel, E. S. (2002). Sex differences in stress responses: social rejection versus achievement stress. *Biological Psychiatry, 52*, 318-327.
- Tops, M., Riese, H., Oldenhinkel, A. J., Rijdsdijk, F. V., & Ormel, J. (2008). Rejection sensitivity relates to hypocortisolism and depressed mood state in young women. *Psychoneuroendocrinology, 33*, 551-559.
- van Eck, M. M. M., Nicolson, N. A., Berkhof, H., & Sulon, J. (1996). Individual differences in cortisol responses to a laboratory speech task and their relationship to responses to stressful daily events. *Biological Psychology, 43*, 69-84.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS Scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.
- Williams, K., Cheing, C., & Choi, W. (2000). Cyberostracism: effects of being ignored over the internet. *Journal of Personality and Social Psychology, 79*, 748-762.
- Young, E. A., & Nolen-Hoeksema, S. (2001). Effect of ruminations on the saliva cortisol response to a social stressor. *Psychoneuroendocrinology, 26*, 319-329.

Supporting Material

S1 Use of contraception and alcohol as controlled for in the cortisol analyses split by group

	Deceived sample	Undeceived sample
Contraception (yes/no)	43/198 (17.84%)	8/17 (32%)
Alcohol use in the last 12 months (yes/no)	170/65 (6 missing) (72.34%)	18/5 (2 missing) (21.74%)
Mean alcohol use per week in units if alcohol was consumed	4.26 units (SD=6.13) (6 missing)	7.88 units (SD=12.80) (1 missing)

**Cognitive bias modification
training in adolescents:
Effects on interpretation
biases and mood**

Lothmann, C.²; Holmes, E.A.³; Chan, S.W.Y.⁴; and Lau, Y.F.J.

Department of Experimental Psychology, Oxford University, UK

² Faculty of Psychiatry and Neuropsychology, Maastricht University, the Netherlands

³ Department of Psychiatry, Oxford University, UK

⁴ Department of Psychiatry, University of Cambridge, UK

Journal of Child Psychology and Psychiatry 2011, 52 (1), 24-32



Abstract

Background: Negative biases in the interpretation of ambiguous material have been linked to anxiety and mood problems. Accumulating data from adults show that positive and negative interpretation styles can be induced through cognitive bias modification (CBM) paradigms with accompanying changes in mood. Despite the therapeutic potential of positive training effects, training paradigms have not yet been explored in adolescents.

Methods: Eighty-two healthy adolescents (aged 13-17 years) were randomly allocated to either positive or negative CBM training. To assess training effects on interpretation bias, participants read ambiguous situations followed by test sentences with positive or negative interpretations of the situation. Participants rated the similarity of these sentences to the previously viewed ambiguous situations. Training effects on negative and positive affect were assessed using visual analogue scales before and after training.

Results: After training, adolescents in the *negative* condition drew more negative and fewer positive interpretations of new ambiguous situations than adolescents in the *positive* condition. Within the positive condition, adolescents endorsed more positive than negative interpretations. In terms of mood changes, positive training resulted in a significant decrease in negative affect across participants, while the negative condition led to a significant decrease in positive affect among male participants only.

Conclusion: This is the first study to demonstrate the plasticity of interpretation bias in adolescents. The immediate training effects on mood suggest that it may be possible to train a more positive interpretation style in youth, potentially helping to protect against anxiety and depressive symptoms.

Keywords: interpretation bias, adolescence, anxiety, mood, cognitive bias modification

Anxiety and mood problems are common in adolescence (Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993) and often persist into adulthood (Pine, Cohen, Gurley, Brook, & Ma, 1998). Prevention of later anxiety and depression relies on a better understanding of early risk mechanisms. Interpretation biases, the tendency to draw negative conclusions from ambiguous situations, are not only prominent characteristics of adult (e.g. Lawson, MacLeod, & Hammond, 2002; MacLeod & Cohen, 1993) but also of adolescent anxiety and mood disorders (e.g. Dearing & Gotlib, 2009; Taghavi, Moradi, Neshat-Doost, Yule, & Dagleish, 2000). Compared to healthy adolescents, highly anxious and depressed adolescents endorse more negative definitions of homophones - words which have both a negative and a neutral meaning, such as *die/dye* - and more often select words that resolve ambiguous scenarios in a negative direction (Dearing & Gotlib, 2009; Taghavi et al., 2000). Despite these findings, fundamental questions over cause and effect remain: do interpretation biases contribute to the development of anxiety and mood symptoms or are they merely mood-congruent effects on cognition?

To address this question, Mathews and Mackintosh (2000) developed a cognitive bias modification (CBM) paradigm to train healthy adults to draw positive or negative interpretations to ambiguous situations. In the training phase, participants read a series of ambiguous situations, one at a time, which ended with a word fragment for the participants to complete. In the positive CBM condition, these word fragments always resolved ambiguity in a positive direction; whereas in the negative CBM condition, these fragments always resolved the ambiguity in a negative direction. Following training, participants in the positive condition drew more positive and fewer negative interpretations of new ambiguous situations than those in the negative condition. Participants from the negative condition also reported higher levels of state anxiety after training (Mathews & Mackintosh, 2000). While the effects of CBM training on interpretation biases have been well-replicated, its effects on reducing anxiety/negative affect and/or increasing positive affect have been somewhat less consistent (e.g. Holmes & Mathews, 2005; Mackintosh, Mathews, Yiend, Ridgeway, & Cook, 2006; Mathews & MacLeod, 2002).

Recent extensions of this research to subclinical and clinical groups yield further results that may inform therapeutic developments. Specifically, these studies suggest that (sub)clinical samples show significant reductions of negative interpretation biases, state anxiety and negative affect following positive training (Blackwell & Holmes, 2010; Hirsch, Mathews, & Clark, 2007; Mathews & Mackintosh, 2000; Mathews, Ridgeway, Cook, & Yiend, 2007; Murphy, Hirsch, Mathews, Smith, & Clark, 2007; Salemink, van den Hout, & Kindt, 2009; Teachman & Addison, 2007; Yiend, Mackintosh, & Mathews,

2005). Not only do these results show that changes in interpretation biases precede changes in mood, they also implicate new targets for psychotherapy.

Despite these encouraging results in adult populations, only three studies have extended this research to youth, more specifically to pre-adolescent children (Muris, Huijding, Mayer, & Hameetman, 2008; Muris, Huijding, Mayer, Remmerswaal, & Vreden, 2009; Vassilopoulos, Banerjee, & Prantzalou, 2009). Muris et al. (2008, 2009) demonstrated in two studies that interpretation biases can be induced in healthy children with a task requiring interpretations of hypothetical situations during an imaginary space journey. Neither study assessed mood. Moreover, the ecological validity of the task is questionable. To address this, Vassilopoulos and colleagues (2009) employed realistic vignettes to induce positive interpretation biases in socially anxious children. Results showed a significant reduction in negative interpretation bias, anxiety, and anticipated anxiety in children who received training versus those who did not. Taken together, these results suggest that interpretation bias modification is possible in pre-adolescent children. Moreover, the latter study also shows the effectiveness of positive training on mood. However, it remains to be answered whether such training paradigms have similar effects in pubertal samples, where the risk for anxiety and mood problems is dramatically increased (Lewinsohn et al., 1993; Lewinsohn, Rohde, & Seeley, 1998). Given developmental differences in cognitive maturation between adolescents, adults, and children (Blakemore, 2006, 2008), we cannot necessarily expect the same pattern of results to emerge in response to training in adults and children as in adolescents.

To address this, the current study aimed to investigate the effects of positive *and* negative CBM on interpretation biases and changes in negative and positive affect in healthy adolescents (aged 13-17) using a modified version of the original adult CBM paradigm (Mathews & Mackintosh, 2000). To increase the ecological validity of training, ambiguous scenarios were developed and modified to increase their relevance to adolescents. Consistent with previous studies (Muris et al., 2008; Muris et al., 2009; Vassilopoulos et al., 2009), we hypothesised that negative bias modification would result in more negative interpretation of new ambiguous situations, increases in negative and decreases in positive affect, while the opposite would be true for positive bias modification.

Methods

Participants

Eighty-two adolescents aged 13 to 17 years (mean age: 15.22 years, *SD*: 1.38, 53.7% females) were recruited from seven mainstream schools in Oxfordshire and Lancashire, UK. The majority of students (97%) had a Caucasian ethnic background. All participants were fluent in English. Before signing informed assent or consent adolescents and their parents were asked whether the adolescent had a severe reading disability, a diagnosed anxiety or a diagnosed mood disorder. If the answer to this question was positive adolescents were excluded from participation. Due to the differences in each school's schedules recruitment methods varied across schools. In some, we recruited through whole school assemblies while in others we targeted specific classes through the class teacher. Informed consents were obtained from adolescents aged 16 years or above and from parents/legal guardians of younger adolescents. Adolescents under the age of 16 also provided written assent. Ethical approval was received from the Central University Research Ethics Committee of the University of Oxford.

Adolescents were randomly allocated to either a positive CBM ($n=41$) or negative CBM ($n=41$) condition. Prior to the first testing session, we generated a random sequence of *ones* and *twos* stratified for gender. We then assigned consecutive participants to either the positive (*one*) or negative (*two*) training condition. Groups did not differ according to gender, age, race, or trait anxiety (all $ps > .11$; **Table 1**). Due to a technical problem with recording participant responses, data are only available for 66 adolescents for the interpretation bias measure after training. Of these 32 received the positive and 34 the negative CBM. These adolescents did not differ from the whole sample ($N=82$) in terms of gender, age, trait anxiety, or baseline negative affect and positive affect (all $ps > .41$; further details on request). Analyses for training effects on change in interpretation bias were therefore based on 66 adolescents whereas training effects on mood were analyzed by including the whole sample ($N=82$).

Table 1. Participant characteristics and training performance.

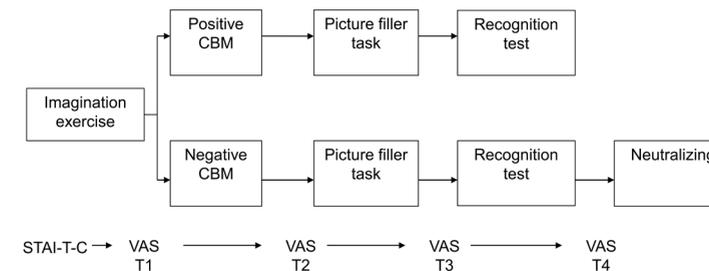
	Means (SD)	
	Positive Group	Negative Group
Demographics		
N	41	41
Age (years)	15.2 (0.23)	15.24 (0.21)
Gender [N (%)]	23 (56.1%) female	21 (51.2%) female
	18 (43.9%) male	20 (48.8%) male
Ethnicity	90.2% white	95% white
	9.8% other (mixed)	5% other (mixed)
Trait anxiety	36.07 (1.00)	34.54 (1.32)
Gender and training performance for results on cognition (N = 66)		
N	32	34
Gender [N(%)]	16 (50.0%) female	16 (40.7%) female
	16 (50.0%) male	18 (59.3%) male
% correct	93.33 (1.12)	86.32 (1.8)*
RT word fragments (ms)	1737.02 (82.21)	2303.88 (161.31)*

Note. * $p < .01$ for between group difference; % correct = percentage of correctly answered comprehension questions during training; RT word fragments = average reaction time to fill in word fragments during training

Overview of task procedures

The overall procedure is shown in **Figure 1**. Prior to training all participants completed the Trait Anxiety Inventory for children (Papay & Spiegelberger, 1986) and engaged in a short ‘imagination’ exercise. Research shows that the use of imagery can enhance the effects of CBM (Holmes, Lang, & Shah, 2009; Holmes & Mathews, 2005). Children and adolescents have good imagery abilities and are able to employ imagery when instructed (Harris, 2000; Kosslyn, 1980; Kosslyn, Margolis, Barrett, Goldknopf, & Daly, 1990). In the exercise we asked participants to close their eyes, imagine coming home after school and describe what they could see, hear, smell, taste, and feel. Once adolescents confirmed that they could ‘picture objects, events and feelings in

their mind’, they were instructed to apply these principles during the computer task. Overall, adolescents reported no problems with image creation. A reminder imagery exercise, during which participants imagined biting into a lemon, was administered immediately before training.

Figure 1. Testing procedures. CBM = cognitive bias modification. STAI-T-C = trait anxiety scale of the State/Trait Anxiety Inventory for Children, VAS = visual analogue scales.

Participants then underwent our adolescent-version of the CBM task which consisted of a training phase and a testing phase. Mood was assessed before training (T1), immediately after training (T2), immediately before the test (T3), and immediately after the test (T4). As residual changes in mood might confound group differences in interpretational style at the test phase, a *picture filler task* lasting 10 minutes was included between training and test phase to remove possible mood differences between groups. This task involved participants rating the pleasantness of 60 neutral pictures on visual analogue scales from ‘*very unpleasant* (0 cm) to ‘*very pleasant*’ (9.8 cm). After completion of all tasks, participants were fully debriefed by the experimenter. Participants in the negative CBM condition also received a brief version of the positive training to neutralize any potential negative effects of negative training.

Cognitive bias modification training

We employed a modified version of the original CBM paradigm for adults (Mathews & Mackintosh, 2000). During the *training* phase participants were presented with a series of ambiguous situations on a computer screen each ending with a word fragment. Participants were asked to complete the word fragment as quickly as possible by typing in the first missing letter. Completing the word fragment resolved the ambiguity of the situation, which led participants in the positive CBM condition to draw positive interpretations and participants in the negative CBM condition to draw negative interpretations. After each item, participants were presented with a comprehension question, which emphasised the emotional meaning of the situation and could only be answered correctly if the ambiguous situation had been interpreted in the intended

direction. This was followed by immediate feedback (for examples from the original paper, please see Mathews & Mackintosh, 2000).

To increase relevance for adolescents, we developed 50 new ambiguous situations relating to peer and romantic relationships and educational and recreational attainments (Collins, Welsh, & Furman, 2009; Pinkerton & Dolan, 2007; Ryan & Shim, 2008; Seeley, Stice, & Rohde, 2009; Sheeber, Hops, Alpert, Davis, & Andrews, 1997). An example of one scenario is: *It is the first day of term. Your new teacher asks everyone to stand up and introduce themselves. After you have finished, you guess the others thought you sounded...* This was followed by a positive or negative word fragment: *cl-v-r (clever)* or *st-p-d (stupid)*. Participants in the positive CBM condition completed the first while participants in the negative condition completed the second fragment. None of the adolescents tested struggled with these fragments. The comprehension question following this item was: *Do you feel unhappy with your introduction?* The correct answer was 'Yes' for participants in the negative condition and 'No' for those in the positive condition, which was followed by feedback about correctness ('Correct!' or 'Wrong!'). Ten 'distractor' scenarios were included to make the purpose of training less explicit. Five of these involved completion of word fragments that resolved ambiguity in the opposite direction of assigned training condition. That is, participants in the positive CBM condition completed five word fragments that resolved ambiguity in a negative direction and *vice versa* for participants in the negative CBM condition. The remaining five distractors involved word fragments that kept the overall valence of the item neutral, for example: *You are about to meet up with your friend. Just before you leave, he phones to say that he can't make it. You think that this is because he is feeling unwell (unwell)*. Thus, a total of 60 scenarios were used during the *training phase*. These were presented in five blocks of 12 items containing 10 items in the direction of assigned condition (i.e. positive or negative), one neutral item, and one item in the opposite direction of assigned condition. Situations were presented randomly in each block. The task was self-paced. Participants were reminded to imagine the situations were happening to them at the beginning of each block.

Assessment of subsequent interpretation bias

During the *testing phase*, a further 10 ambiguous situations were presented, each starting with a title. Again, these items were developed specifically for adolescents. Adolescents were instructed to imagine that each situation was happening to them. They were also explicitly asked to pay attention to the title. This time, however, the word fragment did *not* disambiguate the situation; for example: ***The end-of term prom:*** *As one of the main organizers, you are asked to give a short speech at the end-of term*

prom. When the time comes you get on the stage. As you speak, you notice some of the students in the audience start to l--gh (laugh). Likewise, the comprehension question did not emphasise the emotional meaning of the situation: *'Did you get to the stage to speak?'* Immediately after, participants were given a *recognition test*. The title of the previously presented ambiguous situation ('The end-of-term prom') was presented at the top of the screen with four recognition statements. None of these statements used the exact words of the situations but yielded similar meanings. Two sentences comprised 'targets' depicting either a positive ('As you speak, students in the audience start to laugh approvingly') or a negative interpretation ('As you speak, students in the audience find your efforts laughable'). The remaining two sentences comprised 'foils', statements that were related to the emotional valence of the situations but also included information that was not explicitly given in the ambiguous situations (positive foil: *As you speak, students in the audience start to applaud your comments*, negative foil: *As you speak, students in the audience start to yawn*).

Participants were instructed to rate the similarity of each of these sentences to those presented during the test phase on a scale from 1 (very different in meaning) to 4 (very similar in meaning). Successful interpretation bias modification would be indicated if adolescents in the positive condition gave higher similarity ratings for positive targets than for negative targets; and likewise if adolescents in the negative training rated negative targets as more similar to previously viewed situations than positive targets (Mathews & Mackintosh, 2000). In the original paradigm foils were included to investigate whether training effects specifically altered interpretation styles or whether it facilitated general response biases to valenced material (i.e. the tendency to give higher similarity ratings to positive or negative information more generally). All items and questions were presented on a portable computer Acer 'Travelmate 4720' on a 14" screen with E-Prime 2.0 (Psychology Software Tools, 2007), which also recorded responses and reaction times.

Assessments of mood changes

To assess self-rated mood, we developed 12 visual analogue scales (VAS), measuring negative and positive affect. Items with high face-validity for each emotion were chosen from the state anxiety scale of the State/Trait Anxiety Inventory for children (Papay & Spiegelberger, 1986), and the Positive and Negative Affect Scales for children (PANAS-C) (Laurent et al., 1999), both of which have satisfactory psychometric properties (Laurent et al., 1999; Papay & Spiegelberger, 1986). Consistent with the finding that the high co-occurrence of anxiety and depression (Lewinsohn et al., 1993) is mediated through nonspecific negative affect (Clark & Watson, 1991), four items for state anxiety (nervous, worried, anxious, and scared) and four items for depressed

affect (sad, upset, miserable, and gloomy) were combined to index negative affect. Another four items assessed positive affect (happy, calm, cheerful, and energetic). The VAS were 9.8 cm long between 'not [emotion] at all' (0 cm) and 'very [emotion]' (9.8 cm). Participants indicated how they were feeling at the current moment by marking the line.

Statistical analyses

A 2x2x2 repeated measures ANOVA explored the effects of two between subject factors (Training group: positive vs. negative, Gender: male vs. female) and one within-subject factor (Valence: positive vs. negative) on similarity ratings for targets and for foils in SPSS 16.0 (Chicago, IL). Analyzing targets and foils separately is consistent with previous studies (e.g. Mathews & Mackintosh, 2000). Trait anxiety and age were included as covariates. Concerning targets, one participant was excluded as an outlier (similarity ratings >3SD above the mean) (female, positive condition). These analyses were thus conducted on 65 adolescents. Likewise, a 2x2x2 repeated measures ANOVA was conducted on changes in negative and positive affect from before to after training with two between-subject factors (Training group, Gender) and one within-subjects factor (Time: T1 vs. T2). Again, trait anxiety and age were included as covariates. Significant interactions were followed-up with independent sample t-tests for between-group differences and paired sample t-tests, assessing within-group differences from before to after training for each training group separately. Trait anxiety, age, and gender were only included in follow-up analyses if they significantly modified the 2-way training-group-by-time interaction. Between-group differences in mood at all four time points (T1, T2, T3, and T4) and training performance were assessed with independent sample t-tests. Whenever assumptions of normality or sphericity were violated, the non parametric equivalent for the independent sample t-test, the Mann-Whitney-U-test (test statistic U), was employed. These analyses were based on the whole sample ($N=82$).

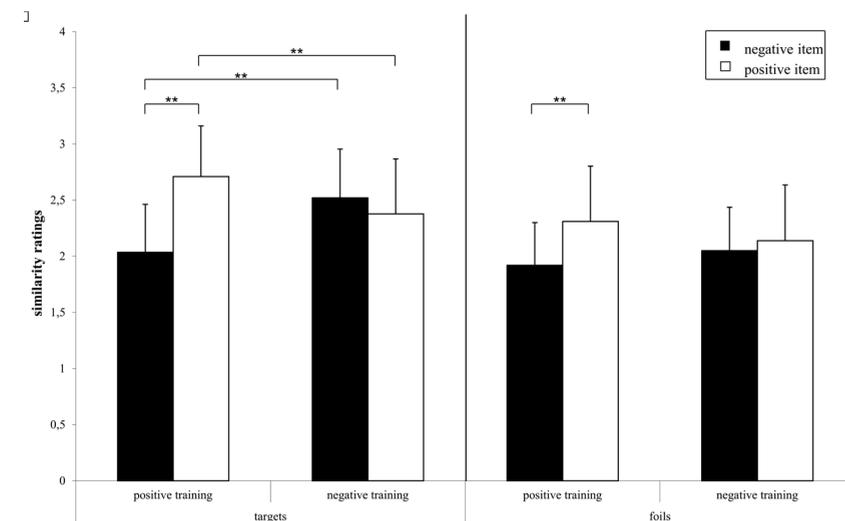
Results

Group differences in interpretation bias

For targets, the full model with two between-subject factors (Training condition, Gender), one within subject factor (Valence), and two covariates (Trait anxiety, Age) revealed a main effect of Gender, $F(1, 59)=4.4, p<.04, \eta^2=.07$, and crucially, a significant Training-Group-by-Valence interaction, $F(1, 59)=24.24, p<.01, \eta^2=.41$. No other main or interaction effects were significant ($ps>.53$). Girls assigned higher similarity to all test items than boys. Compared to the negative training group, the positive training group rated positive targets as more similar, $t(63)=2.84, p<.01, r=.34$, and negative targets as less similar, $t(63)=-4.54, p<.01, r=-.5$, to ambiguous situations in the test phase. The positive group also rated positive targets as more similar to the situations than negative targets, $t(30)=-6.03, p<.01, r=.74$. This within-group effect did not emerge for the negative training group ($t(33)=1.14, ns$) (**Figure 2**).

For foils, the full model revealed only a significant Training-Group-by-Valence interaction, $F(1, 60)=5.57, p<.05, \eta^2=.09$. All other main and interaction effects were not significant ($ps>.53$). Post hoc analyses showed that the positive training group rated positive foils as more similar to the ambiguous situations than negative foils, $t(31)=-3.67, p<.01, r=.55$. None of the other between- or within group t-tests yielded significant findings (all $ps>.16$) (**Figure 2**).

Figure 2. Similarity ratings for positive and negative targets and foils across training groups in the testing phase. ** $p < .01$. Higher similarity ratings indicated greater resemblance of the item to the ambiguous situation.



Group differences in mood changes

For negative affect, the full model revealed a significant main effect for trait anxiety, $F(1, 76)=14.43$, $p<.001$, $\eta^2=.16$, a significant Trait-anxiety-by-Time interaction, $F(1, 76)=6.14$, $p<.05$, $\eta^2=.08$, and finally, the predicted Training-Group-by-Time interaction, $F(1, 76)=4.12$, $p<.05$, $\eta^2=.05$ (**Table 3**). Those with higher levels of trait anxiety reported more negative affect overall. However, individuals with higher trait anxiety also showed larger reductions of negative affect from before to after training, $r=.34$, $p<.01$. Post-hoc tests for the Training-Group-by-Time interaction revealed that for adolescents in the positive group, negative affect significantly decreased from before to after training, $t(40)=2.73$, $p<.01$, $r=.40$, whereas it did not change for adolescents in the negative condition, $t(40)=-.64$, ns (**Table 2**).

The complete model for positive affect revealed no main effects (all $ps>.05$) but a significant Gender-by-Group-by-Time interaction, $F(1, 76)=5.03$, $p<.05$, $\eta^2=.07$ (**Table 3**). Splitting this up for gender groups revealed a significant Group-by-Time interaction in boys, $F(1, 36)=5.24$, $p<.05$, $\eta^2=.15$: positive affect only decreased in those receiving negative training, $t(19)=2.76$, $p<.05$, $r=.54$ but did not change in the positive CBM group, $t(17)=-.24$, ns . This Group-by-Time interaction did not emerge for girls, $F(1, 76)=1.2$, $p=.28$, ns . Instead, only a Trait anxiety-by-Time interaction emerged, $F(1, 40)=4.24$, $p<.05$, $\eta^2=.11$. Here, girls with higher trait anxiety showed smaller reductions of positive affect across time ($r=-.29$, $p=.06$, ns).

Overall, independent sample t-tests for mood differences between the groups at T1, T2, T3, and T4 were conducted to assess whether fundamental differences in mood could have confounded results from the test-phase. None of the comparisons revealed significant effects (all $ps>.24$) (**Table 3**).

Table 2. Negative and positive affect at time point 1 (T1), time point 2 (T2), time point 3 (T3), and time point 4 (T4).

State	Positive Group	T1	T2	T3	T4
anxiety	All	2.67 (1.75)	2.07 (1.71)	1.53 (1.61)	1.36 (1.40)
	Males	2.97 (1.57)	2.33 (1.80)	1.73 (2.15)	1.59 (1.75)
	Females	2.44 (1.87)	1.86 (1.65)	1.38 (1.04)	1.18 (1.05)
	Negative Group				
	All	2.33 (1.69)	1.96 (1.62)	1.39 (1.36)	1.34 (1.48)
	Males	2.44 (1.59)	2.04 (1.76)	1.50 (1.39)	1.25 (1.17)
	Females	2.23 (1.82)	1.87 (1.52)	1.27 (1.36)	1.42 (1.76)
Negative affect	Positive Group				
	All	1.53 (1.74)	1.51 (1.72)	1.13 (1.48)	1.42 (1.95)
	Males	1.64 (1.89)	1.92 (2.01)	1.33 (1.89)	1.86 (2.52)
	Females	1.44 (1.56)	1.18 (1.42)	0.97 (1.06)	1.08 (1.30)
	Negative Group				
	All	1.01 (1.07)	6.23 (1.72)*	1.33 (1.46)	1.27 (1.34)
	Males	1.09 (1.08)	1.92 (1.24)	1.60 (1.21)	1.37 (1.07)
	Females	0.93 (1.08)	1.11 (1.26)	1.07 (1.54)	1.18 (1.57)
Positive affect	Positive Group				
	All	6.39 (1.47)	1.51 (1.72)	6.17 (1.74)	6.12 (1.96)
	Males	6.14 (1.20)	6.20 (1.74)	6.14 (1.77)	6.14 (1.86)
	Females	6.59 (1.65)	6.26 (1.72)	6.20 (1.76)	6.11 (2.08)
	Negative Group				
	All	6.63 (1.59)	6.23 (1.72)	6.12 (1.68)	6.37 (1.86)
	Males	6.62 (1.49)	5.77 (1.74)*	5.46 (1.36)	5.84 (1.84)
	Females	6.62 (1.72)	6.68 (1.61)	6.75 (1.75)	6.88 (1.78)

Note. * $p < .05$ for change from T1 to T2, T1 = before training, T2 = after training, T3 = after picture filler task, T4 = after recognition test. All values are given in cm on visual analogue scales. Higher values indicate that this emotion was experienced more.

Table 3. Main effects and interactions for differences in negative and positive affect from time point 1 (T1) to time point 2 (T2) and between group differences in mood at T1, T2, T3, and T4

	NA	PA
Complete analysis for differences between T1 and T2 (main effects and interactions¹)		
time	p=ns	p=ns
Group	p=ns	p=ns
Gender	p=ns	p=ns
time x group	F(1, 76) = 4.12*	p=ns
time x gender	p=ns	p=ns
time x group x gender	p=ns	F(1, 76) = 4.26*
Independent sample t-tests for differences between groups		
T1	p=ns	p=ns
T2	p=ns	p=ns
T3	p=ns	p=ns
T4	p=ns	p=ns

Note *p<.05, NA=negative affect, PA=positive affect. ¹Trait anxiety and age were entered as covariates

Group differences on training task performance

Adolescents in the positive condition were significantly faster in completing word-fragments than were adolescents in the negative condition, $U=247.00$, $p<.01$, $r=-.42$. Adolescents in the positive training condition were also more accurate ($Mdn = 95%$) in response to comprehension questions than adolescents in the negative training condition ($Mdn = 88%$), $U=277.00$, $p<.01$, $r=.47$ (Table 1). To eliminate the confound that between-group differences in accuracy of responses to comprehension questions and therefore receipt of positive feedback during training would influence training-specific changes, percentage of correct responses was included as a covariate in the above analyses for interpretation bias. Percentage of correct responses to comprehension questions did not correlate with any of the mood measures, suggesting that it did not confound training specific mood changes.

Discussion

This study is the first to examine interpretation bias modification training in adolescents. As adolescent anxiety and mood problems predict later episodes in adulthood (Pine et al., 1998), studying the plasticity of cognitive biases in this age has important implications for understanding early risk mechanisms and developing targets for cognitive therapies. Our study yielded two key findings. First, our modified CBM training paradigm for adolescents successfully induced interpretation biases

in the intended direction. Adolescents in the positive training condition endorsed more positive and less negative interpretations of new ambiguous situations than adolescents in the negative condition. This effect was not merely attributable to differences in mood as there was no group difference in mood immediately before the testing phase (T3). Second, positive training resulted in decreased *negative affect* across participants and negative training decreased *positive affect* but for boys only.

Our training effects on interpretations biases in adolescents extend previous data on healthy adults and children. Numerous studies have shown that positive and negative training induces group differences in the interpretation of new ambiguous situations presented immediately and even 24-hours after training (Holmes & Mathews, 2005; Mathews & Mackintosh, 2000; Muris et al., 2008; Muris et al., 2009; Salemink, van den Houdt, & Kindt, 2007; Vassilopoulos et al., 2009; Yiend et al., 2005). This is the first study to replicate findings in a pubertal sample, suggesting that interpretation biases are plastic in this age range. While a significant training-group-by-valence interaction also characterized foils, these effects were weaker, perhaps suggesting that training affects interpretation specifically and not response bias in general. We also found that modification training alters positive as well as negative mood but like adult data, these trends were less consistent and more work will be needed to verify these.

The current results speak to the apparent effectiveness of the positive CBM relative to the negative CBM. First, adolescents in the positive condition were faster in completing word fragments and more accurate in their responses to comprehension questions during training. Second, within-group differences emerged among individuals receiving positive training such that significantly more positive than negative interpretations were drawn over new ambiguous situations. These within-group differences raise questions over the mechanisms through which positive training exerts its effects. Positive training could either *increase* the salience of positive relative to negative interpretations or *decrease* the salience of negative interpretations relative to positive interpretations, both of which would result in a greater difference between positive and negative interpretations. Disentangling these alternatives relies on baseline measures of interpretation bias. Consistent with the adult literature positive training effects on reduced negative affect were significant and promising, albeit smaller than effects on cognition.

These results need to be viewed in light of several limitations. First, we did not include a baseline measure of interpretation bias, because pre-exposure to the outcome variable (i.e. assessment of interpretation bias via ambiguous situations and recognition

items *before* training) may encourage adolescents to anticipate the purpose of the test phase. However, this has limited our ability to draw conclusions on whether training conditions actually induced *changes* in interpretation biases. Second, to ensure that we completed testing in an appropriate time frame, we assessed mood changes using simple visual analogue scales. While these measures offer an initial indication of whether interpretation bias modification *can* affect mood in adolescents, they may be less valid than standardized questionnaires such as the PANAS-C (Laurent et al., 1999) or STAI-C (Papay & Spiegelberger, 1986). Use of these questionnaires in future studies will allow further extrapolations of adult findings to those in adolescents (e.g. Holmes et al., 2009; Holmes, Mathews, Dagleish, & Mackintosh, 2006; Mathews & Mackintosh, 2000). Third, because of a need to minimize disruption to each school's routines and schedules, our methods of recruitment across schools ranged from brief presentations at whole school assemblies to more targeted invitations aimed at specific classes organised by a school teacher. Therefore, obtaining an accurate and meaningful rate of participation was problematic. However, the range of schools that we approached was diverse, suggesting that the recruited sample, though self-selecting is likely, was representative. Finally, timing constraints prevented more elaborate assessments of other background factors such as social class and life stress, which may also inform the sample's representativeness in future studies.

Regardless of these caveats, our results carry exciting implications for understanding the role of interpretation biases on subsequent pathological emotional development. Adult data have already begun to address the effects of training on stress reactivity (Hirsch, Hayes, & Mathews, 2009; Holmes et al., 2009; Mackintosh et al., 2006; Wilson, Macleod, Mathews, & Rutherford, 2006). These data suggest that vulnerability associated with negative cognitive biases may be expressed by increasing *risks* for anxiety and mood problems in the presence of stress. These questions await investigation in youth.

More importantly, these findings also carry clinical implications, raising the questions of whether modification of interpretation biases using positive training can attenuate future negative outcomes. Promising findings in adults (Hirsch et al., 2009; Holmes et al., 2009; Wilson et al., 2006) and children (Vassilopoulos et al., 2009) demonstrate that vulnerable individuals show reduced negative interpretation biases, anxiety, and depression (Blackwell & Holmes, 2010) after positive training. Clearly, more research is warranted on the long-term effects of these changes, but potentially CBM could serve as a powerful adjunct to treatment, as a relatively simple and cost-effective tool (Koster, Fox, & MacLeod, 2009). More specifically, a computerized delivery of CBM may

be more appealing to youth. Thus, the possibility that CBM could be used to modify early-emerging negative biases as a preventative intervention among high-risk youth is especially intriguing.

In summary, this study has demonstrated that interpretation bias modifications in healthy adolescents is not only possible but also affects mood. These findings replicate previous research and suggest that a positive interpretation style immediately decreases negative affect. Future research should consider including standardized mood measures, pre-training assessment of baseline interpretation bias, and assessment of stress reactivity after training to gain further insight about the developmental and maintenance mechanisms underlying anxiety and mood problems in this high risk age group.

Acknowledgements

The authors are very grateful to schools and students who participated in this study. We greatly appreciated the help with data collection of Katie Lockwood, Kerrie Channer, and Emma Molyneaux and earlier discussion with Bundy Mackintosh about CBM.

References

- Blackwell, S. E., & Holmes, E. A. (2010). Modifying interpretation and imagination in clinical depression: a single case series using cognitive bias modification. *Applied Cognitive Psychology, 24*, 338-350.
- Blakemore, S. J. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*(4), 296-312.
- Blakemore, S. J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience, 9*, 267-277.
- Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*(3), 316-336.
- Collins, W. A., Welsh, D. P., & Furman, W. (2009). Adolescent romantic relationship. *Annual Review of Psychology, 60*, 631-652.
- Dearing, K. H., & Gotlib, I. H. (2009). Interpretation of ambiguous information in girls at risk for depression. *Journal of Abnormal Child Psychology, 37*, 79-91.
- Harris, P. (2000). *The work of the imagination*. Oxford: Wiley-Blackwell.
- Hirsch, C. R., Hayes, S., & Mathews, A. (2009). Looking at the bright side: accessing benign meanings reduces worry. *Journal of Abnormal Psychology, 118*(1), 44-54.
- Hirsch, C. R., Mathews, A., & Clark, D. M. (2007). Inducing an interpretation bias changes self-imagery: a preliminary investigation. *Behavior Research and Therapy, 45*, 2173-2181.
- Holmes, E. A., Lang, T. J., & Shah, D. M. (2009). Developing interpretation bias modification as a "cognitive vaccine" for depressed mood: imagining positive events makes you feel better than thinking about them verbally. *Journal of Abnormal Psychology, 118*(1), 76-88.
- Holmes, E. A., & Mathews, A. (2005). Mental imagery and emotions: a special relationship? *Emotion, 5*, 489-497.
- Holmes, E. A., Mathews, A., Dagleish, T., & Mackintosh, B. (2006). Positive interpretation training: Effects of mental imagery versus verbal training on positive mood. *Behavior Therapy, 37*, 237-247.
- Kosslyn, S. M. (1980). *Image and mind*. Cambridge, Massachusetts: Harvard University Press.
- Kosslyn, S. M., Margolis, J. A., Barrett, A. M., Goldknopf, E. J., & Daly, P. F. (1990). Age differences in imagery abilities. *Child Development, 61*(4), 995-1010.
- Koster, E. H. W., Fox, E., & MacLeod, C. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *Journal of Abnormal Psychology, 118*(1), 1-4.
- Laurent, J., Catanzaro, S. J., Joiner, T. E., Rudolph, K. D., Potter, K. I., Lambert, S., et al. (1999). A measure of positive and negative affect for children: scale development and preliminary validation. *Psychological Assessment, 11*, 326-338.
- Lawson, C., MacLeod, C., & Hammond, G. (2002). Interpretation revealed in the blink of an eye: depressive bias in the resolution of ambiguity. *Journal of Abnormal Psychology, 111*, 321-328.
- Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & Andrews, J. A. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology, 102*(1), 133-144.
- Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1998). Major depressive disorder in older adolescents: prevalence, risk factors, and clinical implications. *Clinical Psychology Review, 18*(7), 765-794.
- Mackintosh, B., Mathews, A., Yiend, J., Ridgeway, V., & Cook, E. (2006). Induced biases in emotional interpretation influences stress vulnerability and endure despite changes in context. *Behavior Therapy, 37*, 209-222.
- MacLeod, C., & Cohen, L. (1993). Anxiety and the interpretation of ambiguity: a text comprehension study. *Journal of Abnormal Psychology, 102*, 238-247.

- Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology, 109*, 602-615.
- Mathews, A., & MacLeod, C. (2002). Inducing processing biases have causal effects on anxiety. *Cognition and Emotion, 16*(3), 331-354.
- Mathews, A., Ridgeway, V., Cook, E., & Yiend, J. (2007). Inducing a benign interpretational bias reduces trait anxiety. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 225-236.
- Muris, P., Huijding, J., Mayer, B., & Hameetman, M. (2008). A space odyssey: experimental manipulation of threat perception and anxiety-related interpretation bias in children. *Child Psychiatry & Human Development, 39*(4), 469-480.
- Muris, P., Huijding, J., Mayer, B., Remmerswaal, D., & Vreden, S. (2009). Ground control to Major Tom: Experimental manipulation of anxiety-related interpretation bias by means of the "space odyssey" paradigm and effects on avoidance tendencies in children. *Journal of Anxiety Disorders, 23*(3), 333-340.
- Murphy, R., Hirsch, C. R., Mathews, A., Smith, K., & Clark, D. M. (2007). Facilitating a benign interpretation bias in a high socially anxious population. *Behavior Research and Therapy, 45*, 1517-1529.
- Papay, J. P., & Spiegelberger, C. D. (1986). Assessment of anxiety and achievement in kindergarten and first- and second-grade children. *Journal of Abnormal Child Psychology, 14*, 279-286.
- Pine, D. S., Cohen, P., Gurley, D., Brook, J., & Ma, Y. (1998). The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depression disorders. *Archives of General Psychiatry, 55*, 56-64.
- Pinkerton, J., & Dolan, P. (2007). Family support, social capital, resilience and adolescent coping. *Child and Family Social Work, 12*, 219-228.
- Psychology Software Tools, I. (2007). *Psychology Software Tools*. Pittsburgh.
- Ryan, A. M., & Shim, S. S. (2008). An exploration of young adolescents' social achievement goals and social adjustment in middle school. *Journal of Educational Psychology, 100*(3), 672-687.
- Salemink, E., van den Houdt, M., & Kindt, M. (2007). Trained interpretive bias and anxiety. *Behavior Research and Therapy, 45*, 329-340.
- Salemink, E., van den Houdt, M., & Kindt, M. (2009). Effects of positive interpretive bias modification in highly anxious individuals. *Journal of Anxiety Disorders, 23*(5), 676-683.
- Seeley, J. R., Stice, E., & Rohde, P. (2009). Screening for depression prevention: identifying adolescent girls at high risk for future depression. *Journal of Abnormal Psychology, 118*(1), 161-170.
- Sheeber, L., Hops, H., Alpert, A., Davis, B., & Andrews, J. (1997). Family support and conflict: prospective relations to adolescent depression. *Journal of Abnormal Child Psychology, 25*(4), 333-344.
- Taghavi, M. R., Moradi, A. R., Neshat-Doost, H. T., Yule, W., & Dagleish, T. (2000). Interpretation of ambiguous emotional information in clinically anxious children and adolescents. *Cognition and Emotion, 14*, 809-822.
- Teachman, B. A., & Addison, L. M. (2007). Training non-threatening interpretation in spider fear. *Cognitive Therapy Research, 32*, 448-459.
- Vassilopoulos, S. P., Banerjee, R., & Prantzalou, C. (2009). Experimental modification of interpretation bias in socially anxious children: Changes in interpretation, anticipated interpersonal anxiety, and social anxiety symptoms. *Behavior Research and Therapy, 47*, 1085-1089.
- Wilson, E. J., Macleod, C., Mathews, A., & Rutherford, E. M. (2006). The causal role of interpretive bias in anxiety reactivity. *Journal of Abnormal Psychology, 115*, 103-111.
- Yiend, J., Mackintosh, B., & Mathews, A. (2005). Enduring consequences of experimentally induced biases in interpretation. *Behavior Research and Therapy, 43*, 779-797.

How to boost positive interpretations? A meta-analysis of the effectiveness of cognitive bias modification for interpretation

Menne-Lothmann, C.¹; Viechtbauer, W.¹; Höhn, P.¹; Kasanova, Z.¹; Haller, S.P.²; Drukker, M.¹; van Os, J.^{1,3}; Wichers, M.¹; and Lau, J.Y.F.⁴

¹ Department of Psychiatry and Psychology, South Limburg Mental Health Research and Teaching Network, EURON, Maastricht University, The Netherlands

² Department of Experimental Psychology, Oxford University, UK

³ Visiting Professor of Psychiatric Epidemiology King's College London, King's Health Partners Department of Psychosis studies Institute of Psychiatry, UK

⁴ Psychology Department, Institute of Psychiatry, Kings College London, UK



Abstract

The current meta-analysis explores the strength of effects of cognitive bias modification training for interpretation bias (CBM-I) on positive (i.e., adaptive) interpretations and mood as well as the training and sample characteristics influencing these effects.

Data-bases were searched with the key words “interpret* bias AND training” and “interpret* bias AND modif*”. Reference lists of identified articles were checked and authors of identified articles were contacted for further relevant articles and unpublished data. Studies were reviewed for inclusion with eligibility criteria being that the study (a) aimed to target interpretation biases through any kind of training, (b) assessed mood and/or interpretation bias as outcome measures, (c) allocated individuals to training conditions at random, and (d) recruited adult samples. A meta-analytic multilevel mixed-effects model was employed to assess standardized mean changes in interpretation bias, negative mood, and emotional reactivity. In addition, several training and sample characteristics were explored for their potential to enhance benign training effectiveness.

On average, benign CBM-I resulted in an increase in positive interpretation bias ($p < .01$) and a decrease in negative mood state ($p < .001$), but did not affect emotional reactivity. These effects were not consistently different from control conditions with no or neutral training. However, within benign training conditions imagery instructions and more training sessions were related to larger cognitive and mood effects, whereas feedback about training performance and inclusion of non-benign training items (instead of including benign items only) boosted cognitive effects only. Finally, training was more effective in women (cognitive *and* mood effects) and presumably samples with symptomatic emotional dysregulation (cognitive effects).

Although the effects of emotional dysregulation and number of training sessions could not well be distinguished, there is an indication that when used with imagery instructions and more training sessions, benign CBM-I can be employed as a useful complementary treatment to usual psychotherapies.

Clinically relevant anxiety and depression are common, affecting up to 30% of individuals during their lifetime (Furmark et al., 2002; Gopinath, Katon, Russo, & Ludman, 2007). These conditions are distressing, disruptive, and costly (Furmark et al., 2002), prompting calls for more research into their treatment and prevention. Cognitive bias modification (CBM) training may be able to partially address this need. CBM training aims to modify information-processing biases linked to anxiety and depression through computerized, repeated practice that reinforces more adaptive styles of processing. An added benefit of CBM is that it can be implemented with relative ease. While recent individual studies appear promising, questions remain on the degree to which CBM can actually ‘correct’ biases and reduce symptomatology.

CBM for interpretation biases (CBM-I) focuses on modifying interpretation biases and draws on extant data from cognitive psychology showing that anxiety and depression are characterized and maintained by negative interpretation biases. This is the tendency to draw negative interpretations from ambiguous stimuli (Koster, Fox, & MacLeod, 2009; Mogg & Bradley, 2005). CBM-I capitalizes on these findings by aiming to ‘train’ more adaptive interpretational styles. There are three main training methods: the homograph paradigm (SGrey & Mathews, 2000), the word-sentence association task (WSAT) paradigm (Beard & Amir, 2008), and, perhaps most widely-used, the ambiguous situations (AS) paradigm (Mathews & Mackintosh, 2000). In each paradigm, ambiguous stimuli are presented to participants across multiple trials – with trials ending with a response from the participant that resolves the stimulus in a benign direction. For example, stimuli can be ambiguous homographs, which are words that have both a negative and a benign meaning (e.g., ‘patient’, which can be interpreted as a doctor’s patient implying something negative like illness, but can also be interpreted as the human capacity of endurance implying a positive character trait), or ambiguous sentences and situations, which can be interpreted in either a negative or benign manner (e.g., “*Having finished painting the lounge, you invite friends around to dinner. As they walk into the room, you can see that they are surprised. Their reaction is one of...*” can be interpreted to mean that the friends are pleased or displeased). Stimulus presentations are then followed by a “probe” to which the participants must respond. Usually these “probes” are word fragments, which must be completed by the participant by indicating the first missing letter (e.g., in case of the ambiguous situation described above, this could be pl—sure for pleasure or h—rr—r for horror). By correctly completing these words, the ambiguity of the stimulus is resolved in a benign direction for the benign CBM-I condition. During comparison conditions, the stimuli are resolved negatively or neutrally. Usually, the word fragments are presented in such a way that there is only one possible solution. To further encourage a valenced interpretation

of the ambiguous stimulus, trainings have made use of adding a comprehension question (e.g., in case of the ambiguous situation described above, this could be “Did your friends like your lounge?” with a correct ‘Yes’-response for the benign and ‘No’-response for the negative training condition) and feedback about response accuracy (Mathews & Mackintosh, 2000).

Training effectiveness is typically assessed with either change in measures of interpretation biases and mood or after training only. Additionally, assessing subsequent mood-reactivity to emotional challenges aims to address whether CBM-I is able to impact on exuberant increase in negative mood in response to stressful stimuli. Several emotional challenges have been employed to investigate whether CBM-I can impact on responses to ‘stress’ stimuli. Stressors have included videos showing accidents, e.g., see (Mackintosh, Mathews, Yiend, Ridgeway, & Cook, 2006), unsolvable anagram tasks, e.g., see (Salemink, van den Hout, & Kindt, 2009), or negative mood inductions, e.g., see (Holmes, Lang, & Shah, 2009). While earlier CBM-I studies used training methods to investigate whether generating benign and negative interpretations would, in unselected/healthy samples, alter mood and/or mood reactivity, more recent studies have assessed the suitability of these training methods to *improve* anxious and depressive moods in subclinical and clinical samples.

Other than individual studies, there is initial meta-analytic evidence for the effectiveness of benign CBM-I. Hallion and Ruscio (2011) investigated the effects of CBM-I in combination with another training package, CBM for attention (CBM-A). They reported significant and large post-training differences in interpretation biases and mood states of benign CBM-I as compared to control conditions (Hallion & Ruscio, 2011). Although these findings are very encouraging, some further elaboration is required to address the potential clinical effectiveness of benign CBM-I for several reasons. First, based on the results above, the possibility remains that a pre-existing tendency to select positive over negative interpretations (i.e., an optimistic bias) can account for the large post-training differences between the benign and control conditions. Although participants were randomly assigned to training conditions, significant post-training differences between training conditions could also have arisen solely from *trained* differences in the opposite direction among the comparison groups. Comparison groups have included ‘no training’ to control for natural fluctuations for that sample on outcome measures; ‘neutral training’ (where there is roughly an equal number of benign and negative resolutions across training items) to control for possible ‘placebo’ effects of CBM-I training; and ‘negative training’ (where consistently negative resolutions of ambiguity are presented) to explore whether post-training differences between

benign and negative training are driven by negative changes in the negative training, positive changes in the benign training, or both. If post-training differences between groups (Hallion & Ruscio, 2011) were mainly driven by negative *changes within* the comparison groups, the clinical benefit of benign CBM-I would be questionable.

Second, although the existing meta-analysis (Hallion & Ruscio, 2011) suggests significant post-training differences between groups on anxiety and depression symptoms, the effects were rather small after an emotional challenge (Hallion & Ruscio, 2011). Anxiety and depression symptoms were conceptualized by merging mood-measures and more elaborate symptom-measures such as the Hamilton Anxiety Rating Scale or the Beck Depression inventory. However, several symptom scales were developed for use *within* clinical samples only and may vary minimally in undiagnosed participants (Beck, Steer, & Garbin, 2011; Hamilton, 1960) – with whom the majority of CBM-I studies have been conducted (Hallion & Ruscio, 2011). Furthermore, findings on training effects on symptoms have been mixed in individual studies, e.g. see (Mackintosh et al., 2006; Yiend, Mackintosh, & Mathews, 2005), and finally, changes in symptoms such as rumination immediately after CBM-I – even in clinical samples – may be considered unlikely. In contrast, continuous variation in mood states not only may be used to characterize the common symptom ‘low mood’ of anxiety and depression but also can be measured reliably in general population samples (APA, 2000; Csikszentmihalyi & Hunter, 2003), particularly in response to emotional challenges (Blackhart, Nelson, Knowles, & Baumeister, 2009). Clearly addressing the effects on CBM-I on mood-states only would help drawing a clearer picture about its potential clinical effectiveness.

Besides expanding Hallion & Ruscio’s pioneering meta-analysis by addressing both the change within the benign condition and the effects on mood only, we most of all need to know how to optimize effects and pin down the factors that make it effective (Hallion & Ruscio, 2011). It is currently unclear whether one of the various CBM-I paradigms used (i.e., homograph, AS, and WSAT) is more effective than others. It has also been suggested by individual studies that training is more effective when participants have to (a) *generate* words and meaning of ambiguity themselves instead of simply being exposed to them (Mathews & Mackintosh, 2000), and (b) *imagine* the situations happening to themselves instead of processing stimuli more passively (Holmes et al., 2009; Holmes, Mathews, Dalgleish, & Mackintosh, 2006). Furthermore, it has been proposed that the administration of feedback about response accuracy will reinforce participants “for making valenced interpretations” (p.606) (Mathews & Mackintosh, 2000, p. 606). However, to the best of the authors’ knowledge, this has not yet been systematically tested.

Additional training characteristics, such as the mode of stimuli presentation (visual or auditory), and the *ratio* of training items in the benign training direction to the total number of presented items (i.e., does the fraction of benign items matter?) also frequently vary between CBM-I studies but have not yet been systematically investigated. Recent studies have also specifically assessed the potential clinical use of CBM-I across several training sessions (Amir & Taylor, 2012; Beard & Amir, 2008; Williams, Blackwell, Mackenzie, Holmes, & Andrews, 2013), suggesting that multiple sessions will increase effectiveness. Finally, it is commonly accepted that people with emotional symptoms suffer from negative interpretation biases, e.g., see (Koster et al., 2009; Mogg & Bradley, 2005) and women present more often with emotional problems than men, e.g. see (Husky, Mazure, Maciejewski, & Swendsen, 2009). It is therefore of clinical relevance to investigate whether benign CBM-I is particularly effective in symptomatic and female samples.

In the current meta-analysis, we first explore benign CBM-I as a possible clinical tool by examining the within group *change* in interpretational style and mood state and second, we assess factors that increase the effectiveness of benign CBM-I. Because we are primarily interested in the potential benefits of benign CBM-I, we conceptualized changes in interpretation bias as increases in positive interpretations (i.e., defined as an adaptive interpretation bias encompassing both benign and explicit positive interpretation styles) as opposed to decreases in negative interpretations. Indeed, previous research has shown that healthy individuals more likely draw positive than negative interpretations of ambiguous situations, whereas individuals with a current anxiety disorder are more likely to draw almost as many positive as negative interpretations (Eysenck, Mogg, May, Richards, & Mathews, 1991). Therefore, the difference between positive and negative interpretation bias after training has been of great interest to researchers and these indices were reported in the majority of studies. We therefore also assessed the post-training endorsement of positive versus negative interpretations.

Next, we assessed changes in negative mood from pre- to post-training as well as in response to an emotional challenge. In a second step, we compared these changes/differences in the benign CBM-I group with various control conditions and explored the degree to which the *change* in interpretation bias was associated with the *change* in mood in response to benign CBM-I training. To the extent that benign CBM-I was effective at changing interpretational style and mood within the benign condition, we investigated whether this effect varied across various training characteristics but also across sample characteristics, such as age, sex, and the inclusion of high symptomatic individuals.

Methods

Protocol

The protocol for reviewing the articles was developed by the first author based on the suggested strategy by Lipsey and Wilson (Eysenck et al., 1991) and adhered to the PRISMA guidelines (**Table S1**). Overall, the protocol consisted of a detailed description of the criteria to be employed for the search strategy, deciding on study eligibility, coding of the necessary variables, and procedures for resolving disagreements in coding (please see below ‘eligibility criteria’, ‘information sources and search’, ‘study selection’, ‘data collection process’, and ‘data items’ for the detailed description).

Eligibility criteria

Eligibility was assessed based on five criteria. First, the study aimed to target interpretation biases through any kind of training. Second, for studies with more than one training group, allocation to training groups was done at random to prevent confounding of the data by any third variables such as personal preference. Third, interpretation biases and/or mood state had to be assessed as the outcome measure to allow the assessment of training effectiveness. Fourth, the sample consisted of healthy adults, adults with subclinical/high-trait symptoms, or adults with clinical diagnoses of any anxiety or major depressive disorder. We did not include studies of child and adolescent samples as the relationship between cognition and mood may vary across development (Neshat-Doost, Taghavi, Moradi, Yule, & Dagleish, 1998). Fifth, studies had to be published in English, German, or Dutch. All identified publications including articles, conference abstracts, and dissertation abstracts were considered eligible.

Information sources and search

In November and December 2010, databases (i.e., EMBASE, Medline, PsychArticles, Psychology & Behavioral Science Collection, PsycINFO, Science Direct, and Web of Science) were searched simultaneously with the key words “interpret* bias AND training” and “interpret* bias AND modif*” for publications in this area between 1992 and 2010. To the authors’ knowledge, CBM-I was first introduced in 2000; additional searches for the years 1992 through 1999 were conducted to ensure that no earlier and possibly less popular accounts were missed. All reference lists of identified articles were cross-checked for further relevant articles. All authors of identified articles were also contacted with a request to send any additional relevant literature and/or unpublished data that might be appropriate for inclusion into the meta-analysis. Finally, follow-up literature searches were conducted in October 2011 and June 2013 for studies published since our first and second search.

Study selection

All hits were screened in a standardized fashion adhering to the PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-analyses; Moher, Liberati, Tetzlaff, Altman, & The-PRISMA-Group, 2009) involving: screening the titles, abstracts, and finally the whole article for eligibility criteria. The whole screening procedure was conducted independently by two researchers (CML and PH for the first two searches, and CML and SH for the last search). If researchers did not agree on inclusion, the article was nevertheless included in the next screening stage to minimize chances of premature exclusion. All articles fulfilling eligibility criteria were included in the meta-analysis if sufficient data were available (either provided in the text or by researcher on request).

Data collection process

All included articles were independently coded by at least two researchers (CML, PH, ZK, and SH) using a standardized coding protocol and appropriate forms (available upon request from the first author; also see 'protocol'). The summary statistics necessary to compute the effect sizes (see below) were retrieved from the articles by at least two independent coders (CML, PH, ZK, and SH), compared, and if necessary adapted. All missing data were requested from authors via e-mail. As most primary articles report between-group differences at post-training rather than within-group comparisons, almost all authors were contacted. All but five out of 35 of those contacted responded positively (response rate = 86%).

Data items

Articles were first coded on the basis of inclusion of the different training conditions used. We distinguished between four conditions: (i) Benign training, where the majority of ambiguous stimuli were resolved in a positive or non-negative direction; (ii) Negative training, where the majority of ambiguous stimuli were resolved in a negative, threatening, or harmful direction; (iii) Neutral training, where ambiguous stimuli were resolved in an overall neutral direction (either by presenting the same amount of stimuli in benign and negative directions, or by resolving stimuli in a neutral direction); and (iv) no training, in which participants were not exposed to any form of interpretation bias training but were simply tested and re-tested on selected outcome measures. Outcome measures were categorized into one of three categories, namely cognition (i.e., any kind of interpretation bias assessment), mood, or other. For each available outcome measure, (a) the sample size (N or df), (b) the mean before and after training or the mean difference between before and after training, (c) the standard deviation, variance, or standard error for before and after training or for the difference

between before and after training, and (d) the correlation or the dependent sample t-value between before and after training were recorded.

Articles were also coded with regard to potential moderators, consisting of sample and training characteristics. Sample characteristics included age (mean age of whole sample), sex (percentage of men in the whole sample), and presence of high levels of anxiety and depressive symptoms (including clinical diagnosis) in participants. Training characteristics included the type of training paradigm used (AS, homograph, WSAT, or other), format of training (generation of the meaning of words and situations versus simple exposure), pre-training instructions (presence or absence of the use of mental imagery), modality of training (visual or auditory), use of feedback (presence or absence of feedback about participant's response (correct/incorrect)), the training ratio (ratio between the number of stimuli in the training direction to the total number of stimuli), and frequency (the total number of training sessions).

Risk of bias in individual studies

Risk of bias in individual studies was attempted to be kept at a minimum by making randomization to training condition an inclusion criteria for those reports including more than one training condition. For those studies that administered training across several sessions, attrition (percentage of drop-out) within each training condition was coded as a proxy for risk of selected attrition in the benign condition.

Summary measures

Individual effect size estimates were computed for each study and for each training condition separately across the following four (within-group) contrasts: (i) post-training endorsements of positive versus negative interpretations; (ii) pre-training versus post-training endorsements of positive interpretations; (iii) pre-training versus post-training negative mood ratings; and (iv) pre-emotional challenge versus post-emotional challenge negative mood ratings. Please note that 'positive' interpretations were defined as including both benign and 'non-threat' as well as explicit positive interpretations. As the effect size measure for the meta-analysis, we used the standardized mean difference for the difference between positive versus negative interpretation endorsements post-training (contrast i) and the standardized mean change for the pre- to post-training/emotional challenge contrasts (contrasts ii through iv). Standardization was based on the differences and change scores, respectively (Gibbons, Hedeker, & Davis, 1993).

In particular, the standardized mean change for each condition was computed with $d = \text{Mean}_{\text{diff}} / \text{SD}_{\text{diff}}$, where $\text{Mean}_{\text{diff}}$ denotes the mean of the change scores, $\text{SD}_{\text{diff}} = \sqrt{(\text{SD}_{\text{pre}}^2 + \text{SD}_{\text{post}}^2 - 2r\text{SD}_{\text{pre}}\text{SD}_{\text{post}})}$, and r denotes the correlation between the pre- and post-training/emotional challenge assessments (for the contrast of the endorsements of the positive versus negative interpretations, $\text{Mean}_{\text{diff}}$ denotes the mean of the endorsement differences, $\text{SD}_{\text{diff}} = \sqrt{(\text{SD}_{\text{positive_bias}}^2 + \text{SD}_{\text{negative_bias}}^2 - 2r\text{SD}_{\text{positive_bias}}\text{SD}_{\text{negative_bias}})}$, and r denotes the correlation between the endorsements of the positive and negative interpretations). If SD_{diff} was not reported, the paired-samples t-test value was employed to calculate the effect size with $d = t/\sqrt{n}$. The sampling variance of the d -values was calculated with $v = 1/n + d^2/2n$, where n denotes the group size.

Sometimes a study would provide sufficient information to compute multiple d -values for the same sample for a particular contrast (e.g., when more than one mood scale was used to assess mood state pre- and post-training). To avoid the problem of non-independent effect size estimates in these samples, we selected only one measure based on an *a priori* established preference list (Cooper, 2010). In general, measures employed more often and assessing the outcome construct more directly were preferred over other measures (see **Table S2**). Moreover, if the same group of subjects underwent more than one training (whether it be a different training condition or a variation of the same training condition) we only computed d for the first training condition the group was exposed to. On the other hand, if different groups of subjects underwent slight variations of the same training condition within the same study (e.g., when one group was exposed to benign training with *visual* stimuli presentation and another group was exposed to benign training with *auditory* stimuli presentation), then multiple d -values for that training type (e.g., benign training) could be extracted (while still preserving the statistical independence of the d -values).

Correlations between the pre- and post-training/challenge measurements (or post-training positive and negative interpretations) were calculated for all studies that reported the necessary values. Of note, r can be inferred when only SD_{diff} , SD_{pre} , and SD_{post} are known. For studies where r was unknown and SD_{diff} had to be computed in order to obtain d , the mean correlation was employed to impute r (this was done separately for each of the four training conditions per contrast). Therefore, instead of leaving out these studies, this approach allowed us to include more samples in the meta-analysis, namely another 17 samples for post-training positive versus negative endorsements, 12 samples for change in negative mood, and two samples for the increase in negative mood in response to an emotional challenge.

To summarize, for each of the four key contrasts (i.e., interpretation bias as assessed with post-training positive versus negative interpretation endorsements; interpretation bias as assessed with change in positive bias from pre- to post-training; immediate change in mood state pre- to post-training; and change in mood state in response to an emotional challenge) a set of effect size estimates across studies was obtained, describing the degree of the difference or the amount of change for each training condition. Depending on the number of independent training conditions employed by a primary study, one, two, or more d -values could be extracted for a particular contrast from each study.

Synthesis of results

Changes within groups and comparison between groups

Due to the multilevel structure of the data (with multiple effect size estimates nested within the studies), we used a meta-analytic multilevel mixed-effects model analogous to the one described by Salanti et al. (2008) for the analyses. In particular, random effects were added at the study level (to account for the fact that the size of the effects may generally be larger or smaller across all conditions examined within a study) and at the effect size level (to account for heterogeneity in the size of the treatment effects). Dummy variables for the four different training conditions were added to the model, so that we could estimate the (average) standardized mean change/difference for each training condition. Furthermore, we could then compare these changes/differences *within* each training condition *between* the four training conditions, yielding six pairwise comparisons (benign-neutral, benign-no training, benign-negative, neutral-no training, neutral-negative, no-training-negative) to control for natural fluctuations for that sample on outcome measures (benign-no training comparison), for possible ‘placebo’ effects of CBM-I training (benign-neutral training comparison), and to explore whether post-training differences between benign and negative training are driven by negative changes in the negative training, by actual positive changes in the benign training, or by both (benign-negative training comparison).

An omnibus Wald-type test was used to test for any differences between the four training conditions. Similarly, the average standardized mean change for each training condition and the pairwise contrasts were tested for significance at $\alpha = .05$ (two-sided). We also report 95% confidence intervals for the estimated averages and pairwise contrasts. Finally, likelihood ratio tests were conducted to test whether the variance in the random effects at the study and the effect size level was significantly greater than zero.

Significant heterogeneity at the study level indicates that, due to nonspecific study characteristics, effect sizes for all training conditions can be larger or smaller in one study than another. Significant heterogeneity at effect size level indicates that the effect of a particular training condition is not constant across studies. Therefore, if significant heterogeneity at the study or effect size level for a particular outcome is found, then the pooled effect sizes based on the meta-analytic models need to be interpreted as depicting the *average* size of the effects.

Factors influencing changes within benign CBM-I

Given our interest in factors that enhance the effects of benign training, secondary meta-regression analyses were conducted to examine potential moderating effects of the various sample and training characteristics. To reduce the number of tests, these analyses were only conducted for the benign training condition, given our *a priori* interest in the effectiveness of this particular training type. Moreover, only contrasts that yielded a significant effect in primary analyses were explored further in these secondary analyses.

The first moderator was training paradigm, in which we distinguished between the ambiguous situations (AS) task versus all other paradigms combined. We did this for two reasons: (i) the AS task was most widely employed, whereas few studies employed the WSAT and homograph task for particular outcomes, and (ii) the AS task has the highest ecological validity, as it describes common everyday life ambiguous situations and might therefore more directly relate to people's real lives than words or sentences (that are employed in the other training paradigms).

Next, we explored training characteristics as moderators, including use of *imaging* instructions (present/absent), instructions to *generate* meaning (present/absent), inclusion of *feedback* (present/absent), and mode of *presentation* (auditory/visual). Sample characteristics that were also examined included: participant *status* (healthy/symptomatic), coded dichotomously, and *age*, *sex* (proportion men), training *frequency* (nr. of training sessions), and *ratio* of training items in the intended training direction (benign or negative) to the total number of presented items, included as continuous variables in the meta-regression models.

For dichotomous variables, we compared effect sizes for the presence and absence of the dichotomous variables (auditory as compared to visual, symptomatic as compared to healthy, for *presentation* and *status*, respectively), reflecting the difference in effect size for the two levels of the moderator. For continuous variables, we assessed

whether greater levels of potential moderators enhanced or attenuated the effects of the benign training condition, reflected by the change in the size of the effect for a one-unit increase in the moderator.

Each of the moderators was added to the meta-regression separately (we were unable to enter the moderators simultaneously as the majority of data would be lost due to missing values (also see table 1), which would have drastically reduced the power to find any relationships). However, since training paradigm (AS, homograph, and WSAT) is a rather complex variable possibly differentially influencing the effect of other training characteristics, we controlled for type of paradigm (whenever training paradigm resulted in a significant moderating effect) when analyzing the more specific training characteristics. Therefore, we investigated whether any given training-characteristic moderated training effectiveness *above* and *beyond* the training paradigm employed. Additionally, we examined the size of the correlations for all other moderators to explore whether any of these were strongly associated. If two moderators were strongly correlated and both revealed a significant effect, it would be difficult to conclude which of the two moderators underlies this effect.

The analyses were carried out for each of the four contrasts separately. Restricted maximum likelihood (REML) estimation was used to fit the models. All of the analyses were carried out with R and S-Plus, using the metafor (Viechtbauer, 2010) and the nlme packages (Pinheiro & Bates, 2000).

Risk of bias across studies

We also examined the presence of publication biases visually (by inspecting the presence of asymmetry in funnel plots) and by including and testing the inverse of the sample size as a potential covariate in the models (Peters, Sutton, Jones, Abrams, & Rushton, 2006). A significant relationship between the inverse of the sample size and the observed *d*-values may be suggestive of publication bias.

Additional analyses

Finally, correlational analyses were employed across all training conditions to investigate whether the change/difference in interpretation bias was associated with changes in negative mood in response to training and in response to an emotional challenge.

To address the potential risk of bias within studies, the main analyses were repeated excluding studies with only one training condition and therefore no random assignment to training condition and excluding conditions that had an attrition rate of > 15%.

Interpretation of Effect Sizes

The standardized mean change values were computed in such a way that more positive (or less negative) values indicate more preferable outcomes (i.e., stronger endorsement of the positive instead of the negative interpretations after the training, stronger endorsement of positive interpretations post- versus pre-training, decreased negative mood post- versus pre-training, and a less pronounced decrease in negative mood post- versus pre-challenge).

Letting $\Phi()$ denote the cumulative density function of a standard normal distribution, the interpretation of the standardized mean change can also be facilitated by noting that $\Phi(d)$ estimates the proportion of individuals for which the difference or change scores reflect a preferable outcome (Morris & DeShon, 2002). For example, an effect size of 0 implies that $\Phi(0) = .50$ (i.e., 50%) of individuals should have a larger positive interpretation bias after the training than before (while 50% have a smaller positive interpretation bias) or that 50% of individuals have a decreased negative mood after the training (while 50% have an increase in negative mood). For an effect size of 0.2 (a “small” effect), the positive interpretation bias is expected to increase (and negative mood is expected to decrease) for $\Phi(0.2) = .58$ (i.e., 58%) of individuals. Effect sizes of 0.5 (a “medium” effect) and 0.8 (a “large” effect) correspond to an increase in positive interpretation bias (and a decrease in negative mood) in 69% and 79% of individuals, respectively (Morris & DeShon, 2002).

Results

Study selection

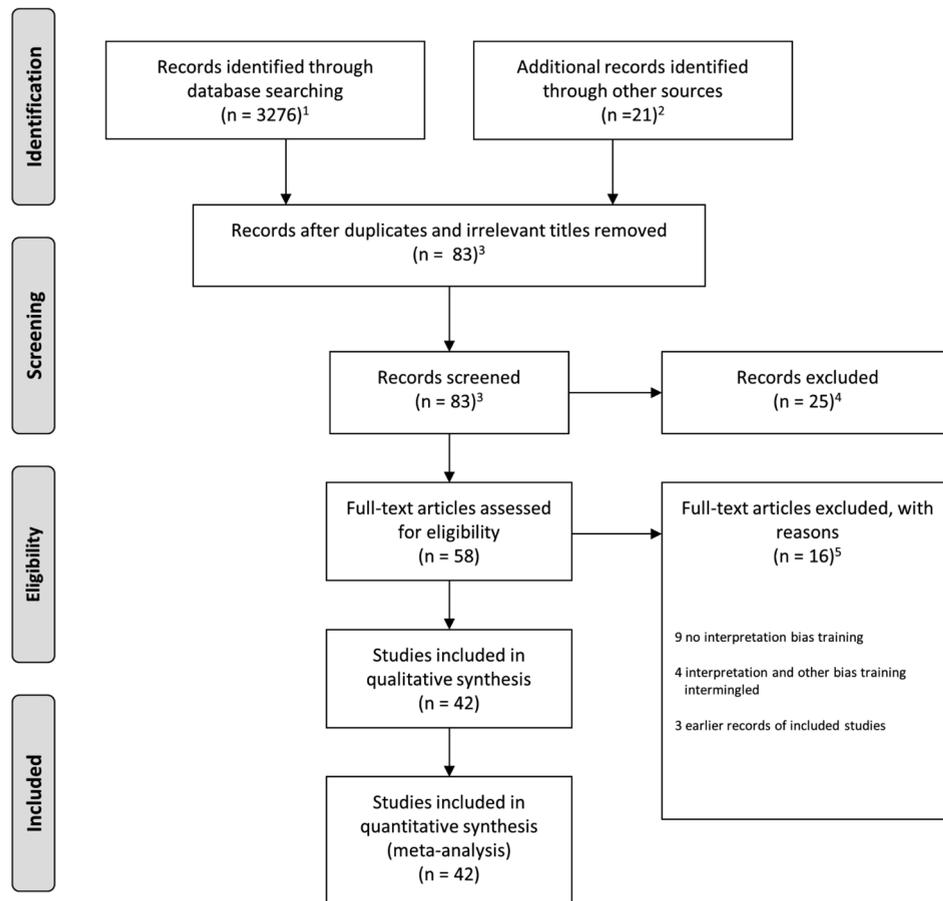
Articles were retrieved according to the PRISMA guidelines (Moher et al., 2009). The numbers of articles screened, and included (and reasons for exclusion) can be found in **figure 1**.

Study characteristics

All study characteristics are reported in table 1 (Nader Amir, Bomyea, & Beard, 2010; Amir & Taylor, 2012; Beard & Amir, 2008; Blackwell & Holmes, 2010; Bowler et al., 2012; Clerkin & Teachman, 2011; Grey & Mathews, 2000; Grey & Mathews, 2009; Hayes, Hirsch, Krebs, & Mathews, 2010; Paula T. Hertel, Mathews, Peterson, & Kintner, 2003;

Hertel, Vasquez, Benbow, & Hughes, 2011; Hirsch, Hayes, & Mathews, 2009; Hirsch, Mathews, & Clark, 2007; Holmes et al., 2009; Holmes et al., 2006; Hoppitt, Mathews, Yiend, & Mackintosh, 2010a; Hoppitt, Mathews, Yiend, & Mackintosh, 2010b; Lang, Blackwell, Harmer, Davison, & Holmes, 2012; Lang, Moulds, & Holmes, 2009; Lange et al., 2010; MacDonald, Koerner, & Antony, 2013; Mackintosh et al., 2006; Mathews & Mackintosh, 2000; Mathews, Ridgeway, Cook, & Yiend, 2007; Murphy, Hirsch, Mathews, Smith, & Clark, 2007; Saleminck & van den Hout, 2010a, 2010b; Saleminck, van den Hout, & Kindt, 2007a, 2007b; Saleminck et al., 2009; Saleminck, van den Hout, & Kindt, 2010c; Standage, Ashwin, & Fox, 2009, 2010; Steel et al., 2010; Steinman & Teachman, 2010; Teachman & Addison, 2008; Tran, Hertel, & Joormann, 2011a; Tran, Siemer, & Joormann, 2011b; Turner et al., 2011; Williams et al., 2013; Wilson, MacLeod, Mathews, & Rutherford, 2006; Yiend et al., 2005). In total 42 articles met inclusion criteria (**figure 1**), of which 28 articles administered the ambiguous situations (AS), six the homograph, four the word-sentence association task (WSAT), and four a combination of two training paradigms or another training paradigm. Combined these articles reported on a total of 59 independent studies (also see **table 1**). Thirty-eight studies were conducted in healthy participants, and 21 studies in symptomatic individuals. The great majority of studies assessed interpretation bias ($k=50$) and mood ($k=48$) as outcome measures, whereas fifteen studies assessed also reactivity to an emotional challenge as an additional outcome measure. Within these studies a total of 125 independent samples were included. Sixty-three samples received a benign, 16 a neutral, 38 negative, and eight no training (for more detail please see **table 1**). In 80 samples imagery instructions were administered, in 60 samples participants were instructed to generate words or meaning of ambiguity, 94 samples received training in a visual format, 18 in an auditory format, and six samples received a combination of both, and finally, 90 samples received feedback during task administration. In total 2526 individuals were included in the analysis.

Figure 1. PRISMA flow chart of article retrieval and selection. ¹the search in June 2013 was conducted on separate searches of the data-bases as the software to do simultaneous searches was no longer available. In total, 109 hits were identified in June 2013. ²zero records were excluded for the search in 2013. ³ten records were screened for the search in 2013. ⁴four articles were excluded for the search in 2013. ⁵three articles were excluded for the search in 2013.



Risk of bias within studies

To address the question whether random assignment and attrition affects results these variables were coded for each study. In total, five included only one training condition and therefore had no random assignment to training group (see also **table 1**): Blackwell and Holmes (2010) trained seven individuals to adopt benign interpretations, Grey and Mathews (2000) administered neutral interpretation bias training to 20 individuals, Holmes and colleagues (2009) trained 20 individuals to adopt benign interpretations, Mathews and Mackintosh (2000) assessed the interpretation bias in 12 individuals in a ‘no training’ condition, and finally Turner & colleagues (2011) trained eight individuals to adopt benign interpretations.

Seven independent studies administered interpretation bias training more than once (see also ‘no of sessions (frequency)’ in **table 1**) and had the following rates of attrition per training condition: Amir and Taylor (2012) 15% in the benign and 13.04% in the neutral training condition, Beard and Amir (2008) 0% in both the benign and the neutral training condition, Blackwell and Holmes (2010) 12.5 % in the one benign training condition, Bowler and colleagues (2013) 12.29% in the benign and 0% in the no training condition, Lang and colleagues (2012) 7.10% in both the benign and the neutral training condition, Mathews and colleagues (2007) 0% in both the benign and no training condition, Salemink and colleagues (2009) 5.56% in both the benign and neutral training condition, and Williams and colleagues (2012) 31.58% in the benign and 12.90% in the no training condition.

Results of individual studies

Effect size estimates for the change in positive bias and negative mood from before to after training as well as for the change in mood from before to after the emotional challenge are reported per study and training condition in the forest plots in figures 2-5.

Table 1. Study descriptives

Article	study	General						Outcome measures			Training conditions					
		health status ¹	training paradigm ²	mean age	% male (sex)	training/total items (ratio) ³	No of sessions (frequency)	interpretation bias	mood-state	emotional challenge	condition	N	imagery instructions (y/n)	word generation (y/n) ⁴	mode of presentation ⁵	Feedback administration (y/n) ⁶
Amir et al (2010)	1	anxious	WSAT	19,50	47	1,00	1	IB questionnaire	STAI-S	NA	benign	29	n	n	visual	Y
											neutral	28	n	n	visual	Y
Amir & Taylor (2012)	1	anxious	WSAT	31,00	29	1,00	12	WSAT	STAI-T/Libowitz Anxiety Scale/SPAI	NA	benign	20	NA	n	visual	Y
											neutral	23	NA	n	visual	Y
Beard & Amir (2008)	1	anxious	WSAT	20,00	7	0,69	8	WSAT	SPAI, STAI-T	NA	benign	13	n	n	visual	Y
											neutral	14	n	n	visual	Y
Blackwell & Holmes (2010)	1	depressed	AS	37,70	29	1,00	7	VAS depressive bias, SST	PAS, NAS	NA	benign	7	Y	n	auditory	n
Bowler et al (2012)	1	anxious	AS	22,70	32	1	4	SST	STAI-T/FNES	NA	benign	21	Y	Y	visual	Y
											no training	21	NA	NA	NA	NA
Clerkin et al (2011)	1	anxious	AS	18,76	35	0,92	1	similarity ratings	NAS	car accident sentence about best friend	benign	50	Y	Y	visual	NA
											neutral	49	Y	Y	visual	NA
Grey & Mathews (2000)	1	healthy	homograph	NA	45	0,33	1	RT word fragment	NA	NA	negative	20	n	Y	visual	Y
											benign	20	n	Y	visual	Y
	2	healthy	homograph	NA	NA	0,40	1	lexical decision task	NA	NA	negative	20	n	Y	visual	Y
											benign	17	n	Y	visual	Y
	3	healthy	homograph	NA	NA	0,40	1	lexical decision task	NA	NA	negative	20	n	n	visual	Y
											benign	20	n	n	visual	Y
	4	healthy	homograph	NA	NA	0,00	1	lexical decision task	NA	NA	neutral	20	n	n	visual	Y
Grey et al (2009)	1	healthy	homograph	NA	45	0,40	1	lexical decision task	NA	NA	benign	18	n	n	visual	Y
											negative	19	n	n	visual	Y
Hayes et al (2010)	1	anxious	homograph & AS	42,00	23	0,80	1	NA	VAS anxiety/ depression/ happy	worry intrusion	benign	20	n	both	both	Y
											neutral	20	n	both	both	Y
Hertel et al (2003)	1	healthy	homograph	NA	33	0,40	1	form image	NA	NA	no training	18	n	n	visual	Y
											negative	18	n	n	visual	Y
											benign	17	n	n	visual	Y
	2	healthy	homograph	NA	36	0,40	1	form image	NA	NA	benign	22	n	n	visual	Y
											negative	22	n	n	visual	Y

Article	study	General						Outcome measures			Training conditions					
		health status ¹	training paradigm ²	mean age	% male (sex)	training/total items (ratio) ³	No of sessions (frequency)	interpretation bias	mood-state	emotional challenge	condition	N	imagery instructions (y/n)	word generation (y/n) ⁴	mode of presentation ⁵	Feedback administration (y/n) ⁶
Hertel et al (2011)	1	healthy	AS	NA	50	0,78	1	NA	VAS happy/distress/tension/ pessimism	NA	negative	16	y	Y	visual	NA
											neutral	16	Y	Y	visual	NA
	2	anxious	AS	NA	50	0,78	1	NA	VAS happy/distress/tension/pessimism	video stressor	benign	20	Y	Y	visual	NA
											neutral	20	Y	Y	visual	NA
Hirsch et al (2007)	1	healthy	AS	NA	17	0,90	1	NA	STAI-S, NAS	NA	negative	12	n	n	visual	Y
Hirsch et al (2009)	1	anxious	homograph & AS	35,65	20	0,86	1	NA	VAS anxiety/ depression/ happy	worry intrusion	neutral	20	n	both	both	Y
Holmes et al (2006)	1	healthy	AS	38,85	35	1,00	1	emotionality ratings	STAI-S, PAS	NA	benign	13	n	n	auditory	n
Holmes et al (2009)	1	healthy	AS	30,98	45	1,00	1	emotionality ratings	STAI-S, STAI-T, PAS	negative mood induction	benign	20	Y	n	auditory	n
Hoppitt et al (2010a)	1	healthy	AS	35,61	39	1,00	1	emotionality ratings	STAI-S	NA	negative	14	Y	Y	visual	Y
											negative	14	Y	n	visual	Y
											RT word fragment	STAI-S	video stressor	negative	22	n
Lang et al (2009)	1	healthy	AS	27,67	48	0,90	1	similarity ratings	PAS, NAS	video stressor, picture from video	benign	24	Y	Y	visual	Y
Lang et al (2012)	1	depressed	AS & other	28,45	23	1,00	6	SST, RIQ	STAI-T	NA	negative	24	Y	Y	visual	Y
											neutral	13	Y	both	both	n
Lange et al (2010)	1	anxious	AS	20,71	12	0,69	1	similarity ratings	STAI-S, Libowitz Anxiety Scale	NA	benign	34	Y	Y	visual	Y
											negative	34	Y	Y	visual	Y
MacDonald et al. (2013)	1	anxious	WSAT	32,79	4	1	1	BBSIQ	Anxiety sensitivity index	NA	benign	17	n	NA	visual	Y
											neutral	17	n	NA	visual	Y

Article	study	General						Outcome measures			Training conditions					
		health status ¹	training paradigm ²	mean age	% male (sex)	training/total items (ratio) ³	No of sessions (frequency)	interpretation bias	mood-state	emotional challenge	condition	N	imagery instructions (y/n)	word generation (y/n) ⁴	mode of presentation ⁵	Feedback administration (y/n) ⁶
Mackintosh et al (2006)	1	healthy	AS	NA	29	0,90	1	similarity ratings	STAI-S	NA	negative	24	Y	Y	visual	Y
											benign	27	Y	Y	visual	Y
	2	healthy	AS	NA	28	0,90	1	similarity ratings	STAI-S	accident video	benign	10	Y	n	visual	Y
											benign	10	Y	n	auditory	Y
Mathews & Mackintosh (2000)											negative	10	Y	n	visual	Y
	1	healthy	AS	NA	NA	0,69	1	similarity ratings	STAI-S	NA	negative	10	Y	Y	visual	Y
											benign	10	Y	Y	visual	Y
	2	healthy	AS	NA	NA	0,69	1	similarity ratings	STAI-S	NA	negative	10	Y	n	visual	Y
											benign	10	Y	n	visual	Y
	3	healthy	AS	NA	NA	NA	0	similarity ratings	NA	NA	no training	12	NA	NA	NA	Y
	4	healthy	AS	NA	NA	0,88	1	similarity ratings	STAI-S	NA	negative	13	Y	Y	visual	Y
											benign	13	Y	Y	visual	Y
	4a	healthy	AS	NA	NA	0,88	1	similarity ratings	STAI-S	NA	negative	13	Y	Y	visual	Y
											benign	13	Y	Y	visual	Y
Mathews et al (2007)	5	healthy	AS	NA	NA	0,60	1	similarity ratings	STAI-S	NA	negative	14	Y	Y	visual	Y
											negative	14	Y		visual	Y
											benign	14	Y	Y	visual	Y
	1	anxious	homograph	40,65	31	0,80	4	reason for events, similarity ratings	STAI-S, STAI-T	NA	benign	19	Y	n	visual	Y
Murphy et al (2007)											no training	20	NA	NA	NA	NA
	1	anxious	AS	20,60	26	1,00	1	similarity ratings	STAI-S	NA	benign	22	Y	n	auditory	Y
											benign *	22	Y	n	auditory	Y
Salemink et al (2007a)											neutral	22	Y	n	auditory	Y
	1	healthy	AS	20,80	24	0,69	1	similarity ratings	STAI-S	unsolvable anagram	benign	60	Y	Y	visual	Y
Salemink et al (2007b)											negative	58	Y	Y	visual	Y
	1	healthy	AS	21,10	NA	0,69	1	similarity ratings, EAST, IB questionnaire	STAI-S, STAI-T	NA	benign	40	Y	Y	visual	Y
Salemink et al (2009)											negative	41	Y	Y	visual	Y
	1	anxious	AS	21,30	18	0,69	8	similarity ratings, IB questionnaire	STAI-S, STAI-T, FNES	unsolvable anagram	benign	17	Y	Y	visual	Y
Salemink et (2010a)											neutral	17	Y	Y	visual	Y
	1	healthy	AS	21,10	9	0,69	1	NA	VAS positive/negative	mood induction	benign	21	Y	Y	visual	Y
											negative	20	Y	Y	visual	Y

Article	study	General						Outcome measures			Training conditions					
		health status ¹	training paradigm ²	mean age	% male (sex)	training/total items (ratio) ³	No of sessions (frequency)	interpretation bias	mood-state	emotional challenge	condition	N	imagery instructions (y/n)	word generation (y/n) ⁴	mode of presentation ⁵	Feedback administration (y/n) ⁶
Salemink et al (2010b)	1	healthy	AS	20,05	25,3	0,75	1	NA	STAI-S	NA	negative	38	Y	Y	visual	Y
Salemink et al (2010c)	1	healthy	AS	20,35	18	0,69	1	similarity ratings, IB questionnaire	NA	NA	benign	37	Y	Y	visual	Y
											negative	20	Y	Y	visual	Y
											no training	51	NA	NA	NA	NA
Standage et al (2009)	1	healthy	AS	22,08	21	1,00	1	similarity ratings	VAS anxiety/ depression	anxiety anticipation, speech	benign	12	Y	n	visual	Y
											benign	12	Y	n	auditory	Y
											negative	12	Y	n	visual	Y
Standage et al (2010)	1	healthy	other	17,60	17	1,00	1	SST	VAS anxiety/ depression	NA	negative	14	n	n	visual	n
											benign	14	n	n	visual	n
											benign	14	n	n	visual	n
Steel et al (2010)	1	anxious	AS	43,00	71	1,00	1	emotionality ratings	STAI-S	NA	benign	11	Y	n	auditory	n
Steinman et al (2010)	1	anxious	AS	18,93	31	0,69	1	NA	PANAS FS	anxiety-symptoms provoking breathing task	no training	25	NA	NA	NA	NA
											benign	25	Y	Y	visual	Y
											neutral	25	Y	Y	visual	Y
Teachman et al (2008)	1	anxious	AS	18,60	26	0,69	1	similarity ratings	NAS	NA	benign	20	Y	Y	visual	Y
											neutral	20	Y	Y	visual	Y
											no training	20	NA	NA	NA	NA
Tran et al (2011a)	1	healthy	AS	NA	52	0,69	1	similarity ratings	NAS	NA	benign	29	Y	Y	visual	NA
Tran et al (2011b)	1	healthy	AS	NA	62	0,69	1	similarity ratings	PAS, NAS	emotional faces/incorrect feedback	negative	29	Y	Y	visual	NA
											benign	25	Y	Y	visual	NA
											negative	25	Y	Y	visual	NA
Turner et al (2011)	1	anxious	AS	24,75	88	1,00	1	NA	VAS mood	NA	benign	8	Y	Y	visual	NA
Williams et al (2013)	1	depressed	AS	44,80	7,5	1,00	7	SST	NA	NA	benign	26	Y	n	auditory	n
Wilson et al (2006)	1	healthy	homograph	18,15	50	1,00	1	RT word fragment	VAS anxiety/ depression	video stressor	no training	27	NA	NA	NA	NA
											negative	24	n	Y	visual	Y
											benign	24	n	Y	visual	Y

Article	study	health status ¹	General					Outcome measures			Training conditions					
			training paradigm ²	mean age	% male (sex)	training/total items (ratio) ³	No of sessions (frequency)	interpretation bias	mood-state	emotional challenge	condition	N	imagery instructions (y/n)	word generation (y/n) ⁴	mode of presentation ⁵	Feedback administration (y/n) ⁶
Yiend et al (2005)	1	healthy	AS	32,50	40	0,90	1	similarity ratings	STAI-S, STAI-T	NA	benign	10	Y	Y	visual	Y
											negative	10	Y	Y	visual	Y
	2	healthy	AS	42,90	29	0,90	1	similarity ratings	STAI-S, STAI-T	NA	benign	12	Y	n	visual	Y
											negative	12	Y	n	visual	Y
	3	healthy	AS	39,60	37	0,90	1	similarity ratings	STAI-S, STAI-T	NA	benign	10	Y	Y	visual	Y
											negative	9	Y	Y	visual	Y

Note. y-YES, n-No; VAS = visual analogue scale; IB questionnaire = interpretation bias questionnaire, STAI-T/S = State Trait Anxiety Inventory – Trait/State; SST = Scrambled Sentence Test; PANAS = positive and negative affect scales; PAS = positive affect scale (from PANAS); NAS = negative affect scale (from PANAS); PANAS FS = fear subscale from PANAS; RT = reaction time; RIQ = Response to intrusions questionnaire, FNES = Fear of Negative Evaluation Questionnaire; EAST = extrinsic affective Simon task; NA = not available; * Murphy et al (2007) employed two independent benign groups that differed in that one group focused on emotional outcomes that were non-negative whereas the other group focused

on pure positive outcomes. As this distinction was not made in the current meta-analysis both these training conditions were included under the ‘benign’ training condition. ¹health status = healthy, anxious, depressed (note: this considers symptoms and clinical diagnoses); ²AS = ambiguous situations paradigm, WSAT = word sentence association task, ‘other’ paradigms included picture word interpretation, sentence completion (Lang et al, 2011), positive or negative valenced statements (Standage, 2010); ³training/total items (ratio) = number of items in training direction divided by total number of items in task; ⁴word generation = were participants required to actively generate emotional meaning?; ⁵mode of presentation = were stimuli presented visually or auditory?; ⁶feedback administration = was feedback about response accuracy administered.

Figure 2. Forest plot of post training difference between positive and negative interpretation bias. Order of same conditions within one study follow the order of table 1.

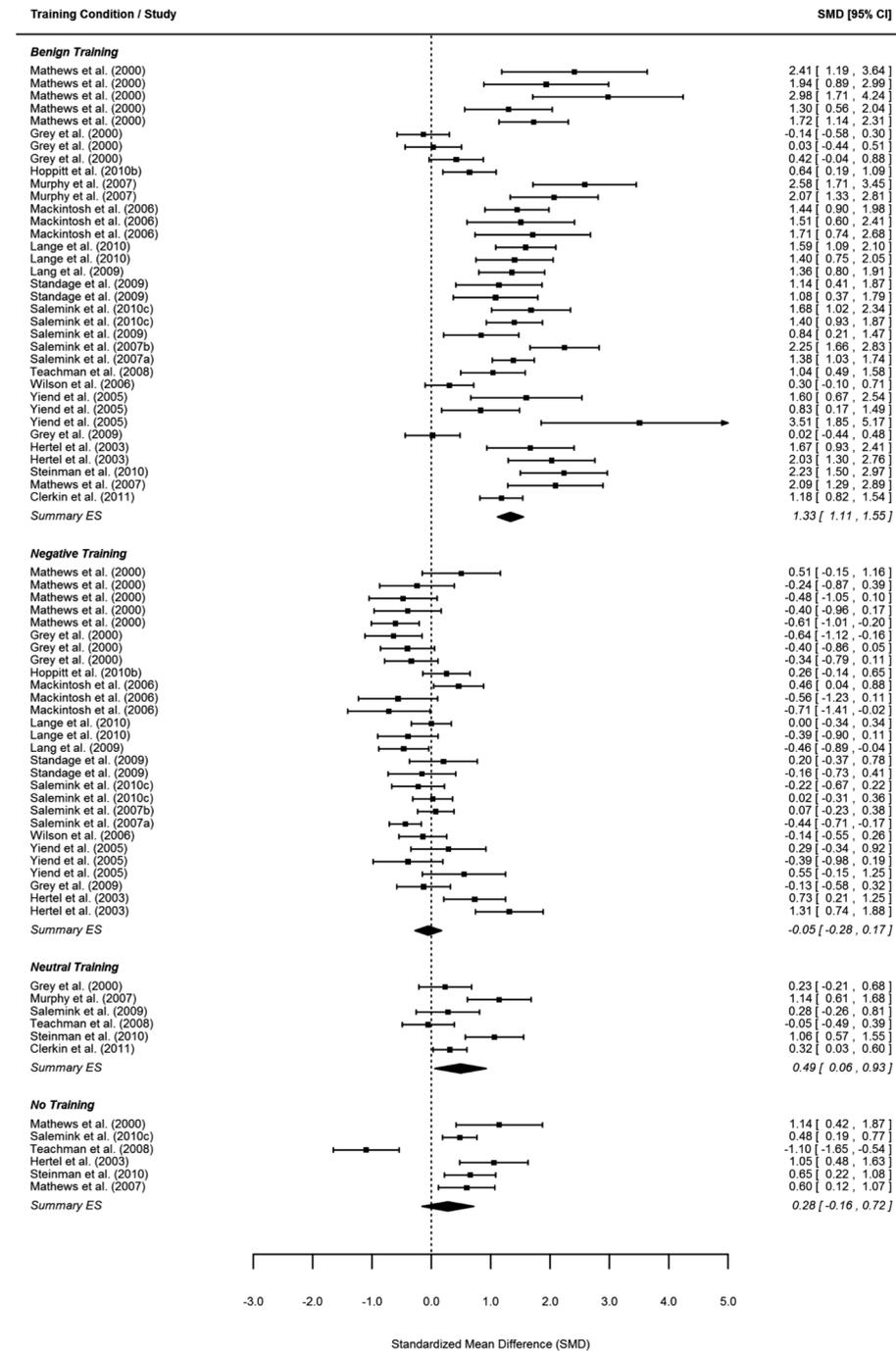


Figure 3. Forest plot of change in positive interpretation bias. Order of same conditions within one study follow the order of table 1.

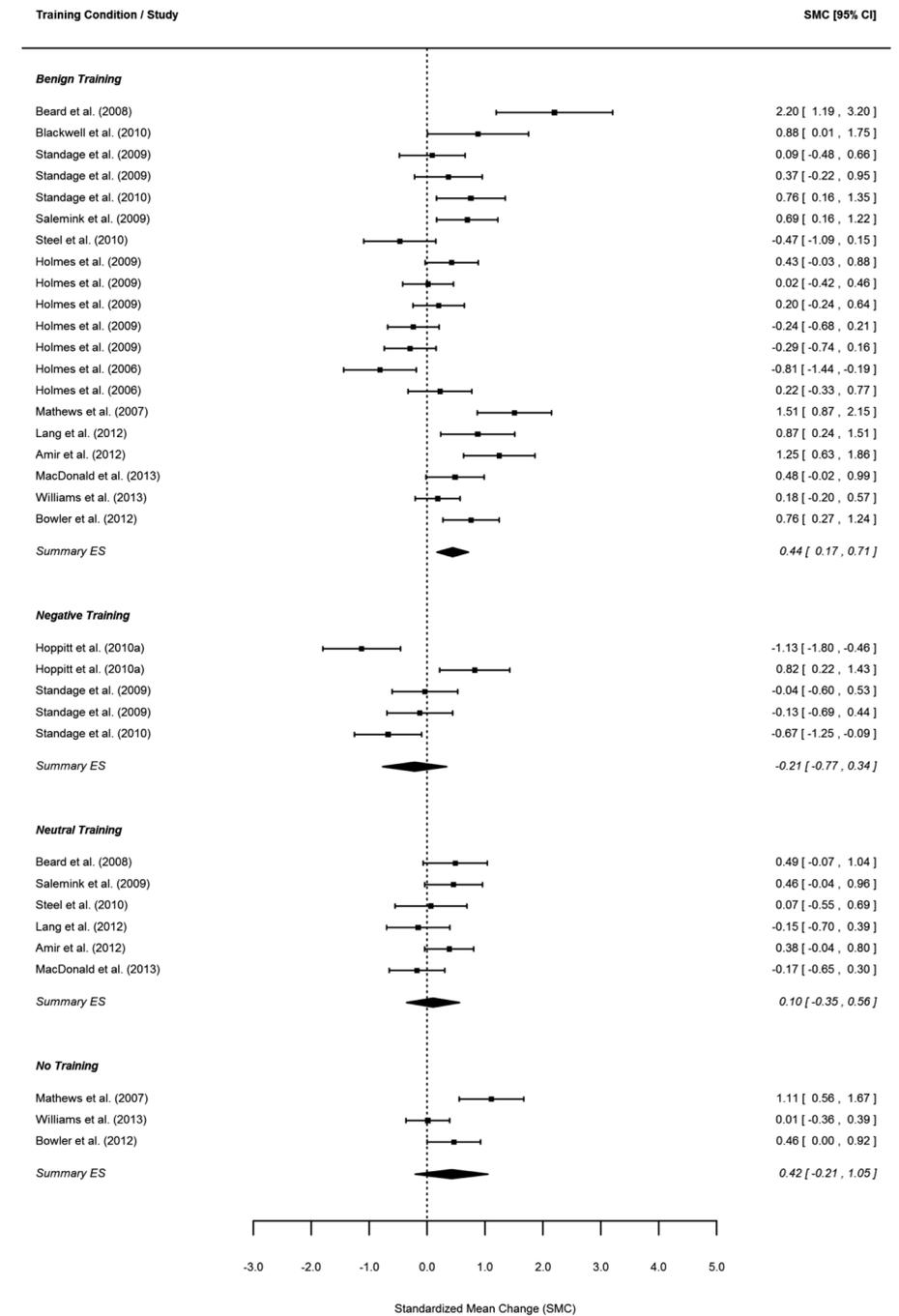


Figure 4. Forest plot of change in negative mood. Order of same conditions within one study follow the order of table 1.

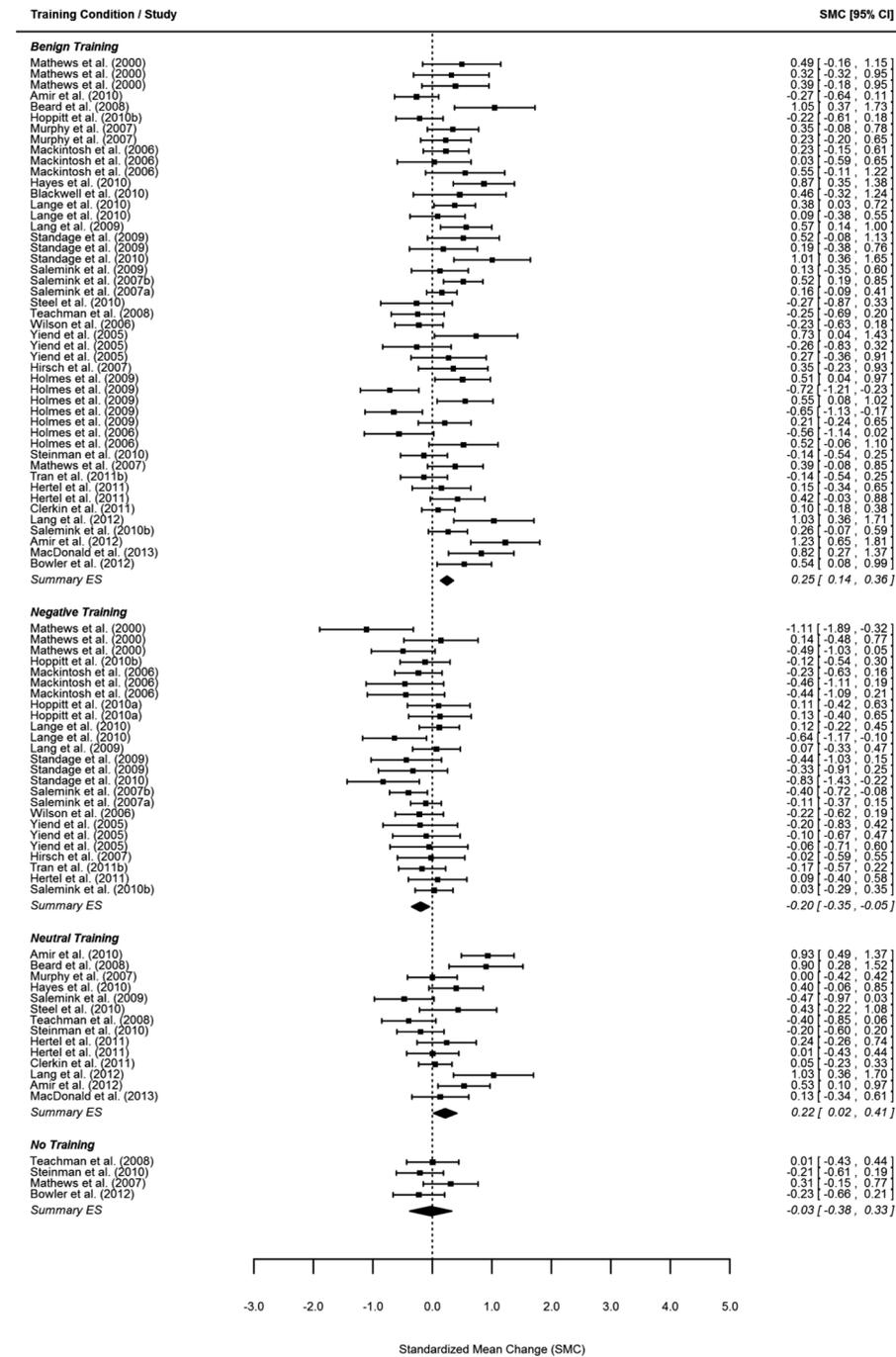
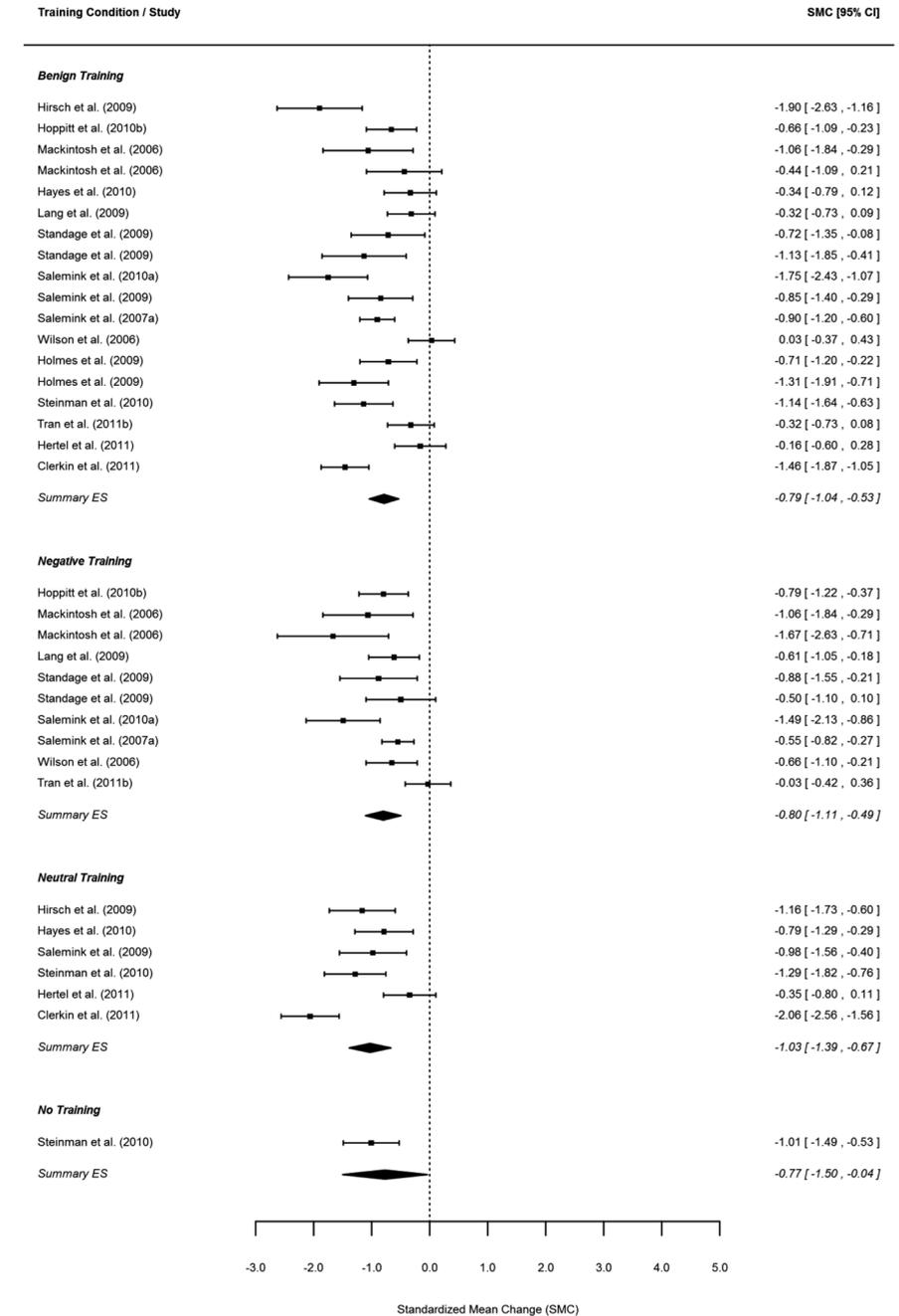


Figure 5. Forest plot of change in negative mood in response to an emotional challenge. Order of same conditions within one study follow the order of table 1. The summary effect size for the no-training condition is not identical to the effect size reported for the sole study in this condition as the model took multiple nesting (within study and within one article) into account and 'corrected' for such effects in the summary effect size.



Synthesis of results I: Does benign CBM-I training alter interpretational styles and mood states?

Table 2 reports the pooled within group effects separately for the four training conditions (benign, negative, neutral, no training) on post-training differences in interpretational style (endorsement of positive versus negative interpretations), on pre- to post-training changes in positive interpretation style, on pre- to post-training changes in negative mood state, and on pre- to post-emotional challenge changes in negative mood state. **Table 3** contains the pairwise comparisons of the within-group effects across the four training conditions.

Post-training endorsement of positive versus negative interpretations

As shown in **Table 2**, across 75 independent samples ($k=75$) (distributed over a total of 34 studies), benign and neutral training demonstrated a significant difference between endorsement of positive and negative interpretations after training. The effect size of this difference for the benign training was substantial (i.e., 1.33 with 95% CI 1.11 – 1.55). For the same difference in the neutral condition this effect was also significant but considerably smaller (i.e., 0.49, 95% CI 0.06 – 0.93). This implies that, on average, 91% of individuals who had received benign training and 69% of individuals who had received neutral training endorsed positive interpretations more strongly than negative interpretations for new ambiguous stimuli after training. Importantly, the size of the effect in the benign condition was significantly larger than in the neutral and any other condition using pairwise comparisons of effect sizes (all pairwise $p<.001$; see **table 3**).

As described earlier, the effect sizes were analyzed with a multilevel model that included a variance component for study level variability (allowing for shifts in the effects across studies irrespective of training condition) and a variance component at the effect size level (accounting for differences in the effects across studies for the various training conditions). Both the study level variance component and the effect size level variance component did reach statistical significance ($p=.002$ and $p<.001$, respectively – not depicted in **table 2**). Therefore, the findings reported above should be viewed as the *average* effects across the different studies.

Table 2. Main results – differences within conditions

Outcome	K (total)	K (condition)	Training condition	ES	95% CI	p
difference between positive and negative interpretation bias after training ¹	75	35	Benign	1.33	1.11; 1.55	<.001
			Negative	-0.05	-0.28; 0.17	.66
			Neutral	0.49	0.06; 0.93	.03
			No training	0.28	-0.16; 0.86	.72
change in positive interpretation bias from pre- to post training ²	34	20	Benign	0.43	0.17; 0.69	<.01
			Negative	-0.22	-0.75; 0.32	.43
			Neutral	0.12	-0.33; 0.57	.59
			No training	0.32	-0.30; 0.93	.31
change in negative mood from pre- to post training ³	90	47	Benign	0.25	0.14; 0.36	<.001
			Negative	-0.20	-0.35; -0.05	<.01
			Neutral	0.22	0.02; 0.41	.03
			No training	-0.03	-0.38; 0.33	.88
difference in negative mood from pre- to post emotional challenge ⁴	35	18	Benign	-0.79	-1.04; -0.53	<.001
			Negative	-0.80	-1.10; -0.49	<.001
			Neutral	-1.03	-1.39; -0.67	<.001
			No training	-0.77	-1.49; -0.05	.04

Note. ES=effect size, CI=confidence interval, k=nr of independent samples; ¹positive values reflect higher positive than negative bias; ²positive values reflect increase in positive bias; ³positive values reflect decrease in negative mood; ⁴negative values reflect increase in negative mood

Pre-post training change in positive interpretation bias

Next, we assessed pre- to post-training changes in the endorsement of positive interpretations for each training condition. Data were available from 34 independent samples ($k=34$) (from a total of 17 studies). Benign training was the only condition to result in a significant change in the selection of positive interpretations of ambiguous stimuli from before to after training ($p<.01$). This effect was of a much smaller size (i.e., 0.43 with 95% CI 0.17 – 0.69; **table 2**) than the one found for post training differences between positive and negative interpretation biases in the benign condition, and can be interpreted as indicating that, on average, 67% of individuals receiving benign training showed an *increase* in the selection of positive interpretations. This change differed significantly ($p<.05$; see **table 3** comparing ‘negative’ and ‘benign’ training

condition under ‘change in positive interpretation bias from pre- to post training’) from the change in the opposite direction for the negative training condition (i.e., -0.22 with 95% CI -0.75 – 0.32; **table 2**). While the effect size level variance component was significantly larger than zero ($p<.001$), we did not find significant study-level heterogeneity ($p=.58$) for this outcome (again, not depicted in table 2). However, since at least one component was significant, this indicates again that the pooled effects need to be viewed as average effect sizes across the different studies.

Pre-post training changes in mood state

Data from 90 samples (from 42 studies) were available to assess training effects on changes in negative mood state. Overall negative mood *decreased* significantly ($p<.001$) in the benign training condition (a small effect size of 0.25 with 95% CI 0.14 – 0.36; **table 2**) indicating that, on average, 60% of individuals showed reductions in negative mood in the benign condition. In contrast, negative mood *increased* significantly ($p<.01$) for those receiving negative training (a small effect size of -0.20 with 95% CI -0.35 – -0.05; **table 2**), indicating that, on average, 58% of individuals showed increases in negative mood after negative training (see **table 3** under ‘change in negative mood from pre- to post training’). In the neutral training condition negative mood decreased significantly ($p=.03$) from before to after training ($ES = 0.22$; 95% CI 0.02 - 0.41; **table 2**). The comparison of the within-group changes *between* the benign and the negative training condition was significant ($p<.001$; **table 3**). Moreover, while the change in the negative training condition varied significantly from the change in the neutral condition ($p<.01$; **table 3**), there was no significant difference in the amount of change in mood state between the benign and neutral or no-training conditions (both p 's $> .05$) (see **table 3** under ‘change in negative mood from pre- to post training’). Again, only heterogeneity at the effect size level ($p<.001$) and not the study level ($p=.86$) was found suggesting that the pooled effects above need to be viewed as average effect sizes (not depicted in table 2).

Table 3. Main results – differences between conditions.

Outcome	K (total)	K (condition)	Training condition			Omnibus Test	
			Training condition	Neutral	Negative		Benign
difference between positive and negative interpretation bias after training ¹	75	35	Benign			QM (df=3)=137.70, $p<.001$	
		28	Negative		1.39 *** (1.15; 1.62)		
		6	Neutral		-0.54* (-1.01; -0.07)		0.84 *** (0.40; 1.29)
		6	No training	0.21 (-0.35; 0.77)	-0.33 (-0.79; 0.31)		1.05 *** (0.60; 1.50)
change in positive interpretation bias from pre- to post training ²	34	20	Benign			QM (df=3)=5.53, $p=.14$	
		5	Negative		0.65* (0.07; 1.23)		
		6	Neutral		-0.34 (-1.04; 0.36)		0.31 (-0.19; 0.81)
		3	No training	-0.19 (-0.95; 0.57)	-0.53 (-1.35; 0.28)		0.12 (-0.53; 0.76)
Change in negative mood from pre- to post training ³	90	47	Benign			QM (df=3)=24.13, $p<.001$	
		25	Negative		0.44 *** (0.26; 0.63)		
		14	Neutral		-0.41 ** (-0.66; -0.17)		0.03 (-0.19; 0.25)
		4	No training	0.24 (-0.16; 0.65)	-0.17 (-0.56; 0.21)		0.27 (-0.10; 0.64)
difference in negative mood from pre- to post emotional challenge ⁴	35	18	Benign			QM (df=3)=2.11, $p=.50$	
		10	Negative		0.01 (-0.25; 0.27)		
		6	Neutral		0.23 (-0.17; 0.63)		0.24 (-0.09; 0.57)
		1	No training	-0.26 (-0.98; 0.46)	-0.03 (-0.77; 0.71)		-0.02 (-0.73; 0.69)

Note. ES=effect size, CI=confidence interval, * $p<.05$, ** $p<.01$, *** $p<.001$. Results are depicted in such a way that positive numbers indicate a positive change (that is (i) a higher positive than negative bias after training, (ii) an increase in positive bias from pre- to post-training, and (iii) a decrease in negative mood from pre- to post-training; ; ¹positive values reflect higher positive than negative bias; ²positive values reflect increase in positive bias; ³positive values reflect decrease in negative mood; ⁴negative values reflect increase in negative mood

Pre-post emotional challenge change in mood state

Mood state before and after an emotional challenge was assessed in 35 samples (15 studies). Negative mood increased significantly in all conditions from before to after the emotional challenge (with estimated effects ranging from -0.77 to -1.03, see **table 2** under 'differences in negative mood from pre- to post emotional challenge'). However, none of the pairwise differences comparing the amount of change between conditions was significant (see **table 3** under 'difference in negative mood from pre- to post emotional challenge'). In contrast to the earlier outcomes, we found significant heterogeneity at the study ($p < .01$), but not at the effect size level ($p = .18$). As before, the pooled effects reflect averages across studies.

Synthesis of results II: Do training- and sample characteristics enhance the effects of benign training?

Next, we assessed the influence of potential moderators on the effects of *benign* training on the various outcome measures that reached significance for the *within-group* differences (**table 2**) in our primary analysis. **Table 4** shows the influence of the moderator variables on the size of the effects for the different contrasts and their significance (for dichotomous moderators, the values reflect the difference between the two levels of the moderator, for continuous moderators, the values reflect the change in the size of the effect for a one-unit increase in the moderator) based on the meta-regression analyses. Due to the differential number of studies using AS, homograph, WSAT and other paradigms, we were only able to compare the AS task with all other paradigms combined. Additionally, we explored the association between the moderators that revealed significant effects within one outcome measure (except for paradigm as this variable was controlled for in the relevant analyses, also see methods). This was done to inform the interpretation of results as we were only able to assess each moderator separately (see **Table S3** for all correlations between moderators per outcome measure).

Post-training endorsement of positive versus negative interpretations

The paradigm employed significantly impacted on the post-training difference between positive and negative interpretation bias in that the AS paradigm ($k=26$) was significantly more effective (**Table 4**). All other paradigms combined ($k=9$) yielded a significant effect size of .69 ($p < .001$), indicating that there is a 75% chance (on average) for endorsing positive interpretations more strongly than negative interpretations. For AS-paradigms this effect increased by $ES = .90$ (see **table 4**), yielding an overall effect of $.69 + .90 = 1.59$, which corresponds to a 94% chance for endorsing positive interpretations more strongly than negative interpretations. In absolute terms, that is

a $94\% - 75\% = 19\%$ percentage points difference. Relative, that is $94/75 = 1.25$, that is a 25% higher chance of endorsing positive interpretations more strongly than negative interpretations. Similar findings emerged for the use of imagery instructions. No use of imagery instructions ($k=8$) resulted in an effect size of .54 ($p < .01$) indicating a chance of about 71% for endorsing more positive than negative interpretations. This effect increased by $ES = 1.05$ ($p < .001$), when imagery instructions were employed ($k=27$) (see **Table 4**), resulting in an overall ES of 1.59 indicating a chance of almost 94% of endorsing more positive than negative interpretations, that is a 33% higher chance than when no imagery instructions were employed.

However, it needs to be noted that paradigm and imagery-use were almost perfectly confounded; all AS-paradigms employed imagery instructions, whereas only one of the nine homograph paradigms did. It is therefore impossible to tell whether paradigm or the use of imagery is driving these results. Finally, while in all-female samples the difference between positive and negative interpretation biases was large and significant ($ES = 2.03$, $p < .001$, 98% chance of endorsing more positive than negative interpretations after training), this difference decreased by $ES = -.02$ ($p < .05$) per percentage point more males in the sample (see **table 4**). Thus, in a sample with for example 40% males this converges to a total effect of 1.28, corresponding to an absolute chance of 90%, which is a 9% decreased chance of endorsing more positive than negative interpretations after training.

For the endorsement of positive versus negative interpretations after training, the use of imagery instructions and male sex within the sample correlated by $r = -.60$ indicating that samples including fewer male participants were more often instructed to use imagery.

Table 4. Moderators

Outcome	Moderator									
	Training characteristics					Sample characteristics				
	training paradigm (AS vs other)	imagery instructions (y/n)	word generation (y/n)	mode of presentation (visual/auditory)	feedback administration (yes/no)	training/total items (ratio)	# sessions (frequency)	health status (symptoms/healthy)	mean age	% males (sex)
cognition										
difference between positive and negative interpretation bias after <i>benign</i> training	0.90 ***	1.05^a ***	-0.34	0.15	0.06	0.17	n.a ^b	0.30	-0.00	-0.02 *
change in positive interpretation bias from pre- to post <i>benign</i> training	-0.95 ***	0.50 **	n.a ^c	-0.50	0.46 *	-2.78 **	0.08 **	0.69 **	-0.02	-0.03 **
mood										
change in negative mood from pre- to post <i>benign</i> training	-0.29^d *	0.50 **	Interaction with paradigm ^e	-0.14	0.20	-0.34	0.07 *	0.15	-0.00	-0.01 **
emotional challenge										
change in negative mood from pre- to post emotional challenge	0.23	-0.07	0.16	-0.15	0.05	0.46	-0.01	-0.20	0.01	0.02 **

Note. Regression coefficients of the meta-regression models including the respective moderator. If paradigm resulted in a significant moderation of training effect, it was included in the tests of the other training-characteristic variables. * $p < .05$, ** $p < .01$, *** $p < .001$; ^a perfect confounding with paradigm; ^b there were 33 samples with one training session, one sample with four, and one sample with eight training sessions, it could thus not reasonably be tested; ^c there were only two samples asking individuals to generate word fragments and 15 that exposed participants to training stimuli, therefore no reasonable testing was possible; ^d $p = .05$; ^e there was a significant interaction with paradigm ($\beta = -.75$, $p = .02$) indicating that the effect of generation depends on the paradigms used, for all combined other paradigms (i.e. homograph, WSAT, or other; $k = 6$) the effectiveness is reduced when generation is employed as compared to when no generation is employed, whereas this is not true for the AS paradigms.

Pre-post training change in positive interpretation bias

For the change in positive interpretation bias from pre- to post-training, again paradigm, imagery instructions, and sex emerged as significant moderators.

However, this time the AS-paradigm ($k = 14$) yielded a non-significant increase in positive interpretations ($ES = .14$, $p = .26$, 56% chance of an increase in positive interpretation bias). This effect was then, however, significantly increased by $ES = .95$, $p < .001$ when 'other' paradigms ($k = 6$; 1 homograph; 3 WSAT, and 2 other paradigms) were employed. This results in an overall $ES = 1.08$ corresponding to a 54% higher chance of an increase in positive interpretation bias. In line with the findings for the difference between positive and negative interpretations after training, again, imagery instructions yielded a significant effect. Whereas no use of imagery instructions ($k = 6$) did not yield a significant effect ($ES = -.20$, $p = .21$, 42% chance of increase in positive interpretation bias), this effect was significantly enhanced by $ES = .50$ ($p < .01$) when imagery instructions were employed ($k = 12$). This results in a net-effect of imagery instructions of $ES = .30$ (62% absolute chance and 48% higher chance of increase in positive interpretation bias). Additionally, while no use of feedback ($k = 12$) yielded a non-significant small effect ($ES = .01$, $p = .96$, corresponding to a 54% chance for an increase in positive interpretations), the use of feedback ($k = 8$) increased this effect significantly by $ES = .46$ ($p < .05$). This results in a total effect of feedback administration of $ES = .46$ (68% absolute chance and 26% higher chance of increase in positive interpretation bias). Also, the ratio of benign training items versus all training items (ratio 0.69 ($k = 2$), 0.80 ($k = 1$), 1.00 ($k = 17$)) showed a significant effect. While the ratio of .69 yielded a large effect ($ES = 0.94$, corresponding to an 83% chance of an increase in interpretation bias) this effect was significantly reduced when the ratio became 1.00 ($p < .01$). This results in an absolute effect of $ES = .07$ (53% absolute chance and 57% lower chance of increase in positive interpretation bias). However, as only three samples employed a ratio < 1 , this result needs to be viewed with caution. Finally, adding one training session significantly increased the effect of training by $ES = .08$ ($k_{\text{one session}} = 12$; $k_{\text{four sessions}} = 2$; $k_{\text{six sessions}} = 1$; $k_{\text{seven sessions}} = 2$; $k_{\text{eight sessions}} = 2$; $k_{\text{twelve sessions}} = 1$, $p < .01$; total $ES = .01$; absolute chance of 51% and higher chance of 6%). Although this may seem a very small effect, the effect of 10 sessions of training would already result in an $ES = .80$, with a higher chance of approximately 49%. Exploring the potential clinical benefits of benign CBM-I, it was particularly interesting to see that while the effect for healthy samples was non-significant ($k = 10$, $ES = .07$, $p = .65$, c.a. 53% chance of increase), this effect increased by $ES = .69$ ($p < .01$) in samples with mood-symptoms ($k = 10$) resulting in a total effect of $ES = .76$, absolute chance of 78% and higher chance of 47% of increase in positive bias. Conform above results, while all-female samples showed a large and significant effect ($ES = 1.30$, $p < .001$, 90% chance of increase), this effect was significantly

reduced by $ES=.03$ ($p<.01$) per percentage point males in the sample. Thus, a sample with for example 40% males, the chance for an increase in positive interpretation bias would be lowered by 61%.

None of the moderators demonstrating a significant effect correlated strongly with another (all $r's \leq .50$), except for the number of training sessions and health status of individuals ($r=.73$) (also see **Table S3**). This indicates that particularly samples with symptoms of depression and/or anxiety were exposed to a repeated number of training sessions. Post-hoc analyses including both 'number of training sessions' and 'health status' revealed the following effects: $ES=.09$, $p<.05$ for 'number of sessions' and $ES= -.01$, $p=.99$, for 'health status' suggesting that the effect for 'no of sessions' depicts a rather robust effect. However, it is important to mention that more than one training session was only administered to individuals with symptoms whereas one training session was administered to both healthy and symptomatic samples (also see **table 1**) making it difficult to meaningfully interpret this result.

Changes in mood state

The change in negative mood from pre- to post-training in benign training conditions was again moderated by paradigm, imagery, and sex.

The AS paradigm resulted in a significant decrease in negative mood ($k=37$; $ES=.19$, $p<.01$, 58% chance of decrease in negative mood). This effect was significantly increased by $ES=.29$ ($p=.05$) when other paradigms were employed ($k=10$). The total effect when other paradigms were employed therefore was $ES=.48$, with an absolute chance of 68% and a higher chance of 17% of decrease in negative mood. The use of no imagery instructions resulted in a non-significant effect ($k=11$; $ES= -.025$, $p=.083$, 40% chance of decrease in negative mood). However, adding imagery instructions ($k=33$) increased this effect significantly by $ES=.50$ ($p<.01$), resulting in a total $ES=.25$, which corresponds to a 60% absolute and 50% higher chance of decrease in negative mood. Adding one session ($k_{\text{one session}} = 40$; $k_{\text{four sessions}} = 2$; $k_{\text{six sessions}} = 1$; $k_{\text{seven sessions}} = 2$; $k_{\text{eight sessions}} = 2$; $k_{\text{twelve sessions}} = 1$) resulted in a significant larger decrease in negative mood by increasing the effect by $ES= .07$, $p<.05$, resulting in a total $ES=.16$, which corresponds to a 56% absolute and 4% higher chance of decrease in negative mood when administering for example two instead of one session. Finally, while all-female samples showed a significant decrease in negative mood ($ES=.70$, $p<.001$, 76% chance), this effect was significantly decreased per one percentage male by $ES=-.01$, $p<.01$. Thus, for 40% males in the sample, the total ES would reduce to $ES=.30$, corresponding to a 23% lower chance. The correlations between the significant moderators here were small (all $r's < .36$) (also see **Table S3**).

To sum up, training paradigm, imagery instructions, the number of training sessions, and sex moderated the impact of benign training on at least two different outcome measures. Feedback administration, the ratio between benign training items and the total number of training items, and health status moderated the impact of benign training on the change in positive interpretation bias from pre- to post training only. Although most correlations between the significant moderators were far from perfect ($\leq .5$), the effects of repeated number of training sessions and mood symptoms for the increase in positive interpretation bias were strongly associated and therefore need to be viewed with caution.

Risk of bias across studies

The p -values of the tests for funnel plot asymmetry for the four outcomes were .72, .79, .46, and .89, respectively. Therefore, based on these tests and the visual examination of the funnel plots (Figure S4), there was no indication of publication bias for any of the main outcomes, as indicated by the absence of an association between the inverse of the sample sizes and the effect sizes (Peters et al., 2006) (see **Figure S4** for some additional analyses possibly hinting at some asymmetry in the pre-training versus post-training endorsements of positive interpretations outcome).

Additional analyses I: Do training effects on interpretation bias and on mood state correlate?

Based on $k=49$ pairs of effect size estimates, the correlation between the endorsement of positive versus negative bias after training and the decrease in negative mood was positive ($r=.60$; CI 0.39 - 0.76; $p<.001$). The correlation between the change in positive interpretation bias from before to after training and decrease in negative mood was also positive ($r=.58$; CI 0.30 - 0.78; $k=32$ pairs of estimates; $p<.001$). There was no significant correlation between the endorsement of positive versus negative bias after training or the change in positive bias from before to after training and the increase in negative mood in response to an emotional challenge ($r=-.0001$, CI -0.41 - 0.41, $k=23$ and $r=-.15$, CI -0.77 - 0.62, $k=8$, respectively).

Additional analyses II: Does randomization and percentage of attrition affect main results?

Excluding all studies with only one training condition (and therefore no random training allocation) and training conditions with an attrition rate above 15% did not change the main results.

Discussion

Cognitive bias modification training for interpretation bias (CBM-I) has recently been considered a promising clinical tool, e.g., see (Blackwell & Holmes, 2010) with potential for boosting positive thoughts (i.e., positive or benign interpretations of ambiguous situations) and thereby improving emotional symptoms. The current meta-analysis addressed some outstanding questions of relevance for the potential use CBM-I in a clinical setting.

Does benign CBM-I boost positive interpretations with improvements in negative mood?

In the current analyses, benign CBM-I resulted in large post-training endorsements of positive relative to negative interpretations, in small to medium changes in positive interpretational style from pre- to post training, and small decreases in negative mood states from pre- to post training. The correlation between these indices of improvements (i.e., change in interpretational style and mood state) was positive and significant. While these benign training effects differed reliably from negative training effects, the difference relative to neutral or no-training conditions remained insignificant for the change in positive bias and mood. Furthermore, benign training did *not* attenuate relative increases in negative mood in response to emotional challenges.

These findings partly support but also extend results from the earlier meta-analysis which combined assessment of CBM-I effects with another training program, CBM-A (Hallion & Ruscio, 2011). Across both meta-analyses (the current one including additional 24 articles) benign CBM-I training reliably 'boosted' positive interpretations post-training. Benign CBM-I resulted in a large post-training difference *between* positive and negative interpretation bias, which was also significantly larger than in any comparison condition. However, the neutral condition also presented with a significantly larger positive as compared to negative interpretation bias after training. As these effects are based on post-training differences only, it is impossible to tell whether the *change* in positive interpretation bias due to training was significant in the benign as well as in the neutral group. We therefore additionally systematically investigated the increase in positive interpretation bias from pre- to post-training. The benign condition showed significant but small changes, which differed from negative training only. This raises the question of whether benign training significantly enhances changes that are attributable to naturally occurring fluctuations in interpretational style (as those demonstrated by individuals receiving no training) and changes that may be explained by placebo effects (as those demonstrated by individuals receiving neutral training). However, it also needs to be taken into account that compared to the benign

training no-training and neutral training were administered less frequently (three and six samples as compared to 20 samples), hence power for these comparisons might have been reduced. Including these comparative training conditions in primary studies might address these questions more reliably in the future. Yet, it is still notable that only benign training showed a significant *change* in interpretational style, indicating its potential as an interventive tool.

Whereas benign CBM-I seems to be the only condition to significantly affect interpretational biases, mood was affected in all but the no-training condition. Benign CBM-I training resulted in significant, but small decreases in negative mood, which significantly differed from changes in the opposite direction in the negative condition, but not from changes in neutral and no-training conditions. The significant positive change in the neutral condition but not in the no-training condition may suggest that neutral training is not completely emotionally 'neutral' after all. These findings may help to explain the inconsistencies in the primary literature, e.g. see (Mathews, 2004), as it shows that the significance of the post training differences in mood probably only becomes obvious when benign and negative groups are compared.

Furthermore, contrary to predictions that the tendency to benignly interpret ambiguous situations should decrease the negative mood response particularly under conditions of emotional provocation (Wilson et al., 2006), we did not find this in the present data. To understand what this implies it is crucial to inspect the diverse emotional challenges employed to investigate negative mood reactivity after CBM-I. Those included overall stressful videos (Hertel et al., 2011; Hoppitt et al., 2010b; Lang et al., 2009; Mackintosh et al., 2006; Wilson et al., 2006), unsolvable anagram tasks (Salemink et al., 2007a, 2009), worry intrusions or negative mood inductions (Hayes et al., 2010; Hirsch et al., 2009; Holmes et al., 2009; Salemink & van den Hout, 2010a), symptom provoking tasks (Clerkin & Teachman, 2011; Steinman & Teachman, 2010), speech anticipation (Standage et al., 2009), and exposure to emotional faces combined with 'incorrect' feedback (Tran et al., 2011b). It can be suggested that most of those did not provoke ambiguity specifically, but were distressing more generally. For example, watching stressful videos (Hertel et al., 2011), trying to solve an unsolvable anagram task under time pressure 'knowing' that 'most people have no problems solving it in time', e.g., see (Salemink et al., 2009) or giving a speech, e.g. see (Standage et al., 2009) should be stressful for most individuals. The results of the current meta-analysis therefore only suggest that benign CBM-I does not alter overall mood-reactivity in response to these universally stressful events or situations. Very likely, emotional challenges currently employed were arguably not the most suitable for assessment of the more

subtle effects of interpretation biases on negative mood-reactivity as the reaction to these stressors are not under the direct influence of interpretation biases. Future studies should consider including emotional challenges that more directly activate the manipulated cognitive mechanism allowing clinically relevant conclusions about benign CBM-I's effects on daily-life stressors, such as having a group of colleagues laughing when you enter the room.

Nevertheless, the current data support the hypothesis that trained differences in interpretational style and changes in mood state are correlated, implying that benign interpretation biases are related with feeling less negative. The manipulation of interpretation style not only resulted in significant *changes* in interpretation style, but these changes were also significantly associated with decreases in negative mood states. Although these mood changes appeared to be of rather small effect size, assuming that individuals are confronted with ambiguous situations repeatedly in their daily lives, the cumulative effect of benignly interpreting these situations might result in clinically significant improvements in the long run.

How can we boost benign CBM-I's effectiveness?

To amplify these positive effects of benign CBM-I, it is of utmost clinical relevance to know *how* training effects can be maximized. The probably most obvious factor, training paradigm, showed somewhat unexpected findings. We distinguished between two paradigm groups: the ambiguous situations paradigm (AS) and the combination of all other employed tasks. While the chances for a larger positive than negative interpretation bias after training were higher for individuals receiving the ambiguous situations paradigm (AS), the chances for an increase in positive interpretation bias from pre-to post training were larger for individuals receiving other training paradigms. A possible explanation is that the assessment of the post-training endorsement of positive versus negative bias (but not of the pre-post training differences in positive interpretation bias) shows great similarity to the training method employed in AS-paradigms whereas studies investigating changes in positive interpretations typically use questionnaire measures that do not resemble the employed training program. In support of this suggestion, previous research has shown that, while the endorsement of positive versus negative interpretation style after training showed significant training effects with the AS-paradigm, other interpretation bias assessment tools did not (Salemink et al., 2007b). So, perhaps training effects are more visible when training and test are of increased similarity. Alternatively, the findings for the post-training endorsement of positive versus negative bias may have been solely driven by the use of imagery instructions which were (almost) only employed in studies with the

AS-paradigm. From the current results it is impossible to tell whether the AS-paradigm would still be superior to the homograph paradigm if no imagery instructions had been employed. Overall, the AS-paradigm was developed to increase stimuli realism (Mathews, 2004; Mathews & Mackintosh, 2000) and is a task with high face validity. In order to overcome this methodological issue, future studies should consider adding different outcome measures.

Despite this imprecision about which paradigm provides the best results, several procedural factors were found to significantly increase benign CBM-I's effectiveness *above and beyond the type of training paradigm*. Most convincingly, imagery instructions were able to significantly boost improvements in both positive interpretation bias and negative mood. Imagery has long been assumed to have a special link with emotions (Holmes et al., 2006). The use of imagery in CBM-I has been proposed to increase training effects and primary studies have supported this idea (Holmes et al., 2009; Holmes et al., 2006). This is the first meta-analysis indicating that imagery instructions can have moderate to large effects on increasing the effect of benign CBM-I on interpretation biases and accompanying mood. It needs to be considered however, that the use of imagery was associated with percentage females in the sample for post training endorsement of positive versus negative interpretations. This may indicate that the effect of imagery for this particular outcome is partly attributable to the effect of sex and vice versa. However, both sex and imagery yielded significant effects for each of the other outcome measures as well (where only very weak correlations were observable) supporting the robustness of the imaging effect. We therefore suggest that the general finding of the value of imagery instructions in benign CBM-I is valid.

Likewise, the observation that provision of repeated training sessions increased training effectiveness further confirms theory and findings of primary studies focusing on the clinical use of CBM-I. Although the effect sizes were small (.08 for increase in positive interpretation bias and .07 for decrease in negative mood), it is important to realize that this is the increase in effect size for just adding *one* additional training session. Therefore, this effect increases to a certain degree with additional training sessions. However, it is crucial to understand here that repeated training sessions were most often employed in samples with mood-symptoms (with the latter also showing significant influence on the increase in positive interpretation bias). This makes it difficult to determine which of the two moderators drove this effect although post-hoc analyses including both factors simultaneously supported the robustness of the influence of more training sessions. In the future, using internet based CBM applications in combination with momentary assessment techniques (Csikszentmihalyi

& Larson, 1987; Wichers et al., 2011) may enhance the feasibility of having people ‘train’ their positive interpretation biases more frequently. For example, people may download a simple training program and carry out exercises on a weekly or daily basis, thereby integrating positive interpretation tendencies to their daily life situations. One pioneering study (Bowler et al., 2012) already demonstrated the feasibility and effectiveness of a multi-session internet-based CBM-I. While not influencing any of the other outcome measures, administering feedback about response accuracy and having a moderate ratio of benign training items and total items within the training was further related to a larger increase in positive interpretation bias with moderate (feedback) to large (ratio) effect sizes. Speculatively, both these factors foster engagement with the training task. That is, feedback is only administered if participants have to respond in some way or another to the stimuli presented (not necessarily by completing word-fragments as was assessed with the ‘generate’ variable that showed no effect) and may encourage them to concentrate and elaborate on the information presented. Having not 100% of the items benign may have a similar effect. It has been shown in studies of instrumental conditioning that responses learned under partial reinforcement are much harder to extinguish than responses learned under continuous reinforcement (Gleitman, Fridlund, & Reisberg, 2004, p. 161). The same may be true for the current task where participants had to learn on most but not all occasions to interpret situations as benign – which is presumably also somewhat more reflective of real-life interactions (Mathews et al., 2007).

Do people who are most vulnerable actually respond to CBM-I?

We also sought to examine whether benign CBM-I was universally effective, or if it was of particular benefit to those who were most vulnerable. Our findings indicated that particularly women (who are more vulnerable to develop mood symptoms) tend to benefit more from benign CBM-I, both cognitively (interpretation bias) and emotionally (mood) with significant and large effect sizes. This is in agreement with theory, as it has long been recognized that women are more emotionally reactive than men, e.g., see (Husky et al., 2009).

The increase in positive interpretation bias from pre- to post-training was particularly prominent in samples with anxiety and mood symptoms/diagnoses as compared to healthy (non-symptomatic) controls (showing a non-significant change). However, as already mentioned above, this effect of mood-symptoms cannot completely be torn apart from the effect of repeated training sessions. It is possible that both effects add a significant part but are smaller than suggested here. Although post-hoc analyses did not support the influence of emotional symptoms, it is important to notice that all

studies administering more than one training session were conducted in symptomatic samples, possibly suggesting that repeated sessions are of added value within symptomatic samples. To resolve this issue, research administering repeated sessions to both healthy and symptomatic samples are needed. Exposure to one or more sessions of benign CBM-I may result in only moderate changes in healthy populations as they already possess a substantial positive interpretation bias to begin with, resulting in a ceiling effect. However, for people with a potentially lowered positive interpretation bias, benign CBM-I can cause a significant and large increase in positive interpretation bias. Therefore, benign CBM-I may add therapeutic benefit to general Cognitive Behavioral Therapy (CBT) paradigms by targeting dysfunctional cognitions more directly and specifically without much additional effort from both the therapist and the patient. Very recently, two studies intriguingly demonstrated that benign CBM-I was indeed associated with clinical relevant improvements (Bowler et al., 2012; Williams et al., 2013).

Limitations and future directions

Several issues need consideration in interpreting the current findings. First, a substantial threat to conclusive meta-analyses is the problem of unpublished data. Often, studies that yield significant effect-sizes are more likely to be published than studies not yielding significant effect sizes, a phenomenon referred to as publication bias (Peters et al., 2006). To draw firm conclusions about the effectiveness of CBM-I, it is necessary to include as many studies as possible (preferably all). To this end, we contacted all authors in the field to request unpublished data, and also verified that any poster-abstract findings were included in the analysis. Moreover, as there was no indication for publication bias (Morris & DeShon, 2002), we cautiously suggest that even had there been other unpublished data, these would not have changed the results.

Second, in the current meta-analysis we specifically focused on mood state in order to be as precise as possible in what CBM-I actually does, based on what has actually been assessed. Low mood is a common symptom of mood- and anxiety disorders (APA, 2000). However, it does not encompass the whole spectrum of possible mood symptoms. Based on the current findings, it may be concluded that the symptom of ‘low mood’ can be influenced by benign CBM-I. However, this effect is small and whether other symptoms like feelings of guilt or rumination can also be influenced cannot be answered. It needs to be considered though that CBM-I effects on general psychopathology may not immediately become apparent after a single session of CBM-I training for two reasons. First, interpretation biases only start to play a role in

reducing mood when people are confronted with ambiguous situations in their own life. Second, interpretation biases represent only one potential bias in information processing (Mathews, 2004), which, in turn, represents only one possible cause of emotional disturbance (Hallion & Ruscio, 2011). Studies are currently emerging showing that training more than one benign cognitive bias (e.g., interpretation and attention bias) (Beard, Weisberg, & Amir, 2011; Brosan, Hoppitt, Shelfer, Sillence, & Mackintosh, 2011) and repeated training sessions of CBM-I (Bowler et al., 2012; Williams et al., 2013) result in clinically relevant improvements.

Finally, it needs to be acknowledged that the current moderators were all investigated separately, preventing conclusions about the sum of their effects when included simultaneously. We did control all analyses for the paradigm employed as this was a rather complex construct encompassing stimulus presentation and processing. Therefore, the current results depict the effects of the moderators above and beyond the effect of paradigm but not above and beyond the presence of any other moderators. However, testing the effects of all possible moderators to investigate their combined additive and interactive effects would have drastically reduced the power to draw any meaningful conclusions. Most of the correlations between the moderators were small to moderate reducing the likelihood that the results were confounded by collinearity. However, - and as already elaborated above - , the effect of repeated training sessions and mood symptoms on the increase in positive interpretation bias in benign training conditions remains indistinguishable so far. Although it needs to be noted that both are intuitively valid and have been supported by primary research, the currently found effects may depict overestimations. In the future, the advent of internet access and momentary assessment technology, like currently developed apps for smartphones, can be used to implement benign CBM-I more efficiently. Participants can download programs to their private computers and train more frequently than possible in the lab. Having a training program on their smartphones would even enable participants to train benign interpretations in their daily lives. Because imagery seems to play a crucial role in increasing benign CBM-I effectiveness, it might be particularly interesting to examine whether novel technical tools such as virtual reality can enhance the computerized training even more. Furthermore, as adolescents are particularly vulnerable to emotional disturbance (Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993) and initial primary studies have demonstrated effectiveness of CBM-I in these samples (Lothmann, Holmes, Chan, & Lau, 2011; Muris, Huijding, Mayer, & Hameetman, 2008; Salemink & Wiers, 2012), future research needs to demonstrate whether the current results for adult samples translate to adolescent samples.

Conclusions

Although results were less strong as previously suggested, it may be concluded that benign CBM-I has clinical potential. It was associated with an increase in positive interpretation bias and a decrease in negative mood. These effects (i.e., increase in positive interpretation bias and decrease in negative mood) were significantly correlated indicating that an improvement in cognitive style goes hand in hand with immediate improvements in mood. Although effects did not consistently differ from control training conditions and the effects of repeated number of training sessions was not well distinguishable from the effects of mood-symptoms for the change in positive interpretation bias, employing imagery instructions in particular and to a lesser extent repeated training sessions, a moderate ratio of benign training items and total items, and administering feedback during training can significantly increase these effects. Most encouragingly, there are indications that benign CBM-I paradigms appear particularly effective in vulnerable samples.

Acknowledgements

The authors thank Katie Lockwood and Machteld Telman for data-collection at the early stages of this research. The authors also thank Emily Holmes and Michael Browning for helpful discussions at the beginning of this project and support in the process. Last but not least we are very grateful to all authors of the numerous primary studies that provided us with the within-group data necessary to conduct the current meta-analysis. We very much appreciate their effort and time that went into supporting our research

References

- Amir, N., Bomyea, J., & Beard, C. (2010). The effect of single-session interpretation modification on attention bias in socially anxious individuals. *Journal of Anxiety Disorders, 24*(2), 178-182.
- Amir, N., & Taylor, C. (2012). Interpretation training in individuals with generalized social anxiety disorder: a randomized controlled trial. *Journal of Consulting and Clinical Psychology, 80*(3), 497-511.
- APA. (2000). *Quick reference to the diagnostic criteria from DSM-IV-TR*. Arlington, VA: American Psychiatric Association.
- Beard, C., & Amir, N. (2008). A multi-session interpretation modification program: Changes in interpretation and social anxiety symptoms. *Behaviour Research and Therapy, 46*(10), 1135-1141.
- Beard, C., Weisberg, R., & Amir, N. (2011). Combined cognitive bias modification treatment for social anxiety disorder: a pilot trial. *Depression & Anxiety, 0*, 1-8.
- Beck, A. T., Steer, R. A., & Garbin, M. G. (2011). Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77-100.
- Blackhart, G., Nelson, B. C., Knowles, M. L., & Baumeister, R. F. (2009). Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: a meta-analytic review of 192 studies on social exclusion. *Personality and Social Psychology Review, 13*(4), 269-309.
- Blackwell, S. E., & Holmes, E. A. (2010). Modifying interpretation and imagination in clinical depression: a single case series using cognitive bias modification. *Applied Cognitive Psychology, 24*(3), 338-350. doi: 10.1002/acp.1680
- Bowler, J. O., Mackintosh, B., Dunn, B. D., Mathews, A., Dagleish, T., & Hoppitt, L. (2012). A comparison of cognitive bias modification for interpretation and computerized cognitive behavior therapy: effects on anxiety, depression, attentional control, and interpretive bias. *Journal of Consulting and Clinical Psychology, 80*(6), 1021-1033.
- Brosan, L., Hoppitt, L., Shelfer, A., Sillence, A., & Mackintosh, B. (2011). Cognitive bias modification for attention and interpretation reduces trait and state anxiety in anxious patients referred to an out-patient service: results from a pilot study. *Journal of Behavior Therapy and Experimental Psychiatry, 42*, 258-264.
- Clerkin, E. M., & Teachman, B. A. (2011). Training interpretation bias among individuals with symptoms of obsessive compulsive disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 42*, 337-343.
- Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4 ed.). Thousand Oaks, CA: Sage.
- Csikszentmihalyi, M., & Hunter, J. (2003). Happiness in everyday life: the use of experience sampling. *Journal of Happiness Studies, 4*(185-199).
- Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *J Nerv Ment Dis, 175*, 526-536.
- Eysenck, M. W., Mogg, K., May, J., Richards, A., & Mathews, A. (1991). Bias in interpretation of ambiguous sentences related to threat in anxiety. *Journal of Abnormal Psychology, 100*(2), 144-150.
- Furmark, T., Tillfors, M., Merteinsdottir, I., Fischer, H., Pissiota, A., Langstrom, B., & Fredrikson, M. (2002). Common changes in cerebral blood flow in patients with social phobia treated with citalopram or cognitive-behavior therapy. *Arch Gen Psychiatry, 59*, 425-433.
- Gibbons, R. D., Hedeker, D. R., & Davis, J. M. (1993). Estimation of effect size from a series of experiments involving paired comparisons. *Journal of Educational Statistics, 18*(3), 271-279.
- Gleitman, H., Fridlund, A. J., & Reisberg, D. (2004). *Psychology*. New York: W.W. Norton & Company.

- Gopinath, S., Katon, W. J., Russo, J. E., & Ludman, E. J. (2007). Clinical factors associated with relapse in primary care patients with chronic or recurrent depression. *Journal of Affective Disorders, 101*, 57-63.
- Grey, S., & Mathews, A. (2000). Effects of training on interpretation of emotional ambiguity. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 53A*(4), 1143-1162.
- Grey, S. J., & Mathews, A. M. (2009). Cognitive bias modification: priming with an ambiguous homograph is necessary to detect an interpretation training effect. *Journal of Behavior Therapy and Experimental Psychiatry, 40*(2), 338-343.
- Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin, 137*(6), 940-958.
- Hamilton, M. (1960). A rating scale for depression. *J. Neurol. Neurosurg. Psychiat., 23*, 56-62.
- Hayes, S., Hirsch, C. R., Krebs, G., & Mathews, A. (2010). The effects of modifying interpretation bias on worry in generalized anxiety disorder. *Behaviour Research and Therapy, 48*(3), 171-178.
- Hertel, P. T., Mathews, A., Peterson, S., & Kintner, K. (2003). Transfer of training emotionally biased interpretations. *Applied Cognitive Psychology, 17*(7), 775-784.
- Hertel, P. T., Vasquez, E., Benbow, A., & Hughes, M. (2011). Recollection is impaired by the modification of interpretation bias. *Journal of Abnormal Psychology, Advance online publication. doi:10.1037/a0023974*.
- Hirsch, C. R., Hayes, S., & Mathews, A. (2009). Looking on the bright side: Accessing benign meanings reduces worry. *Journal of Abnormal Psychology, 118*(1), 44-54.
- Hirsch, C. R., Mathews, A., & Clark, D. M. (2007). Inducing an interpretation bias changes self-imagery: A preliminary investigation. *Behaviour Research and Therapy, 45*(9), 2173-2181.
- Holmes, E. A., Lang, T. J., & Shah, D. M. (2009). Developing interpretation bias modification as a 'cognitive vaccine' for depressed mood: Imagining positive events makes you feel better than thinking about them verbally. *Journal of Abnormal Psychology, 118*(1), 76-88.
- Holmes, E. A., Mathews, A., Dalgleish, T., & Mackintosh, B. (2006). Positive Interpretation Training: Effects of Mental Imagery Versus Verbal Training on Positive Mood. *Behavior Therapy, 37*(3), 237-247.
- Hoppitt, L., Mathews, A., Yiend, J., & Mackintosh, B. (2010a). Cognitive bias modification: the critical role of active training in modifying emotional responses. *Behavior Therapy, 41*(1), 73-81.
- Hoppitt, L., Mathews, A., Yiend, J., & Mackintosh, B. (2010b). Cognitive mechanisms underlying the emotional effects of bias Modification. *Applied Cognitive Psychology, 24*(3), 312-325. doi: 10.1002/acp.1678
- Husky, M. M., Mazure, C. M., Maciejewski, P. K., & Swendsen, J. D. (2009). Past depression and gender interact to influence emotional reactivity to daily life stress. *Cognitive Therapy and Research, 33*, 264-271.
- Koster, E. H. W., Fox, E., & MacLeod, C. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *Journal of Abnormal Psychology, 118*(1), 1-4.
- Lang, T. J., Blackwell, S. E., Harmer, C., Davison, P., & Holmes, E. A. (2012). Cognitive bias modification using mental imagery for depression: developing a novel computerized intervention to change negative thinking styles. *European Journal of Personality, 26*(2), 145-157.
- Lang, T. J., Moulds, M. L., & Holmes, E. A. (2009). Reducing depressive intrusions via a computerized cognitive bias modification of appraisals task: Developing a cognitive vaccine. *Behaviour Research and Therapy, 47*(2), 139-145.
- Lange, W. G., Salemink, E., Windey, I., Keijsers, G. P. J., Krans, J., Becker, E. S., & Rinck, M. (2010). Does modified interpretation bias influence automatic avoidance behaviour? *Applied Cognitive Psychology, 24*(3), 326-337. doi: 10.1002/acp.1679
- Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & Andrews, J. A. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology, 102*(133-144).
- Lothmann, C., Holmes, E. A., Chan, S., & Lau, J. Y. (2011). Cognitive bias modification in adolescents: effects on interpretation biases and mood. *The Journal of Child Psychology and Psychiatry, 52*(1), 24-32.
- MacDonald, E. M., Koerner, N., & Antony, M. M. (2013). Modification of interpretive bias: impact on anxiety sensitivity, information processing and response to induced bodily sensations. *Cognitive Therapy and Research, 37*, 860-871.
- Mackintosh, B., Mathews, A., Yiend, J., Ridgeway, V., & Cook, E. (2006). Induced biases in emotional interpretation influence stress vulnerability and endure despite changes in context. *Behavior Therapy, 37*(3), 209-222.
- Mathews, A. (2004). On the malleability of emotional encoding. *Behaviour Research and Therapy, 42*(9), 1019-1036.
- Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology, 109*(4), 602-615.
- Mathews, A., Ridgeway, V., Cook, E., & Yiend, J. (2007). Inducing a benign interpretational bias reduces trait anxiety. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(2), 225-236.
- Mogg, K., & Bradley, B. P. (2005). Attentional bias in generalized anxiety disorder versus depressive disorder. *Cognitive Therapy and Research, 29*(1), 29-45.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The-PRISMA-Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med, 6*(6), e1000097.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-group designs. *Psychological Methods, 7*(1), 105-125.
- Muris, P., Huijding, J., Mayer, B., & Hameetman, M. (2008). A space odyssey: experimental manipulation of threat perception and anxiety-related interpretation bias in children. *Child Psychiatry and Human Development, 39*, 469-480.
- Murphy, R., Hirsch, C. R., Mathews, A., Smith, K., & Clark, D. M. (2007). Facilitating a benign interpretation bias in a high socially anxious population. *Behaviour Research and Therapy, 45*(7), 1517-1529.
- Neshat-Doost, H. T., Taghavi, M. R., Moradi, A. R., Yule, W., & Dagleish, T. (1998). Memory of emotional trait adjectives in clinically depressed youth. *Journal of Abnormal Psychology, 107*(279-286).
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2006). Comparison of two methods to detect publication bias in meta-analysis. *Journal of the American Medical Association, 295*(6), 676-680.
- Pinheiro, J. C., & Bates, D. (2000). *Mixed-effects models in S and S-Plus*. New York: Springer.
- Salanti, G., Higgins, J. P. T., Ades, A. E., & Ioannidis, J. P. A. (2008). Evaluation of networks of randomized trials. *Statistical methods in medical research, 17*(3), 279-301.
- Salemink, E., & van den Hout, M. (2010a). Trained interpretive bias survives mood change. *Journal of Behavior Therapy and Experimental Psychiatry, 41*(3), 310-315.

- Salemink, E., & van den Hout, M. (2010b). Validation of the “recognition task” used in the training of interpretation biases. *Journal of Behavior Therapy and Experimental Psychiatry*, *41*(2), 140-144.
- Salemink, E., van den Hout, M., & Kindt, M. (2007a). Trained interpretive bias and anxiety. *Behaviour Research and Therapy*, *45*(2), 329-340.
- Salemink, E., van den Hout, M., & Kindt, M. (2007b). Trained interpretive bias: Validity and effects on anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(2), 212-224.
- Salemink, E., van den Hout, M., & Kindt, M. (2009). Effects of positive interpretive bias modification in highly anxious individuals. *Journal of Anxiety Disorders*, *23*(5), 676-683.
- Salemink, E., van den Hout, M., & Kindt, M. (2010c). Generalisation of modified interpretive bias across tasks and domains. *Cognition and Emotion*, *24*(3), 453-464.
- Salemink, E., & Wiers, R. W. (2012). Adolescent threat-related interpretive bias and its modification: the moderating role of regulatory control. *Behavior Research and Therapy*, *50*, 40-46.
- Standage, H., Ashwin, C., & Fox, E. (2009). Comparing visual and auditory presentation for the modification of interpretation bias. *Journal of Behavior Therapy and Experimental Psychiatry*, *40*(4), 558-570.
- Standage, H., Ashwin, C., & Fox, E. (2010). Is manipulation of mood a critical component of cognitive bias modification procedures? *Behaviour Research and Therapy*, *48*(1), 4-10.
- Steel, C., Wykes, T., Ruddle, A., Smith, G., Shah, D. M., & Holmes, E. A. (2010). Can we harness computerised cognitive bias modification to treat anxiety in schizophrenia? A first step highlighting the role of mental imagery. *Psychiatry Research*, *178*(3), 451-455.
- Steinman, S. A., & Teachman, B. A. (2010). Modifying interpretations among individuals high in anxiety sensitivity. *Journal of Anxiety Disorders*, *24*(1), 71-78.
- Teachman, B. A., & Addison, L. M. (2008). Training non-threatening interpretations in spider fear. *Cognitive Therapy and Research*, *32*(3), 448-459.
- Tran, T. B., Hertel, P. T., & Joormann, J. (2011a). Cognitive bias modification: induced interpretive biases affect memory. *Emotion*, *11*(1), 145-152.
- Tran, T. B., Siemer, M., & Joormann, J. (2011b). Implicit interpretation biases affect emotional vulnerability: a training study. *Cognition and Emotion*, *25*(3), 546-558.
- Turner, R., Hoppitt, L., Hodgekins, J., Wilkinson, J., Mackintosh, B., & Fowler, D. (2011). Cognitive bias modification in the treatment of social anxiety in early psychosis: a single case study. *Behavioural and Cognitive Psychotherapy*, *39*, 341-347.
- Viechtbauer, W. (2010). Conducting meta-analysis in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1-48.
- Wichers, M., Simons, C., Kramer, I. M. A., Hartmann, J. A., Lothmann, C., Myin-Germeys, I., . . . van Os, J. (2011). Momentary assessment technology as a tool to help patients with depression help themselves. *Acta Psychiatrica Scandinavica*, *124*(4), 262-272.
- Williams, A. D., Blackwell, S., Mackenzie, A., Holmes, E. A., & Andrews, G. (2013). Combining imagination and reason in treatment of depression: a randomized controlled trial of internet-based cognitive-bias modification and internet-CBT for depression. *Journal of Consulting and Clinical Psychology*. doi: 10.1037/a0033247
- Wilson, E. J., MacLeod, C., Mathews, A., & Rutherford, E. M. (2006). The causal role of interpretive bias in anxiety reactivity. *Journal of Abnormal Psychology*, *115*(1), 103-111.
- Yiend, J., Mackintosh, B., & Mathews, A. (2005). Enduring consequences of experimentally induced biases in interpretation. *Behaviour Research and Therapy*, *43*(6), 779-797.

Supporting Material

S1 Table PRISMA Checklist

S2 Measurement order

S3 Correlations between moderators per outcome measure

S4 Funnel plots

Table S1

Section/topic	#	Checklist item	Reported in section
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Title
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	Abstract
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	Introduction
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	Introduction
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Methods
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	Methods
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	Methods
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Methods
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Methods
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	Methods

Section/topic	#	Checklist item	Reported in section
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Methods
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	Methods
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	Methods
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	Methods
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	Methods
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	Methods
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	Results (Figure 1)
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Results (Table 1)
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Results
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Results (Figures 2-5)
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Results (Table 2,3)
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Results (Figure S4)
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Results

Section/topic	#	Checklist item	Reported in section
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	Discussion
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	Discussion
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	Discussion
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	n.a. (Online submission system)

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed100009

S2 Order of single measurements per outcome category

Order was determined based on measured concept, face validity of measurement tool employed, and frequency of usage in collection of primary studies

COGNITION MEASUREMENTS ORDER

ORDER	CODE	NAME
1	C01	Recognition ratings
2	C02	RT to word fragments
3	C03	Word sentence association task
4	C04	Lexical decision task
5	C14	Emotionality ratings, ambiguous description
6	C06	Scrambled sentence test
7	C05	Form and describe image
8	C07	Interpretation bias questionnaire
9	C13	Video ratings
10	C12	Reason for events
11	C18	BBSIQ (Bodily sensation symptoms)
12	C17	VAS depressive bias
13	C19	RIP (response to intrusions questionnaire)

MOOD MEASUREMENT ORDER

ORDER	CODE	NAME
1	M01	STAI-state
2	M02	STAI-trait
3	M06	NAS (negative affect scale, PANAS)
4	M08	VAS anxiety
5	M09	VAS depression
6	M10	FNES (fear of negative evaluations scale)
7	M21	PANAS FS (PANAS fear subscale)
8	M13	Libowitz Social Anxiety Scale
9	M22	Anxiety Sensitivity Index
10	M11	SPAI (social phobia and anxiety inventory)
11	M20	VAS sad
12	M27	VAS negative

Table S3 Correlations between moderators per outcome measure

1) difference between positive and negative interpretation bias after training

	IMAGE	GENERATE	PRESENT	FEEDBACK	RATIO	FREQ	STATUS	AGE	SEX
IMAGE	1.00	0.12	0.20	0.31	0.56	0.12	0.32	-0.19	-0.60
GENERATE	0.12	1.00	-0.46	-0.15	0.02	0.05	0.08	-0.22	0.03
PRESENT	0.20	-0.46	1.00	0.06	0.39	-0.08	0.20	-0.20	-0.19
FEEDBACK	0.31	-0.15	0.06	1.00	-0.22	0.04	0.10	-0.45	-0.18
RATIO	0.56	0.02	0.39	-0.22	1.00	-0.03	0.13	0.29	-0.07
FREQ	0.12	0.05	-0.08	0.04	-0.03	1.00	0.39	0.06	-0.20
STATUS	0.32	0.08	0.20	0.10	0.13	0.39	1.00	-0.30	-0.35
AGE	-0.19	-0.22	-0.20	-0.45	0.29	0.06	-0.30	1.00	0.37
SEX	-0.60	0.03	-0.19	-0.18	-0.07	-0.20	-0.35	0.37	1.00

2) change in positive interpretation bias from pre- to post training

	IMAGE	GENERATE	PRESENT	FEEDBACK	RATIO	FREQ	STATUS	AGE	SEX
IMAGE	1.00	0.27	-0.03	0.25	0.04	0.23	0.40	0.29	0.12
GENERATE	0.27	1.00	-0.49	0.49	-0.36	0.42	0.44	-0.32	-0.16
PRESENT	-0.03	-0.49	1.00	-0.78	0.50	-0.41	-0.47	0.41	0.54
FEEDBACK	0.25	0.49	-0.78	1.00	-0.50	0.36	0.41	-0.34	-0.47
RATIO	0.04	-0.36	0.50	-0.50	1.00	-0.45	-0.41	0.22	0.42
FREQ	0.23	0.42	-0.41	0.36	-0.45	1.00	0.73	0.04	-0.39
STATUS	0.40	0.44	-0.47	0.41	-0.41	0.73	1.00	0.28	-0.12
AGE	0.29	-0.32	0.41	-0.34	0.22	0.04	0.28	1.00	0.54
SEX	0.12	-0.16	0.54	-0.47	0.42	-0.39	-0.12	0.54	1.00

3) change in negative mood from pre- to post training

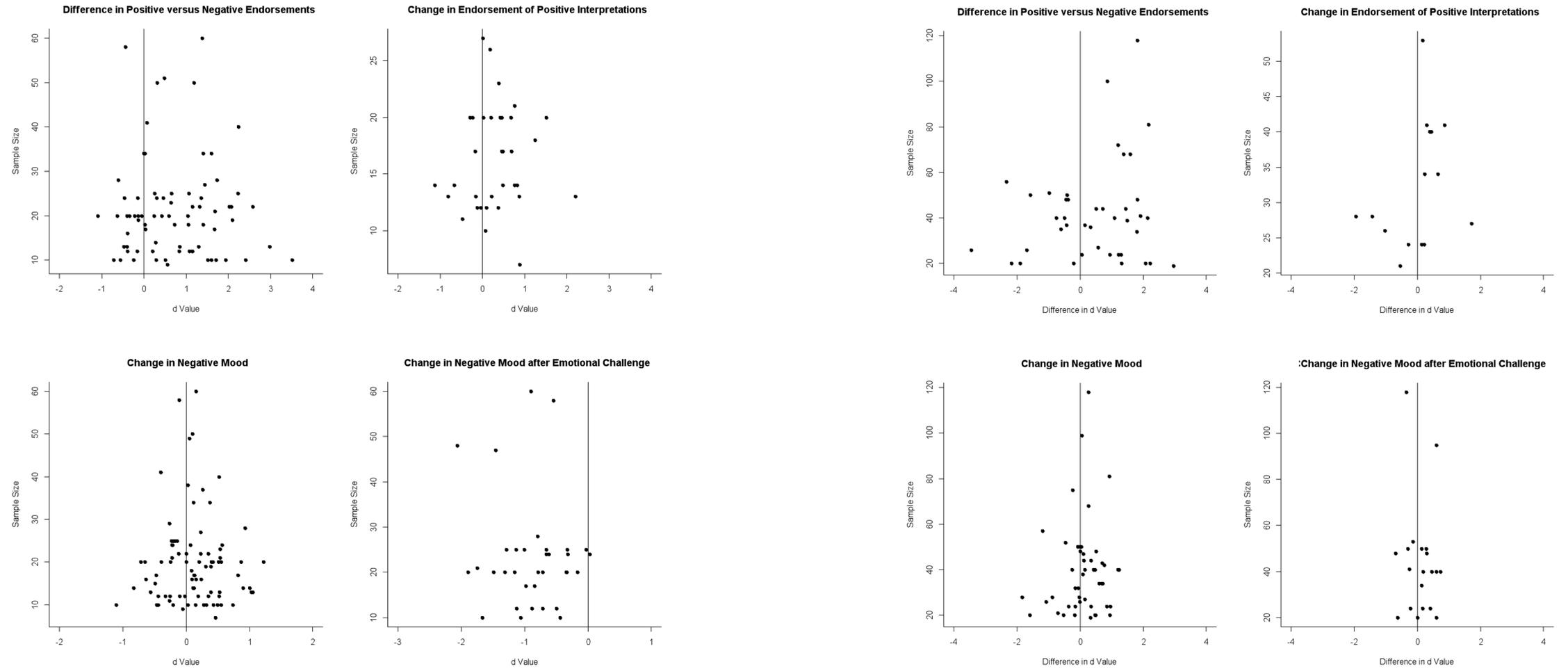
	IMAGE	GENERATE	PRESENT	FEEDBACK	RATIO	FREQ	STATUS	AGE	SEX
IMAGE	1.00	0.30	-0.11	0.33	-0.30	0.02	0.13	-0.04	0.05
GENERATE	0.30	1.00	-0.64	0.43	-0.55	-0.06	0.09	-0.27	0.13
PRESENT	-0.11	-0.64	1.00	-0.65	0.55	-0.10	-0.09	0.24	0.15
FEEDBACK	0.33	0.43	-0.65	1.00	-0.53	0.02	0.18	-0.35	-0.35
RATIO	-0.30	-0.55	0.55	-0.53	1.00	0.02	-0.07	0.29	0.21
FREQ	0.02	-0.06	-0.10	0.02	0.02	1.00	0.46	0.04	-0.28
STATUS	0.13	0.09	-0.09	0.18	-0.07	0.46	1.00	-0.11	-0.19
AGE	-0.04	-0.27	0.24	-0.35	0.29	0.04	-0.11	1.00	0.38
SEX	0.05	0.13	0.15	-0.35	0.21	-0.28	-0.19	0.38	1.00

4) difference in negative mood from pre- to post emotional challenge

	IMAGE	GENERATE	PRESENT	FEEDBACK	RATIO	FREQ	STATUS	AGE	SEX
IMAGE	1.00	-0.04	-0.09	0.35	-0.35	0.15	-0.09	-0.66	-0.10
GENERATE	-0.04	1.00	-0.75	0.23	-0.59	0.20	0.45	-0.20	0.18
PRESENT	-0.09	-0.75	1.00	-0.43	0.48	-0.15	-0.33	0.29	0.01
FEEDBACK	0.35	0.23	-0.43	1.00	-0.50	0.13	0.30	-0.50	-0.54
RATIO	-0.35	-0.59	0.48	-0.50	1.00	-0.33	-0.41	0.25	0.20
FREQ	0.15	0.20	-0.15	0.13	-0.33	1.00	0.34	-0.18	-0.27
STATUS	-0.09	0.45	-0.33	0.30	-0.41	0.34	1.00	0.06	-0.19
AGE	-0.66	-0.20	0.29	-0.50	0.25	-0.18	0.06	1.00	0.12
SEX	-0.10	0.18	0.01	-0.54	0.20	-0.27	-0.19	0.12	1.00

Note. IMAGE = imagery instructions (yes/no); GENERATE = participants had to generate disambiguating words (yes/no); PRESENT = stimulus presentation (visual/auditory); RATIO = ratio of items in benign direction relative to all items; FREQ = number of training sessions; STATUS = health status (healthy/symptomatic); AGE = mean age of sample in years; SEX = male/female.

Figure S4 Funnel Plots for All Condition Level Effect Sizes and for All Possible Pairwise Differences Effect Sizes.



CHAPTER

8

General Discussion



Affect dynamics, that is momentary affective states and how they change from one moment to the next and in response to situations and appraisals, have been proposed to constitute a basic mechanism of mental health (Wichers, 2013). The current thesis aimed at increasing the understanding about positive and negative affect dynamic processes. Several affect dynamic processes were investigated in various ways. More specifically, the studies of this thesis addressed the influence of genetic variation on positive affect dynamics (Chapter 3, 4), stress as a contributor to negative affect dynamic processes (Chapter 5) as well as a target for change (Chapter 6, 7), and the relation between positive and negative affect dynamics (Chapter 2,3).

In this chapter the results of these studies will be discussed. First, I will focus on the role of genetic variation in positive affect dynamics (momentary positive affect and reward sensitivity) and argue that genetic variation may play a minor role. Second, I will summarize the current findings on stress and show that both social evaluative situations as well as trainings aimed at reducing stress appraisals influence negative and positive affect dynamic processes. Finally, I will combine the results from various chapters of this thesis relating to positive *and* negative affect dynamics and argue that they are distinguishable but related constructs. Potential implications for the treatment of mood-disorders will be discussed. Throughout this chapter I will indicate which questions still need to be addressed in future studies and how mood symptoms may alter affect dynamic processes.

Positive affect dynamics and genes

Genetic variation has been shown to be associated with individual differences in negative affect dynamics, including negative affect and stress sensitivity (i.e. negative affect in the context of unpleasant appraisals) as experienced in the flow of daily life (Jacobs et al., 2012; Jacobs et al., 2005; Wichers et al., 2007b). Such knowledge was scant for momentary positive affect dynamics so far. The current thesis addressed this gap and finds no strong indication for genetic variation in positive affect or reward sensitivity (i.e. positive affect in the context of pleasant appraisals) (Chapter 3, Chapter 4). Together these findings suggest that while individual differences in access to downward spiral processes is subject to genetic variation, individual differences in access to upward spiral dynamics is not. How these findings fit with previous research is discussed separately for positive affect and reward sensitivity in more detail below.

Positive affect

As mentioned in the introduction wellbeing can be separated into a cognitive (e.g., life satisfaction) and an affective component (Kuppens, Realo, & Diener, 2008; Nes, 2010;

Shmotkin, 2005). Positive affect constitutes the affective part of wellbeing (Kuppens et al., 2008; Nes, 2010; Shmotkin, 2005) and the current findings suggest that this part of wellbeing does not vary as a function of genetic variation. Some studies have previously examined the heritability of wellbeing. For example, Bartels and Boomsma (2009) investigated in a large genetically informative twin-sibling sample to what degree individual differences in the evaluation of quality of life in general, quality of life at the moment, satisfaction with life, and overall estimations of subjective happiness were subject to genetic variation. They not only found that individual differences on each of these measurements varied as a function of genetic variation but that all constructs shared an underlying genetic factor (Bartels & Boomsma, 2009). In another large genetically informative sample, Nes and colleagues found that subjective wellbeing, defined as a sum of evaluations on life satisfaction and general positive and negative affect, was to a large degree influenced by genetic variation (Nes, Czajkowski, & Tambs, 2010). While these results clearly show that individual differences in the cognitive part of wellbeing (e.g., life satisfaction, quality of life) is subject to genetic variation, the assessed affective parts like overall evaluations of subjective happiness (Bartels & Boomsma, 2009) and general positive and negative affect (Nes et al., 2010) need to be viewed with more caution. They were either assessed in a very general manner (Bartels & Boomsma, 2009) allowing for recall biases (Shiffman, Stone, & Hufford, 2008) and influences of cognitive wellbeing (Diener, 1994) or were directly combined with cognitive wellbeing assessments (Nes et al., 2010). To prevent such intertwining of positive affect with cognitive aspects of wellbeing, the studies of this thesis assessed positive affect prospectively in the flow of daily life. The current results indicate that genetic variation plays no role in individual differences in such assessed positive affect and expand previous findings on retrospectively assessed positive affect (Baker, Cesa, Gatz, & Mellins, 1992). In sum these findings suggest that the role of genetic variation on individual differences in wellbeing seems to differ between the different parts of wellbeing. While genetic variation may make some individuals more prone to cognitively evaluate their lives as more satisfying, such genetic variation appears not to matter for the actual experience of momentary positive affect in the flow of daily life. Future research needs yet to define what this implies for the interaction and association between the affective and the cognitive parts of wellbeing.

Reward sensitivity

The experience of reward has been associated with resilience against and recovery from mood-symptoms (Forbes, 2009; Geschwind et al., 2010; Wichers et al., 2009). Contrasting numerous previous findings (Bogdan & Pizzagalli, 2009; Camara et al., 2010; Dreher, Kohn, Kolachana, Weinberger, & Berman, 2009; Schmack et al., 2008;

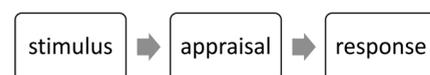
Wichers et al., 2008; Yacubian et al., 2007) the current results demonstrated only a weak – if any – influence of genetic variation on reward sensitivity in both a behavior genetic (Chapter 3) and a molecular genetic study (Chapter 4). How can this discrepancy in results be explained?

All but one (Wichers et al., 2008) of the previous studies constituted of an experimental design, that is, they included controlled reward-stimuli (Bogdan & Pizzagalli, 2009; Camara et al., 2010; Schmack et al., 2008; Yacubian et al., 2007) and defined reward sensitivity as a *stimulus* – response association (Figure 1a). Analog to the appraisal model for affective experiences (Kuppens & Tong, 2010; Roseman & Edvokas, 2004; Smith & Kirby, 2009; also see figure 4 of introduction), the appraisal of the stimulus occurs *after* the stimulus but *before* the response. This indicates that the stimulus – response association can be split into a stimulus – appraisal and an appraisal – response association (Figure 1b). The current thesis employed the latter association to define reward sensitivity, i.e. the *appraisal* – response association.

Figure 1a.



Figure 1b.



However, one previous experience sampling study already indicated that at least one genetic variant affecting experimentally defined reward (stimulus – response association) (Dreher et al., 2009) also affects experience sampling defined reward (appraisal – response association) (Wichers et al., 2008). It is therefore unlikely that these differences in reward conceptualization are a valid explanation for the discrepancy in results.

Rather, it needs to be considered that all previous studies considered one or maximally two genetic variants at the same time (e.g., Dreher et al., 2009; Wichers et al., 2008; Yacubian et al., 2007). The studies of the current thesis significantly differ in that aspect as they considered either the sum of the effects of genetic variation (Chapter 3, behavioral genetic design) or numerous genetic variants implicated in the brain's reward circuitry (Chapter 4, molecular genetic design). While the effect of one or two variants may become obvious when investigated in isolation, this effect may be too weak to be detected as significant influences once the influence of other genetic

variants is considered as well. When the influence of multiple genetic variants is tested in the same sample, the statistical chances for finding an effect even though there is none rise by the number of tests employed, i.e. a Type I error (Field, 2005). In order to prevent such a false-positive finding, the level of significance needs to be adapted according to the number of tests (multiple testing correction) as explained in more detail in Chapter 4. Such multiple testing correction was employed in the current thesis (Chapter 4) and the effects of hardly any genetic variant survived this. Such multiple testing correction is not necessary when single genetic variants are tested in isolation. Therefore, studies investigating the effect of only one or two genetic variants – as most previous studies did – may have higher chances to find small genetic effects than studies investigating the effect of numerous genetic variants. The current behavior genetic study also indicated some trend for a genetic effect; however, this seemed to be too weak to reach significance in this sample (Chapter 3), which was rather small for a genetic study. In sum this suggests that genetic variation may play a role in individual differences in reward experiences, however, the effects are presumably very small and may only be detected when explored in large samples or when single genetic variants are tested in isolation.

Moreover, the finding of the current thesis that one genetic variant survived multiple testing correction in the sample of individuals with residual depressive symptoms but not in the general population sample (Chapter 4) may indicate that it is not the genetic variant *alone* but in *interaction* with the sample that impacts on reward sensitivity. It has been emphasized that genetic influences may only become apparent when all other potentially influential factors become more stable (Nes, 2010). Nes (2010) illustrates this with a comprehensive example: individual differences in body height are well accepted to be highly influenced by genetic variation. Individuals in industrialized countries have a fairly similar diet and hence environmental influences on individual differences in body height become negligible. However, individuals in poor countries may differ in their diet due to their access to resources and therefore environmental influences on individual differences in body height increase. In the latter group environmental influences are diverse and the degree of the genetic effect may therefore be more obscured. Similarly, the sample with residual depressive symptoms of the current thesis may have been more similar on particular factors impacting on reward sensitivity (e.g., the mere experience of depressive symptoms, therapy, reduced social contacts) than individuals from the general population. Future research needs to consider investigating the role of genes in interaction with other potentially influential factors in order to understand the role of genes in reward sensitivity.

The experience of stress and an opportunity for change

Next to understanding the influence of genetic variation on positive affect dynamics another aim of the current thesis was to increase the understanding about processes in the experience of stress. After all, the experience of stress plays a central role in psychopathology (Garland et al., 2010). The current thesis reports on the development of a tool for investigating immediate stress reactions in response to peer evaluations as they may occur on online social networks. Furthermore, it was examined whether and how the tendency to stressfully appraise ambiguous situations can be reduced through training.

The stress of peer evaluations via online social networks

Social network interactions are common, particularly among adolescents (Donnerstein, 2012; Pempek, Yermolayeva, & Calvert, 2009), who are vulnerable to develop symptoms of psychopathology (Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993). One of the main features of online social networks is to comment on others' personal characteristics and to receive others' comments on own personal characteristics (Gonzales & Hancock, 2011). As the pure threat of receiving social evaluations has been demonstrated to be experienced as stressful (Dickerson & Kemeny, 2004; Pempek et al., 2009), it is very important to understand the immediate reactions to slightly negative online social evaluations. This thesis presents a new paradigm to assess such reactions (Chapter 5). In this paradigm participants are told that they are required to evaluate five unknown peers (matched in age and gender) on their looks, intelligence, and congeniality and that the same five peers will evaluate them in the same manner. To this end, all participants are asked to make a short video-profile where they provide information about personal characteristics (e.g. hobbies, stance in life). Then participants receive mild (pre-programmed) negative feedback on their looks, intelligence, and congeniality. The results support the validity of the stressor by demonstrating that the experience of social peer evaluations increased negative affect and reduced positive affect. Importantly, affective responses were particularly pronounced in individuals with symptoms of psychopathology during the last week. This indicates that the experience of stress appears to potentiate once symptoms of psychopathology are present. This is in line with previous research that experimentally induced stress (Maner, Miller, Schmidt, & Eckel, 2010) and indicates that access to downward spiral processes is particularly enhanced in these individuals.

Having established that slightly negative peer evaluation as experienced in online social networks fuels negative affect and hampers positive affect, it is now interesting to investigate whether these reactions relate to an increased risk for developing future

psychopathology and which factors may prevent such a development. The current thesis presents a suitable tool for exploring these important questions in the future.

Manipulating stress appraisals

It is important to understand the immediate affective reactions to the very specific stressful situation of social peer evaluations (Figure 2a). The current results demonstrate that such a stressor not only fuels the experience of momentary negative affect but also hampers the experience of momentary positive affect. However, understanding the affect dynamic processes of stress requires - analog to the experience of reward - that the experience of stress be split in two associations according to the appraisal model of affective experience: the stimulus-unpleasant appraisal association and the unpleasant appraisal-negative affect association (Kuppens & Tong, 2010; Roseman & Edevokas, 2004; Smith & Kirby, 2009; also see figure 4 of introduction) (Figure 2b).

Figure 2a.



Figure 2b.



Chapter 6 and Chapter 7 of this thesis more deeply focused on the stimulus-appraisal association; more specifically, on how manipulation of the ambiguous stimulus-unpleasant appraisal association may contribute to a reduction of experienced stress as a potential clinical mechanism.

Chapter 6 and 7 of the current thesis focused on the manipulation of negative interpretation biases, which includes that ambiguous situations are appraised as threatening, stressful, or unmanageable. Such negative interpretation biases are typical for individuals suffering from mood-symptoms (Beck, 1970; Beck, Emery, & Greenberg, 1985; Koster, Fox, & MacLeod, 2009; MacLeod & Cohen, 1993). Chapter 6 first demonstrates that it is possible to train young adolescents to adopt both a negative interpretation (i.e., unpleasant appraisals) and a positive interpretation (i.e., pleasant appraisals) bias. These results are in line with numerous studies in adult samples (e.g., see Amir, Bomyea, & Beard, 2010; Salemink, van den Hout, & Kindt, 2010c) and have now been replicated in adolescent samples as well (e.g., see Salemink & Wiers, 2012). This lends support to the clinical use of such trainings as particularly adolescents are vulnerable to develop mood symptoms (Lewinsohn et al., 1993) and

are particularly stress-sensitive in social situations (Sommerville, 2013), which are often ambiguous. Further supporting the clinical relevance of these trainings, the results of Chapter 7 show that training a benign interpretation bias is effective in inducing adaptive interpretation styles (i.e., benign and/or positive interpretation bias) across studies and pinned down factors increasing training effectiveness (Chapter 7). Most prominently, these included instructing participants to imagine being in the depicted situations themselves and administering more training sessions.

Despite these interesting results, it needs to be acknowledged that across studies the effects of benign cognitive bias modification for interpretation (CBM-I) was not reliably different from control conditions. Additionally, affect directly after training and in response to stressful stimuli was not reliably influenced (Chapter 7). This raises the questions whether these trainings are potent enough to decrease psychopathology. Indicating that they are, it has been shown that trainings were particularly effective in vulnerable samples (Chapter 7) and that such trainings do result in clinically significant improvements in mood symptoms in the long-run (Brosan, Hoppitt, Shelfer, Sillence, & Mackintosh, 2011; Williams, Blackwell, Mackenzie, Holmes, & Andrews, 2013). How do CBM-Is do this if their immediate effects on interpretation bias, affect, and stress sensitivity are so inconsistent?

It has been proposed previously that small adaptations of appraisal style should already have significant impact on symptoms (Kuppens et al., 2008). The findings of this thesis indeed suggest that the mechanism of action of benign CBM-Is neither lies in the immediate improvements of affect nor in decreased reactivity to stressful stimuli. Rather, the change in appraisals – even though sometimes too small to differ from control conditions (Chapter 7) – may be sufficient to crucially impact on affect dynamics as they occur in the flow of daily life. In the flow of daily life interpretation styles are employed repeatedly to make sense of daily life *ambiguous* situations. Each ambiguous situation can only be interpreted once in the flow of daily life. The effects of benign CBM-I on daily life affect dynamics would therefore naturally include increased numbers of pleasant and decreased numbers of unpleasant appraisals. Previous experience sampling research has shown that individuals with depressive symptoms – who likely embrace a negative interpretation style – indeed experience fewer pleasant appraisals than healthy controls (Bylsma, Taylor-Clift, & Rottenberg, 2011; Peeters, Nicolson, Berkhof, Delespaul, & De Vries, 2003). Combined the findings show that the tendency to interpret ambiguous situations as benign or pleasant can be increased through benign CBM-I (Chapter 7), that the number of pleasant appraisals are particularly low in individuals with depressive symptoms (Bylsma et al., 2011; Peeters et al., 2003), and

that benign CBM-I is associated with reduced mood symptoms (Brosan et al., 2011; Williams et al., 2013) suggests that benign CBM-I results in *more* pleasant appraisals of ambiguous situations in the flow of daily life and therefore increases the opportunities for experiencing positive affect.

Another effect of benign CBM-I on daily life affect dynamics should be the decreased number of unpleasant appraisals and hence fewer opportunities for triggering negative affect. Experiencing negative affect in response to negative appraisals can be summarized as stress sensitivity. Many researchers have tried to capture this proposed effect of benign CBM-I on stress sensitivity by exposing individuals to stressful stimuli and assessing affective reactions (e.g., accident videos; Mackintosh, Mathews, Yiend, Ridgeway, & Cook, 2006; giving a speech; Standage, Ashwin, & Fox, 2009). The current results (Chapter 7) show that benign CBM-I does *not* affect such defined stress sensitivity. However, stress sensitivity defined as an increase in negative affect in response to unambiguous stressful stimuli completely bypasses the expected effect of benign CBM-I. Benign CBM-I reduces negative interpretation tendencies. Thus, benign CBM-I should result in *fewer* unpleasant appraisals of *ambiguous* situations. Unambiguous stressors, however, leave little room for interpretation and very likely result in unpleasant appraisals *independent* of interpretation style. Future research needs to combine CBM-Is with experience sampling methodology to validate that CBM-I act through both increasing the number of pleasant and decreasing the number of unpleasant appraisals in the flow of daily life.

In sum, the findings of the current thesis on stress show that peer evaluative situations are experienced as stressful and not only fuel negative but also hamper positive affect dynamics. The findings also show that reducing unpleasant appraisals (i.e. negative dynamics) and enhancing pleasant appraisals (positive dynamics) of ambiguous daily life situations is possible through simple training methods. This suggests that the experience of stress as well as the reduction of stress affects both: positive and negative affect dynamic processes. In the following it will be discussed that positive and negative affect dynamics are separate constructs that, however, show important relations.

Positive and negative affect dynamics: ultimately distinguishable yet related affective constructs

As has been elaborated upon in the introduction upward and downward spirals are separable constructs that are not simple opposites of one another (Garland et al., 2010). The current thesis explored positive as well as negative affect dynamic processes that

may facilitate access to either spiral. In this section, it will be argued that also these positive and negative affect dynamics constitute two separable constructs, which – however – are still connected and able to influence one another. Clinical implications of these findings will be discussed.

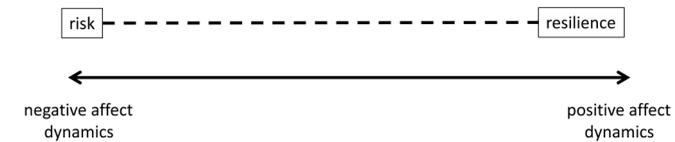
Distinguishable constructs

Previous researchers have emphasized that positive and negative affect dynamic processes are clearly distinguishable constructs that serve different functions (Fredrickson, 1998; Geschwind, Peeters, Drukker, van Os, & Wichers, 2011; Geschwind et al., 2010; Watson et al., 1995; Wichers et al., 2009; Wichers et al., 2007b; Wichers et al., 2010). The findings of the current thesis add to this knowledge in three ways. First, the finding that positive and negative affective experiences are but weakly associated (Watson et al., 1995) was replicated and extended to positive and negative affect in the flow of daily life (Chapter 2). If momentary positive and negative affect were two ends of one affective dimension we would have expected a much stronger – if not a perfect – association. Second, while previous research has shown that individual differences in negative affect dynamics are subject to genetic variation, the findings of this thesis suggest that positive affect dynamics are not (Chapter 3, Chapter 4). If both positive and negative affect dynamics were two ends of one construct we would expect that the same factors which influence negative affect dynamics also influence positive affect dynamics. At least for genes this seems not to be the case. Third, in this thesis the association between momentary reward and stress sensitivity, that is positive affect in the context of pleasant appraisals and negative affect in the context of unpleasant appraisals, was first directly tested and no association could be detected (Chapter 3). This also indicates that reward and stress sensitivity are not two ends of one “sensitivity”-construct but separate entities. Together these findings suggest that positive and negative affect dynamics are two co-existing entities of affective experience in the flow of daily life. Thus experiencing high levels of negative affect dynamics is not the same as experiencing low levels of positive affect dynamics. What does this imply for the role of both affect dynamics in mental health?

Positive and negative affect dynamics have been shown to play an important role in mental health. For example, stress sensitivity have been associated with risk for mood disorders (e.g., Wichers et al., 2007b). Reward sensitivity, on the other side, have been associated with resilience against and recovery from mood disorders (e.g., Geschwind et al., 2010; Wichers et al., 2009). In line with the current results that positive and negative affect dynamics are separate constructs it has also been demonstrated that reduced stress sensitivity did not predict improvements in mood-symptoms while

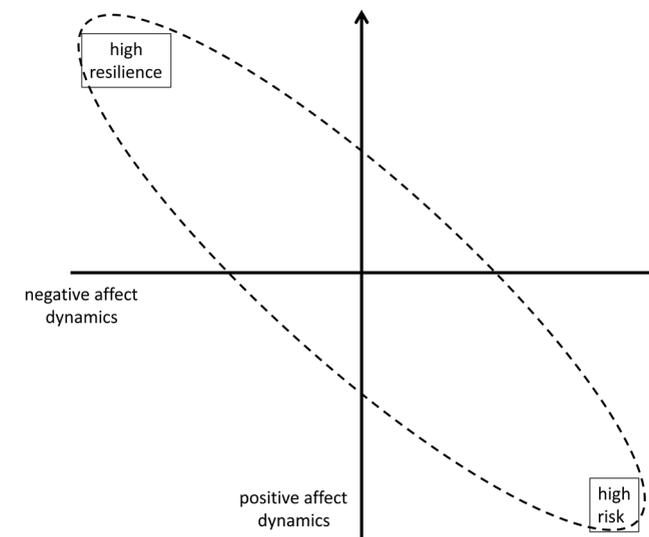
increased reward sensitivity did (Wichers et al., 2009). Combined these findings strongly contradict the idea that risk and resilience as captured with negative and positive affect dynamics are two ends of one bipolar continuum as depicted in Figure 3a.

Figure 3a.



Rather, the data suggest a two-dimensional framework for risk and resilience along two separate axes for positive and negative affect dynamics (Figure 3b).

Figure 3b.



In such a framework, risk and resilience would depict along both axes. That is, experiencing high levels of negative affect dynamics including stress sensitivity and negative affect *only* may not be a sufficient risk for developing mood disorders. Only when *high* levels of negative affect dynamics combine with *low* levels of positive affect dynamics including reward sensitivity and positive affect downward spiral processes may take over (Garland et al., 2010). As long as positive affect dynamics are high enough access to downward spiral processes may be obstructed. In that case, it would be a matter of balance between positive and negative affect dynamics rather than high or

low levels of one *or* the other. Albeit still hypothetical for the experience of momentary positive and negative affect dynamics (including momentary positive affect, negative affect, reward sensitivity, and stress sensitivity), this idea is not new. Affect balance theories propose that it is the relative predominance of overall positive over negative affect that matters for mental health (Baker et al., 1992; Lewis, McCollam, & Joseph, 2000; Shmotkin, 2005; van Schuur & Kruijtbosch, 1995). In support for this ratios of 3:1 positive over negative affect as assessed on a daily basis have been associated with optimal functioning (Frederickson & Losada, 2005). Future research needs to combine data about momentary positive and negative affect dynamics with mental health status in order to potentially expand affect balance theories to other momentary affect dynamic processes.

Related constructs

That positive and negative affect dynamics constitute two separate constructs does not allow for the conclusion that they are inherently independent from one another. For example, the experience of positive affect has been demonstrated to counteract consequences of negative affect (Fredrickson & Levenson, 1998; Garland, Gaylord, & Frederickson, 2011; Wichers et al., 2007a). Similarly, the exposure to chronic stress has been associated with a reduction in reward experiences (Berenbaum & Connelly, 1993; Pizzagalli, Bogdan, Ratner, & Jahn, 2007; Willner, 2005). Thus, even though both positive and negative affect dynamics exist next to each other, they are not unrelated. The current thesis supports this in two ways. First, as discussed in the previous section, particular factors influence both positive and negative affect dynamic processes, e.g. slightly negative social evaluation and cognitive bias modification for interpretation. Second, the current thesis zoomed in on the immediate effects of changes in positive affect on subsequent change in negative affect in the flow of daily life and vice versa. In Chapter 2 healthy controls were compared to individuals suffering from major depression. First of all, it could be replicated that individuals with a major depression presented the expected affect imbalance (APA, 2000; Watson et al., 1995); that is, higher levels of negative affect and lower levels of positive affect than healthy controls. More importantly, individuals with a major depressive episode that responded well to future treatment presented some interesting positive affect dynamics. First episode responders demonstrated a similar strong increase in positive affect as healthy controls. Recurrent episode responders demonstrated a particular ability to use their (weaker) increases in positive affect to downregulate subsequent negative affect. Individuals with a current depressive episode but without intact levels of positive affect or strong inhibitory control of positive on negative affect did not respond well to treatment. These results emphasize the importance of positive affect in the recovery from mood

symptoms. On top of that, the results indicate that this does not necessarily include increased levels of positive affect per se, but may also consist of a particular inhibitory effect on negative affect.

This is conform previous studies on daily life affect dynamics in individuals suffering from depression. For example, it has been reported that recovery from depressive symptoms is associated with the ability to holding on to positive affect (Höhn et al., 2013) and reward sensitivity (Wichers et al., 2009). Interestingly, particularly strong inhibitory control of pleasant appraisals on negative affect has also been reported in individuals with depressive symptoms (Bylsma et al., 2011; Peeters et al., 2003). In sum this indicates that particularly positive affect dynamic processes are associated with reduced mood symptoms in the future. What may these findings implicate for the treatment of mood disorders?

Clinical implications

Above findings demonstrate that standard treatment for depression appears effective in individuals that demonstrated either some intact or enhanced positive affect dynamics (i.e. similar increases in positive affect as healthy controls, particular ability to hold on to positive affect, ability to experience reward sensitivity) or an increased inhibitory control of positive affect dynamics on negative affect dynamics (i.e. decrease in subsequent negative affect after increases in positive affect, decrease in subsequent negative affect after pleasant appraisals). As negative affect dynamics, like negative affect and stress sensitivity, are particularly enhanced in individuals with mood symptoms (e.g., APA, 2000; Wichers et al., 2007b), standard treatment for depression comprehensibly mainly focuses on downregulating these negative affect dynamics (Seligman, Rashid, & Parcs, 2006). In above figure 3b this would indicate moving from right (high negative affect dynamics) to left (low negative affect dynamics). For individuals with some intact levels of positive affect dynamics, i.e. individuals situated more to the center or above on the vertical axis, this may be sufficient to recover from mood symptoms (and move them to the upper left quadrant). For individuals without intact levels of positive affect but with a strong inhibitory control of positive on negative affect dynamics this may also be sufficient as this particular ability may boost treatment effects enough to re-establish a healthy affect balance. However, individuals without intact levels of positive affect and no strong inhibitory control of positive on negative affect dynamics (i.e. individuals situated at the lower right quadrant of figure 3b) standard treatment may still be sufficient to move them further left but not necessarily further up in figure 3b.

Particularly this last group but arguably all individuals suffering from mood symptoms may therefore benefit from interventions that also upregulate positive affect dynamics. Treatments have been developed to particularly enhance positive affect dynamics and empirical evidence is emerging about their effectiveness (Geschwind et al., 2011; Seligman et al., 2006). While treatments usually consist of different components, it has also been indicated that very specific ‘interventions’ can boost positive affect dynamics. For example, it has been indicated that physical activity is associated with an increase in momentary positive affect (Wichers et al., 2012). Furthermore, the current thesis could demonstrate that manipulating a single process, i.e. fostering benign or positive interpretations of ambiguous situations, is possible in individuals suffering from mood symptoms. Other studies have demonstrated that manipulation of this process indeed results in reduced mood symptoms (Brosan et al., 2011; Williams et al., 2013). In the future it might be worthwhile to investigate whether a combination of standard treatments and treatments focusing on the enhancement of positive affect dynamics – particularly in individuals with reduced abilities to experience or use positive affect dynamics – can increase the effectiveness of treatment for mood symptoms.

References

- Amir, N., Bomyea, J., & Beard, C. (2010). The effect of single-session interpretation modification on attention bias in socially anxious individuals. *Journal of Anxiety Disorders, 24*(2), 178-182.
- APA. (2000). *Quick reference to the diagnostic criteria from DSM-IV-TR*. Arlington, VA: American Psychiatric Association.
- Baker, L. A., Cesa, I. L., Gatz, M., & Mellins, C. (1992). Genetic and environmental influences on positive and negative affect: support for a two-factor theory. *Psychology and Ageing, 7*(1), 158-163.
- Bartels, M., & Boomsma, D. I. (2009). Born to be happy? The etiology of subjective well-being. *Behavior Genetics, 39*, 605-615.
- Beck, A. T. (1970). *Depression. Causes and Treatment*. Philadelphia: University of Pennsylvania Press.
- Beck, A. T., Emery, G., & Greenberg, R. (1985). *Anxiety Disorders and Phobias. A Cognitive Perspective*. New York: Basic Books.
- Berenbaum, H., & Connelly, J. (1993). The effects of stress on hedonic capacity. *Journal of Abnormal Psychology, 102*(3), 474-481.
- Bogdan, R., & Pizzagalli, D. A. (2009). The heritability of hedonic capacity and perceived stress: a twin study evaluation of candidate depressive phenotypes. *Psychological Medicine, 39*, 211-218.
- Brosan, L., Hoppitt, L., Shelfer, A., Silence, A., & Mackintosh, B. (2011). Cognitive bias modification for attention and interpretation reduces trait and state anxiety in anxious patients referred to an out-patient service: results from a pilot study. *Journal of Behavior Therapy and Experimental Psychiatry, 42*, 258-264.
- Bylsma, L. M., Taylor-Clift, A., & Rottenberg, J. (2011). Emotional reactivity to daily events in major and minor depression. *Journal of Abnormal Psychology, 120*(1), 155-167.
- Camara, E., Krämer, U. M., Cunillera, T., Marco-Pallarés, J., Cucurell, D., Nager, W., . . . Münte, T. F. (2010). The effects of COMT (Val108/158) and DRD4 (SNP - 521) dopamine genotypes on brain activations related to valence and magnitude of rewards. *Cerebral Cortex, 20*, 1985-1996.
- Dickerson, S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*(3), 355-391.
- Diener, E. (1994). Assessing subjective well-being: progress and opportunities. *Social Indicators Research, 31*, 103-157.
- Donnerstein, E. (2012). Internet bullying. *Pediatric Clinics of North America, 59*(3), 623-633.
- Dreher, J.-C., Kohn, P., Kolachana, B., Weinberger, D. R., & Berman, K. F. (2009). Variation in dopamine genes influences responsivity of the human reward system. *PNAS, 106*(2), 617-622.
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London, Thousand Oaks, New Delhi: SAGE publications.
- Forbes, E. E. (2009). Where's the fun in that? Broadening the focus on reward function in depression. *Biological Psychiatry, 66*, 199-200.
- Frederickson, B. L., & Losada, M. F. (2005). Positive affect and the complex dynamics of human flourishing. *American Psychologist, 60*(7), 678-686.
- Frederickson, B. L. (1998). What good are positive emotions? *Review of General Psychology, 2*(3).
- Frederickson, B. L., & Levenson, R. W. (1998). Positive emotions speed recovery from the cardiovascular sequelae of negative emotions. *Cognition and Emotion, 12*(2), 191-220.
- Garland, E. L., Frederickson, B. L., Kring, A. M., Johnson, D. P., Meyer, P. S., & Penn, D. L. (2010). Upward spirals of positive emotions counter downward spirals of negativity: insights from the broaden/and/build theory and affective neuroscience on the treatment of emotion dysfunctions and deficits in psychopathology. *Clinical Psychology Review, 30*, 849-864.

- Garland, E. L., Gaylord, S. A., & Frederickson, B. L. (2011). Positive reappraisal mediates the stress-reductive effects of mindfulness: an upward spiral process. *Midfulness*, 2(59-67).
- Geschwind, N., Peeters, F., Drukker, M., van Os, J., & Wichers, M. (2011). Mindfulness training increases momentary positive emotions and reward experience in adults vulnerable to depression: a randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 79(5), 618-628.
- Geschwind, N., Peeters, F., Jacobs, N., Delespaul, P., Derom, C., Thiery, E., . . . Wichers, M. (2010). Meeting risk with resilience: high daily life reward experience preserves mental health. *Acta Psychiatrica Scandinavica*, 122(129-138).
- Gonzales, A. L., & Hancock, J. T. (2011). Mirror, mirror on my facebook wall: effects of exposure to facebook on self-esteem. *Cyberpsychology, Behavior, and Social Networking*, 14(1-2), 79-83.
- Höhn, P., Menne-Lothmann, C., Peeters, F., Nicolson, N. A., Jacobs, N., Derom, C., . . . Wichers, M. (2013). Moment-to-moment transfer of positive emotions in daily life predicts future course of depression in both general population and patient samples. *PLOS ONE*, 8(9), e75655.
- Jacobs, N., Menne-Lothmann, C., Derom, C., Thiery, E., van Os, J., & Wichers, M. (2012). Deconstructing the familiarity of variability in momentary negative and positive affect. *Acta Psychiatrica Scandinavica*, 1-10.
- Jacobs, N., Rijdsdijk, F., Derom, C., Vlietinck, R., Delespaul, P., van Os, J., & Myin-Germeys, I. (2005). Genes making one feel blue in the flow of daily life: a momentary assessment study of gene-stress interaction. *Psychosomatic Medicine*, 68, 201-206.
- Koster, E.H.W., Fox, E., & MacLeod, C. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *Journal of Abnormal Psychology*, 118(1), 1-4.
- Kuppens, P., Realo, A., & Diener, E. (2008). The role of positive and negative emotions in life satisfaction judgment across nations. *Journal of Personality and Social Psychology*, 95(1), 66-75.
- Kuppens, P., & Tong, E. M. W. (2010). An appraisal account of individual differences in emotional experience. *Social and Personality Compass*, 4(1138-1150).
- Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & Andrews, J. A. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology*, 102(133-144).
- Lewis, C. A., McCollam, P., & Joseph, S. (2000). Convergent validity of the depression-happiness scale with the Bradburn affect balance scale. *Social Behavior and Personality*, 28(6), 579-584.
- Mackintosh, B., Mathews, A., Yiend, J., Ridgeway, V., & Cook, E. (2006). Induced Biases in Emotional Interpretation Influence Stress Vulnerability and Endure Despite Changes in Context. *Behavior Therapy*, 37(3), 209-222.
- MacLeod, C., & Cohen, L. (1993). Anxiety and the interpretation of ambiguity: a text comprehension study. *Journal of Abnormal Psychology*, 102, 238-247.
- Maner, J. K., Miller, S. L., Schmidt, N. B., & Eckel, L. A. (2010). The endocrinology of exclusion: rejection elicits motivationally tuned changes in progesterone. *Psychological Science*, 21(4), 581-588.
- Nes, B. R. (2010). Happiness in behaviour genetics: findings and implications. *Journal of Happiness Studies*, 11, 369-381.
- Nes, B. R., Czajkowski, N., & Tambs, K. (2010). Family matters: happiness in nuclear families and twins. *Behavioral Genetics*, 40, 577-590.
- Nes, R. B. (2010). Happiness in behavior genetics: findings and implications. *Journal of Happiness Studies*, 11, 369-381.
- Peeters, F., Nicolson, N. A., Berkhof, J., ., Delespaul, P., & De Vries, M. (2003). Effects of daily events on mood states in major depressive disorder. *Journal of Abnormal Psychology*, 112(2), 203-211.
- Pempek, T. A., Yermolayeva, Y. A., & Calvert, S. L. (2009). College students' social networking experiences. *Journal of Applied Developmental Psychology*, 30(227-238).
- Pizzagalli, D. A., Bogdan, R., Ratner, K. G., & Jahn, A. L. (2007). Increased perceived stress is associated with blunted hedonic capacity: potential implications for depression research. *Behavior Research and Therapy*, 45, 2742-2753.
- Roseman, I., & Edvokas, A. (2004). Appraisals cause experienced emotions: experimental evidence. *Cognition and Emotion*, 18(1), 1-28.
- Salemink, E., van den Hout, M., & Kindt, M. (2010c). Generalisation of modified interpretive bias across tasks and domains. *Cognition and Emotion*, 24(3), 453-464.
- Salemink, E., & Wiers, R. W. (2012). Adolescent threat-related interpretive bias and its modification: the moderating role of regulatory control. *Behavior Research and Therapy*, 50, 40-46.
- Schmack, K., Schlagenhauf, F., Sterzer, P., Wrase, J., Beck, A., Dembler, T., . . . Gallinat, J. (2008). Catechol-O-methyltransferase val158met genotype influences neural processing of reward anticipation. *NeuroImage*, 42, 1631-1638.
- Seligman, M., Rashid, T., & Parcs, A. C. (2006). Positive Psychotherapy. *American Psychologist*, 61(8), 774-788.
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4(1), 1-32.
- Shmotkin, D. (2005). Happiness in the face of adversity: reformulating the dynamic and modular bases of subjective well-being. *Review of General Psychology*, 9(4), 291-325.
- Smith, A. C., & Kirby, L. D. (2009). Putting appraisals in context: toward a relational model of appraisal and emotion. *Cognition and Emotion*, 23(7), 1352-1372.
- Sommerville, L. H. (2013). The teenage brain: sensitivity to social evaluation. *Psychological Science*, 22(2), 121-127.
- Standage, H., Ashwin, C., & Fox, E. (2009). Comparing visual and auditory presentation for the modification of interpretation bias. *Journal of Behavior Therapy and Experimental Psychiatry*, 40(4), 558-570.
- van Schuur, W. H., & Kruijtbosch, M. (1995). Measuring subjective well-being: unfolding the Bradburn affect balance scale. *Social Indicators Research*, 36(1), 49-74.
- Watson, D., Clark, L. A., Weber, K., Assenheimer, J. S., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: II. Exploring the symptom structure of depression and anxiety in student, adult, and patient samples. *Journal of Abnormal Psychology*, 104(1), 15-25.
- Wichers, M. (2013). The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine*, 1-12.
- Wichers, M., Aguilera, M., Kenis, G., Krabbendam, L., Myin-Germeys, I., Jacobs, N., . . . van Os, J. (2008). The catechol-O-methyl transferase val158met polymorphism and experience of reward in the flow of daily life. *Neuropsychopharmacology*, 33, 3030-3036.
- Wichers, M., Barge-Schaapveld, D. Q. C. M., Nicolson, N. A., Peeters, F., de Vries, M., Mengelers, R., & van Os, J. (2009). Reduced stress-sensitivity or increased reward experience: the psychological mechanism of response to antidepressant medication. *Neuropsychopharmacology*, 34, 923-931.
- Wichers, M., Myin-Germeys, I., Jacobs, N., Peeters, F., Kenis, G., Derom, C., . . . van Os, J. (2007a). Evidence that moment-to-moment variation in positive emotions buffer genetic risk for depression: a momentary assessment twin study. *Acta Psychiatrica Scandinavica*, 115, 451-457.
- Wichers, M., Myin-Germeys, I., Jacobs, N., Peeters, F., Kenis, G., Derom, C., . . . van Os, J. (2007b). Genetic risk for depression and stress-induced negative affect in daily life. *British Journal of Psychiatry*, 191, 218-223.

- Wichers, M., Peeters, F., Geschwind, N., Jacobs, N., Simons, C. J. P., Derom, C., . . . van Os, J. (2010). Unveiling patterns of affective responses in daily life may improve outcome prediction in depression: a momentary assessment study. *Journal of Affective Disorders, 124*, 191-195.
- Wichers, M., Peeters, F., Rutten, B., Jacobs, N., Derom, C., Thiery, E., . . . van Os, J. (2012). A time-lagged momentary assessment study on daily life physical activity and affect. *Health Psychology, 31*(2), 135-144.
- Williams, A. D., Blackwell, S. E., Mackenzie, A., Holmes, E. A., & Andrews, G. (2013). Combining imagination and reason in the treatment of depression: a randomized controlled trial of internet-based cognitive-bias modification and internet-CBT for depression. *Journal of Consulting and Clinical Psychology, Advance online publication*. doi: 10.1037/a0033247.
- Willner, P. (2005). Chronic mild stress (CMS) revisited: consistency and behavioral-neurobiological concordance in the effects of CMS. *Neuropsychobiology, 52*, 90-110.
- Yacubian, J., Sommer, T., Schroeder, K., Gläscher, J., Kalisch, R., Leuenberger, B., . . . Büchel, C. (2007). Gene-gene interaction associated with neural reward sensitivity. *PNAS, 104*(19), 8125-8130.



Summary

The current thesis presents research about affect dynamic processes that remained ill-understood so far. Affect dynamics include the momentary experience of positive and negative affect, changes in these experiences from one moment to the next, and changes in response to appraisals and situations. The first part of this thesis (Chapter 2, 3, and 4) examines affect dynamics in daily life with the experience sampling methodology. More specifically, the interaction of positive and negative affect in relation to depressive symptoms and the role of genes in positive affect and reward sensitivity were investigated. The second part (Chapter 5, 6, and 7) focuses on the experience of stress and the opportunity for changing stressful appraisals. To this end, experimental designs are reported.

Chapter 1 introduces the affect dynamic processes that were investigated in the current thesis. It is argued that both positive and negative affect dynamic processes are important building blocks of mental health and that the interaction between positive and negative affect dynamics need further investigation. Genes and stress are introduced as crucial influential factors for affective experiences. Furthermore, two suitable methods to investigate momentary affective experience are discussed: experience sampling and experimental designs. The Chapter concludes with the aims and an overview of the current thesis.

Chapter 2 explores the dynamic interplay in momentary positive and negative affect (PA, NA, respectively). More specifically, subsequent PA in response to NA change as well as subsequent NA in response to PA-change in the flow of daily life was examined. It was explored whether any of these processes were able to differentiate between healthy controls (n=39) and individuals suffering from a current major depressive episode (n=47) or were able to predict future recovery from major depression. The results indicated that individuals with a recurrent depressive episode that later responded to treatment differed in their NA-PA dynamics from individuals that did not respond well to treatment. The first group showed a stronger reduction of NA following PA-booster as compared to individuals that did not respond well to treatment and as compared to controls. This suggests that individuals with a recurrent episode of depression that will recover in the future are particularly effective in counteracting the processes of the downward spiral of negative emotions by using PA-booster to suppress subsequent NA. These results also suggest that subtle affect dynamics of daily life may help predicting treatment response.

Chapter 3 presents a behavioral genetic twin-design. It was investigated whether and to what degree the momentary experience of positive affect and reward sensitivity (the experience of positive affect in the context of pleasant appraisals) is heritable and

whether stress sensitivity (negative affect in the context of unpleasant appraisals) and reward sensitivity (positive affect in the context of pleasant appraisals) were associated. Positive affect was solely environmentally influenced, whereas the results for reward sensitivity were less conclusive. A very small – if any – genetic effect on reward sensitivity was indicated. Stress sensitivity was not correlated with reward sensitivity. This may indicate that at least some affect dynamics indicating vulnerability (stress sensitivity) are inherently independent from affect dynamics indicating resilience and wellbeing (reward sensitivity) in the general population.

Chapter 4 explored whether any of the genetic variations previously shown to impact on the brain's reward-system would be associated with reward sensitivity in daily life. To this end the impact of 24 genetic variations on reward sensitivity were investigated in a general population sample (n=621) and a sample of individuals with residual depressive symptoms (n=130). While none of the genes became significant in the general population sample, the genetic variant 'OPRM1 rs1799971' was significantly associated with the level of reward sensitivity in the sample of individuals with residual depressive symptoms. These results indicate that within a sample with residual depressive symptoms particular genes may influence reward sensitivity in daily life.

Chapter 5 reports on the development and validation of a new social stressor mimicking peer interactions in online social networks. Two-hundred-and-forty-one young participants were exposed to slightly negative peer evaluations on looks, intelligence, and congeniality. Results showed that negative affect and cortisol-levels increased and positive affect and implicit self-esteem decreased in response to evaluation exposure, indicating significant stress. Furthermore, female participants and participants with mood-symptoms showed a stronger affective reaction, whereas young adolescent participants showed an increase in cortisol release. These findings indicate that online social evaluative situations influence daily life affect dynamics – particularly in individuals vulnerable to mood-symptoms.

While social evaluative situations may be unequivocally stressful for everyone, it has been demonstrated that individuals suffering from mood symptoms start to experience stress in situations that are ambiguous. This is because they start to interpret ambiguous situations as stressful (i.e., they demonstrate a negative interpretation bias). It has repeatedly been shown that this negative interpretation bias can be modified through cognitive bias modification training in adults. **Chapter 6** reports on the first study showing that these trainings also affect interpretation biases in adolescents (n=82). Immediate effects on positive and negative affect were less obvious and became only significant when individuals were trained to draw negative

interpretations of ambiguous situations. These results suggest that a crucial part influencing affect dynamics in daily life – the appraisal of stressful situations – can be modified in adolescents if ambiguous situations are concerned.

While it has been demonstrated that cognitive bias modification trainings for interpretation biases (CBM-I) are generally effective in changing interpretation biases and associated affect, it remained unknown what the clinical potential of training a benign interpretation bias would be and which factors make CBM-I trainings effect. **Chapter 7** therefore reports on a meta-analysis on the effectiveness of benign CBM-I in increasing an adaptive, i.e., positive, interpretation bias and decreasing negative affect. It could be shown that benign CBM-I generally results in an increase in a positive interpretation bias and a decrease in negative affect although these effects were not consistently different from various control conditions. This effect could partially be boosted by instructing participants to imagine the ambiguous situations of the training were actually happening to them, by administering more training sessions, and by administering feedback about response correctness during training. Most encouragingly, benign CBM-I trainings were most effective in individuals with mood-symptoms and females. Surprisingly, benign training did not result in reduced stress-reactivity to a subsequent stressor. However, most of these stressors did not provoke ambiguity but were stressful more generally and may therefore not have captured the effect of benign CBM-I. Overall, these results show that and how the tendency to negatively appraise situations can be altered. Future research needs to demonstrate whether this can actually influence affect dynamics in daily life.

Chapter 8 provides a general discussion of the studies included in this thesis. Also implications for future research and mental health care are discussed.



Samenvatting

Dit proefschrift presenteert onderzoek over processen in 'affect dynamics' dit tot nu toe niet goed werden begrepen. Affect dynamics omvat het beleven van positieve en negatief affect zoals ze in het moment voorkomen, veranderingen in deze belevingen van het éne moment naar het andere, en veranderingen hierin ten opzichte van inschattingen en situaties. Het eerste deel van dit proefschrift (Hoofdstuk 2, 3, en 4) onderzoekt affect dynamics in het dagelijks leven met de experience sampling method (ESM). Hier werden de interactie van positief en negatief affect met betrekking tot depressieve symptomen en de rol van genen in positief affect en beloningsgevoeligheid onderzocht. Het tweede deel (Hoofdstuk 5, 6, en 7) richt zich op het beleven van stress en de mogelijkheid stressvolle inschattingen te veranderen. Dit deel rapporteert over experimentele designs.

Hoofdstuk 1 introduceert affect dynamische processen die in dit proefschrift onderzocht werden. Het beargumenteert dat zowel positief als negatief affect belangrijke bouwstenen zijn van geestelijke gezondheid en dat de interactie tussen positief en negatief affect meer onderzocht moet worden. Genen en stress worden als belangrijke beïnvloedende factoren voor affectieve belevingen geïntroduceerd. Bovendien, twee geschikte methodes voor het onderzoeken van affectieve belevingen in het moment worden bediscussieerd: experience sampling en experimentele designs. Dit hoofdstuk concludeert met de doelen van dit proefschrift.

Hoofdstuk 2 onderzoekt het dynamisch samenspel in positief en negatief affect zoals ze in het moment ervaren worden (PA, NA, respectievelijk). Dat is, navolgend PA in reactie op veranderingen in NA zowel als navolgend NA in reactie op veranderingen in PA werden in het dagelijks leven onderzocht. Er werd nagegaan of één van deze processen in staat is een onderscheid te maken tussen gezonde controles (n=39) en individuen met een depressieve episode (n=47), of in staat is toekomstig herstel van depressieve episodes te voorspellen. De resultaten laten zien dat patiënten met een recidiverende depressie die achteraf goed op behandeling reageerden verschilden in hun NA-PA dynamiek van soortgelijke patiënten die achteraf niet goed reageerden. De eerste groep vertoonde namelijk een sterkere daling in NA als gevolg van een momentane verhoging in PA dan de tweede groep. Dus een PA ervaring lijkt bij de eerste groep een sterker effect te hebben. Deze groep bleek zelfs nog een groter effect te hebben van een PA ervaring dan gezonde controles. Dit suggereert dat mensen met een recidiverende depressieve episode die in de toekomst van deze episode herstellen bijzonder effectief zijn in het tegenwerken van de downward spiral van negatieve emoties door het gebruiken van PA-verhogingen om NA te onderdrukken. De resultaten suggereren bovendien dat subtiele affect dynamics in het dagelijks leven helpen om response op behandeling te voorspellen.

Hoofdstuk 3 presenteert een genetisch informatief tweeling-design. Er werd onderzocht of en tot welke mate het in-het-moment beleven van positief affect en beloningsgevoeligheid (het beleven van positief affect in de context van plezierig ervaren situaties) erfelijk zijn en of stressgevoeligheid (negatief affect in de context van onaangename ervaren situaties) samenhangen. Positief affect werd alleen door omgevingsfactoren beïnvloed, terwijl de resultaten voor beloningsgevoeligheid minder duidelijk waren. Er werd maar een heel klein genetisch effect op beloningservaring aangegeven. Stressgevoeligheid hing niet samen met beloningsgevoeligheid. Dit zou kunnen betekenen dat in de algemene bevolking affect dynamics die kwetsbaarheid markeren (stressgevoeligheid) inherent onafhankelijk zijn van affect dynamics die weerbaarheid en well-being markeren (beloningsgevoeligheid).

Hoofdstuk 4 onderzoekt of genetische varianten die eerder aangetoond werden een rol te spelen in het beloningssysteem van de hersenen samenhangen met beloningsgevoeligheid in het dagelijks leven. De invloed van 24 genetische variaties op beloningsgevoeligheid werd in de algemene populatie (n=621) en een sample met residuele depressieve klachten (n=130) onderzocht. Terwijl in de algemene populatie geen van de genetische variaties een significante uitslag gaf, hing in de groep met residuele depressieve klachten de genetische variatie "OPRM1 rs1799971" duidelijk met beloningsgevoeligheid samen. Dus de resultaten tonen aan dat binnen een groep mensen met depressieve klachten bepaalde genetische variaties mogelijk samenhangen met beloningsgevoeligheid in het dagelijks leven.

Hoofdstuk 5 rapporteert over de ontwikkeling en validatie van een nieuwe sociale stressor die peer interacties in online social networks nabootst. Twee-honderd-eneenveertig jonge deelnemers ontvingen licht negatieve evaluaties van peers op hun uiterlijk, intelligentie, en sympathie. De resultaten lieten zien dat in reactie op blootstelling aan deze evaluatie negatief affect en cortisol-levels omhoog en positief affect en impliciet zelfwaardering omlaag gingen. Dit geeft aan dat de sociale stressor duidelijk effect had op de deelnemers. Bovendien vertoonden vrouwelijke deelnemers en deelnemers met stemmingsklachten een sterkere affectieve reactie dan mannen en deelnemers zonder stemmingssymptomen. Ook lieten jongere adolescente deelnemers een sterkere cortisol reactie zien dan relatief oudere deelnemers. Deze bevindingen geven aan dat online sociaal evaluatieve situaties affect en cortisol reacties teweeg kunnen brengen – en dat in bijzondere mate in jonge adolescenten, vrouwen en mensen met bestaande stemmingsklachten.

Terwijl sociaal evaluatieve situaties waarschijnlijk door iedereen als stressvol ervaren worden, is aangetoond dat mensen die last hebben van stemmingstoornissen ook

volledig ambigue situaties als stressvol te ervaren. Dit is omdat ze ambigue situaties negatief *interpreteren* (i.e., ze vertonen een negatieve interpretatie bias). Het is herhaaldelijk aangetoond dat deze negatieve interpretatie bias door 'cognitive bias modification training' in volwassenen kan worden aangepast. **Hoofdstuk 6** rapporteert de eerste studie die laat zien dat dit soort trainingen ook interpretatie biases in adolescenten (n=82) kunnen veranderen. Onmiddellijke effecten op positief en negatief affect waren minder duidelijk en werden alleen significant als een negatieve interpretatie bias werd aangeleerd. Deze resultaten suggereren dat een belangrijk factor voor affect dynamics in het dagelijks leven – het inschatten van situaties als negatief – in adolescenten veranderd kan worden in de gevallen waarin deze situaties ambigue zijn.

Terwijl het aangetoond is dat cognitive bias modification trainingen (CBM-I) over het algemeen effectief zijn in het veranderen van interpretatie biases en geassocieerde affect, was het tot nu toe onduidelijk welke klinische potentie deze trainingen hebben en welke factoren CBM-Is bijzonder effectief maken. Daarom rapporteert **hoofdstuk 7** over een meta-analyse over de effectiviteit van goedaardige CBM-I trainingen in het verhogen van een adaptieve, i.e. positieve, interpretatie bias en het verlagen van negatief affect. De resultaten tonden aan dat licht positief CBM-I over het algemeen in een verhoging van een positieve interpretatie bias en in een verlaging van negatief affect resulteerde, alhoewel deze effecten niet altijd van controle condities verschilden. Verder tonden de resultaten aan dat dit effect door imaginatie-instructies, meerdere training sessies en het geven van feedback over de gegeven antwoorden, nog extra versterkt werd. Belangrijk was de bevinding dat licht positief CBM-I trainingen vooral effectief waren in mensen met stemmingsproblemen en vrouwen. Tegen de verwachting leidde licht positief CBM-I niet tot een gereduceerde stressgevoeligheid op een navolgende stressor ten opzichte van andere training condities. Echter, de meeste stressoren provoocerden niet noodzakelijk ambiguiteit maar waren over het algemeen stressvol. Daarom bestaat de mogelijkheid dat deze stressoren het eigenlijke effect van CBM-I, i.e. de interpretatie van ambigue situaties als niet-stressvol, niet hebben kunnen vangen. Samengevat tonen deze resultaten aan dat en in welke mate de neiging ambigue situaties negatief in te schatten veranderd kan worden. Toekomst moet laten zien of deze training ook affect dynamics in het dagelijks leven kan beïnvloeden.

Hoofdstuk 8 geeft een algemene discussie over de studies van dit proefschrift. Ook worden de implicaties voor toekomstig onderzoek en geestelijke gezondheid bediscussieerd.



Valorization

Valorization is “[a] process [that] facilitates the practical application of scientific knowledge. By enhancing the suitability and availability of research findings, valorization increases the possibility that third parties can apply these findings” (AWT, 2007). More specifically, for me, valorization translates to answering the question of why I did this research and how the findings of my thesis can be used for fostering mental wellbeing.

In my thesis I was driven to find out what makes us feel happy and even-tempered in our day-to-day living. I addressed this from different perspectives and with various methods. One very important factor hampering feeling well is the experience of stress. The experience of stress is not necessarily solely caused by exposure to more stressful events. Rather the *appraisal* of particular situations as stressful results in feeling stressed. For example, individuals suffering from mood symptoms are oftentimes plagued by the stressful appraisal of naturally ambiguous situations (Beck, 1970; Beck, Emery, & Greenberg, 1985). Even though the situation is ambiguous, these individuals appraise it as negative and consequently they experience more stress. In the current thesis I elaborate on the possibility of changing the tendency to negatively appraise situations in order to decrease the experience of stress.

In the last two chapters of my thesis (Chapter 6 and 7) I report on computerized trainings for benign or positive interpretation biases. Interpretation biases consist of cognitive tendencies to interpret ambiguous situations in daily life in either a negative or threatening, or a benign or positive manner. For example, imagine you enter a room at work. Some of your colleagues are standing in a corner. As you enter the room they are looking at you laughing. This situation is inherently ambiguous as it does have neither a positive nor a negative implication for you. However, our brains are preprogrammed to make sense of the world and are therefore tempted to interpret this situation to be either good or bad. Imagine this interpretation is that your colleagues are laughing because they were talking about that something you said earlier was ridiculous. This would be a negative interpretation bias and transfer the inherently ambiguous situation into a stressful or even threatening situation, most likely causing you to feel stressed and negative about yourself. Contrary, your interpretation could also have been that your colleagues are happy seeing you or are in a good humor. This then would be a positive/benign interpretation bias. This transfers the inherently ambiguous situation into a pleasant one and probably makes you feel good and happy. Such ambiguous situations occur frequently in our daily lives. Having a tendency to interpret these situations negatively makes the world a stressful and hostile place for us and leaves us depressed or anxious. Contrary, having a tendency

to interpret these situations as benign or positive, would make us feel at ease and more happy. Researchers have explored the possibility of training benign/positive interpretation biases via computerized trainings. Various training paradigms have been developed and it has very recently been shown that practicing these trainings is associated with decreased symptoms in depression and anxiety (Williams, Blackwell, Mackenzie, Holmes, & Andrews, 2013). In the current thesis I add to the indicated use of such trainings in two ways. First, I showed that these trainings are not only effective in adults (in whom all trainings for inducing a benign interpretation bias had been conducted thus far) but also in adolescents. This is important because we know that adolescents are particularly vigilant to social interactions (Sommerville, 2013), which are very often ambiguous. Furthermore, adolescents are at increased risk to develop mood symptoms (Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993). Second, I was able to pin down the factors that increase the effectiveness of interpretation bias trainings. For example, I found that when people are instructed to really imagine being in the ambiguous situations presented to them during the training, it was much more effective in inducing a benign interpretation bias than when individuals were not instructed to do this. Also, when more training sessions were conducted effectiveness was increased.

Relevance and Target groups

These findings are of relevance for the treatment and prevention of mood symptoms. Depression and anxiety cause the most years lived with disability from all mental health disorders (WHO, 2013) and negative interpretation biases are common in these disorders (Beck, 1970; Beck et al., 1985). However, large percentages of individuals with such symptoms are not treated (WHO, 2013). The findings of the current thesis indicate that these trainings are effective in a broad age-range including adolescents and adults. Furthermore, the findings of the current thesis show how such trainings need to be implemented in order to effectively reduce negative interpretation biases. One way of valorizing these findings for society is through making effective trainings broadly and – ideally – freely available. Giving free access to such training methods to both adults and adolescents suffering from mood symptoms or feel particularly stressed may be a promising new avenue for self-help and supporting psychotherapy. For example, individuals may realize that they are always alert to potentially threatening or negative situations and suffer from the resulting stress. Other individuals may be receiving treatment for their mood symptoms already. Those that receive psychopharmacological treatment may feel somewhat passive about their own role in remission and wish to engage more. Those that receive psychotherapeutic treatment may feel that it can sometimes be (too) demanding for their cognitive resources to disengage from

automatized maladaptive cognitive strategies, i.e. negative interpreting ambiguous situations. For these individuals (and their therapists) it may be of high value to have a tool that supports them modifying automatized negative interpretation biases.

What do we have to offer and what does still need to be done?

The results of the current thesis clearly indicate how training needs to be conceptualized to effectively manipulate negative interpretation biases. The neatest thing about it is that it can be fully computerized and accessed on the internet. For this purpose a website needs to be created and sustained. Individuals interested in training their positive interpretation biases could log-in to this page by creating their own username. As a start individuals could assess the existence and the strength of a negative interpretation bias through a short experiment. As our findings suggest that trainings are more effective when conducted repeatedly a username would enable the program to offer slightly different versions of the training every time the individual logs in. Individuals can then 'train' repeatedly at own pace without getting used to the situations within the training.

Even though we demonstrated in this thesis that trainings for benign interpretation biases are effective and how they should be conceptualized to gain most effectiveness in adapting interpretation biases, two important issues need attention. First, it may be necessary to quantify the effects of such a free-paced self-help program. For this purpose, individuals can be asked whether they agree to provide their data anonymously for research purposes quantifying the effects of such trainings. For individuals agreeing to take part, additional information including basic demographic variables (age, sex), brief mood symptom measurements, and engagement in other treatment, including psychopharmacological and psychotherapeutic treatment, needs to be collected. For a fixed number of individuals (e.g., one hundred) and after a number of training sessions (e.g., 6) interpretation bias and symptoms would be assessed again and the effectiveness of training on both interpretation style and symptoms can be quantified. The second important issue is the costs. The website needs to be developed and sustained, data need to be managed, assessment and training of interpretation biases need to be developed, and server space needs to be available. It is possible to ask individuals to pay a short amount of money when they register on the website. Ideally, however, insurances or universities may sponsor such an endeavor to fulfill a social commitment. The latter option is to be preferred because it prevents that only those individuals with sufficient money have access to the website. Adolescents and individuals with low socio-economic status may be at a disadvantageous position.

Outreach

In order to successfully establish such a website, it is important to make known to several groups of individuals that fully computerized interpretation bias modification trainings exist and that they can effectively change interpretation biases. To this end, patient organizations, insurances, associations of psychotherapists, psychiatrists, and teachers need to be personally targeted. This might be achieved by personally presenting the results of the current thesis and the website to them. Furthermore, it would be important to explore their judgment about the feasibility and interest in such a program. If necessary, their concerns might be used in optimizing the program and website.

References

- Adviesraad voor het wetenschaps- en technologiebeleid (AWT) (2007). *Summary Advice 70. A radiant future – policies for 'valorisation' of the humanities and social sciences*. Retrieved 21.10.2014 from <http://www.awti.nl/publicaties/alfa-en-gamma-stralen/item94>.
- Beck, A. T. (1970). *Depression. Causes and Treatment*. Philadelphia: University of Pennsylvania Press.
- Beck, A. T., Emery, G., & Greenberg, R. (1985). *Anxiety Disorders and Phobias. A Cognitive Perspective*. New York: Basic Books.
- Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & Andrews, J. A. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology, 102*(133-144).
- Sommerville, L. H. (2013). The teenage brain: sensitivity to social evaluation. *Psychological Science, 22*(2), 121-127.
- WHO. (2013). Depression. Key facts. Retrieved 24.07.2013 from <http://www.who.int/mediacentre/factsheets/fs369/en/2013>
- Williams, A. D., Blackwell, S. E., Mackenzie, A., Holmes, E. A., & Andrews, G. (2013). Combining imagination and reason in the treatment of depression: a randomized controlled trial of internet-based cognitive-bias modification and internet-CBT for depression. *Journal of Consulting and Clinical Psychology, Advance online publication*. doi: 10.1037/a0033247.



Dankwoord

De omgevingsfactoren die de totstandkoming van dit proefschrift hebben bevorderd zijn vooral de mensen die me zowel op professioneel gebied alsook in mijn persoonlijk leven hebben voorbereid, begeleid, en gesteund.

Ten eerste wil ik graag mijn **promotieteam** danken, voor de mogelijkheid om met jullie begeleiding mijn promotieonderzoek te doen. **Jim**, bedankt dat jij ongeacht jouw drukke agenda altijd beschikbaar was om input voor mijn papers te leveren; dat je luisterde naar mijn vragen en hier altijd snel op reageerde; en vooral voor jouw enthousiasme voor ons werk. **Marieke**, bedankt voor alles wat ik van jou heb mogen leren. Bedankt dat je me altijd de ruimte hebt gegeven om op mijn eigen manier te werken en ideeën te ontwikkelen; dat je er altijd beschikbaar was met waardevol, opbouwend en enthousiast feedback; dat je me hebt gesteund in mijn professionele maar ook persoonlijke keuzes. Tijdens al die tijd heb ik nooit het gevoel moeten hebben dat jouw belangen te zeer aan mijn keuzes gebonden waren, wat me een heel prettig gevoel van zelfstandigheid en vrijheid gaf. Dank! **Nele**, zonder jou was ik mogelijk niet door de laatste maanden heen gekomen. Jouw vaardigheid achter het werk de mens te zien en hindernissen met weinige empatische woorden de schrik te nemen (soms leek me deze vaardigheid wel wat metafysisch als ik bv na een een-zins-email over handtekeningen een antwoord over volhouden en vertrouwen in mijn vaardigheden terug kreeg ;)). Bedankt voor jouw geduld om uren met mij achter het scherm te zitten om Mx modellen uit te leggen en mee te puzzelen; ik heb hier enorm veel van jou mogen leren! En vooral bedankt voor je gezelligheid en je manier niet alles te serieus te nemen. Dit heeft ons vooral in het twinssCan team en hier omheen de een of andere deur geopend.

Bedankt aan alle leden van de **leescommissie** en **corona** voor jullie tijd en inzet. **Inez**, bedankt dat je voorzitter bent en voor de goede gesprekken - vooral over het leven met kinderen - in de loop van de afgelopen vijf jaren.

Jennifer, thanks for the great time in Oxford and the successful collaboration throughout the last 5 years.

Het **twinsCan** team in Maastricht: vooral Jeroen en Debora, maar ook Marieke, Nele, Ruud, Jim, Bart, Jill, Katrien, en Davinia. Bedankt voor de heel waardevolle en belonende ervaring wat alles mogelijk is als men in een goed team mag werken! **Jeroen**, het was een groot geluk voor mij met jou op een project samen te mogen werken wat ons allebei erg aan het hart lag. Jouw enthousiasme, jouw vermogen vol te houden en dingen tegen weerstand door te zetten, en jouw gezellige manier hebben me altijd met veel plezier aan het project laten werken. Bedankt voor een fantastische

tijd vol met onderzoekersgeest en vriendschap, lachen en vloeken, hard werken en af en toe lekker eten, maar zeker altijd een (of meestal meer) kopje(s) koffie. **Debora**, voor het helemaal enthousiast mee denken en werken, voor jouw vaardigheid leiding te geven aan het team in Gent, voor de gezellige reizen op de trein, voor de leuke dagen in Gent, voor je vriendschap op Vijverdal als ik er nog bijna niemand kende. Het was en is een enorme geruststelling jou in het team te hebben omdat je alles goed in handen had en nog steeds heeft! **Jill**, **Katrien**, en **Davinia**, bedankt voor al jullie tijd en inbrengst. Zonder jullie was twinssCan nooit 'gebeurd'; dat jullie de tweelingen hebben weten te overtuigen; dat jullie zo zelfstandig in Gent gewerkt hebben; en dat jullie bereid waren door heel Oost-Vlaanderen te reizen. Ook voor de gezellige dagen in Gent wil ik jullie van harte bedanken. **Marieke**, **Nele**, en **Ruud** voor de frequente en productieve begeleiding vooral tijdens de opstartfase. **Catherine** en **Evert**, bedankt dat jullie ons welkom heten in het EFPTS en we jullie infrastructuur hebben mogen gebruiken voor twinssCan. En natuurlijk: bedankt aan alle **tweelingen** die bereid waren aan ons intensief onderzoek mee te werken!

Mijn kamergenootje, **Martine**. Bedankt voor de enorm gezellige tijd op kamer SN 2.055 ;), voor al de leuke wetenschappelijke maar vooral de persoonlijke gesprekken; het deed vaak deugd positieve en negatieve nieuws meteen met iemand te kunnen delen. Ik zal deze tijd nooit vergeten en er altijd met veel positief affect aan terug denken J!

Mijn psyMate-lotgenotje, **Jessy**. Danke für deine Freundschaft, für die Möglichkeit stressvolle Phasen mit dir zu besprechen und schöne Dinge mit dir zu teilen. Du fehlst mir hier!

Dina, es ist schön, dass wir uns kennengelernt haben. Vielen Dank für dein offenes Ohr und deine Freundschaft. Vielleicht schaffen wir es ja nach den gescheiterten Anläufen der Vergangenheit in der (weiteren) Zukunft doch mal zusammen ein Projektchen auf die Beine zu stellen ;).

Sanne, voor de taalrijke bijna-lotgenoten gesprekken. Voor je altijd opbouwende manier en vaardigheid te lachen J!

Lieve **AiO-collega's**, vooral **Nicole L.**, **Nicole G.**, **Petra**, **Zuzana**, **Feikje**, **Annelie**, en **Jindra**: voor de gesprekken, grapjes en de gezelligheid; voor de wetenschappelijke brainstorm discussies; jullie hebben ervoor gezorgd dat ik altijd graag naar Vijverdal kwam. Petra en Zuzana, dat jullie mij zo veel met de meta hebben geholpen. Dank!

Ron, dat ik altijd met mijn onbegrip voor computerdingen bij jou terecht kon en je altijd een snelle oplossing voor me had. **Truda**, dat je altijd aanspreekbaar was voor allerlei vragen rond data en vragenlijsten. **Trees**, dat je me altijd hebt verder geholpen met administratieve dingen. **Leni**, voor de leuke gesprekken en je goede humeur!

Wendy, voor je steun, hulp, en vooral gezelligheid bij het team REMOD Maastricht. **Claudia** en **Ingrid**, voor de tijd in Eindhoven en Maastricht, de leuke etentjes en een goede samenwerking aan een moeilijk project. Het was enorm leerzaam voor me.

Wolfgang, was hätte ich nur ohne dich gemacht? Jedenfalls keine Meta-Analyse, mit der ich gut leben kann, keine Gen-Analysen, bei denen ich ein akzeptables Gefühl hab. Ausserdem hätte ich wesentlich weniger gelacht und hätte meine Doktorarbeit auch wesentlich dümmmer abgeschlossen ;). Vielen, vielen Dank für die Stunden (Tage), die du mit mir an der Meta-Analyse gearbeitet hast, und all das Wissen, was dank Deiner Fähigkeit, Statistik begreiflich zu erklären, seinen Weg in mein Hirn gefunden hat. **Marjan**, voor je steun en input bij de meta-analyse en andere ingewikkelde statistische vragen!

Steffi, schön, dass du trotz der Kilometer zwischen uns eine Konstante in meinem Leben bist, die mir ab und zu den Kopf zurecht rückt und mir so oft zugehört hat. Danke, dass du meine Freundin bist!

Christiane, Uta, Jenny, Ani, Britta, Maxi, Frank; schön, dass es euch gibt. Ihr habt mich immer daran erinnert, dass promovieren nicht alles ist. Dass hat mir oft geholfen meine Problemchen zu relativieren! Danke!

Ruth und **Sarah**, danke für die vielen schoenen Donnerstagnachmittage; dass wir diese besondere Lebensphase teilen konnten; für euer offenes Ohr und eure guten Ratschläge.

Meine lieben Geschwister, **Christian, Nina**, und **Marion**. Dass wir trotz Entfernungen eine so gute Beziehung zueinander haben, ist nicht zuletzt Eurem Einsatz in den letzten Jahren zu verdanken. Dafür bin ich Euch sehr dankbar. Für die schönen Urlaube und Tage in der Friedrichstrasse, die ein gutes Gegengewicht zu meiner Arbeit darstellten.

Meine lieben Omas. **Oma Annemarie**, du hast mir so viel über ein zufriedenes und glückliches Leben gezeigt. Dein Vorbild trage ich in mir. **Oma Kläre**, dass ich immer bei dir willkommen war; dass ich bei dir immer eine Auszeit nehmen konnte von einem lauten Zuhause und wir beide uns so gut verstanden haben. Für die Freitagabende und – nächte und Samstagvormittage, und dass du mich immer unterstützt hast.

Mama und **Papa**, ohne Euch und Eurem festen Glauben an mich und meine Fähigkeiten hätte ich weder das Studium, noch den Doktor gewagt. Danke, dass es euch egal war, was ich mache, solange ich etwas mache, was mir Freude bereitet. Dass Ihr mich mit der Fähigkeit ausgestattet habt, meinen Weg zu erkennen und mit dem Selbstvertrauen, diesen auch gegen (äusseren) Widerstände und Widrigkeiten zu gehen. Mama, ohne dich hätte meine Doktorarbeit in den letzten 2,5 Jahren wohl brach gelegen. Danke, dass du den Weg mit Daniel und mir gehst und für Nora eine so liebevolle und empatische Bezugsperson bist.

Daniel, mein Engel. Dir gebührt mein grösster Dank. Nicht nur für die zahlreichen Stunden, die du mir in meinen Krisen mit der Doktorarbeit zur Seite gestanden hast, sondern vor Allem dafür, dass ich bei dir einfach 'sein' kann. Dass du den Mut hast, mich wirklich in meiner Arbeit zu unterstützen und mit mir den geteilten "Familien-Arbeits-Weg" zu wagen. Ich liebe dich. **Nora**, danke, dass du den Weg zu uns gefunden hast, dass du mir in deinem kurzen Leben schon so viele essentielle Dinge beigebracht hast und soviel Freude und Glück verbreitest. **Kleiner Zwerg**, dass du zu uns unterwegs bist, ich freue mich schon sehr auf dich!



Curriculum Vitae

Claudia Menne-Lothmann was born on September 22nd, 1983, in Düren, Germany. In 2003 she graduated from secondary school (Abitur) at the Gymnasium am Wirteltor, Düren. After spending 8 months in Canada studying English, she started studying Psychology at Maastricht University in 2004. As part of her bachelor she took part in the Honors program at Maastricht University and spent one semester at Bond University, Queensland, Australia. She received her Bachelor of Science in Cognitive Psychology in 2007. In the summer of 2007 she conducted a clinical internship as a psychologist for two months at the Landeskliniken Düren, Germany. She then started the Research Master in Cognitive Neuroscience, Neuropsychology, and Psychopathology at Maastricht University. During this time she worked as a research assistant at the Faculty of Psychology, Maastricht University. She furthermore spent three months at the psychiatric clinic of the University Hospital Aachen, Germany, to conduct a clinical internship and research about empathy in depression. The research for her master-thesis entitled "Effects of interpretation bias modification training on emotional interpretation, anxiety, and affect in healthy adolescents" took place at Oxford University, United Kingdom, with Dr. Jennifer Lau from January to July 2009. In September 2009 she received her Master of Science (cum laude). From September 2009 until January 2015 she worked on her PhD project at the Department of Psychiatry and Psychology, Maastricht University. In 2014 she started clinical training to become a psychological psychotherapist in Cologne, Germany.



List of publications

Publications

- Lothmann, C.**, Holmes, E.A., Chan, S.W.Y., & Lau, J. (2011). Cognitive bias modification training in adolescents: Effects on interpretation biases and mood. *The Journal of Child Psychology and Psychiatry/Child and Adolescent Mental Health*, 52 (1), 24-32.
- Wichers, M., Simons, C.J.P., Kramer, I.M.A., Hartmann, J.A., **Lothmann, C.**, Myin-Germeys, I., van Bemmelen, A.L., Peeters, F., Delespaul, P., & van Os, J. (2011). Momentary assessment technology as a tool to help patients with depression help themselves. *Acta Psychiatrica Scandinavica*, 124 (4), 262-272.
- Wichers, M., Hartmann, J.A., Kramer, I.M.A., **Lothmann, C.**, Peeters, F., van Bemmelen, L., Myin-Germeys, I., Delespaul, P., van Os, J., Simons, C.J.P. (2011). Translating assessments of the film of daily life into person-tailored feedback interventions in depression. *Acta Psychiatrica Scandinavica*, 123 (5), 402-3.
- Wichers, M., **Lothmann, C.**, Simons, C.J.P., Nicolson, N.A., & Peeters, F. (2012). The dynamic interplay between negative and positive emotions in daily life predicts response to treatment in depression: A momentary assessment study. *British Journal of Clinical Psychology*, 51(2): 206-22.
- Menne-Lothmann, C.**, Jacobs, N., Derom, C., Thiery, E., van Os, J., & Wichers, M. (2012). Genetic and environmental causes of individual differences in daily life positive affect and reward experience and its overlap with stress-sensitivity. *Behavior Genetics*, 42, 778-786.
- Jacobs, N., **Menne-Lothmann, C.**, Derom, C., Thiery, E., van Os, J., Wichers, M. (2013). Deconstructing the familiarity of variability in momentary negative and positive affect. *Acta Psychiatrica Scandinavica*, 127, 318-327.
- Rutten, B., Hammels, C. *, Geschwind, N.* **Menne-Lothmann, C.***, Pishva, E., Schruers, K., van den Hove, D., Kenis, G, van Os, J., Wichers, M. (2013). Resilience in mental health: linking psychological and neurobiological perspectives. *Acta Psychiatrica Scandinavica*, 1-18.
- Höhn, P., **Menne-Lothmann, C.**, Peeters, F., Nicolson, N. A., Jacobs, N., Derom, C., Thiery, E., van Os, J., Wichers, M. (2013). Moment-to-moment transfer of positive emotions in daily life predicts future course of depression in both general population and patient samples. *PLoS ONE*, 8(9): e75655.
- Menne-Lothmann, C.**, Viechtbauer, Höhn, P., Kasanova, Z., Drukker, M., van Os, J., Wichers, M., & Lau, J. (2014). How to boost positive interpretations? A meta-analysis of the effectiveness of cognitive bias modification for interpretation. *PLoS ONE*, 9(6): e100925.
- Kramer, I.* **Simons, C.J.P.***, De Wild-Hartmann, J.A., **Menne-Lothmann, C.**, Viechtbauer, W., Peeters, F., Schruers, K., Van Bemmelen, A.L., Myin-Germeys, I., Delespaul, Ph., Van Os, J., Wichers, M. (2014). A therapeutic application of the experience sampling method in the treatment of depression: A randomized controlled trial. *World Psychiatry*, 13 (1), 68-77.
- Bakker, J.M., Lieveise, R., **Menne-Lothmann, C.**, Viechtbauer, W., Pishva, E., Kenis, G., Geschwind, N., Peeters, F., van Os, J., Wichers, M. (2014). Therapygenetics in mindfulness-based cognitive therapy: do genes have an impact on therapy-induced change in real-life positive affective experiences? *Translational Psychiatry*, 4:e384.

Submitted

- Hartmann, J.A., Wichers, M., **Menne-Lothmann, C.**, Kramer, I., Viechtbauer, W., Peeters, F., Schruers, K., van Bemmelen, L., Myin-Germeys, I., Delespaul, P., van Os, J., Simons, C. Personalized feedback on positive affect in daily life: A randomized controlled trial using the Experience Sampling Method as intervention in depression. Submitted.
- Simons, J.P.C., Hartmann, J.A., Kramer, I., **Menne-Lothmann, C.**, Höhn, P., van Bemmelen, A.L., Myin-Germeys, I., Delespaul, P., van Os, J., Wichers, M. Effects of self-monitoring of mental states in daily life on empowerment in patients with depression: a randomized controlled trial.
- Lataster, T., Viechtbauer, W., Oorschot, M., Collip, D., Janssens, M., van Nierop, M., **Menne-Lothmann, C.**, Lardinois, M., Lataster, J., Jacobs, N., Thewissen, V., van der Steen, Y., Wagner, M., Gimpel-Drees, J., Geschwind, N., Simons, C., Nicolson, N., Delespaul, P., Wichers, M*, Myin-Germeys, I*. Evidence for a two-factor positive and negative affect structure in daily life: presenting the Maastricht Momentary Mood Questionnaire (3MQ).
- Menne-Lothmann, C.**, Viechtbauer, W., Bakker, J., Kenis, G., Jacobs, N., Geschwind, N., Collip, D., Pishva, E., Derom, C., Thiery, E., van Os, J., Wichers, M. A modest impact of genes on real-world reward experience: a special role for the OPRM1 gene.
- Menne-Lothmann, C.***, Decoster, J*, van Winkel, R#, Collip, D.#, van Os, J., De Hert, M., Derom, C., Thiery, E., Jacobs, N. & Wichers, M. Psychological and biological validation of a novel digital social peer evaluation experiment (digi-SPEE).

Presentations

Poster presentation at the 18th World Congress on Psychiatric Genetics in Athens, Greece, October 2010. The heritability of resilience-phenotypes and their association with depression vulnerability.

Invited lecture in the seminar series 'New trends in human genetics' at the Institute of Human Genetics, Bonn University, Germany, November 2010. Topic: Gene-environment interaction and momentary assessment of reactive phenotypes in psychiatry.

Poster presentation at the 2nd World Congress on Positive Psychology in Philadelphia, the United States of America. The heritability of momentary positive emotions and the association with stress-sensitivity in daily life.

Oral presentation at the 2nd World Congress on Positive Psychology in Philadelphia, the United States of America. Gene-environment interaction in daily-life happiness.

Oral presentation at the Emotions Congres in Tilburg, the Netherlands, October 2011. Gene-environment interaction in daily life positive affect.

Oral presentation at the 40th Voorjaarscongres der Nederlandse Vereniging voor Psychiatrie (NVvP) in Maastricht, the Netherlands, April 2012. De erfelijkheid van dagelijks positief affect en het verband et stress-gevoeligheid.

Oral presentation at the 41st Voorjaarscongres der Nederlandse Vereniging voor Psychiatrie (NVvP) in Maastricht, the Netherlands, April 2013. Spelen genen een rol bij beloningservaringen in het dagelijks leven?

Oral presentation at the 2nd conference of the Society of Ambulatory Assessment in Amsterdam, the Netherlands, June 2013. The role of genes in daily life reward experience.

Oral presentation at the Najaarscongres der Vereniging voor Gedragstherapie en Cognitieve Therapie in Veldhoven, the Netherlands, November 2013. Zijn trainingen voor het versterken van positieve interpretaties effectief? Een meta-analyse.

Oral presentation at the 42nd Voorjaarscongres der Nederlandse Vereniging voor Psychiatrie (NVvP) in Maastricht, the Netherlands, April 2014. Risico en weerbaarheid: de rol van jeugdtrauma, geboortegewicht, en opvoeding in sociale stress reactiviteit.