# HAPPY FACES AND OTHER REWARDS

Different perspectives on a bias away from positive and toward negative information as an underlying mechanism of depression

CHARLOTTE VRIJEN

# HAPPY FACES
## AND OTHER REWARDS

Different perspectives on a bias away from positive and toward negative information as an underlying mechanism of depression

CHARLOTTE VRIJEN

The research reported in this dissertation (with the exception of Chapters 3 & 4) is part of the No Fun No Glory project. This project was funded by a Vici grant (016.001/002) from the Netherlands Organization for Scientific Research, which was awarded to Professor A.J. Oldehinkel. We are grateful to everyone who participated in our research and to the educational institutes that facilitated the recruitment of participants.

The research presented in Chapters 3 & 4 of this dissertation is based on the TRacking Adolescents' Individual Lives Survey (TRAILS). Participating centers of TRAILS include various departments of the University Medical Center and University of Groningen, the University of Utrecht, the Radboud Medical Center Nijmegen, and the Parnassia Bavo group, all in the Netherlands. TRAILS has been financially supported by various grants from the Netherlands Organization for Scientific Research (NWO), ZonMW, GB-MaGW, the Dutch Ministry of Justice, the European Science Foundation, the European Research Council, BBMRI-NL, and the participating universities. We are grateful to everyone who participated in this research or worked on this project to make it possible.



| | |
|---|---|
| Lay-out & cover design: | www.designyourthesis.com |
| Printing: | Ridderprint BV, Ridderkerk |
| ISBN (print): | 978-94-034-1456-0 |
| ISBN (digital): | 978-94-034-1455-3 |

rijksuniversiteit
groningen

# Happy faces and other rewards

## Different perspectives on a bias away from positive and toward negative information as an underlying mechanism of depression

**PhD thesis**

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus prof. E. Sterken
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Monday 13 May 2019 at 16.15 hours

by

**Charlotte Vrijen**

born on 11 March 1978
in Rolde

**Supervisors**

Prof. A.J. Oldehinkel

Dr. C.A. Hartman

Prof. P. de Jonge

**Assessment Committee**

Prof. P. Cuijpers

Prof. M.C. Wichers

Prof. C. MacLeod

**Paranimfen**

Barbara Michiels

Eeske van Roekel

## TABLE OF CONTENTS


| | | |
|---|---|---|
| **Chapter 1** | General introduction | 9 |
| **Chapter 2** | Lower sensitivity to happy and angry facial emotions in young adults with psychiatric problems | 31 |
| **Chapter 3** | Slow identification of facial happiness in early adolescence predicts onset of depression during 8 years of follow-up | 59 |
| **Chapter 4** | Reward-related attentional bias at age 16 predicts onset of depression during nine years of follow-up | 85 |
| **Chapter 5** | Spread the joy: How high and low bias for happy facial emotions translate into different daily life affect dynamics | 119 |
| **Chapter 6** | An exploratory randomized controlled trial of personalized lifestyle advice and tandem skydives as a means to reduce anhedonia | 155 |
| **Chapter 7** | Alpha-amylase reactivity and recovery patterns in anhedonic young adults performing a tandem skydive | 193 |
| **Chapter 8** | Measuring BDNF in saliva using commercial ELISA: Results from a small pilot study | 213 |
| **Chapter 9** | General discussion | 227 |
| | References | 251 |
| | Nederlandse samenvatting (Dutch summary) | 279 |
| | Dankwoord (Acknowledgements) | 289 |
| | List of recent SHARE publications | 293 |
| | Short biography and list of publications | 297 |

# CHAPTER 1

General introduction

## INTRODUCTION

The aim of this dissertation was to achieve a better understanding of unipolar depression[1] in adolescents and young adults. Depression is a concept defined in the DSM (American Psychiatric Association, 2013) as a combination of five symptoms from a larger list of nine symptoms. Two symptoms, namely sad or depressed mood and anhedonia, have been defined as the core symptoms of depression, and to receive a depression diagnosis one has to suffer from at least one of these two core symptoms (American Psychiatric Association, 2013). Anhedonia refers to loss of interest or pleasure in things or activities one used to enjoy before, and is characterized by low levels of pleasure and a lack of motivation to actively pursue potentially rewarding activities (American Psychiatric Association, 2013; Treadway & Zald, 2011).

Adolescent depression constitutes a major mental health problem. The incidence of depression increases during adolescence and at the end of adolescence the estimated cumulative incidence is 16-28% in community based samples (Lewinsohn, Clarke, Seeley, & Rohde, 1994; Ormel et al., 2015). For adolescents aged 10-24, depression is the first leading cause of disability as measured in disability-adjusted life-years (DALYS) (Gore et al., 2011). Adolescence is a period in which it becomes increasingly important to build friendships and other social networks outside the family (Collins & Laursen, 2004), and a period in which important decisions are made about future schools or careers. Therefore, in adolescence the effects of symptoms of depression may be particularly detrimental. Adolescent depression has been associated with negative psychosocial consequences (Lewinsohn, Rohde, & Seeley, 1998), suicide risk (Brent et al., 1993), and recurrence of depression in adulthood (Pine, Cohen, Cohen, & Brook, 1999; Wilcox & Anthony, 2004). Depression in adolescence is often not recognized (Lewinsohn et al., 1998; Zuckerbrot & Jensen, 2006). It is therefore important to uncover mechanisms underlying adolescent depression and susceptibility to adolescent depression, as these may ultimately inform strategies for prevention and early treatment.

Evidence is mounting that healthy individuals show a bias toward positive or rewarding stimuli (Leppänen & Hietanen, 2004; Pool, Brosch, Delplanque, & Sander, 2016), and it has been suggested that a relative bias away from positive and toward negative information has a central and causal role in the development of depression (Clark, Chamberlain, & Sahakian, 2009; Roiser, Elliott, & Sahakian, 2012). There is evidence that this so called low positive bias is associated with current depression (Disner, Beevers, Haigh, & Beck, 2011; Roiser et al., 2012) and can be modified by pharmacological treatments (Geschwind et al., 2011; Harmer, Goodwin, & Cowen, 2009) and cognitive therapy (Roiser et al., 2012). Additionally, there is preliminary evidence that a low positive bias is already present prior to the actual manifestation of depression (Clark et al., 2009; Mathews & MacLeod, 2005; Roiser et al., 2012). Thus, this bias may play a central role in the development and maintenance of depression as well as in treatment and prevention.

1 Henceforth, depression without further specification refers to unipolar depression.

In this dissertation I studied a low positive bias as an underlying mechanism of depression. I explored this bias with different instruments and from different perspectives. Facial emotion identification tasks and a reward task were used to assess a relative bias toward positive stimuli (happy faces, reward and non-punishment) and away from negative stimuli (negative facial emotions, non-reward and punishment). Relatively fast identification of other people's happy facial emotions compared to their negative facial emotions (high happy bias) and relatively great attention to reward (high reward responsiveness) both reflect a bias toward positive and away from negative information. Low happy bias and low reward responsiveness have independently been associated with depression, but because of small sample sizes and heterogeneous patient groups, particularly the results for low happy bias have remained inconclusive so far. I studied low positive bias in the perspective of current depression, vulnerability for depression, and treatment of depression.

Below, I will first discuss previous studies that support the idea that a low positive bias plays an important and causal role in the development and maintenance of depression as well as in the treatment of depression, and offer a theoretical framework. I will continue with a more detailed discussion of the parts of the framework studied in this dissertation. These are, first, the role of positive bias, that is, happy bias and reward responsiveness, in the development and maintenance of depression, followed by the effect of one specific treatment, that is, behavioral activation, on depressive symptoms, pleasure and positive bias.

## A BIAS AWAY FROM POSITIVE AND TOWARD NEGATIVE INFORMATION IN DEPRESSION: AVAILABLE EVIDENCE AND THEORETICAL MODEL

The idea that depression is characterized by negative biases in virtually every domain of information processing is part of Beck's cognitive model (Beck, 1967b, 1967a), which has been dominating research for decades and has remained influential ever since. New behavioral tasks have been developed and new methods, for example, from the field of neuroscience, have become available to examine different parts of the cognitive model. Recently, there have been attempts to update Beck's original cognitive model with neuroscientific evidence (Clark et al., 2009; Disner et al., 2011; Roiser et al., 2012). Information processing biases have been assessed on different levels (attention, perception, memory, etc.), with different instruments, for example, facial emotion identification tasks and tasks measuring responsiveness to gain (reward) and loss (punishment), and behavioral tasks in combination with neuroscience (Disner et al., 2011; Roiser et al., 2012). Overall, there is evidence that currently depressed patients do not show the bias toward positive or rewarding stimuli that characterizes healthy individuals (Leppänen & Hietanen, 2004; Pool et al., 2016), but show information processing biases toward negative and away from positive information (Bourke, Douglas, & Porter, 2010; Disner et al., 2011; Roiser et al.,

2012). Whereas healthy individuals require greater cognitive effort to disengage from positive information, depressed individuals require greater effort to disengage from negative information (Disner et al., 2011). Treatments have been found to target information processing biases. It has been reported that antidepressant medication modifies affective biases in healthy (Harmer et al., 2009; Harmer, Shelley, Cowen, & Goodwin, 2004) as well as in depressed individuals (Fu et al., 2007; Harmer et al., 2009), and that only after this bias modification depressive symptoms gradually start to diminish in depressed patients (Harmer et al., 2009, 2004; Roiser et al., 2012). The notion that early modification of affective biases precedes improvements in mood has been proposed as a possible explanation as to why it takes relatively long for antidepressants to diminish depressive symptoms (Harmer et al., 2009). There is also first evidence that adolescents *at risk* for depression show a bias toward negative stimuli and do not show the bias toward positive or rewarding stimuli which characterizes healthy individuals (Joormann, Talbot, & Gotlib, 2007; Lopez-Duran, Kuhlman, George, & Kovacs, 2013; Luking, Pagliaccio, Luby, & Barch, 2016). This suggests that a biased processing of positive and negative information may reflect a vulnerability for depression.

I have integrated the cognitive model of depression (Beck, 1967b) with the recent behavioral and neuropsychological evidence and preliminary evidence of the central and causal role of a low positive bias in depression (Beck & Haigh, 2014; Clark et al., 2009; Disner et al., 2011; Roiser et al., 2012) in a theoretical framework. This framework is used as the starting point of this dissertation and the basis of my hypotheses. The model that forms the backbone of this theoretical framework (see Figure 1) is based on two largely similar models which were presented by Disner and colleagues (2011) and Roiser and colleagues (2012) to integrate the cognitive model of depression with neuropsychological evidence.

The general idea behind the model is that genetic vulnerabilities together with environmental triggers, and in combination with a biological sensitivity to stress, may lead to information processing biases. Subsequently, these biases may cause depressive behavior and symptoms. Thus the information processing biases precede the depressive symptoms and have a central and causal role in their development. In turn, the depressive symptoms and behavior strengthen the information processing biases and in this way depressed individuals end up in a self-maintaining vicious cycle (Beck & Haigh, 2014; Disner et al., 2011; Roiser et al., 2012). Empirical studies have found at least some support for most associations depicted in the model, but were mostly uninformative about the direction of the associations. There are, for example, only preliminary indications that information processing biases precede onset of depression.

On a more detailed level of low positive bias, two central and interacting subsystems in the development and maintenance of depression are the automatic information processing system and the reflective, or voluntary (i.e., controlled), information processing system (Beck & Haigh, 2014; Roiser et al., 2012). The automatic system has the function of rapid processing of stimuli that may signal personal threat, gain, or loss. The more reflective, voluntary, system is slower but

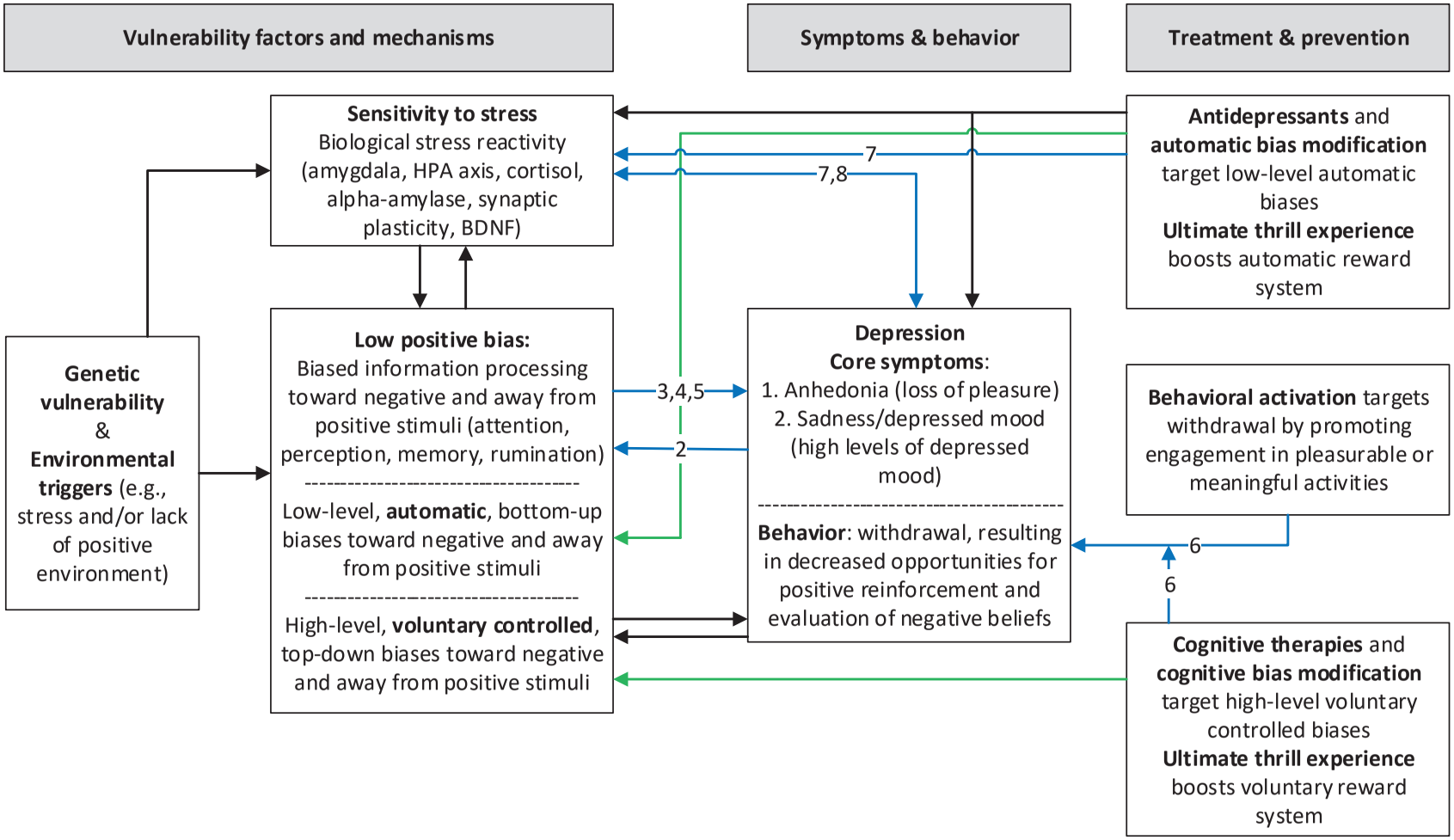



**FIGURE 1.** Theoretical model underlying this dissertation, based on the available evidence and preliminary evidence at the start of my dissertation. The blue arrows reflect associations that were studied as part of the papers presented in this dissertation; the numbers indicate the corresponding chapters. In Chapters 2, 5 and 7 associations were investigated without the specific direction because no data were available on the direction of the associations. The green arrows represent associations that were not tested in the papers but have been (partly) explored in additional analyses for the purpose of this dissertation (see addendum to Chapter 6).

also more accurate. Both can be biased and can mutually influence each other. An automatic bias toward negative and away from positive information may exert its influence in a bottom-up manner by processing disproportionally more negative than positive information and this may in the end contribute to the development of dysfunctional negative cognitions and beliefs, which, in turn, reinforce the biases. When the voluntary system is not negatively biased it may be able to correct automatic biases. On the other hand, negatively biased voluntary systems, that is, dysfunctional negative cognitive biases, are considered to bias information processing in a top-down manner. Negatively biased information processing ultimately results in withdrawal behavior and decreased opportunities for positive reinforcement, resulting in even more withdrawal and stronger negative biases.

Different treatments for depression target different mechanisms to break the self-maintaining vicious depressive cycle (Roiser et al., 2012). Antidepressants and cognitive therapy differentially target the information processing biases; antidepressant medicines primarily modify automatic information processing biases and cognitive therapies primarily modify the more voluntary and higher level biases (Roiser et al., 2012). Behavioral activation is aimed at increasing pleasurable activities. As such this treatment method does not directly target information processing biases, but broadens people's behavioral repertoires, thereby increasing opportunities for positive reinforcement (Jacobson, Martell, & Dimidjian, 2001). Behavioral activation may have an indirect effect on information processing biases, but, to my best knowledge, no empirical studies have investigated whether behavioral activation in the end also results in bias modification. It should be noted that bias modification is most likely not the only pathway from treatment to improvement of depressive symptoms. There is also evidence that several antidepressants target a biological sensitivity to stress and mood (Millan, 2006), but these other pathways were not investigated as part of this PhD project, nor were cognitive therapy and antidepressants.

Although, for the sake of simplicity, a bias toward negative and away from positive information is presented as a single mechanism in the model, there is evidence that a bias toward negative information and a bias away from positive information are independently associated with depression, through partly different mechanisms. These systems, the so-called avoidance system and the approach system (Carver, 2006; Ernst & Fudge, 2009), may be associated with different depressive symptoms. Depressed individuals who suffer from the core symptom of anhedonia but not from the core symptom of sadness are expected to show primarily a bias away from positive information (i.e., happy faces and rewards), whereas those who only suffer from sadness but not anhedonia are expected to show primarily a bias toward negative information (Forbes & Dahl, 2005; Luking, Pagliaccio, Luby, & Barch, 2015). In individuals with both biases these may mutually influence one another, that is, a strong focus on negative information probably results in ignoring positive information and *vice versa* (Beck & Haigh, 2014; Disner et al., 2011).

## HAPPY BIAS AND REWARD BIAS AS PREDICTORS OF DEPRESSION

The two measures of a bias away from positive and toward negative information on which I focus in my dissertation are happy bias and responsiveness to reward. Although findings from these tasks were already discussed above in the general context of a bias away from positive and toward negative information, I will now briefly describe the specific evidence available from previous studies for each of these measures, as well as the gaps in the literature.

### Happy bias as a bias toward positive and away from negative information

Happy bias is measured by means of facial emotion identification tasks and is often operationalized as a higher accuracy or speed in identifying other people's happy facial emotions or a lower accuracy or identification speed for negative facial emotions. Several studies found that depressed individuals experienced more difficulties in identifying happy emotions than non-depressed individuals (Joormann & Gotlib, 2006; Surguladze et al., 2004), others that depressed individuals performed better in identifying sad emotions (Gotlib, Krasnoperova, Yue, & Joormann, 2004; Leppänen, 2006). Results of a systematic review suggested that depressed individuals interpret happy, neutral or emotionally ambiguous facial expressions as more sad or less happy than non-depressed individuals, and show biased attention towards sad facial emotion and away from happy facial emotions (Bourke et al., 2010). Another study found evidence that increased recognition of fear reflects a vulnerability to depression (Bhagwagar, Cowen, Goodwin, & Harmer, 2004). Furthermore, two meta-studies found evidence for a general low ability of discerning between different facial emotions in depressed patients, rather than emotion-specific deficits (Dalili, Penton-Voak, Harmer, & Munafò, 2014; Kohler, Hoffman, Eastman, Healey, & Moberg, 2011). Finally, several studies did not find any facial emotion identification and processing impairments in depressed individuals (Archer, Hay, & Young, 1992; Gaebel & Wölwer, 1992). The inconsistencies in outcomes from different studies may partly be explained by small sample sizes (Bediou et al., 2012; Bourke et al., 2010; Dalili et al., 2014) and the heterogeneity of depressed samples (Bourke et al., 2010; Kohler et al., 2011).

In comparison with the large number of studies devoted to investigating emotion processing in currently depressed individuals, only few studies explored facial emotion processing as a potential vulnerability marker for depression. One study found that boys (but not girls) with a familiar risk for depression were able to identify more subtle sad emotions than low-risk boys (Lopez-Duran et al., 2013). Another that, after a negative mood induction, adolescent girls at risk for depression showed an attentional bias toward negative (versus neutral) facial expressions compared to low-risk girls, and did not show the attentional bias toward positive (versus neutral) facial expressions that was characteristic for low-risk girls (Joormann et al., 2007).

Most of the previous studies tested facial emotion identification separately for the different facial emotions, most commonly happiness, sadness, anger and fear, but did not test relative happy bias, that is, how accurate or fast people are in identifying happy facial emotions compared to negative (or neutral) facial emotions. This latter relative, within-subject happy bias is the operationalization that most closely resembles a bias toward positive and away from negative stimuli, which has been proposed to have a central and causal role in the onset and maintenance of depression (Roiser et al., 2012). There is evidence that healthy individuals show a relative happy bias, that is, they identified other people's happy facial emotions faster than their negative facial emotions (Leppänen & Hietanen, 2004), but only one study attempted to test whether relative happy bias was associated with depression (Wright et al., 2009). In this study a difference score was used to operationalize happy bias (mean reaction time for correctly identified sad, fearful and angry facial emotions minus mean reaction time for correctly identified happy facial emotions). Wright and colleagues (2009) reported that females (but not males) with MDD were slower in identifying negative facial emotions (sadness, fear and anger) relative to positive ones (happiness) than healthy controls. This effect is the opposite of what would be expected if healthy individuals show a relative bias toward happy faces and depressed individuals a relative bias toward negative faces. However, because only the aggregated difference score was reported, and the reaction times (RTs) for happy, sad, angry, and fearful faces were not reported separately, it is unclear whether this finding reflects a real difference in relative happy bias. Because depressed individuals are often slower in reaction time tasks, the effect reported by Wright and colleagues could also be driven by depressed women being slower in identifying all four facial emotions, thus both the positive and the negative ones, than control women[2]. The evidence thus remains inconclusive. From Chapter 3 onward I used a relative, within-subject operationalization of happy bias, representing how fast people could identify other people's happy facial emotions compared to how fast they identified negative facial emotions (e.g., sadness, anger, fear).

## Reward bias as a bias toward positive and away from negative information

There is compelling evidence that depressed adolescents show a lower reward responsiveness than their non-depressed peers (Forbes et al., 2006, 2009; Forbes & Dahl, 2012; McCabe & Gotlib, 1995; Pizzagalli, Iosifescu, Hallett, Ratner, & Fava, 2008). This is reflected in decreased activity

2 For example, if healthy women have an average RT of 1000 ms for negative faces and an average RT of 500 ms for happy faces, and depressed women have an average RT of 1500 ms for negative faces and an average RT of 750 ms for happy faces, their respective difference scores would be 500 ms (healthy) and 750 ms (depressed). Thus, if difference scores are used this would result in a higher score for depressed than for control women, as was reported in the paper by Wright and colleagues, even though the proportional bias score would be exactly the same for both groups (1000/500 for healthy and 1500/750 for depressed women).

in reward-related brain areas (Forbes et al., 2006, 2009), decreased reward learning (Pizzagalli et al., 2008), and the absence of the protective attentional bias toward positive stimuli that characterizes healthy individuals (McCabe & Gotlib, 1995).

Several studies explored low reward responsiveness as a potential vulnerability marker for depression. Adolescents with a high familial risk for depression showed a lower reward responsiveness than their low risk peers, as reflected in decreased activity in reward-related brain areas (Luking et al., 2016). There are also first indications that low reward responsiveness predicts both an increase in future depressive symptoms (Forbes, Shaw, & Dahl, 2007; Morgan, Olino, McMakin, Ryan, & Forbes, 2013; Nelson, Perlman, Klein, Kotov, & Hajcak, 2016; Rawal, Collishaw, Thapar, & Rice, 2013; Telzer, Fuligni, Lieberman, & Galván, 2014) and first onset of depressive disorder in adolescents (Bress, Foti, Kotov, Klein, & Hajcak, 2013; Forbes et al., 2007; Nelson et al., 2016; Pan et al., 2017; Rawal et al., 2013; Stringaris et al., 2015). Particularly the prospective association between reward responsiveness with onset of depressive disorder is important to investigate further because of its high clinical relevance and possible implications for prevention. The evidence of such a prospective association is based is based on only few individuals who were healthy at baseline and became depressed during follow-up, ranging from 3 (Forbes et al., 2007) to 44 (Pan et al., 2017).

Unlike the facial emotion identification studies, reward studies commonly use tasks that are specifically designed to assess relative reward bias, that is, how one responds to rewards compared to how one responds to non-rewards. The non-rewards can be either neutral or negative and several tasks allow testing gains (reward) and losses (punishment) separately, for example, monetary incentive delay tasks (Gotlib et al., 2010; Knutson, Bhanji, Cooney, Atlas, & Gotlib, 2008). It has been suggested that not only reward processing is altered in individuals suffering from depression, but punishment processing (e.g., responses to stress, punishment, or loss) is affected as well. There is evidence that depressed individuals may be more sensitive to negative feedback and show impaired functioning following negative feedback (Elliott, Sahakian, Herrod, Robbins, & Paykel, 1997; Elliott et al., 1996; Eshel & Roiser, 2010). It has been proposed that the two may interact (Forbes & Dahl, 2012); chronic as well as acute stress may affect responsiveness to rewards and *vice versa*. Acute stress has been found to reduce reward responsiveness in highly stress-reactive (Berghorst, Bogdan, Frank, & Pizzagalli, 2013) and in anhedonic individuals (Bogdan & Pizzagalli, 2006). Reward sensitivity has been found to protect against acute stress, and seems to be particularly important in stress recovery (Corral-Frías, Nadel, Fellous, & Jacobs, 2016).

## Gaps in the literature on positive bias and depression

There are three important gaps in the literature on positive bias and depression: (1) previous studies have largely neglected to investigate prospective associations between a low positive bias and depression; (2) disorder-specificity and symptom-specificity were rarely addressed, and (3) previous studies did not investigate why a low positive bias would make a person more vulnerable to depression.

### *Prospective associations*

In an attempt to identify underlying mechanisms of depression, I was most interested in whether a bias toward negative and away from positive information precedes first onset of depression. If this is the case it may be feasible to screen for low reward responsiveness in early adolescence and enhance reward responsiveness to prevent adolescents to become depressed in the first place. Previous studies have only provided first indications that a low positive bias may precede first onset of depression and hence be a vulnerability marker for depression. Prospective associations between positive bias and depression are described in **Chapter 3** and **Chapter 4** of this dissertation.

### *Symptom-specificity and disorder-specificity*

Depression is a concept defined in the DSM (American Psychiatric Association, 2013) as a combination of a number of symptoms from a larger list of symptoms, which introduces heterogeneity. Two people can receive a diagnosis of Major Depressive Disorder (MDD) while not having more than one overlapping symptom; they do not even have to have the same core symptom according to the DSM, as experiencing at least one of them, that is, either anhedonia or sadness, is a necessary condition for an MDD diagnosis. There is evidence that the etiology of the two core symptoms is partly different (Carver, 2006; Ernst & Fudge, 2009). Heterogeneity among depressed individuals can be caused by differences in symptoms, but also by comorbidity with other psychiatric disorders. Comorbidity between depression and other psychiatric disorders, particularly anxiety disorders, is high (Lamers et al., 2011; Lewinsohn et al., 1998).

Previous studies investigating associations between depression and happy bias often focused on depressed patients without investigating the two core symptoms of depression separately, and without adjusting for or excluding participants with comorbid disorders. Therefore, it is unclear whether the results reported in these studies reflect associations with one of the two core symptoms or both, and whether the results are specific to depression or are driven by co-occurring psychiatric problems (e.g., anxiety) or by more general problem characteristics (e.g., severity). Several studies on reward responsiveness did investigate the two core symptoms separately and found that low reward responsiveness was most convincingly associated with anhedonia (Chase et al., 2010; Luking et al., 2015; Pizzagalli et al., 2008). Similar to studies on happy bias, previous studies investigating associations between depression and

reward responsiveness commonly did not address disorder-specificity and a question remains whether findings are specific for depression or are driven by comorbid psychiatric disorders (Forbes & Dahl, 2012). Symptom-specificity is addressed in **Chapter 3** of this dissertation, and disorder-specificity in **Chapter 2**, **Chapter 3**, and **Chapter 4**.

***Why would a low positive bias make a person more vulnerable to depression?***

Happy bias and reward bias are implicit, positive biases of which people are probably unaware themselves, therefore they are commonly assessed with standardized laboratory tasks. Given the (preliminary) evidence that low happy bias and low reward bias may mark vulnerability to depression, and the accumulating evidence that the smallest building blocks of an individual's adaptive and maladaptive affect patterns are found in daily life affect dynamics (Pe et al., 2015; Trull, Lane, Koval, & Ebner-Priemer, 2015; Wigman et al., 2015), I would expect that these laboratory-based positive biases also reflect differences in daily life affect dynamics. However, to date this has not been investigated and the importance and scope of laboratory-based positive bias in people's daily life has remained an important gap in the literature. It is important to address this gap to facilitate the interpretation of laboratory measures of positive bias, and I expected that an investigation of adaptive and maladaptive affect dynamics associated with positive bias might also provide clues to why a low happy bias is associated with an increased risk for depression.

What types of affect dynamics are adaptive and what types are maladaptive and may ultimately lead to depression? The function of emotions is to prepare and guide action for dealing with important events in our lives (Frijda, 2007). Emotions provide strong motivations for action, either to approach (for example in the case of positive emotions or expected reward) or avoid (for example in the case of fear and anxiety). According to the broadening-and-built theory of Fredrickson positive affect broadens attention whereas negative affect narrows attentional scope (Basso, Schefft, Ris, & Dember, 1996; Fredrickson, 2001); and a broad attentional scope allows a person to approach and explore the world, and so to build valuable resources for self-development, well-being and mental health (Cohn, Fredrickson, Brown, Mikels, & Conway, 2009; Fredrickson, 1998, 2004; Kashdan, Rose, & Fincham, 2004). Evidence from laboratory studies and Ecological Momentary Assessment (EMA) studies suggest that the following types of affect dynamics are adaptive and promote mental health: (1) the ability to sustain positive affect and positive experiences over time (Heininga, Van Roekel, Ahles, Oldehinkel, & Mezulis, 2017; Heller et al., 2009; Höhn et al., 2013; Horner et al., 2014; McMakin, Santiago, & Shirk, 2009); (2) the ability to use positive experiences to generate positive affect and *vice versa* (Geschwind et al., 2010; Wichers et al., 2010); (3) the ability to use positive affect and positive experiences to dampen negative affect, negative thoughts (i.e., rumination), and negative experiences (Fredrickson & Levenson, 1998; Hilt & Pollak, 2013; Tugade & Fredrickson, 2004).

I expected individuals with a laboratory-based high positive bias to show the more adaptive daily life affect dynamics and the ones with a low positive bias the more maladaptive affect dynamics. **Chapter 5** of this dissertation describes the results of a study of the daily life correlates of laboratory measures of happy bias. I investigated whether high and low positive bias are associated with different daily life affect dynamics, and whether these differences can provide first clues to why people with a low positive bias may be more vulnerable to develop depression.

## TREATMENT & PREVENTION

Because of the hypothesized directed associations from information processing biases to depressive symptoms and *vice versa* (as illustrated in Figure 1), it seems plausible that treatments and preventions would, after the initial direct influence on one of the components, ultimately spread through the entire model. For example, for two common treatments for depression, antidepressants and cognitive therapy, there is preliminary empirical evidence that they directly target, respectively, automatic and voluntary information processing biases, and subsequently lead to improvement of depressive symptoms (Harmer et al., 2009, 2004; Roiser et al., 2012). In this dissertation another common treatment for depression is described, behavioral activation. In this treatment, depressed patients are instructed to engage in pleasurable or meaningful activities to provide opportunities for positive reinforcement (Jacobson et al., 2001; Lewinsohn, Sullivan, & Grosscup, 1980). Behavioral activation is often used as an important component of cognitive therapy (cognitive behavioral therapy; CBT), but in this dissertation I focus specifically on the behavioral activation component. Behavioral activation attempts to target avoidance, withdrawal, and inactivity, and help depressed patients to engage in positive, reinforcing experiences (Jacobson et al., 2001). The effects of information processing biases on depressive symptoms and *vice versa* (depicted in Figure 1) suggest that behavioral activation may in the end, through improvement of depressive behavior and symptoms, result in a modification of information processing biases. However, this has not been investigated and thus there is no empirical evidence available to corroborate or contradict this.

There is evidence that behavioral activation can successfully reduce depressive symptoms and that the effects of behavioral activation on depression in adults are comparable to the effects of antidepressants and cognitive therapy (Cuijpers, van Straten, & Warmerdam, 2007; Ekers, Richards, & Gilbody, 2008; Jacobson et al., 1996; Mazzucchelli Trevor, Kane Robert, & Rees Clare, 2009; Shinohara et al., 2013). Note that this is in accordance with the hypothesized model (Figure 1), because regardless of the component that is targeted initially, in the end all components of the model are expected to be affected. Part of the attractiveness of behavioral activation is the relative simplicity of the procedure; it has been suggested that no complex skills are required from either therapists or patients (Lejuez, Hopko, LePage, Hopko, & McNeil, 2001; Mazzucchelli et al., 2009). For this reason behavioral activation may also be suitable for adolescents and children,

and may not require a fully trained therapist. Behavioral activation may be a promising alternative to antidepressants or more complex therapeutic methods, although it should be noted that similar treatment effects on a group level do not imply that the treatments are interchangeable on the level of an individual.

Behavioral activation may be a particularly promising approach for individuals suffering from anhedonia (Treadway & Zald, 2011). Compared to peers without this symptom, adolescents suffering from anhedonia do not show the natural bias toward information signaling positive or rewarding outcomes that is common to healthy individuals, in other words, they lack a positive bias or show a lower positive bias than their peers. The reason why behavioral activation is a promising approach to anhedonia is that an increase in activities is encouraged without waiting until someone "feels like" engaging in activities. In this way the lack of motivation characteristic to anhedonia is bypassed (Treadway & Zald, 2011). However, difficulties in the motivational dimension may still make it more difficult for anhedonic individuals to be motivated to comply with any type of therapy. Additionally, a remaining problem is that anhedonic symptoms may make it particularly difficult to identify one's own pleasurable or meaningful activities.

Teaching depressed individuals to acquire more insight in associations between daily activities and positive and negative mood has been an important component of behavioral activation programs. In the program developed decades ago by Lewinsohn and colleagues (1980), for example, the most pleasant and most unpleasant activities for a particular patient were identified based on their scores on the Pleasant Events Schedule and the Unpleasant Events Schedule, which assessed the frequency and level of pleasantness during the past month. These most frequent most pleasant and most unpleasant activities were subsequently included in the patient's activity schedule and the patient was instructed to monitor these activities and mood on a daily basis, in order to learn to recognize associations between them. In this way patients did not need to have knowledge of their daily life associations between activities and pleasure beforehand, but learned about them during treatment, which may be particularly important for patients suffering from anhedonia.

## Gaps in the literature on treatment and prevention of depression

### *Behavioral activation and anhedonia*

A first gap in the literature is that previous studies did not investigate whether behavioral activation is feasible and successful for individuals suffering from anhedonia. The behavioral activation approach may bypass the problems these individuals have with identifying pleasurable activities and may help them increase their pleasure and decrease their depressive symptoms. A second gap is that although many studies investigated whether behavioral activation improves depressive symptoms in general (Cuijpers et al., 2007; Ekers et al., 2008; Mazzucchelli Trevor et al., 2009; Shinohara et al., 2013), only few investigated the effects of behavioral activation on increasing pleasure (Jacobson et al., 1996; Zeiss, Lewinsohn, & Muñoz, 1979). They found that

behavioral activation (as well as other therapies) increased pleasure. **Chapter 6** describes the results of a tailored behavioral activation intervention in young adults suffering from anhedonia, including effects on depressive symptoms in general and on pleasure.

***Does behavioral activation modify low positive bias?***

A third gap is the lack of empirical evidence on whether the influence of behavioral activation is limited to a modification of behavior and subsequent improvement of depressive symptoms, or whether in the end information processing biases are modified as well. This is important to establish, since a reduction of the biases may render the individual less vulnerable to developing depressive symptoms again in the future and implies that behavioral activation may be used to prevent first onset of depression. The effects of behavioral activation on positive bias are reported in the addendum of **Chapter 6**.

***Tailored approach based on multiple momentary assessments per day***

A disadvantage of the earlier behavioral activation programs is that the original activity schedule that determined which activities would be monitored daily was still based on patients' own account of how pleasant or unpleasant a wide range of activities had been in the past month. It is also unclear whether monitoring activities and mood once per day is the right time scale to identify the most relevant associations between activities and mood as targets for treatment. Since both activities and mood fluctuate throughout the day, one daily measure may not capture all the relevant information. Furthermore, it may be very difficult to recall all activities and moods during the day when asked to do so once at the end of the day, and there is evidence that recall may be influenced by an individual's emotional and motivational state during perceiving and reporting (Kihlstrom, Eich, Sandbrand, & Tobias, 2000; Shiffman, Stone, & Hufford, 2008). Until recently, behavioral activation programs were limited by practical constraints. It was not feasible to have patients monitor many different activities and mood multiple times per day. However, with the current techniques and statistical methods it is feasible to use electronic questionnaires multiple times per day for a longer period of time to collect data about lifestyle activities, pleasure, and mood, and to analyze these data for each individual separately. This method no longer depends on individuals' own knowledge about their associations between activities and pleasure, because activities and pleasure can, from the start, be reported separately and associations between them can be tested statistically for each individual separately. The method reported in **Chapter 6** makes use of recent technical advancements and consists of a tailored behavioral activation approach based on three momentary assessments of activities, pleasure, and mood per day.

***An ultimate thrill experience as an additional motivational boost***

Because low motivation makes it difficult to carry out behavioral changes, regardless of whether it is a personalized advice, an exploration of new methods that increase motivation to make lifestyle changes could be a fruitful approach to increase the effect of behavioral activation programs. An example of a potentially promising method to increase motivation is a free fall experience, for example, a skydive. Such an extreme thrill experience may boost the reward system and make it easier for participants to carry out lifestyle advice. It is known that skydiving results in strong physiological effects, for example, increases in heart rate, blood pressure, cortisol levels and alpha-amylase levels (Chatterton, Vogelsong, Lu, & Hudgens, 1997; Hare, Wetherell, & Smith, 2013; Meyer et al., 2015), and strong psychological effects, that is, extreme fear before and during the free fall (Hare et al., 2013), followed by euphoria afterwards (Meyer et al., 2015). Additionally, animal research revealed that mice who experienced a free fall showed an increase in dopamine neuron firing in the ventral tegmental area, a brain area associated with reward motivation (Wang & Tsien, 2011). Altogether, this evidence suggests that the experience of a free fall might give a boost to the reward system. **Chapter 6** describes the results of a study aimed to investigate whether a tandem skydive combined with tailored behavioral activation increases pleasure and PA more and decreases depressive symptoms and NA more than tailored behavioral activation without the skydive.

The hypothesis that a skydive will give a boost to the reward system is based on the assumption that a skydive is such an extreme experience that it evokes universal physiological and psychological responses. There is evidence of uniformly extreme responses in non-depressed populations (Chatterton et al., 1997; Hare et al., 2013; Meyer et al., 2015), but it is unclear if these can be generalized to individuals who suffer from anhedonia. Anhedonia has been characterized as blunted responsiveness, either only to rewards or to rewarding as well as to negative contexts (Rottenberg, Gross, & Gotlib, 2005), and a blunted response to a tandem skydive might prevent it from evoking the hypothesized boost in motivation. I investigated whether anhedonic young adults show the expected, universal, stress reactivity and recovery patterns in response to the skydive, and examined individual differences (**Chapter 7**). Stress reactivity and recovery were measured by means of alpha-amylase, an enzyme that is often used as a biomarker for stress (Nater & Rohleder, 2009; Takai et al., 2004). The original plan was not only to assess salivary alpha-amylase during the tandem skydive, but also another biomarker, namely, brain-derived neurotrophic factor (BDNF). BDNF is a protein that regulates neuronal plasticity and has been found to be associated with depression and susceptibility to stress (Autry & Monteggia, 2012), and has been suggested as a biomarker for successful treatment of depression. Because it was unclear whether BDNF could be measured reliably in saliva with the commercially available ELISA kits, this was tested in a small pilot study (**Chapter 8**).

## MAIN AIMS OF THE STUDIES DESCRIBED IN THE EMPIRICAL CHAPTERS

### Chapter 2

To investigate (1) associations between emotion identification (happy, sad, angry and fearful faces) and five psychiatric problem domains, that is, depressive problems, anxiety problems, avoidance problems, ADHD problems and antisocial problems; (2) the domain-specificity of these associations.

### Chapter 3

To investigate whether (1) a lack of bias toward happy facial emotions in early adolescence predicts first onset of depression during 8 years of follow-up; (2) findings are specific for depression or are driven by comorbid anxiety; (3) findings are driven by one of the two core symptoms of depression, that is, anhedonia or sadness.

### Chapter 4

To investigate whether (1) reward-related attentional biases on an automatic and on a voluntary level of information processing predict onset of depression during nine years of follow-up; (2) attention to reward and attention to loss differentially predict onset of depression; (3) findings are specific for depression or are driven by other psychiatric disorders.

### Chapter 5

To investigate whether young adults with a high bias toward happy facial emotions during a laboratory task and those with a low bias toward happy facial emotions show different daily life affect dynamics.

### Chapter 6

To investigate whether (1) a tailored lifestyle advice based on a person's own specific associations between lifestyle factors and pleasure during one month of momentary assessments (3 times per day) helps young adults who suffered from anhedonia to increase their pleasure; (2) a tandem

skydive in addition to the tailored lifestyle advice results in a larger increase in pleasure than the lifestyle advice alone; *(3) these interventions result in increases in positive bias (i.e., higher happy bias and higher reward responsiveness).*[3]

### Chapter 7

To investigate whether in anhedonic young adults (1) a tandem skydive elicits extreme biological responses; (2) there are individual differences in alpha-amylase reactivity to and recovery from a tandem skydive; (3) trait depressive and anxiety problems, trait positive affect (i.e., level of pleasure and reward responsiveness), and state anxiety, positive affect and self-esteem prior to the skydive are associated with alpha-amylase reactivity and recovery patterns; and (4) alpha-amylase reactivity and recovery patterns are associated with pre- to post-jump changes in state anxiety, positive affect, and self-esteem.

### Chapter 8

To explore whether brain-derived neurotrophic factor (BDNF) could be measured reliably in saliva with the commercially available ELISA kits.

## DESCRIPTION OF THE DATASETS

For two chapters in this dissertation (Chapters 3 and 4) data were used from the TRacking Adolescents' Individual Lives (TRAILS) survey, an ongoing cohort study in which individuals have been followed from early adolescence (age 11) and were assessed on their emotional and behavioral development every two or three years. This cohort study allowed for investigating prospective associations, that is, whether low happy bias and low reward responsiveness at one time point predicted onset of depression at a later time point.

For four chapters (Chapters 2, 5, 6, and 7) data were collected as part of the No Fun No Glory (NFNG) study, which is a biopsychosocial investigation of anhedonia in young adults. The NFNG study consists of a large screening survey among almost 3,000 young adults, followed by an intervention study for which 69 young adults with high levels of persistent anhedonia and 69 matched controls were selected from the screening survey. During the intervention study anhedonic participants completed momentary assessments three times a day for more than three months and controls for one month. Two interventions were offered: a personalized

3 Italics indicate *post hoc* analyses that were not part of the original paper, but were carried out to test additional parts of the model (Fig. 1) presented in this dissertation.

lifestyle advice based on one month of momentary assessments; and this same personalized lifestyle advice combined with a tandem skydive. The screening survey and follow-up surveys also included a facial emotion identification task.

For Chapter 8, data were used from a pilot study conducted by Dr. Maria Schenk and additional data collected by myself.

**TABLE 1.** Characteristics of the Datasets Used in This Dissertation

| Dataset | Design | Type of data used | Sample | Chapter |
|---|---|---|---|---|
| TRAILS | Longitudinal | Facial emotion identification speed at T1, CIDI lifetime unipolar depression diagnosis at T4 | Population and clinical cohort, without onset of depression ≤ T1<br>N=1840 | 3 |
| TRAILS | Longitudinal | Spatial Orienting Task (SOT; attention to reward and punishment) at T3, CIDI lifetime unipolar depression diagnosis at T4, LIDAS lifetime depression self-report at T6 | Selection of population cohort that completed the SOT, without onset of depression ≤ T3<br>N=531 | 4 |
| NFNG | Cross-sectional | Facial emotion identification speed during morph task at T0 and Adult-Self-Report measure of psychiatric problems at T0 | Screening sample<br>N=2577 | 2 |
| NFNG | Cross-sectional | Happy bias based on facial emotion identification speed during morph task at T0 and T2, affect measures during the observation month (30 days of momentary assessments, 3 beeps a day with fixed 6 hour time intervals) | N=25 high happy bias, 90 daily life affect measures per person<br>N=25 low happy bias, 90 daily life affect measures per person | 5 |
| NFNG | Randomized controlled trial | Anhedonia and depression at T0 (screening), T1 (start observation month), T2 (start intervention month), and T3 (end intervention month), two 30 day periods of three momentary assessments per day (observation month & intervention month) | N=69 anhedonic young adults; N=69 matched controls<br>RCT anhedonic participants:<br>N=22 no intervention;<br>N=22 personalized lifestyle advice;<br>N=25 personalized lifestyle advice + tandem skydive | 6 |
| NFNG | Longitudinal | Depression, anxiety, and PA at T0, momentary assessments of PA, anxiety, and self-esteem, alpha-amylase measures from saliva collected 4 times on day of tandem skydive | N=61 | 7 |
| Pilot NFNG & pilot study Maria Schenk | Cross-sectional | BDNF in blood plasma and saliva, collected on 3-5 different days with 1-3 measures per day | N=6, in total 33 blood samples + 33 saliva samples | 8 |

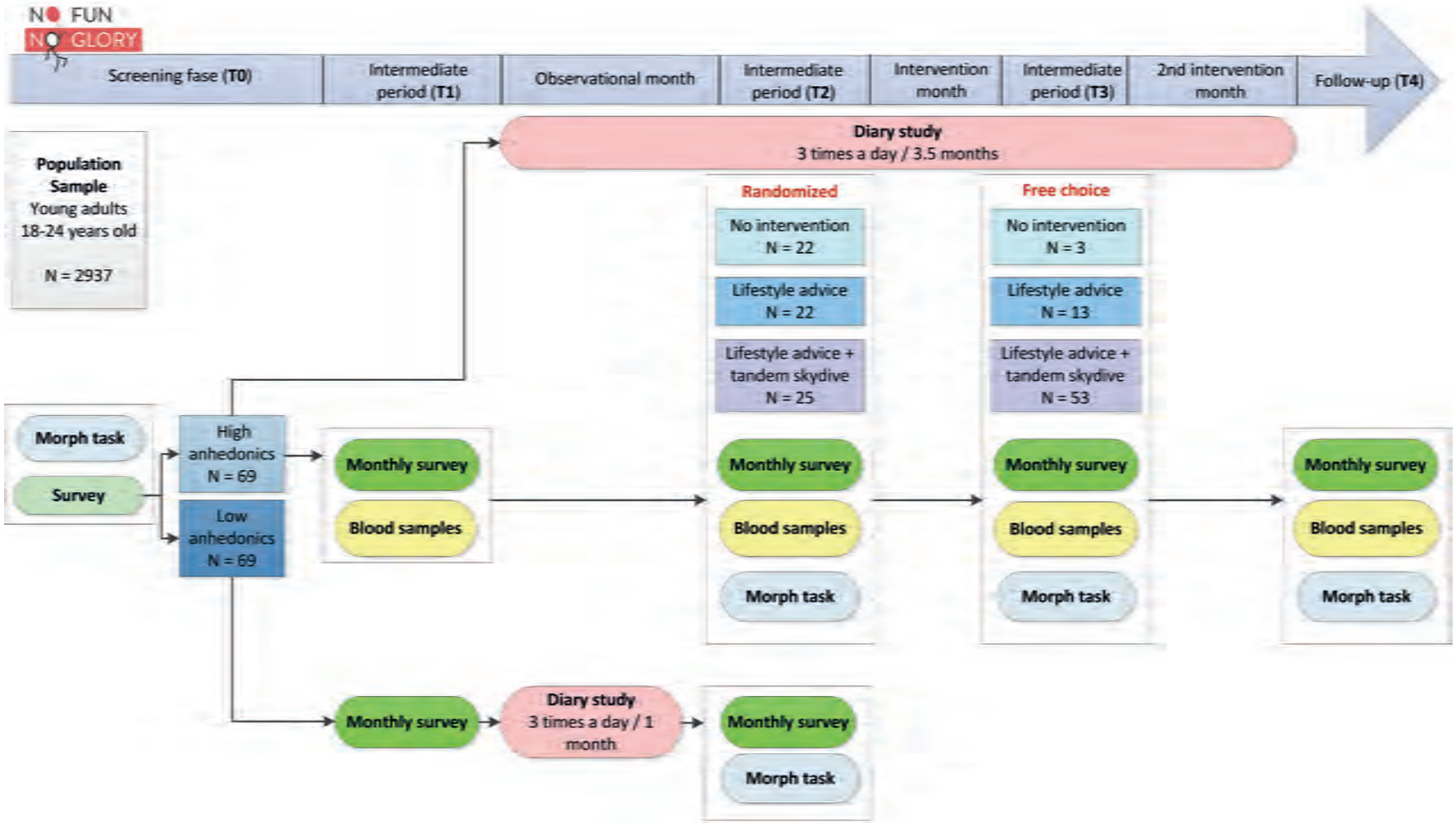



**FIGURE 2.** Flowchart NFNG study

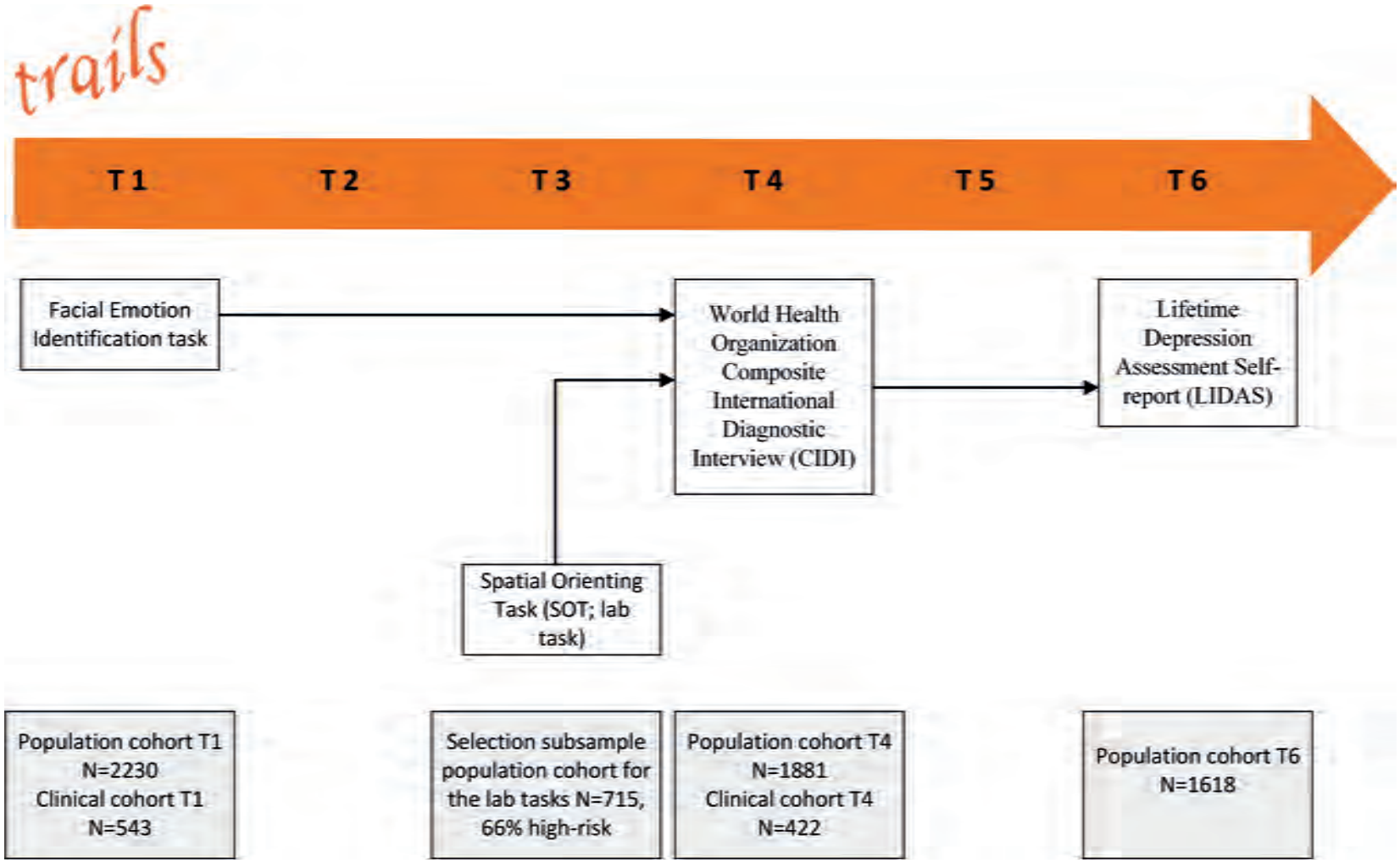



**FIGURE 3.** Flowchart TRAILS study; only measures used in the present dissertation are depicted.

# CHAPTER 2

## Lower sensitivity to happy and angry facial emotions in young adults with psychiatric problems

Published as:
Vrijen, C., Hartman, C. A., Lodder, G. M. A., Verhagen, M., de Jonge, P., & Oldehinkel, A. J. (2016). Lower Sensitivity to Happy and Angry Facial Emotions in Young Adults with Psychiatric Problems. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01797

## ABSTRACT

Many psychiatric problem domains have been associated with emotion-specific biases or general deficiencies in facial emotion identification. However, both within and between psychiatric problem domains, large variability exists in the types of emotion identification problems that were reported. Moreover, since the domain-specificity of the findings was often not addressed, it remains unclear whether patterns found for specific problem domains can be better explained by co-occurrence of other psychiatric problems or by more generic characteristics of psychopathology, for example, problem severity. In this study, we aimed to investigate associations between emotion identification biases and five psychiatric problem domains, and to determine the domain-specificity of these biases. Data were collected as part of the 'No Fun No Glory' study and involved 2,577 young adults. The study participants completed a dynamic facial emotion identification task involving happy, sad, angry, and fearful faces, and filled in the Adult Self-Report Questionnaire, of which we used the scales depressive problems, anxiety problems, avoidance problems, Attention-Deficit Hyperactivity Disorder (ADHD) problems and antisocial problems. Our results suggest that participants with antisocial problems were significantly less sensitive to happy facial emotions, participants with ADHD problems were less sensitive to angry emotions, and participants with avoidance problems were less sensitive to both angry and happy emotions. These effects could not be fully explained by co-occurring psychiatric problems. Whereas this seems to indicate domain-specificity, inspection of the overall pattern of effect sizes regardless of statistical significance reveals generic patterns as well, in that for all psychiatric problem domains the effect sizes for happy and angry emotions were larger than the effect sizes for sad and fearful emotions. As happy and angry emotions are strongly associated with approach and avoidance mechanisms in social interaction, these mechanisms may hold the key to understanding the associations between facial emotion identification and a wide range of psychiatric problems.

## INTRODUCTION

Facial emotion processing is critical for normal emotional development and engagement in social relationships. Social information gained by processing emotional expressions informs people about the attitudes of others and holds cues for behavioral responses (Niedenthal, Halberstadt, Margolin, & Innes-Ker, 2000; Salovey & Mayer, 1990). Therefore, emotion identification is considered to be one of the key elements of successful social interaction. In recent years, many different psychiatric disorders have been associated with emotion-specific biases or general deficiencies in facial emotion identification (Bediou et al., 2012; Kret & Ploeger, 2015). It has been suggested that different problem domains each have their own characteristic condition-specific facial emotion identification biases or deficiencies, which may be useful in early detection and as a target in treatment (Bediou et al., 2012; Isaac, 2012; Penton-Voak, Bate, Lewis, & Munafò, 2012; Penton-Voak et al., 2013; Rinck, 2013). However, both within and between psychiatric problem domains, large variability exists in the types of emotion identification problems that were reported. The heterogeneous results within specific psychiatric problem domains appear to be due, at least partly, to methodological limitations and differences, for example, small sample sizes, the use of different types of facial emotion processing tasks, and diversity in the study populations regarding combinations of symptoms, symptom severity and comorbidity. This limits the comparability of studies within the same problem domain. Furthermore, most studies only focused on emotion identification deficiencies or biases in one problem domain without excluding participants with co-occurring problems, or adjusting for the presence of these co-occurring problems. This means that the specificity of the facial emotion identification patterns found for a psychiatric problem domain was not addressed. It therefore remains unclear whether patterns found for specific problem domains can be better explained by co-occurrence of other psychiatric problems, or by more general characteristics of psychopathology, for example, problem severity.

Among the implicated problem domains are social anxiety, depression, Attention-Deficit Hyperactivity Disorder (ADHD), and antisocial behavior, with tentative evidence for avoidance behavior. These psychiatric problem domains have been associated with an overall problem with identifying emotions as well as with biased identification of emotions, that is, a heightened or lowered ability to identify specific emotions. For depression, meta-analyses indicate evidence for a bias towards sad faces and away from happy faces (Bourke et al., 2010; Joormann & Gotlib, 2006) and, to a lesser extent, for an overall lower facial emotion identification speed (Bourke et al., 2010; Kohler et al., 2011). In individuals with a history of depression relatively rapid fear identification was found as well (Bhagwagar et al., 2004). Regarding anxiety, a recent meta-analysis showed evidence for a small general emotion identification deficiency in people with social phobia and generalized anxiety (Plana, Lavoie, Battaglia, & Achim, 2014), but it should be noted that emotion-specific effects were ignored, and only total accuracy and intensity scores over all emotions were tested. Notably, other studies reported opposite results, in that people

with generalized anxiety tended to perform better at facial emotion identification (Bui et al., 2015). Several studies also found emotion-specific effects: for socially anxious participants a higher sensitivity to angry faces was found (Joormann & Gotlib, 2006), but opposite findings of a lower sensitivity to anger and disgust were also reported (Montagne et al., 2006). ADHD has been associated with overall lower facial emotion identification skills (Rommelse, Geurts, Franke, Buitelaar, & Hartman, 2011; Sinzig, Morsch, & Lehmkuhl, 2008), as well as with more specific problems in identifying sad (Aspan et al., 2014; Pelc, Kornreich, Foisy, & Dan, 2006; Schönenberg, Schneidt, Wiedemann, & Jusyte, 2015), fearful (Aspan et al., 2014; Schönenberg et al., 2015) and angry emotions (Pelc et al., 2006). Antisocial behavior seems to be primarily related to more difficulties with identifying fear, but has also been associated with difficulties in identifying sadness (Blair et al., 2004; Marsh & Blair, 2008) and subtle happy emotions (Kahler et al., 2012). Avoidance behavior has not been thoroughly investigated, but a first preliminary study suggests that people with avoidant personality problems make more errors in classifying fearful emotions (Rosenthal et al., 2011). Thus, previous findings were heterogeneous, both within and between psychiatric problem domains.

Only few studies have addressed domain-specificity to date. Two studies compared depressed participants, socially anxious participants and healthy controls on their facial emotion identification skills (Gotlib, Kasch, et al., 2004; Joormann & Gotlib, 2006), and found that depressed participants were less capable of identifying subtle happy emotions than the other two groups, whereas participants suffering from social phobia were more proficient in identifying subtle angry emotions than the other two groups. To our best knowledge, no studies explicitly addressed the domain-specificity of emotion identification in a wider range of psychiatric domains. Detailed information regarding the domain-specificity of emotion identification is crucial for achieving a better understanding of the mechanisms underlying different psychiatric problem domains, and may ultimately result in the development of more fine-grained diagnostic tools and treatments.

The aim of the first part of the current study was to investigate whether facial emotion identification bias was related to five different psychiatric problem domains, that is, depressive problems, anxiety problems, avoidance problems, ADHD problems and antisocial problems, in a general population sample of young adults. For all of these problem domains there is evidence of an association with facial emotion identification from previous studies, and the occurrence of these problems in a general population of young adults is also quite common, which is why we considered them the most relevant to investigate and expected sufficient power for all analyses. The advantage of testing all associations in one study is that, due to more methodological homogeneity, the findings for the five domains in our study are more comparable than findings from different studies. The aim of the second part of the study was to determine the domain-specificity of the associations. We did not have hypotheses on domain-specificity in advance because of the lack of previous studies addressing this matter. The two parts of the study are complementary. The first part of the study is aimed at providing more insight into the

associations between facial emotion identification and psychiatric problem domains as such, whereas the second part reflects a more mechanistic approach in which unique contributions of single psychiatric problem domains are explored.

We used a so-called 'morph' task in which movie clips were shown of neutral faces which gradually changed into full intensity facial emotions (Joormann & Gotlib, 2006; Lodder, Scholte, Goossens, Engels, & Verhagen, 2015). The task measured at what intensity participants were able to identify the facial emotion. The use of a morph task enabled us to measure identification of more subtle traces of emotions, which is assumed to give a more ecologically valid perspective than the often used static full intensity facial emotion tasks (Joormann & Gotlib, 2006; Niedenthal et al., 2000). In everyday life full intensity facial emotions are rare but we encounter subtle traces of facial emotions all the time. The benefit of the emotion identification morph task we used is that it enabled us to tap into these frequently occurring everyday life social situations which are essential to social functioning. In addition, because the stimuli gradually change from neutral to full intensity emotions, the morph task was considered a more sensitive instrument than full intensity tasks. High task sensitivity was important in light of our participants; they were not patients with severe psychiatric problems and therefore only subtle alterations in emotion identification were to be expected.

## METHODS

### Sample and procedure

This study is based on data collected as part of the 'No Fun No Glory' project, which investigates anhedonia in young adults. The study was approved by the Medical Ethical Committee from the University Medical Center Groningen (no. 2014/508), participants were treated in accordance with the Declaration of Helsinki and indicated their informed consent online prior to enrolment in the study.

We collected the present data as part of an online survey, for which participants in the northern part of the Netherlands were recruited through advertisements on electronic learning environments of university and higher and intermediate vocational education institutes. We also pitched the study during lectures and classes, and invited participants to participate through flyers and advertisements on social media. After subscribing on the study website (www.nofunnoglory.nl), participants received an email with the link to the online survey, containing questionnaires about, for example, pleasure, psychiatric problems and stress. A more detailed description of the questionnaires is available in the 'No Fun No Glory' research protocol (Van Roekel et al., 2016). Upon finishing the final questionnaire, participants were automatically directed to a facial emotion identification task. After completion of the questionnaire and the task, which, in total, took them on average 35 minutes, participants received a gift card of 10 Euro

and participated in a lottery for fashion vouchers, tablets and a 4-day city trip. Most participants completed the online survey and emotion task at their own preferred time and place, but on a few occasions (< 3% of all participants) teachers of intermediate vocational education institutes allowed for the survey and emotion task to be completed in their classroom during regular school hours.

A total of 3,035 participants subscribed to the study website and started the survey. Participants were included in the current study if they had completed both the Adult Self-Report Questionnaire (ASR) and the facial emotion identification task (N = 2,620). The task required installing a plugin and attrition between the questionnaire and the task was mostly due to technical problems regarding the plugin. We excluded 43 participants because of suspiciously high error rates or reaction times on the facial emotion identification task, yielding a sample of 2,577 participants. In the description of the task procedures these exclusion criteria are explained in more detail.

The mean age of the participants was 21.4 years (*SD* 1.9; range 18 - 27 years) and 78% were females. Most participants attended or had attended university (57%) or higher vocational education (31%), followed by intermediate vocational education (10%) and other types of education (2%).

## Measures and procedures

### *Psychiatric problems*

The Adult Self-Report (ASR) was used to assess psychiatric problems. The ASR is a standardized questionnaire of behavioral and emotional problems, which has been shown to have good reliability and validity (Achenbach & Rescorla, 2003). Responses can be summed to form scale scores on six diagnostic domains based on the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV; (American Psychiatric Association, 1994): Depressive problems (14 items), Anxiety problems (7 items), Avoidant personality problems (7 items), Antisocial personality problems (20 items), ADHD problems (13 items) and Somatic problems (9 items). Somatic problems were not included in our study, since there was no theoretical or empirical evidence of the relevance of facial emotion identification for somatic problems. For each problem, answer categories were: 0 = 'Not True'; 1 = 'Somewhat or Sometimes True'; 2 = 'Very True or Often True'. We divided scale scores by the number of items in the scale so that scores from different problem domains were on the same metric and could be compared easily. In addition to these domain-specific scale scores, a total problems score was calculated for each individual by averaging the mean scores of the five problem domains. In our sample, Cronbach's alpha's were .83 for the Depressive problems scale, .74 for Anxiety problems, .79 for Avoidance problems, .80 for ADHD problems, .66 for Antisocial problems and .91 for Total problems.

### *Identification of facial emotions*

A morph task developed at Radboud University Nijmegen, the Netherlands (Lodder et al., 2015), was used to assess the emotional intensity of a facial expression required for participants to identify the expressed emotion. In the version we used, stimuli consisted of 24 movie clips that lasted 10 s and contained 100 frames depicting the gradual change (i.e., 'morph') from a neutral facial expression to one of four full intensity emotional expressions: happiness, sadness, anger or fear (see Figure 1 for examples).

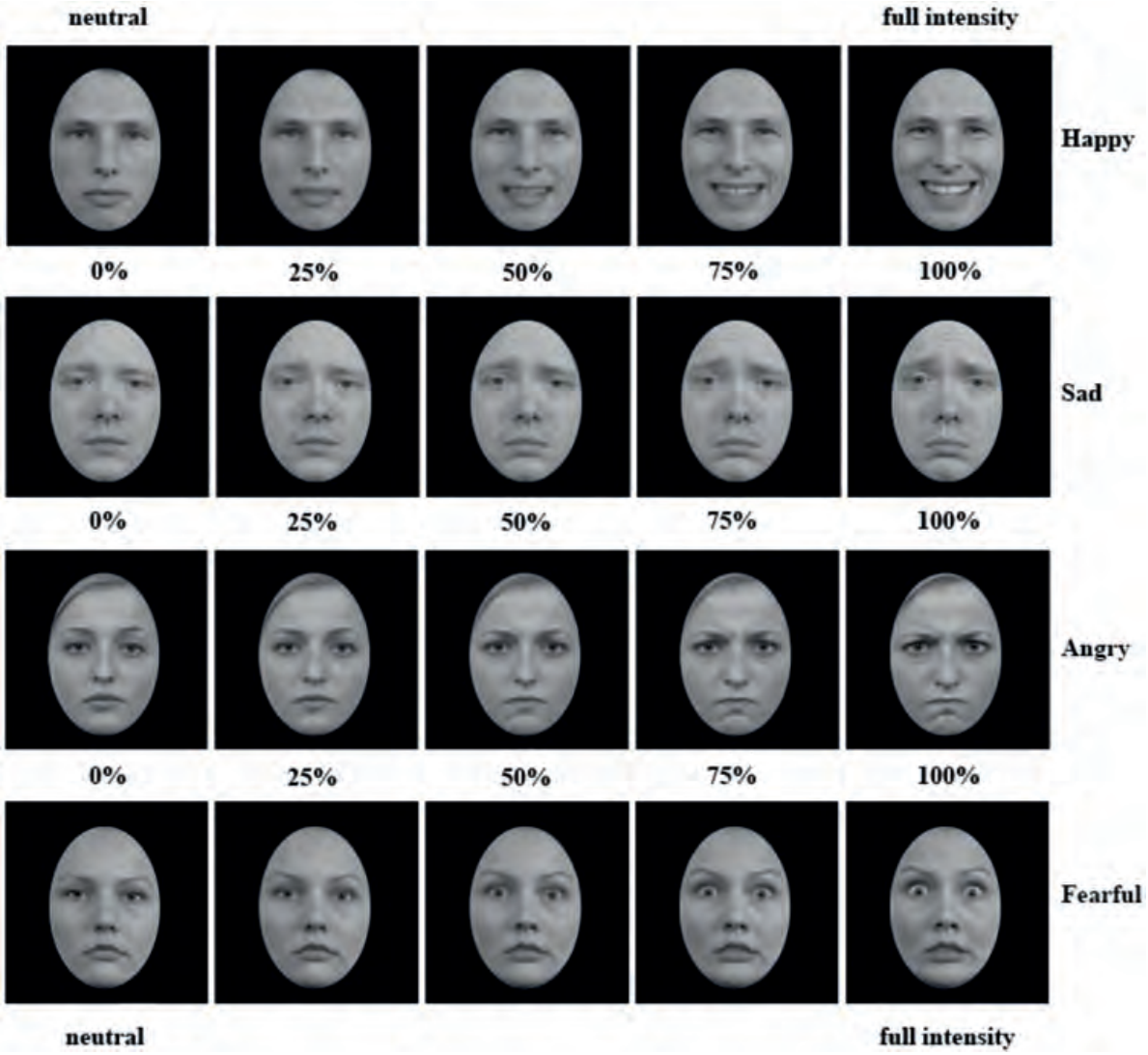



**FIGURE 1.** Examples of the morphs from neutral to full intensity happy, sad, angry and fearful expressions. From each movie clip, five of in total hundred frames are presented in this figure.

The movies had a resolution of 256 x 256 pixels, and were created with FaceMorpher (Luxand Inc., Alexandria, VA, USA) from high quality pictures of six different actors (50% females) from the Radboud Faces Database (Langner et al., 2010). Pictures were cropped with an ovoid frame and converted to gray scale to avoid distracting external cues. Four movies were created of each actor, that is, one for each emotional expression. The original task contained 48 movie clips, that is, twelve per facial emotion, whereas we used a shortened version of 24 movie clips, that is, six for each emotion.

The morph task was programmed in Inquisit 4 (Millisecond, Seattle, USA). The task started with the instruction that participants were about to see movies of faces gradually changing from neutral to emotional expressions. Participants were asked to press the space bar as soon as they were able to identify the emotion. After pressing the spacebar the stimulus movie disappeared and participants indicated the emotion they identified by clicking on one of the four emotion labels. After clicking 'next' a fixation cross appeared in the middle of the screen for 500 ms, followed by a new stimulus. The order of the movie clips was randomized for each participant separately. Before the start of the actual task participants were shown a complete 10-s example movie, followed by two practice trials. After the practice trials the instructions were repeated, followed by the actual task consisting of 24 trials.

We excluded participants with less than 50% correct answers in total (N = 29), because they most likely did not understand the task, were focused on other things, or their strategy consisted of structurally pressing the spacebar before they identified the target emotion. For the remaining participants, the mean reaction time (RT) of correctly identified trials was calculated per emotion. RTs were considered reliable if they were based on at least four movie clips, and hence were calculated only if participants correctly identified at least four out of six movie clips of a specific emotion. This resulted in 6 missing values for RT Happy, 40 for RT Sad, 26 for RT Angry and 52 for RT Fear. Fourteen participants were excluded because they scored the maximum RT (10000 ms) on all four emotions, which indicates that they never pressed the spacebar and just waited until the movie clips stopped. After these spurious RTs were removed, a cut-off for further outliers was defined based on visual inspection of the distribution of the RTs of the four emotions in the complete sample. Because emotion identification differences in the tails of the distribution are likely to be particularly relevant for associations with psychiatric symptoms in a normal population sample, we were rather restrictive and only removed obvious outliers based on visual inspection, that is, RTs that were not connected to the main distribution curve. After the previous steps there were no suspiciously low RTs left upon visual inspection, but we did see indications of high RT outliers and defined mean RT scores higher than or equal to 9800 ms as outliers. The same upper RT limit was used for each emotion to ensure that potential emotion-specific biases would not be removed by removing outliers. Only three participants scored ≥ 9800 ms on one or more emotions; these high scores were considered as missing values.

## Statistical analysis

Using SPSS version 22.0, we set out with calculating descriptive statistics of facial emotion identification RTs and ASR DSM problem scores. Next, we performed a series of regression analyses to determine whether the emotional intensity of a facial expression required for its identification, as measured by RTs, was associated with depressive problems, anxiety problems, avoidance problems, ADHD problems, antisocial problems, or total problems. In the first part of the study we performed these analyses for each psychiatric problem domain separately and in the second part we explored the domain-specificity of the associations that were found by adjusting for scores on psychiatric problem domains that had shown similar emotion identification patterns.

***Part 1: associations between facial emotion identification RTs and psychiatric problems***

We regressed ASR DSM problem scores on facial emotion identification RTs. Including facial emotion RTs as independent rather than dependent variables intuitively made more sense to us, but evidently, since our data are cross-sectional, interpretation can never exceed the level of associations and the choice of emotion identification RTs as independent rather than dependent variables is rather arbitrary. We used standardized (Z-values) RTs and ASR DSM problem scores to be able to compare regression coefficients across different emotions and problem domains. All analyses consisted of two steps: first, for each problem domain the effects of the RTs for happy, sad, angry and fearful emotions were tested separately, adjusted for gender and age (i.e., single-emotion models). Because of tentative evidence that identification patterns across multiple emotions may be more relevant to psychiatric problems than identification of individual emotions (Oldehinkel, Hartman, Oort, & Nederhof, 2015; Vrijen, Hartman, & Oldehinkel, 2016; Wright et al., 2009), in a second step for each problem domain full emotion models were tested including the RTs for all facial emotions (i.e., multi-emotion models), again adjusted for gender and age. Since the dependent variables (i.e., ASR DSM problem scores) and residuals were not normally distributed, we estimated *p*-values from 10,000 bootstrap samples to assure the robustness of our results.

To correct for multiple tests, we used the effective number of tests (Meff) (Li & Ji, 2005) as input for the classical False Discovery Rate (FDR) method (Benjamini & Hochberg, 1995). Combining the Meff and the FDR method, as suggested by Li and Ji (Li & Ji, 2005), enabled us to take into account correlations between tested variables as well as the proportion of significant associations compared to the total number of performed tests. In this way, we corrected for multiple tests without the unwarranted loss of power of more conservative methods. We calculated separate Meffs for the dependent and the independent variables, which were multiplied to obtain the total effective number of tests (i.e., 12; see Supplementary Material for further details). Because we analyzed all emotions separately as well as in full emotion models, and therefore tested all associations twice, we multiplied the effective number of tests by 2 and

used a Meff score of 24 as input for the FDR method. The maximum acceptable FDR was set to 0.05, resulting in an FDR-derived significance threshold of 0.0188 for all analyses. Results were only interpreted as significant for *p*-values below this threshold. Further details of this correction for multiple tests as well as all calculations are presented in the Supplementary Material.

***Part 2: domain-specificity***

Analyses to explore the domain-specificity of associations between facial emotion identification and psychiatric problems were based on patterns found in the results of the analyses as described above. If multiple psychiatric problems domains showed comparable emotion identification patterns (regardless of statistical significance), significant single- and multi-emotion associations from part 1 of the study were re-estimated while adjusting for the problem domains with comparable patterns.

***Sensitivity analyses***

We performed sensitivity analyses to check whether a different method to define outliers changed the results. There is no consensus on what the best method is for defining RT outliers and recommendations vary between not removing outliers, removing outliers based on absolute cut-offs, removing or trimming outliers based on SDs from the mean, use of transformations, and numerous other methods (Erceg-Hurn & Mirosevich, 2008; Ratcliff, 1993; Whelan, 2008; Wilcox, 2012). The advantage of our restrictive approach to defining outliers is that we most likely only removed spurious RT scores and no genuine RTs. A disadvantage may be that we may not have removed all spurious scores, in which case genuine effects may have remained undetected. Following recommendations of Ratcliff (1993), we applied a second outlier approach as a sensitivity check, that is, we defined outliers as three or more SDs from the mean and trimmed them to the closest non-outlier value; this method was for example used in Lodder et al. (2015). All analyses were repeated and compared to the outcomes which were based on the original outlier selection. We also performed sensitivity analyses to investigate if adjusting for education level changed the results.

***Open Science***

Data and syntax have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/pxg7n/.

# RESULTS

## Descriptive statistics (Table 1)

Mean average scores on all ASR DSM-IV problem domains were between 0 ("not true") and 1 ("a little true or sometimes true"), as can be expected in a general population study. Females reported on average more depressive, anxiety, avoidance and total problems, whereas males reported slightly more ADHD and antisocial problems. All problem domains were positively correlated, with correlations varying between $r = .28$ (avoidance and ADHD problems) and $r = .69$ (anxiety and depressive problems). The complete correlation matrix of the psychiatric problem domains is presented in Supplementary Table S-1.

**Table 1.** Descriptive Statistics of the Main Variables in This Study

| | Mean (*SD*) | | |
|---|---|---|---|
| **Variables** | **Total sample N = 2522-2577** | **Males N = 545-566** | **Females N = 1977-2011** |
| Depressive problems[a] | 0.42 (0.33) | 0.32 (0.30) | 0.45 (0.34) |
| Anxiety problems[a] | 0.58 (0.40) | 0.45 (0.36) | 0.61 (0.40) |
| Avoidance problems[a] | 0.43 (0.39) | 0.39 (0.39) | 0.44 (0.39) |
| ADHD problems[a] | 0.44 (0.32) | 0.45 (0.32) | 0.43 (0.32) |
| Antisocial problems[a] | 0.12 (0.13) | 0.15 (0.15) | 0.11 (0.12) |
| Total problems[b] | 0.40 (0.24) | 0.35 (0.23) | 0.41 (0.25) |
| RT Happy[c] | 4113 (848) | 4232 (906) | 4080 (828) |
| RT Sad[c] | 6542 (1059) | 6702 (1093) | 6499 (1045) |
| RT Angry[c] | 5585 (995) | 5816 (1035) | 5521 (975) |
| RT Fearful[c] | 5861 (1039) | 6050 (1107) | 5809 (1014) |
| EP Happy[d] | 1.15 (4.62) | 1.48 (5.43) | 1.06 (4.37) |
| EP Sad[d] | 5.89 (9.59) | 7.36 (11.12) | 5.49 (9.09) |
| EP Angry[d] | 6.06 (9.65) | 7.51 (10.44) | 5.65 (9.38) |
| EP Fearful[d] | 8.37 (11.01) | 9.48 (11.20) | 8.06 (10.94) |

[a] Mean scores on Adult Self Report (ASR) DSM-IV domains: Depressive problems, Anxiety problems, Avoidant personality problems, Attention Deficit/Hyperactivity (ADHD) problems, Antisocial personality problems; Answer categories: 0="not true", 1="a little true or sometimes true", 2="very much true or often true"

[b] Total problems score = average of the sum of mean scores on the above ASR DSM-IV scales

[c] RT = mean reaction time for correct responses measured in milliseconds

[d] EP = mean error proportion

Table 1 contains the mean Reaction Times (RTs) for identification of the different facial emotions. RTs represent the emotional intensity of the facial expressions required to correctly identify the emotions. Happy emotions were generally identified earlier (i.e., before emotional intensity reached 50%) than the other facial emotions, whereas sad emotions were identified later (i.e., after 65% emotional intensity) than all others. Females required less emotional intensity to identify emotions than males. Correlations among RTs (presented in Supplementary Table S-2) varied from $r = .46$ (RT Happy and RT Fearful) to $r = .58$ (RT Sad and RT Fearful).

As can be seen from the lower part of Table 1, the mean Error Proportions (EPs) for all emotions were low. The mean EP of 7.51 for angry emotions for example indicates that on average participants made identification errors in only 7.5% of the cases, which corresponds to less than half an error out of six possible errors in identifying angry emotions, as an average of one error out of six movie clips would represent a mean EP of 16.67.

## Associations between ASR DSM-IV problem domains and emotion identification RTs (Table 2)

No significant associations were found between depressive or anxiety problems and facial emotion identification RTs. Higher scores on avoidance problems were associated with higher RTs for identifying happiness and anger in the single-emotion models, which means that participants with avoidance problems required relatively intense happy and angry facial emotions in order to identify these emotions. In the multi-emotion model only RT Angry remained significant. Higher scores on ADHD problems were associated with higher RTs for anger in the single-emotion as well as in the multi-emotion model, and higher scores on antisocial problems with higher RTs for happiness in the single-emotion, but not in the multi-emotion model. Finally, higher scores on total problems were associated with higher RTs for happiness and anger in the single-emotion models, whereas only RT Angry remained significant in the multi-emotion model.

## Domain-specificity

Statistically significant results were found for the specific psychiatric domains of avoidance, ADHD and antisocial problems (see upper part of Table 2). To explore the domain-specificity of these findings we repeated the analyses of the significant associations as reported in the upper part of Table 2, this time adjusting for the scores on all problem domains that showed similar emotion identification patterns. Regardless of statistical significance, a general emotion identification pattern seemed to be present for all problem domains, that is, the highest absolute *B*s belonged to RT Happy and RT Angry in all single-emotion models for all domains, and all of these *B*s were positive, suggesting patterns in similar directions. In other words, all psychiatric problem domains showed the same pattern in that they were all more strongly associated with RTs for happy and angry emotions than with RTs for sad and fearful emotions. Therefore, to explore

the domain-specificity of the reported significant associations, the single-emotion analyses for avoidance, ADHD and antisocial problems and the multi-emotion analyses for avoidance and ADHD problems were repeated while adjusting for all other specific problem domains.

A comparison of the regression coefficients (*B*s) of the single- and multi-emotion models (upper part of Table 2) to the estimates of the domain-specificity analyses (lower part of Table 2) showed similar patterns for all three problem domains. Between 53% and 67% of each regression coefficient remained after adjusting for the other problem domains.



## Sensitivity analyses

### *Outlier selection method*

The use of an alternative method to handle outliers, with outliers defined as deviating three or more SDs from the mean and trimmed to the closest non-outlier value, did not change the results of the analyses: estimates and p-values of were comparable to Table 2, with only small differences. The same results as before remained significant below the multiple test correction threshold (data available upon request).

### *Education level*

Adjusting for education level, *B*s generally decreased. Fewer *p*-values reached statistical significance at α=.05 compared to the *p*-values in our main analyses, and a new multiple test correction significance threshold which was calculated based on the results adjusted for education level left none of the effects statistically significant. Regardless of statistical significance, the general patterns of the associations between facial emotion identification and psychiatric problems were similar to those without adjusting for education level. For exact regression coefficients and *p*-values, see Supplementary Material, Table S-4.

**Table 2.** Bootstrapping Results of ASR Depressive Problems, Anxiety problems, Avoidance problems, ADHD problems, Antisocial problems and Total Problems Regressed on Facial Emotion Identification Reaction Times

| | | RT Happy | | RT Sad | | RT Angry | | RT Fearful | |
|---|---|---|---|---|---|---|---|---|---|
| | | *B* | *p*-value | *B* | *p*-value | *B* | *p*-value | *B* | *p*-value |
| Single-emotion models | Depressive problems | .037 | .061 | .001 | .993 | .040 | .055 | .021 | .312 |
| | Anxiety problems | .030 | .144 | .001 | .977 | .027 | .188 | .012 | .541 |
| | Avoidance problems | .055 | .009* | .011 | .582 | .061 | .005* | .013 | .516 |
| | ADHD problems | .045 | .035 | .005 | .797 | .056 | .005* | .029 | .153 |
| | Antisocial problems | .057 | .007* | .016 | .405 | .037 | .061 | .029 | .145 |
| | Total problems | .055 | .007* | .007 | .725 | .058 | .006* | .025 | .241 |
| Multi-emotion models[a] | Depressive problems | .034 | .153 | -.044 | .102 | .048 | .074 | .004 | .887 |
| | Anxiety problems | .029 | .225 | -.026 | .343 | .036 | .181 | -.004 | .883 |
| | Avoidance problems | .048 | .060 | -.045 | .117 | .071 | .012* | -.012 | .653 |
| | ADHD problems | .036 | .150 | -.039 | .149 | .070 | .008* | -.006 | .834 |
| | Antisocial problems | .048 | .045 | -.022 | .380 | .027 | .298 | .002 | .934 |
| | Total problems | .049 | .045 | -.047 | .087 | .069 | .013* | -.006 | .844 |
| Domain-specificity of single-emotion models[b] | Avoidance problems | .032 | .040 | | | .038 | .016* | | |
| | ADHD problems | | | | | .033 | .041 | | |
| | Antisocial problems | .030 | .090 | | | | | | |
| Domain-specificity of multi-emotion models[c] | Avoidance problems | | | | | .044 | .035 | | |
| | ADHD problems | | | | | .047 | .025 | | |

ASR = Adult Self-report, ADHD = Attention/Deficit Hyperactivity Disorder; RT = mean reaction time for correct responses in milliseconds; all variables were standardized (Z-values), therefore *B*s can be interpreted as $\beta$s; all effects were adjusted for gender and age; all *p*-values were estimated from 10,000 bootstrap samples

*$p$-value < multiple test correction significance threshold of 0.0188

[a] Multi-emotion models included RTs for all four emotions

[b] Significant single-emotion associations ($p < 0.0188$) adjusted for all other specific psychiatric problem domain scores

[c] Significant multi-emotion associations ($p < 0.0188$) adjusted for all other specific psychiatric problem domain scores

## DISCUSSION

The aim of this study was to examine associations between facial emotion identification and psychiatric problems in young adults, and to explore the domain-specificity of these associations. We used a morph task in which movie clips were presented that gradually morphed from a neutral facial expression to a full intensity facial emotion.

Following expectations, we found associations between facial emotion identification and psychiatric problems. More specifically, young adults with avoidance problems required more intense happy and angry facial emotions to correctly identify these emotions, that is, they were less sensitive to subtle happy and angry emotions. Furthermore, antisocial problems were mainly associated with lower sensitivity to happy facial emotions, and ADHD with lower sensitivity to angry emotions. Contrary to what we expected, we did not find associations between facial emotion identification and depressive or anxiety problems. We found emotion-specific biases, but there was no evidence for overall emotion identification deficiencies. The effects we found could not be fully explained by co-occurring psychiatric problems. Whereas this seems to indicate domain-specificity, inspection of the overall pattern of effect sizes regardless of statistical significance revealed generic patterns as well, in that for all psychiatric problem domains the effect sizes for happy and angry emotions were larger than the effect sizes for sad and fearful emotions. For each problem domain the findings will be discussed in more detail below, followed by a discussion of the domain-specificity, sensitivity analyses, strengths and limitations, and suggestions for future research.

### Facial emotion identification in different psychiatric problem domains

#### *Avoidance problems*

Our finding that experiencing avoidance problems was associated with lower sensitivity to happy and angry emotions may be explained by approach-avoidance mechanisms. It has been argued that happy facial emotions are naturally rewarding social stimuli which elicit approach behavior and that there is a bias towards happy faces in the general population (Evans, Fleming, Dolan, & Averbeck, 2011; Furl, Gallagher, & Averbeck, 2012). Angry facial emotions are commonly regarded as threatening and they evoke avoidance (Seidel, Habel, Kirschner, Gur, & Derntl, 2010). Previous findings further suggest that the avoidance system elicits withdrawal behavior and inhibits goal-directed behavior (Corr, 2001; Leventhal, 2008), and that people suffering from avoidance problems tend to rate other people as more rejecting and less friendly (Meyer, Pilkonis, & Beevers, 2004). More avoidance behavior in response to angry faces and less approach behavior in response to happy faces may have resulted in a lower sensitivity to happy and angry emotions during the emotion identification task. Subtle traces of anger may not be picked up if anger is more strongly avoided and subtle traces of happiness may not be picked up if the urge to approach happiness is weaker.

In the only other study that explicitly addressed the relation between avoidance problems and facial emotion identification, adults meeting full diagnostic criteria of DSM-IV Avoidant Personality Disorder (APD) made more errors in classifying full intensity fearful emotions during a morph task, but did not differ from controls in terms of sensitivity (Rosenthal et al., 2011). This is in contrast with our finding that avoidance problems are associated with lower sensitivity to facial happiness. There is less discrepancy with our finding that avoidance problems are associated with lower sensitivity to anger, since fear and anger are both negative emotions that have both been found to evoke avoidance in socially anxious individuals (Rosenthal et al., 2011; Stirling, Eley, & Clark, 2006). Several methodological differences between our study and the one by Rosenthal and colleagues could be responsible for the contrasting results on sensitivity to happiness: whereas in our study participants were only asked to differentiate between the four basic emotions, the task used by Rosenthal and colleagues (2011) also contained disgust and surprise as facial emotions. Importantly, their sample consisted of only 17 adults with APD and 16 controls. The general pattern of their reported average sensitivity scores per emotion per group suggests that the use a larger sample may in fact have resulted in significant associations for sensitivity to happiness and anger in the same direction as found in our study.

***Antisocial problems***

Antisocial problems have frequently been associated with poor identification of sad or fearful facial emotions (Blair et al., 2004; Marsh & Blair, 2008). Lack of empathy is often used as an explanation for poor identification of sad emotions and an ability to experience fear as an explanation for poor identification of fearful emotions. In our study we found no evidence of associations between antisocial problems and identification of sad or fearful facial emotions. We did find that participants with antisocial problems were less sensitive to happy emotions, which has been occasionally but not consistently found in other studies (Kahler et al., 2012). Lack of empathy or an inability to experience fear cannot explain this finding, however a third characteristic of antisocial problems may. That is, antisociality is also characterized by a lower appreciation of social interactions. As was suggested by Kahler et al. (2012), being unable to identify more subtle traces of happy emotions may contribute to experiencing social interactions as unsupportive and stressful, which may reinforce hostile attitudes. Reversely, low appreciation of social interactions may lead to inattention to rewarding social cues such as smiling faces, thereby blocking the possibility of positive reinforcement. It is more difficult to explain why we did *not* find associations between antisocial problems and identification of sad and fearful emotions. As Blair et al. (2004) and Marsh and Blair (2008) studied people with more severe antisocial problems than we did in the current study, we could speculate that perhaps lack of empathy and an inability to experience fear are symptoms associated with more severe antisocial problems, which would explain why we did not find associations between antisocial problems and identification of sad and fearful emotions.

### *ADHD problems*

In previous studies broad facial emotion processing deficiencies in relation to ADHD have been reported (Rommelse et al., 2011; Sinzig et al., 2008) as well as emotion-specific biases (Aspan et al., 2014; Pelc et al., 2006; Schönenberg et al., 2015). In the current study, which differed from the previous ones in that we investigated young adults instead of children, and assessed ADHD problems in a non-referred sample by means of self-reports rather than in patients with formal diagnoses, we found an association with angry but not the other three emotions. Less proficiency in identifying angry facial emotions has been reported before (Pelc et al., 2006; Singh et al., 1998), but not consistently. A possible explanation could be that children with ADHD have learned to avoid anger when it is expressed to them, or miss important cues necessary for identifying angry expressions altogether (Pelc et al., 2006; Singh et al., 1998). The latter explanation is in line with results from studies using event-related potentials and skin conductance responses, which suggest that children with ADHD show less sensitivity to punishment (Masunami, Okazaki, & Maekawa, 2009; van Meel, Heslenfeld, Oosterlaan, Luman, & Sergeant, 2011). Although speculative, findings may also link to in relation to ADHD reported neural activation (Cubillo, Halari, Smith, Taylor, & Rubia, 2012) and connectivity (Tomasi & Volkow, 2012) differences in lateral orbitofrontal cortex, an area which has been implicated in the evaluation of punishment (Kringelbach & Rolls, 2004; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001). However, there is no convincing body of evidence for these findings. Because we only found evidence that ADHD is associated with problems in identifying anger, it is unlikely that our findings are due to a general problem with concentration on the task for participants with heightened ADHD problems, as this would have resulted in identification problems on all emotions.

### *Depressive and anxiety problems*

Contrary to expectations, no associations were found for depressive and anxiety problems. Although the results were not significant, participants who reported depressive and anxiety problems tended to show patterns similar to those of avoidance problems, that is, experiencing more problems was associated with being less sensitive to happy and angry emotions. Since avoidance behavior plays an important role in depressive and anxiety disorders (Chawla & Ostafin, 2007; Ottenbreit & Dobson, 2004; Ottenbreit, Dobson, & Quigley, 2014), the associations between depressive or anxious problems and emotion identification found in previous studies, which were often conducted in clinical patient populations, may have been at least partly driven by avoidance problems. Though speculative, this explanation would be consistent with our lack of findings for depressive and anxiety problems, since the ASR depressive and anxiety scales contain hardly any avoidance items whereas in clinical patients avoidance problems are part and parcel of depression and anxiety disorders.

## Domain- and emotion-specificity

Our results indicate that co-occurring problems in other psychiatric domains can only partly explain the effects found for avoidance problems, ADHD problems and antisocial problems. Whereas this seems to indicate domain-specificity, for comparing emotion identification patterns between the different domains it makes more sense to compare effect sizes than *p*-values. Inspection of the overall pattern of effect sizes regardless of statistical significance suggests that all psychiatric problem domains were more strongly associated with sensitivity to happy and angry emotions than with sensitivity to sad and fearful emotions. A possible explanation would be that the variation in emotion identification skills within each of the psychiatric problem domains is caused by different mechanisms, and that these mechanisms coincidentally resulted in similar emotion identification patterns. Another explanation would be the presence of a single mechanism that underlies the variation in facial emotion identification across all psychiatric problem domains. We propose that our findings may reflect a generic association of psychiatric problems with less approach tendencies in response to happy facial emotions and more avoidance tendencies in response to angry facial emotions. Variation in the role approach and avoidance problems play in the different problem domains could explain domain-specific variations in effect size as well as why effects for either happy or angry emotions are more pronounced in certain domains.

A closer look at the function of happy and angry facial emotions, and how they are different from sad and fearful emotions, may help to explain our finding that only identification of happy and angry emotions was associated with psychiatric problems. Although happy and angry facial emotions elicit opposite responses, that is, approach and avoidance (Evans et al., 2011; Furl et al., 2012; Marsh, Ambady, & Kleck, 2005; Seidel et al., 2010), there are also commonalities. Happiness and anger are both strong and relatively unambiguous facial emotions associated with direct relevance to the perceiver. Sadness and fear are often less directly aimed at an individual and are more ambiguous in their relevance to the perceiver. These differences are supported by studies showing that happy and angry emotions are identified faster than sad or fearful ones (Calvo & Marrero, 2009; Williams, Moss, Bradshaw, & Mattingley, 2005) and that, in frontal view, directly facing participants, happy and angry emotions are identified more easily, whereas sad and fearful emotions are better identified when gazing in a different direction (Adams & Kleck, 2005). Happy and angry emotions presented in frontal view are also found to elicit stronger emotional responses than when presented in averted gaze direction (Sato, Kochiyama, Uono, & Yoshikawa, 2010). This would imply that, from the perspective of our participants, happy and angry expressions differ from sad and fearful ones in that they are most strongly and unambiguously associated with approach and avoidance tendencies in relation to the person expressing these emotions. If, as we proposed, approach and avoidance mechanisms are indeed underlying the associations between facial emotion identification and psychiatric problems, it seems plausible that differences are found particularly regarding happy and angry expressions,

since these most unambiguously evoke strong approach and avoidance tendencies. Please note that we only offered frontal view stimuli, which potentially increased the salience of happy and angry emotions, but cannot test whether averted gaze directions would have changed the results of this study.



## Education level

Education level explained part of the variation in psychiatric problems, that is, individuals with low levels of education reported more psychiatric problems, and the already small main effects of facial emotion identification on psychiatric problems weakened. Education level is a rough proxy for intelligence, parental socioeconomic status and the social environment combined. Therefore, we can only speculate on why education level and facial emotion identification explained partly the same variation in psychiatric problems. It is possible that education level or intelligence is partly confounding the relationship between facial emotion identification and psychiatric problems, for example if low intelligence or low education level renders an individual vulnerable to psychiatric problems and also explains slower responses to a facial emotion identification task. An alternative explanation is that the presence of psychiatric problems influences educational attainment, as was found in several studies (McLeod & Fettes, 2007; Veldman et al., 2014), as well as facial emotion identification.

## Strengths and limitations

The large sample size of our study enabled us to consider multiple problem domains at once, using similar methods and instruments for all domains, and explore the domain-specificity and emotion-specificity of associations between facial emotion identification and psychiatric problems. A methodological strength of our study is that it gives an indication of the validity of online assessment of the morph task. As far as we know we are the first to use an online facial emotion identification morph task that is completed in participants' own environment without a researcher present. The emotion identification patterns and mean reaction times we found are highly similar to the ones reported in a study among 173 female undergraduate students, in which a 48-video-clip version of the morph task was assessed in a laboratory situation (Lodder et al., 2015). In our study the variation was slightly higher for all emotions, with the largest difference for happy emotions. The comparability of our results suggests that the shortened 24-video-clip version of the facial emotion identification task produces valid results, also when assessed online outside the lab without an instructor present. This is promising for future research since it reduces practical constraints for assessing larger groups of people on multiple occasions.

Our study also has several limitations, warranting that the findings should be interpreted with caution. First, our results apply to a general population of young adults and results cannot be generalized to more severely affected clinical populations. However, the variation in psychiatric

problems in our sample suggests that the small effects we found are not due to accidental recruitment of an overly healthy group. Fifteen percent of the participants experienced psychiatric problems at a clinical level on at least one of the five psychiatric problem domains, 12% experienced problems at a subclinical level and 73% remained within the normal range for all problem domains (Achenbach & Rescorla, 2003). For the separate problem domains, between 1% (antisocial problems) and 8% (avoidance problems) of the participants experienced clinical levels of psychiatric problems and between 1% (antisocial problems) and 9% (depressive problems) experienced psychiatric problems at subclinical levels. We therefore propose that only small effects can be expected in a general population of young adults, but whether this also holds for clinical populations remains to be determined. A second limitation is that psychiatric problems were assessed by means of the Adult Self-Report, and particularly for ADHD problems self-report is not the gold standard. More valid ADHD scores would have been established if parent and teacher reports had been taken into account as well, but these were not collected.

## Future research

First of all, our approach of comparing identification of multiple facial emotions in multiple psychiatric problem domains should be repeated in clinical samples that are sufficiently large to allow adjusting for co-occurring psychiatric symptoms. In more severely affected patient groups, more pronounced emotion identification patterns may be found with clearer implications of underlying mechanisms. A large advantage of investigating multiple problem domains in one study per se is that due to more methodological homogeneity, the findings for different domains are better comparable than findings between different studies. Second, more research is needed to identify the underlying mechanisms of associations between facial emotion identification and psychiatric problems. A potentially viable direction for future research would be a multidisciplinary investigation of the role of approach and avoidance mechanisms in relation to emotion identification in different psychiatric domains, perhaps starting with clinical depression and anxiety, since these have been most consistently associated with approach and avoidance problems (Chawla & Ostafin, 2007; Ottenbreit & Dobson, 2004; Ottenbreit et al., 2014). Combining behavioral measures with neurobiological correlates of approach and avoidance tendencies in response to facial emotions in the different psychiatric domains seems particularly relevant for understanding the underlying mechanisms. The amygdala and associated circuitry have been primarily related to avoidance mechanisms, and the striatum and associated circuitry to approach mechanisms (Ernst & Fudge, 2009). Furthermore, amygdala and striatum abnormalities have been associated with different psychiatric problem domains, for example with depression (Stuhrmann, Suslow, & Dannlowski, 2011), (social) anxiety (Freitas-Ferrari et al., 2010; Shin & Liberzon, 2009), avoidance behavior (Schlund, Hudgins, Magee, & Dymond, 2013), ADHD (Cubillo et al., 2012; Tajima-Pozo et al., 2016) and antisocial behavior (Blair, 2013; Glenn & Yang, 2012). To further link approach and reward mechanisms to facial emotion processing

in the specific psychiatric problem domains, we commend future study of activation and connectivity of the amygdala, striatum and their associated circuits during the processing of conscious as well as subliminally presented facial emotions. It has already been found in several studies that depressed patients showed greater amygdala responses for sad than for happy subliminally presented faces whereas healthy controls showed greater responses to the happy faces (Stuhrmann et al., 2013; Suslow et al., 2010; Victor, Furey, Fromm, Ohman, & Drevets, 2010). Third, consideration of the different types of facial emotion identification errors, for example, mistaking happy faces for sad faces or sad faces for angry faces, may also provide new insights. The morph task, which focuses on the time it takes to identify an emotion, is not equipped to investigate the different types of errors; since errors are rare there is insufficient power to consider the different error types. Studies that use other paradigms, wherein participants are offered pictures or movie clips of specific morphing stages using a forced choice paradigm, are better suitable for this type of research.

More in general, there is a large diversity of findings related to facial emotion identification and psychiatric problems. New studies, as the current one, only seem to increase this heterogeneity, and the small effect sizes reported in our study are a rule rather than exception, especially in general population studies. Although effects reported in clinical populations are generally slightly larger, the diversity of results equally applies and clinical relevance is difficult to establish. Continuing to investigate associations between DSM-based psychiatric domains and facial emotion identification may not be the way to proceed towards clinically relevant findings in the future. In the perspective of recent calls for a revision of the categorical DSM system into a more dimensional approach to psychiatry (Kendler, 2012), it may be more beneficial to focus on smaller units of psychopathology, for example symptoms or clusters of symptoms. In support of this, we previously found that facial emotion identification was a stronger predictor of symptoms of anhedonia than of depression itself (Vrijen, Hartman, & Oldehinkel, 2016). Additionally, a potentially viable way to proceed would be to use facial emotion identification biases or deficiencies rather than psychiatric diagnoses or symptoms as the starting point, and from this perspective investigate associations between *extreme* emotion identification biases or deficiencies and psychiatric symptoms. It seems quite plausible that individual differences in facial emotion identification are not associated with psychiatric problems as long as they remain within a certain range. Comparing a group with extreme deviations in emotion identification to a 'normal range' group may produce more clinically relevant findings and may potentially indicate new mechanisms underlying associations between facial emotion identification and domain-transcendent combinations of psychiatric symptoms.

# SUPPLEMENTARY MATERIAL

## 1. Multiple test correction method: False Discovery Rate combined with Effective number of tests

To correct for multiple tests, the effective number of tests (Meff) (Li & Ji, 2005) was calculated and used as input for the classical False Discovery Rate (FDR) method (Benjamini & Hochberg, 1995). The Meff and the FDR method were both developed as instruments to correct for multiple testing while remaining sufficient power, but they are based on different principles. The main idea behind the Meff method is that the effective number of tests can be determined by means of the correlations between tested variables. If correlations are higher, the number of effective tests decreases and correcting for the total rather than the effective number of tests is too conservative and results in unwarranted loss of power. Characteristic of the FDR method is that power decreases when the number of significant results decreases. The main thought behind this approach is that finding one significant result in 20 tests calls for a more stringent correction than finding 10 significant results in 20 tests. The advantage in power of FDR above more conservative methods increases when more significant results are found and when the total number of tests increases. Combining the Meff and the FDR method, as suggested by a.o. Li & Ji (2005), seemed appropriate for our study since we expected high correlations between our dependent variables, moderate correlations between our independent variables, and based on previous findings we also expected to find multiple significant results.

### *1.1 Meff backgrounds and calculation method*

The Li and Ji Meff is an adaptation of the Cheverud method (Cheverud, 2001). Li and Ji claim that Cheverud's Meff is still too conservative and is only appropriate for two extreme cases, i.e. cases of high correlations between tested variables and cases with hardly any correlations between tested variables, and not for studies with many tests and moderate correlations between tested variables. According to Li and Ji their method gives a more accurate estimate of the Meff that works in the extreme cases as well as in the continuum between these extremes. Especially in the continuum Cheverud's Meff is claimed to be overly large and in this area the Li and Ji Meff would result in more power.

Although the Meff is often calculated only for dependent variables, since we aimed at correcting for the number of independent variables as well, we calculated separate Meffs for the dependent and the independent variables and then multiplied them to establish the total number of effective tests. First, the correlation matrix of our six dependent variables was used to calculate eigenvalues for each of these variables, by using the application offered on https://sites.google.com/site/junningl/software. Subsequently, the following equation, as proposed by Li & Ji (2005), was applied to the eigenvalues:

$$\begin{cases} \text{Meff} = \sum_{i=1}^{M} f(|\lambda i|) \\ f(x) = I(x \geq 1) + (x - \lfloor x \rfloor), x \geq 0 \end{cases}$$

Eigenvalues are decomposed into an **integral part**, with $I(x \geq 1)$ representing what should be counted as 1 test, and a *nonintegral* part $x$–$\lfloor x \rfloor$, representing what counts as a partial test.
The Meff was first calculated for the dependent variables. This resulted in a $\text{Meff}_{\text{dependent}}$ score of 4 (see Table S-1).

**Table S-1.** Correlation Matrix Dependent Variables, Eigenvalues and $\text{Meff}_{\text{dependent}}$

| | Correlations | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Depressive** | **Anxiety** | **Avoidance** | **ADHD** | **Antisocial** | **Total prob.** | **Eigenvalues** | **Meff** |
| Depressive | 1.000 | 0.694 | 0.635 | 0.521 | 0.407 | 0.882 | **3**.*8675* | 1.8675 |
| Anxiety | 0.694 | 1.000 | 0.573 | 0.405 | 0.311 | 0.837 | **0**.*9350* | 0.9350 |
| Avoidance | 0.635 | 0.573 | 1.000 | 0.281 | 0.284 | 0.784 | **0**.*5357* | 0.5357 |
| ADHD | 0.521 | 0.405 | 0.281 | 1.000 | 0.495 | 0.689 | **0**.*3959* | 0.3959 |
| Antisocial | 0.407 | 0.311 | 0.284 | 0.495 | 1.000 | 0.539 | **0**.*2694* | 0.2694 |
| Total prob. | 0.882 | 0.837 | 0.784 | 0.680 | 0.539 | 1.000 | **0**.*0000* | 0.0000 + |
| $\text{Meff}_{\text{dependent}}$ | | | | | | | | **4.0035** |

The same procedure was followed for the four independent variables, resulting in a $\text{Meff}_{\text{independent}}$ score of 3 (see Table S-2).

**Table S-2.** Correlation Matrix Independent Variables, Eigenvalues and $\text{Meff}_{\text{independent}}$

| | Correlations | | | | | |
|---|---|---|---|---|---|---|
| | **RT Happy** | **RT Sad** | **RT Angry** | **RT Fear** | **Eigenvalues** | **Meff** |
| RT Happy | 1.000 | 0.469 | 0.480 | 0.459 | **2**.*5384* | 1.5384 |
| RT Sad | 0.469 | 1.000 | 0.557 | 0.576 | **0**.*5738* | 0.5738 |
| RT Angry | 0.480 | 0.557 | 1.000 | 0.530 | **0**.*4691* | 0.4691 |
| RT Fear | 0.459 | 0.576 | 0.530 | 1.000 | **0**.*4187* | 0.4187+ |
| $\text{Meff}_{\text{independent}}$ | | | | | | **3.000** |

Multiplying $Meff_{dependent}$ and $Meff_{independent}$ resulted in a total of 12 effective tests. Because for all problem domains we analyzed all emotions separately as well as in full emotion models and therefore tested all of them twice, we multiplied the effective number of tests by 2 and used a Meff score of 24 as input for the FDR method.

### *1.2 FDR backgrounds and calculations (Table S-3)*

For the calculation of the classical FDR (Benjamini & Hochberg, 1995), first the p-values of all performed statistical tests are ranked from low to high. Subsequently, with alpha set to 0.05, for each found result an FDR corrected significance threshold is calculated:

$$FDR\ derived\ significance\ treshold\ = \frac{0.05}{number\ of\ tests\ /\ ranking}$$

When these are calculated, it is determined which of the original p-value is still smaller than the FDR corrected significance threshold. Each result with that ranking or higher is still considered significant after multiple testing. We combined FDR and Meff and therefore replaced the number of tests by the Meff-value:

$$FDR\ derived\ significance\ treshold\ = \frac{0.05}{Meff\ /\ ranking}$$

As can be seen in the table below, for the hypothesis with the 9th p-value ranking the p-value is still below the FDR-derived significance threshold, but this is no longer the case for the 10th p-value. The FDR-derived significance thresholds for rank 9 can be calculated as follows:

$$\frac{0.05}{24\ /\ 9}\ =\ 0.01875$$

Since the 9th p-value is the last one to remain below the FDR-derived significance threshold, this significance threshold (0.01875) is the threshold for all tests.

**Table S-3.** Effective Number of Tests (Meff = 24) Applied to FDR Classical Method, With Alpha set to 0.05

| Hypothesis name | *p*-value | Rank | Ascending p-values | Hypothesis name | FDR-derived significance thresholds | FDR-adjusted *p*-values |
|---|---|---|---|---|---|---|
| Happy depres | 0.061 | 1 | 0.005 | Angry adhd* | 0.002083 | 0.030857 |
| Sad depres | 0.993 | 2 | 0.005 | Angry avoi* | 0.004167 | 0.030857 |
| Angry depres | 0.055 | 3 | 0.006 | Angry total* | 0.006250 | 0.030857 |
| Fear depres | 0.312 | 4 | 0.007 | Happy total* | 0.008333 | 0.030857 |
| Happy anx | 0.144 | 5 | 0.007 | Happy antisoc* | 0.010417 | 0.030857 |
| Sad anx | 0.977 | 6 | 0.008 | Angry adhd multi* | 0.012500 | 0.030857 |
| Angry anx | 0.188 | 7 | 0.009 | Happy avoi* | 0.014583 | 0.030857 |
| Fear anx | 0.541 | 8 | 0.012 | Angry avoi multi* | 0.016667 | 0.034667 |
| Happy avoi | 0.009 | 9 | 0.013 | Angry total multi* | 0.018750 | 0.034667 |
| Sad avoi | 0.582 | 10 | 0.035 | Happy adhd | 0.020833 | 0.084000 |
| Angry avoi | 0.005 | 11 | 0.045 | Happy total multi | 0.022917 | 0.090000 |
| Fear avoi | 0.516 | 12 | 0.045 | Happy antisoc multi | 0.025000 | 0.090000 |
| Happy adhd | 0.035 | 13 | 0.055 | Angry depres | 0.027083 | 0.091500 |
| Sad adhd | 0.797 | 14 | 0.060 | Happy avoi multi | 0.029167 | 0.091500 |
| Angry adhd | 0.005 | 15 | 0.061 | Angry antisoc | 0.031250 | 0.091500 |
| Fear adhd | 0.153 | 16 | 0.061 | Happy depress | 0.033333 | 0.091500 |
| Happy antisoc | 0.007 | 17 | 0.074 | Angry depres multi | 0.035417 | 0.104471 |
| Sad antisoc | 0.405 | 18 | 0.087 | Sad total multi | 0.037500 | 0.116000 |
| Angry antisoc | 0.061 | 19 | 0.102 | Sad depres multi | 0.039583 | 0.128842 |
| Fear antisoc | 0.145 | 20 | 0.117 | Sad avoi multi | 0.041667 | 0.140400 |
| Happy total | 0.007 | 21 | 0.144 | Happy anx | 0.043750 | 0.141231 |
| Sad total | 0.725 | 22 | 0.145 | Fear antisoc | 0.045833 | 0.141231 |
| Angry total | 0.006 | 23 | 0.149 | Sad adhd multi | 0.047917 | 0.141231 |
| Fear total | 0.241 | 24 | 0.150 | Happy adhd multi | 0.050000 | 0.141231 |
| Happy depres multi | 0.153 | 25 | 0.153 | Happy depres multi | 0.052083 | 0.141231 |
| Sad depres multi | 0.102 | 26 | 0.153 | Fear adhd | 0.054167 | 0.141231 |
| Angry depres multi | 0.074 | 27 | 0.181 | Angry anx multi | 0.056250 | 0.160889 |
| Fear depres multi | 0.887 | 28 | 0.188 | Angry anx | 0.058333 | 0.161143 |
| Happy anx multi | 0.225 | 29 | 0.225 | Happy anx multi | 0.060417 | 0.186207 |
| Sad anx multi | 0.343 | 30 | 0.241 | Fear total | 0.062500 | 0.192800 |
| Angry anx multi | 0.181 | 31 | 0.298 | Angry antisoc multi | 0.064583 | 0.230710 |
| Fear anx multi | 0.883 | 32 | 0.312 | Fear depress | 0.066667 | 0.234000 |

**Table S-3.** (Continued)

| Hypothesis name | *p*-value | Rank | Ascending p-values | Hypothesis name | FDR-derived significance thresholds | FDR-adjusted *p*-values |
|---|---|---|---|---|---|---|
| Happy avoi multi | 0.060 | 33 | 0.343 | Sad anx multi | 0.068750 | 0.249455 |
| Sad avoi multi | 0.117 | 34 | 0.380 | Sad antisoc multi | 0.070833 | 0.268235 |
| Angry avoi multi | 0.012 | 35 | 0.405 | Sad antisoc | 0.072917 | 0.277714 |
| Fear avoi multi | 0.653 | 36 | 0.516 | Fear avoi | 0.075000 | 0.344000 |
| Happy adhd multi | 0.150 | 37 | 0.541 | Fear anx | 0.077083 | 0.350919 |
| Sad adhd multi | 0.149 | 38 | 0.582 | Sad avoi | 0.079167 | 0.367579 |
| Angry adhd multi | 0.008 | 39 | 0.653 | Fear avoi multi | 0.081250 | 0.401846 |
| Fear adhd multi | 0.834 | 40 | 0.725 | Sad total | 0.083333 | 0.435000 |
| Happy antisoc multi | 0.045 | 41 | 0.797 | Sad adhd | 0.085417 | 0.466537 |
| Sad antisoc multi | 0.380 | 42 | 0.834 | Fear adhd multi | 0.087500 | 0.471070 |
| Angry antisoc multi | 0.298 | 43 | 0.844 | Fear total multi | 0.089583 | 0.471070 |
| Fear antisoc multi | 0.934 | 44 | 0.883 | Fear anx multi | 0.091667 | 0.473067 |
| Happy total multi | 0.045 | 45 | 0.887 | Fear depres multi | 0.093750 | 0.473067 |
| Sad total multi | 0.087 | 46 | 0.934 | Fear antisoc multi | 0.095833 | 0.487304 |
| Angry total multi | 0.013 | 47 | 0.977 | Sad anx | 0.097917 | 0.496500 |
| Fear total multi | 0.844 | 48 | 0.993 | Sad depress | 0.100000 | 0.496500 |

* Significant after multiple test correction

## 2. Post hoc sensitivity analyses: adjusting for education level

**Table S-4.** Bootstrapping Results of ASR Depressive Problems, Anxiety Problems, Avoidance Problems, ADHD Problems, Antisocial Problems and Total Problems Regressed on Facial Emotion Identification Reaction Times, **Adjusted for Education Level**

| | | RT Happy | | RT Sad | | RT Angry | | RT Fearful | |
|---|---|---|---|---|---|---|---|---|---|
| | | *B* | *p*-value | *B* | *p*-value | *B* | *p*-value | *B* | *p*-value |
| Single-emotion models | Depressive problems | .029 | .136 | -.006 | .749 | .032 | .119 | .010 | .634 |
| | Anxiety problems | .026 | .200 | -.003 | .895 | .024 | .250 | .007 | .727 |
| | Avoidance problems | .051 | .015 | .007 | .728 | .056 | .008 | .007 | .753 |
| | ADHD problems | .033 | .118 | -.005 | .809 | .044 | .028 | .011 | .578 |
| | Antisocial problems | .042 | .043 | .003 | .860 | .020 | .298 | .006 | .774 |
| | Total problems | .046 | .026 | -.001 | .953 | .048 | .020 | .011 | .607 |
| Multi-emotion models[a] | Depressive problems | .031 | .203 | -.041 | .123 | .045 | .093 | -.005 | .845 |
| | Anxiety problems | .028 | .254 | -.024 | .368 | .034 | .195 | -.008 | .756 |
| | Avoidance problems | .046 | .071 | -.043 | .133 | .069 | .015 | -.018 | .506 |
| | ADHD problems | .031 | .220 | -.035 | .190 | .065 | .012 | -.020 | .437 |
| | Antisocial problems | .041 | .090 | -.017 | .505 | .021 | .419 | -.017 | .520 |
| | Total problems | .045 | .067 | -.044 | .112 | .065 | .018 | -.017 | .533 |

ASR = Adult Self-report, ADHD = Attention/Deficit Hyperactivity Disorder; RT = mean reaction time for correct responses in milliseconds; all variables were standardized (Z-values), therefore *B*s can be interpreted as *β*s; all effects were adjusted for gender, age and education level; all *p*-values were estimated from 10,000 bootstrap samples

**No *p*-values < multiple test correction significance thresholds**

[a] Multi-emotion models contained all four emotion RT

# CHAPTER 3

## Slow identification of facial happiness in early adolescence predicts onset of depression during 8 years of follow-up

Published as:
Vrijen, C., Hartman, C. A., & Oldehinkel, A. J. (2016). Slow identification of facial happiness in early adolescence predicts onset of depression during 8 years of follow-up. *European Child & Adolescent Psychiatry*, *25*(11), 1255-1266. https://doi.org/10.1007/s00787-016-0846-1

## ABSTRACT

Adolescent onset depression places a high burden on those who suffer from it and is difficult to treat. An improved understanding of mechanisms underlying susceptibility to adolescent depression may be useful in early detection and as target in treatment. Facial emotion identification bias has been suggested as trait marker for depression, but results have been inconclusive. To explore whether facial emotion identification biases may be trait markers for depression, we tested whether the speed with which young adolescents identified happy, sad, angry and fearful facial emotions predicted onset of depression during an eight-year follow-up period. We hypothesized that facial emotion identification speed predicts depression in a symptom-congruent way and differentially predicts symptoms of anhedonia and sadness. Data were collected as part of the TRacking Adolescents' Individual Lives Survey (TRAILS), and involved 1840 adolescents who participated in a facial emotion identification test at age 11 and were subjected to the World Health Organization Composite International Diagnostic Interview (CIDI) at age 19. In a multi-emotion model, slow identification of happy facial emotions tentatively predicted onset of depressive disorder within the follow-up period. Slow identification of happy emotions and fast identification of sad emotions predicted symptoms of anhedonia, but not symptoms of sadness. Our results suggest that the relative speed of identification of happiness in relation to the identification of sadness is a better predictor of depression than the identification of either facial emotion alone. A possible mechanism underlying the predictive role of facial emotion identification may be a less reactive reward system.

## INTRODUCTION

Adolescent unipolar depression is a common and major mental health problem. In mid to late adolescence, the estimated 1 year prevalence is 4-8% worldwide (Costello, Erkanli, & Angold, 2006; Lewinsohn et al., 1994; Ormel et al., 2015), and at the end of adolescence the estimated cumulative incidence is 16-28% in community samples (Lewinsohn et al., 1994; Ormel et al., 2015). For adolescents aged 10-19, depression is the first leading cause of disability as measured in disability-adjusted life-years (DALYS) (Gore et al., 2011), and is strongly linked to suicide risk (Brent et al., 1993). According to the DSM (American Psychiatric Association, 2013), the two core symptoms of depression are anhedonia (loss of pleasure) and sadness; experiencing either one of these symptoms is a necessary condition for receiving the diagnosis Major Depressive Disorder (MDD). The burden of adolescent depression is not limited to adolescence: depression in adolescence is a strong predictor of adult depression. Experiences of subclinical symptoms of anhedonia and sadness during adolescence also predict adult depression, with stronger evidence for anhedonia (Pine et al., 1999; Wilcox & Anthony, 2004) than for sadness (Wilcox & Anthony, 2004). Depression in adolescence is often not recognized and not adequately treated (Zuckerbrot & Jensen, 2006). Therefore, it is important to uncover mechanisms underlying susceptibility to adolescent depression and its core symptoms, to better understand the disorder. Ultimately, improved understanding may be useful in early detection and as target in treatment. Facial emotion processing[4] is critical for normal emotional development and engaging in social relationships, and has been implicated as potential susceptibility factor for depression. The aim of the current study is to explore facial emotion processing bias in early adolescence as a potential trait marker for depression and symptoms of anhedonia and sadness in middle and late adolescence.



The body of knowledge concerning the connection between facial emotion processing and depressive symptoms is rapidly expanding, but results are highly heterogeneous. Research so far has largely been guided by cognitive theories claiming that negative cognitions initiate, maintain and strengthen depressive schemas (Beck, 1967a), and by network models of emotion (Bower, 1981; Ingram, 1984) claiming that mood-congruent stimuli are processed more easily and correctly than mood-incongruent stimuli. According to these theories, depressed individuals are expected to suffer from negative biases in virtually all types of information processing, including perception, attention and memory. From the perspective of contemporary cognitive neuropsychological models of depression, which attempt to reconcile the cognitive theory with neurobiological findings, emotion processing bias is essential in understanding the mechanisms of depression. These biases have been claimed to be present before mood starts to deteriorate,

4 In this paper, 'facial emotion processing' is used for reference to facial emotion processes that are not limited to facial emotion identification, but also entail attention, memory, etc. 'Facial emotion identification' will be reserved for the more specific concept of identification only.

and to be the main operating mechanism of depression treatment. It has been postulated that treatment enhances mood only indirectly through changing emotion processing biases, which in their turn instigate further changes that ultimately lead to improvement of mood (Disner et al., 2011; Roiser et al., 2012). This suggests a direct relevance of measuring and monitoring emotion processing biases. Claims have also been made about the specificity of the emotion biases that are related to depressive symptoms. The content-specificity hypothesis states that depressed persons demonstrate stronger biases for themes that are consistent with depressed disorder, e.g., sadness and loss, than for anxiety-related stimuli such as threat and anger (Beck, 1976).

So far, empirical studies about facial emotion processing have been unable to provide conclusive empirical support for or against the above theoretical frameworks. Many studies have indicated relationships between biases in facial emotion information processing and depressive symptoms, but their findings have been far from consistent (Archer et al., 1992; Bistricky, Ingram, & Atchley, 2011; Joormann & Gotlib, 2006; Kohler et al., 2011; Wright et al., 2009). This can at least partly be explained by small sample sizes and varying measurements, instruments and experimental procedures (Bediou et al., 2012; Dalili et al., 2014). Another plausible cause is the heterogeneity of depressed patient groups (Surguladze et al., 2004). When MDD patients are studied, usually no specific categories of depression are differentiated, and most studies do not address comorbidity issues. Despite evidence that the two core symptoms of depression according to the DSM (American Psychiatric Association, 2013), i.e., anhedonia and sadness, are associated with distinct psychophysiological systems reflecting approach and avoidance tendencies (Carver, 2006; Ernst & Fudge, 2009), so far these core symptoms have not been used to differentiate between different types of depressive patients with respect to facial emotion processing. Because earlier studies did not distinguish between anhedonia and sadness and often not between depressed patients with and without anxiety disorders either, they were not equipped to test symptom- or content-specificity.

Another unresolved issue is that because of a lack of longitudinal studies, it is not clear whether facial emotion processing biases precede depressive symptoms or the other way around, and whether these relationships are trait- or state-dependent. Biases that are only present during depressive episodes suggest state dependency; biases that are also found preceding a depressive episode or after remission rather reflect traits. Only a limited number of studies have been published on facial emotion processing bias as a potential trait marker for depression, by investigating first-degree relatives or recovered depressed patients. Studies in first-degree relatives of depressed patients found that a familiar risk for depression increased 7-13-year-old boys' ability to identify sad emotions (Lopez-Duran et al., 2013), and that negative mood induction generated an attentional bias towards sad faces in 9-14-year-old girls with a familiar risk for depression, while girls without familiar risk showed a bias towards happy faces (Joormann et al., 2007). Studies involving recovered depressed patients showed that biases towards negative facial emotions (Bhagwagar et al., 2004), away from positive emotions

(LeMoult, Joormann, Sherdell, Wright, & Gotlib, 2009), or both (Anderson et al., 2011) persisted after remission. Furthermore, a study in young adults revealed that past depression (trait) was associated with greater salience of sad target faces, whereas current dysphoria (state) was related to a failure to inhibit responses to sad distractor faces (Bistricky, Atchley, Ingram, & O'Hare, 2014). Hence, empirical evidence regarding trait-dependent relationships between facial emotion identification and depression remains inconclusive and further research is needed.

Both facial emotion identification skills and onset of depression have been consistently reported to differ between males and females: the available evidence suggests a small female advantage in facial emotion identification throughout life (McClure, 2000), and a higher prevalence of depression in females than males starting in adolescence (Twenge & Nolen-Hoeksema, 2002). Gender differences have also been suggested in the relationships of facial emotion processing with depression (Lopez-Duran et al., 2013; Wright et al., 2009), but most were not suitable to test these gender differences statistically because of small samples. The one study that did test gender differences (Wright et al., 2009), found that facial emotion identification was associated with depression only in females. Altogether, this advocates considering gender when studying associations between facial emotion identification and depression.



We addressed several of the issues described above by performing a longitudinal study with a large sample size, differentiating between the two core depression symptoms anhedonia and sadness, and taking into account comorbid anxiety diagnoses. Focusing on facial emotion processing as a potential trait marker for depression, we investigated whether differences in emotion identification speed predicted later onset of depressive disorder and whether they differentially predicted symptoms of anhedonia and sadness. Major Depressive Disorder, Minor Depressive Disorder and Dysthymia were taken together in the overarching construct 'Depressive Disorder', since recent findings suggest that Minor Depressive Disorder and Dysthymia represent the same pathology as Major Depressive Disorder, and differences among the disorders concern severity rather than qualitative characteristics (Fils et al., 2010; Judd, Akiskal, Zeller, & et al, 2000; Kessler et al., 2003).

We tested the following hypotheses:

(1) Speed in facial emotion identification at age 11 predicts the onset of later depressive disorder. (1a) Based on theories of mood congruence we expected that depression onset is predicted by slower identification of happy emotions and faster identification of sad emotions. This will be referred to as 'symptom congruence'[5]. We also expected to find (1b) content-specificity, i.e., stronger associations with happiness and sadness than with anger and fear. Considering the core symptoms anhedonia and sadness separately, we hypothesized that: (2a) slower

5 'Mood congruence' more commonly refers to a relationship between an identification bias and a concurrent mood state at one and the same time point. We stretched this use to a different time frame by hypothesizing the presence of a bias at one time-point (age 11) to be congruent with depressive symptoms that are developed later (between age 11 and 19). To avoid misunderstandings, from now on we will no longer use 'mood congruence', but refer to this idea as 'symptom congruence'.

identification of happy facial emotions is more predictive of anhedonia than of sadness; (2b) faster identification of sad emotions is more predictive of sadness than of anhedonia; (2c) fearful and angry face identification predict neither anhedonia nor sadness. Because of the inconsistent results of previous studies and tentative indications of the relevance of considering multiple emotions simultaneously (Wright et al., 2009)), we considered single-emotion models as well as multi-emotion models.

All associations were tested both regardless of comorbid anxiety and after exclusion of individuals with (lifetime) Social Phobia (SP) or Generalized Anxiety Disorder (GAD), to explore to what extent the associations, if any, were depression-specific. SP and GAD were selected because of their social orientation, high comorbidity with depressive symptoms (Kessler, Chiu, Demler, & Walters, 2005) and earlier evidence of associations between SP or GAD and facial emotion processing biases (Bradley, Mogg, White, Groom, & De Bono, 1999; Joormann & Gotlib, 2006). Because of the plausibility of gender-specific relationships between facial emotion identification measures and onset of depression, we also tested gender interactions.

## METHOD

### Sample and procedure

This study is based on data collected as part of the TRacking Adolescents' Individual Lives Survey (TRAILS), an ongoing cohort study investigating mental health and social development from early adolescence into adulthood. The study consists of two prospective cohort studies, a population-based cohort ($N = 2230$) and a clinical cohort ($N = 543$). TRAILS was approved by the Dutch Central Committee on Research Involving Human Subjects (CCMO), participants were treated in accordance with the Declaration of Helsinki, and written consent was acquired from all adolescents and their parents.

The data collection in both cohorts involved largely the same measures and participants were assessed at largely the same ages, every two or three years (Oldehinkel, Rosmalen, et al., 2015). The specific questionnaires and tasks used were described in a previous report (Oldehinkel, Rosmalen, et al., 2015). For the present study, we used data from the first (T1) and fourth (T4) waves of both cohorts. The participants of the population cohort were recruited from primary schools (response rate 90%) in five municipalities in the northern region of the Netherlands. Of all eligible children, 2230 (76%) agreed to participate. For more details on the selection procedure see De Winter and colleagues (2005). At T1, which ran from March 2001 until July 2002, the mean age of the population cohort was 11.1 years (*SD* 0.6), and 51% were females. At T4 (from October 2008 until September 2010), 1881 adolescent participated again (retention rate 84%), the mean age was 19.1 years (*SD* 0.6), and 52% were females. Participants of the clinical cohort had been in contact with a specialized mental health service in the North of the Netherlands before the

age of ten. Of all eligible participants asked, 543 (43%) agreed to participate in the study. As was expected, non-response in this particular group was larger than it was in the population cohort. However, no significant differences were found between responders and non-responders in age, gender, parental education, age of referral to mental health services, teacher reports on mental health and on school achievement (except for lower mathematics performance in non-responders) (Huisman et al., 2008). At T1 (from September 2004 until December 2005), the mean age of the clinical cohort was 11.1 years (*SD* 0.5) and 34% were females; at T4 (from September 2012 until April 2014), 422 adolescents participated again (retention rate 78%), the mean age was 19.1 years (SD 0.7), and 34% were females. The larger proportion of boys compared to girls in the clinical cohort is due to the fact that children with pervasive developmental disorder, attention deficit/hyperactivity and externalizing problems are referred to mental health services more often than those with internalizing problems (Garralda & Bailey, 1988; Gershon, 2002), and these problems are more common in boys than in girls (Bongers, Koot, van der Ende, & Verhulst, 2004; Fombonne, Simmons, Ford, Meltzer, & Goodman, 2001; Gershon, 2002).



From both cohorts, we selected all participants who: (1) completed the facial emotion identification task at T1; (2) had been subjected to the World Health Organization Composite International Diagnostic Interview (CIDI) at T4; and (3) had not had a depressive disorder, i.e., Major Depressive Disorder, Minor Depressive Disorder or Dysthymia, as measured by the CIDI retrospectively at T4, prior to taking the facial emotion identification task. This yielded a sample of 1840 participants (81% of the remaining population cohort at T4, 76% of the remaining clinical cohort at T4).

Since the TRAILS study covered numerous research questions, no a priori power analysis was performed regarding our specific research question. For the present study, a post hoc power analysis for logistic regression (Faul, Erdfelder, Buchner, & Lang, 2009) with an alpha set to 0.05, 1840 included participants, a proportion of lifetime depressed participants of 0.20, and a predefined effect of 20% increased risk of depressive disorder per SD increase in the predictor variable, yielded an estimated power of 0.88. To detect an effect of 10% increased risk the power decreased to 0.37. For outcomes with a proportion of about 0.35, like lifetime symptoms of anhedonia and sadness, the estimated power for effects of 20% and 10% increased risk was, respectively, 0.96 and 0.49.

## Measures

### *Facial emotion identification*

Facial emotion identification was measured by means of the 'Identification of Facial Expressions' (IFE) task at T1. This task was the last of seven tasks selected from the Amsterdam Neuropsychological Tasks program (ANT) (De Sonneville, 1999, p. 3), which in total took approximately 70 min to complete. Detailed information on the ANT testing procedures and the IFE task is provided in Supplement 1 of the Supplementary Material.

Our hypotheses concerned the facial emotions happiness, sadness, anger and fear. Participants were included if they had completed the IFE task on at least one of these four emotions. For each of the four emotions we calculated the Error Proportion (EP) and the Reaction Time (RT). EPs were calculated as the mean proportion of misses and false alarms: EP = ((misses / (misses + hits)) + (false alarms / (false alarms + correct rejections))) / 2. RTs were calculated by the mean RT across hits and correct rejections. Subsequently, EPs and RTs of more than four standard deviations above the mean (Stevens, 2002) or EPs indicating performance at chance level, i.e., of 50% or higher, were considered outliers and treated as missing. Because EP and RT potentially influence each other, outliers in one outcome parameter were also considered missing in the other. The percentage of missing EPs and RTs, including outliers, was less than 1.3% for each facial emotion.

***Depressive disorder and symptoms of anhedonia and sadness***

At wave T4, the World Health Organization Composite International Diagnostic Interview (CIDI) version 3.0 (Kessler & Üstün, 2004) was used to assess onset of psychiatric disorders. The CIDI is a structured diagnostic interview which has been shown to have good reliability and validity in assessing current and lifetime DSM-IV disorders (Andrews & Peters, 1998; Haro et al., 2006; Kessler et al., 2004). The interview started with a screening section for all participants, meant to determine which of the subsequent sections on specific disorders should be included in the interview. For each of these specific disorders age of onset was also registered.

In the present study, we were interested in the following outcome measures: (1) depressive disorder and (2) symptoms of anhedonia and sadness regardless of depressive disorder. The screening part of the CIDI depression section contained three questions: (1) 'Have you ever in your life had a period lasting several days or longer when most of the day you felt sad, empty or depressed?'; (2) 'Have you ever had a period lasting several days or longer when most of the day you were very discouraged about how things were going in your life?'; (3) 'Have you ever had a period lasting several days or longer when you lost interest in most things you usually enjoy like work, hobbies, and personal relationships?'. All participants who endorsed at least one of these symptoms entered the whole depression section of the CIDI, which allowed classifying the participants according to DSM-IV criteria for Major Depressive Disorder, Minor Depressive Disorder, Dysthymia, Recurrent Brief Depression, and Bipolar Disorder. For the present study, depressive disorder was operationalized as the occurrence of at least one of the following disorders: Major Depressive Disorder, Minor Depressive Disorder or Dysthymia, and first onset of depressive disorder as the first onset of any of these three disorders. Since we were interested in predicting the incidence of depressive disorder after T1, we excluded the participants whose age of onset of depressive disorder was lower than or equal to their age at the time they took the facial emotion identification test at T1.

Within depressive disorders, correlations between anhedonia and sadness are high (in our sample, $r = .78$, $p < .001$), leaving little power to test anhedonia and sadness separately. Moreover, subclinical expressions of anhedonia and sadness were considered informative too. Therefore, we used CIDI screening items to determine the presence of symptoms of anhedonia and sadness regardless of the diagnostic status. Anhedonia was measured by the item 'Have you ever had a period lasting several days or longer when you lost interest in most things you usually enjoy like work, hobbies, and personal relationships?', and sadness by the item 'Have you ever in your life had a period lasting several days or longer when most of the day you felt sad, empty or depressed?'. The correlation between these items was .37 ($p < .001$).



Since we were interested in predicting the incidence of symptoms of anhedonia and sadness after T1, we omitted participants who had already reported these symptoms in the Youth Self-Report (YSR) (Achenbach, 1991) at T1, when they were asked to report about the 6 months prior to T1. More specifically, we excluded participants with high scores (i.e., 'clearly or often') on T1 YSR item 'I enjoy very little', and participants with high scores (i.e., 'clearly or often') on T1 YSR item 'I am sad, unhappy or depressed' from the regression analyses of symptoms of anhedonia or sadness on facial emotion identification speed.

## Statistical analysis

Using SPSS version 22.0, we performed a series of logistic regression analyses to determine whether facial emotion identification speed (RTs) predicted onset of depressive disorder, anhedonia and sadness. Facial emotion identification can be assessed by RTs (speed) and by EPs (accuracy). We focused on speed rather than accuracy, because the task used (static facial emotion expressions presented at full intensity) is relatively easy for 11-year-old adolescents, and we therefore expected that identification speed would have more discriminative power than identification accuracy. To account for possible associations between speed and accuracy, we adjusted for accuracy.

Standardized RTs were used to be able to compare Odds Ratios (ORs) across different emotions. All analyses consisted of two steps: first the effects of the RTs for happy, sad, angry and fearful emotions were tested separately, adjusted for the respective EPs, gender, and age at the time of the IFE task. Second, we started with a full model including the EPs and RTs for all facial emotions and ran a backward conditional logistic regression analysis (again adjusting for gender and age), to estimate the combined effect of emotion identification speed of multiple emotions. In the final models we always adjusted for the EPs of all of the RTs in the model, to ensure that found effects could not have been driven by EPs rather than RTs.

In the first step of our analysis, i.e. testing the emotions separately, significance was set at .05 and in the second step, i.e. backward conditional logistic regression analyses, the entry criterion was set at .05 and the removal criterion at .10. The choice of a backward rather than a forward selection procedure was motivated by the idea that forward selection involves a higher risk of

excluding predictors with a suppressor effect (i.e. predictors that are only significant if certain other predictors are included in the model as well), which we did not want to ignore beforehand because of the exploratory nature of this part of our study. The exploratory nature of this study was also the reason for choosing a backward conditional logistic regression removal criterion of .10 and not correcting for multiple tests in our initial analyses. The latter also implies that our results should not be interpreted in a formal discriminatory way, which is why we did not focus on single significant results but on more general patterns. A False Discovery Rate (FDR) method (Benjamini & Hochberg, 1995) was employed post hoc to give an indication which effects meet multiple test correction criteria. The maximum acceptable FDR was set to 0.05.

Since our analyses on the core symptoms anhedonia and sadness were primarily aimed at identifying symptom-specific facial emotion identification patterns, the effects of anhedonia and sadness were corrected for each other in these models. For the purpose of identifying potential gender differences, gender*RT interactions were tested for all separate emotion models. Practical limitations prohibited the inclusion of interactions with gender in the multi-emotion backward selection models.

Several additional specificity and sensitivity analyses were performed. To check if findings pertained specifically to depression all associations were tested both regardless of comorbid anxiety and after exclusion of all participants with T4 retrospective CIDI-based lifetime diagnoses of SP or GAD. Finally, we checked whether adjusting for baseline speed or cohort status (population cohort or clinical cohort) changed the main results.

## RESULTS

### Number of participants excluded per exclusion criterion

From the total sample of 1921 participants who had been subjected to the IFE task at T1 as well as to the CIDI interview at T4, we excluded 81 participants with a CIDI depressive disorder first onset at the same time or prior to taking the IFE task. For the analyses concerning symptoms of anhedonia and sadness, we further excluded 45 participants with high scores (i.e., 'clearly or often') on T1 YSR item 'I enjoy very little', and 35 participants with high scores (i.e., 'clearly or often') on T1 YSR item 'I am sad, unhappy or depressed'. The percentage of missing EPs and RTs, including outliers, was less than 1.3% for each facial emotion: for happy we excluded 14 participants, for sad 10 participants, for angry 20 participants and for fearful 22 participants.

For the regression analyses of depression on facial emotion identification this yielded the following samples sizes: for RT Happy $N = 1826$, for RT Sad $N = 1830$, for RT Angry $N = 1820$, for RT Fearful $N = 1818$ and for the backward selection model $N = 1785$. And for the regression analyses of anhedonia and sadness on facial emotion identification the following: RT Happy $N = 1738$, for

RT Sad *N* = 1742, for RT Angry *N* = 1733, for RT Fearful *N* = 1733 and for the backward selection model *N* = 1701. For the additional specificity analyses we excluded 290 participants with CIDI-based lifetime diagnoses of SP or GAD.

## Descriptive statistics

Descriptive information of the variables used in this study is presented in Table 1. Between age 11 and age 19, 19% of our sample developed a first depressive disorder, 35% experienced symptoms of anhedonia and 42% experienced symptoms of sadness. Female incidence rates were higher than male rates for depression and sadness, but not for anhedonia. Frequencies are in accordance with findings from previous studies that in adolescence more females than males get depressed (Twenge & Nolen-Hoeksema, 2002).



**TABLE 1.** Descriptive Statistics

| | Frequencies (%) / Mean (*SD*) | | |
|---|---|---|---|
| **Variables** | **Total sample *N* = 1818-1840** | **Males *N* = 896-909** | **Females *N* = 922-931** |
| Depression[a] | 344 (19%) | 111 (12%) | 233 (25%) |
| Anhedonia[b] | 618 (35%) | 306 (35%) | 312 (34%) |
| Sadness[c] | 765 (42%) | 318 (36%) | 447 (49%) |
| RT Baseline[d] | 334 (49) | 332 (46) | 336 (52) |
| RT Happy[d] | 878 (206) | 883 (213) | 874 (199) |
| RT Sad[d] | 1210 (287) | 1229 (299) | 1191 (274) |
| RT Angry[d] | 1117 (259) | 1125 (271) | 1108 (245) |
| RT Fearful[d] | 1112 (277) | 1119 (282) | 1105 (272) |
| EP Happy[e] | 3.3 (3.3) | 3.5 (3.4) | 3.1 (3.3) |
| EP Sad[e] | 12.6 (9.2) | 13.4 (9.4) | 11.9 (9.0) |
| EP Angry[e] | 8.4 (6.1) | 8.6 (5.9) | 8.1 (6.3) |
| EP Fearful[e] | 7.5 (6.6) | 7.9 (6.8) | 7.2 (6.4) |

[a] CIDI-based DSM-IV diagnosis of Major Depressive Disorder, Minor Depressive Disorder or Dysthymia, with age of onset between 11 and 19;
[b] Symptoms of anhedonia for at least several consecutive days between age 11 and 19;
[c] Symptoms of sadness for at least several consecutive days between age 11 and 19;
[d] RT = mean reaction time for correct responses measured in milliseconds, assessed at age 11;
[e] EP = mean error proportion, assessed at age 11.

Table 1 shows differences in Reaction Times (RTs) of identifying the different facial emotions. Happy emotions were identified faster (shorter RTs) than the other facial emotions and participants had most difficulties with identifying sad facial expressions (longer RTs). Females

identified facial emotions faster than males. Correlations between RTs (data not presented in table) varied from $r = .61$, $p < .001$ (RT Happy and RT Fear), to $r = .74$, $p < .001$ (RT Angry and RT Sad). See Supplement 2 for descriptive statistics of facial emotion identification RTs and EPs by emotion by diagnostic group.

## Prediction of onset of depressive disorder by facial emotion identification RTs (Table 2, left side)

Facial emotion identification RTs at age 11 did not significantly predict onset of depressive disorder when testing each facial emotion separately. Backward conditional logistic regression analysis resulted in a multi-emotion model in which onset of depression was predicted by longer RTs (OR greater than 1) for happiness in combination with shorter RTs (OR smaller than 1) for sadness, of which only RT Happy reached statistical significance at α=.05. These results are graphically presented in Fig. 1a. Additional analyses revealed that the presence of RT Happy in the multi-emotion model relied specifically on the inclusion of RT Sad and the other way around; neither RT Happy nor RT Sad could be replaced by RT Fearful or RT Angry. Excluding either RT Happy or RT Sad from the model resulted in no longer finding any effect with backward model selection.

**TABLE 2.** Results Logistic Regression Analyses of DSM-IV Depression and Symptoms of Anhedonia and Sadness for at Least Several Days Between Age 11 and Age 19 on Facial Emotion Identification Reaction Times at Age 11

| | | **Depression[a] N = 1785-1830** | | **Anhedonia[b] with sadness as covariate N = 1701-1742** | | **Sadness[c] with anhedonia as covariate N = 1701-1742** | |
|---|---|---|---|---|---|---|---|
| | | **OR** | **P** | **OR** | **P** | **OR** | **P** |
| Emotions tested separately | RT Happy | 1.07 | .25 | 1.05 | .39 | 0.97 | .60 |
| | RT Sad | 0.97 | .67 | 0.93 | .16 | 1.04 | .44 |
| | RT Angry | 1.02 | .77 | 0.99 | .86 | 0.98 | .64 |
| | RT Fearful | 0.96 | .54 | 1.01 | .87 | 1.04 | .51 |
| Models backward selection | RT Happy | 1.19 | .04 | 1.22 | .009 | | |
| | RT Sad | 0.86 | .09 | 0.80 | .005 | | |
| | RT Angry | | | | | | |
| | RT Fearful | | | | | | |
| Post hoc analyses | RT Sad - RT Happy (HS) | 0.89 | .08 | 0.85 | .004 | Not tested | Not tested |

All effects were adjusted for error proportions, gender and age at the time of the facial emotion identification task;
OR = odds ratio; RT = mean reaction time for correct responses; all RTs in this table are standardized (Z-values) with one exception: HS was calculated on unstandardized RT Sad and RT Happy and was standardized afterwards

[a] CIDI-based DSM-IV diagnosis of Major Depressive Disorder, Minor Depressive Disorder or Dysthymia;

[b] Symptoms of anhedonia for at least several consecutive days;

[c] Symptoms of sadness for at least several consecutive days

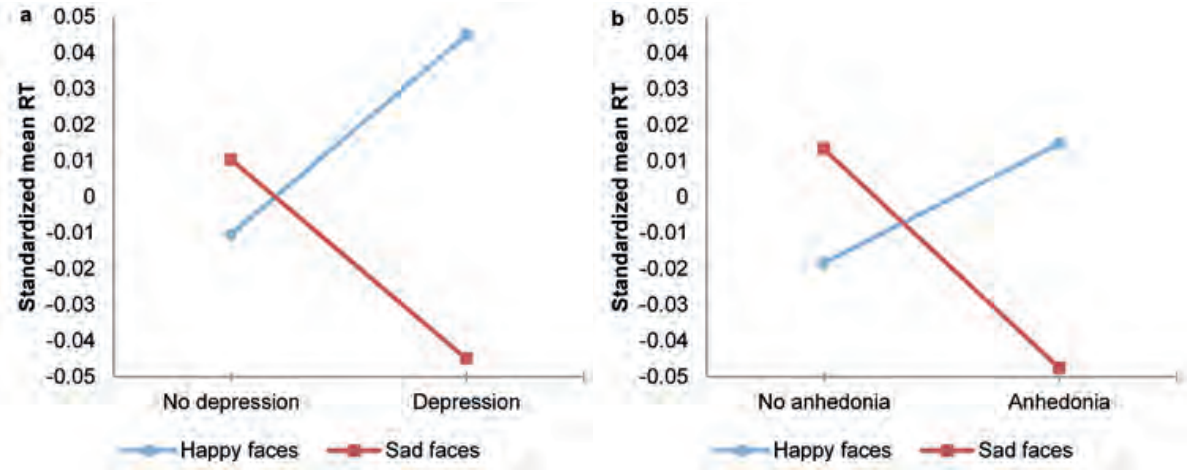



**FIGURE 1.** Standardized Reaction Times (RTs) for the identification of happy and sad facial emotions, for young adolescents with and without onset of depressive disorder during the eight-year follow-up period (Fig. 1a), and for those with and without symptoms of anhedonia during the follow-up period (Fig. 1b) (please note that standardized RTs are presented in this figure. With respect to the absolute values, for those who later develop depression or experience symptoms of anhedonia the group mean of RT Happy is still lower than the group mean of RT Sad.)

## Prediction of symptoms of anhedonia and sadness by facial emotion identification RTs (Table 2, right side)

Facial emotion identification RTs did not predict anhedonia when each facial emotion was tested separately. Backward conditional logistic regression analysis resulted in a multi-emotion model in which anhedonia was predicted by longer RTs for happiness and shorter RTs for sadness (see also Fig. 1b), of which both RT Happy and RT Sad reached statistical significance. Again, additional analyses revealed that the presence of RT Happy in the multi-emotion model relied specifically on the inclusion of RT Sad and the other way around. Excluding either of them from the model resulted in no longer finding any effect with backward model selection.

No significant associations were found between facial emotion identification RTs and later symptoms of sadness, neither when emotion RTs were tested separately, nor when they were tested in a multi-emotion model.

In Table 2 symptoms of anhedonia were corrected for symptoms of sadness and the other way around. Additional analyses without correcting for anhedonia and sadness showed no important differences in results (for exact p-values and ORs see Supplement 3).

## Multiple test correction

After FDR correction for multiple testing, the effect of RT Happy on depressive disorder was no longer significant ($p_{FDR\text{-}corrected}$=.28). Effects of RT happy ($p_{FDR\text{-}corrected}$=.03) and RT Sad ($p_{FDR\text{-}corrected}$=.03) on symptoms of anhedonia remained significant after FDR correction.

## Post hoc analyses (Table 2, bottom)

Post hoc analyses were performed in an attempt to explain the differences in results between emotions tested in separate models and in one multi-emotion model. Combined, the results of the single- and multi-emotion models (see Table 2; Fig. 1a and 1b) suggest that the intra-individual contrast between RT Happy and RT Sad is more relevant to predict onset of depressive disorder and symptoms of anhedonia than each RT individually.

To test the plausibility of this explanation, we constructed the variable: Happy/Sad specialization (HS) = RT Sad - RT Happy. A value of 0 indicated no Happy/Sad specialization, i.e., participants responded equally fast to happy and sad emotions. Positive values indicated a specialization in identifying happiness (in the sense of being faster) and negative values a specialization in identifying sad faces (again, in the sense of being faster). Subsequently, we checked whether the standardized new variable predicted depressive disorder and symptoms of anhedonia. As in previous analyses, we adjusted for relevant EPs, gender and age at the time of the IFE task. Table 2 (bottom) shows the results of the post hoc analyses. HS was found to predict symptoms of anhedonia significantly. The ORs of HS were below 1 which means that more specialization towards identifying happy faces was associated with decreased risk of developing anhedonia. The same pattern was found for depression, but HS did not reach statistical significance at α=.05.

## The effects of excluding SP and GAD participants

After excluding participants with lifetime diagnoses of SP or GAD, the effect of RT Happy on depression became slightly stronger and the effect of RT Sad weakened compared to the findings based on the complete sample. Regarding symptoms of anhedonia and sadness, excluding participants with SP or GAD diagnoses did not result in different patterns compared to the ones found in the whole sample, except for slightly stronger effects for anhedonia. For specific p-values and ORs of the analyses after excluding SP or GAD participants, see Supplement 4.

## Gender differences

We did not find any significant RT Happy*gender, RT Sad*gender, RT Angry*gender or RT Fearful*gender effects for depression, anhedonia or sadness in the single-emotion models.

### Additional sensitivity analyses

Adjusting for baseline speed did not change results (data not presented). Adjusting for cohort status did not make a difference to the effect sizes, i.e., the ORs remained virtually the same, but the decrease in power was reflected in slightly higher *p*-values (see Supplement 5).

## DISCUSSION



The aim of this study was to examine whether facial emotion identification in early adolescence predicts onset of depressive disorder, whether it differentially predicts symptoms of anhedonia and sadness, and whether it does so in a symptom-congruent and content-specific way. Because of the exploratory nature of this study our main focus was on patterns rather than single results. Our results provide tentative evidence in favor of the hypothesis that facial emotion identification in early adolescence predicts onset of depressive disorder and symptoms of anhedonia within eight-year follow-up. In support of the hypothesis of symptom congruence, both risk of depressive disorder and risk of anhedonia were associated with slower identification of happy emotions; risk of anhedonia was also associated with faster identification of sad emotions. However, we found no evidence for the hypothesis that symptoms of sadness are predicted by faster identification of sad emotions. In favor of the content-specificity hypothesis, identification of angry or fearful emotions predicted neither onset of depressive disorder, nor symptoms of anhedonia or sadness.

Our prospective findings suggest that facial emotion identification bias may be a symptom-congruent trait marker for depressive disorder and anhedonia. These associations were only found when considering multi-emotion models or Happy/Sad specialization. It seems primarily relevant how fast young adolescents identify happy facial emotions compared to how fast they identify sad emotions. Our results suggest that those who identify sad expressions faster than happy ones or are only relatively faster in identifying happiness seem more prone to developing depression or symptoms of anhedonia. Although largely similar patterns were found for depressive disorder and anhedonia, effects for depressive disorder did not meet multiple test correction criteria, and should therefore be interpreted with caution. Not finding results for symptoms of sadness implies that facial emotion identification is not a trait marker for sadness, but it could still be a symptom-congruent *state* marker for sadness. The content-specificity of the associations found implies that young adolescents' identification of happy and sad facial emotions is more relevant for onset of depressive disorder and symptoms of anhedonia than identification of facial anger and fear.

Effects of facial emotion identification on depressive disorder seem to be mainly carried by symptoms of anhedonia and not by symptoms of sadness. Happy faces are naturally strongly rewarding stimuli. Since identifying happy faces less fast relative to sad ones seems to predict

depression and anhedonia but not sadness, we propose that the mechanism underlying this vulnerability might be related to the functioning of the reward system. As mentioned in the introduction, symptoms of anhedonia and symptoms of sadness have been associated with partly different psychophysiological systems: the so-called approach (reward-related) and avoidance systems (Carver, 2006; Ernst & Fudge, 2009). Whereas the reward system is assumed to be an important underlying mechanism of the development of anhedonia, there is less evidence that it is also directly involved in the development of sadness. Being able to identify happy emotions much faster than sad ones may point at a more reactive reward system, whereas small differences between identifying these emotions, or even being faster in identifying sad emotions, suggest a more passive reward system. One of the mechanisms underlying depression and anhedonia may be an impaired tendency to modulate behavior as a function of prior rewarding experiences (Pizzagalli et al., 2009, 2008). Whereas individuals without vulnerability for depression have a tendency to approach rewarding stimuli (e.g., happy faces), those with a vulnerability for depression and anhedonia may not have developed this inclination. Blunted responsiveness to rewarding stimuli could contribute to loss of interest in the environment and in this way contribute to the onset of depression (Pizzagalli, Jahn, & O'Shea, 2005). A bias towards sad faces and away from happy ones may also contribute to onset of depression via ineffective emotion regulation strategies. Attitudes of dwelling on negative feelings and avoiding positive cues that could help to overcome negative experiences have been found to predict depression (Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). This pathway is supported by recent findings that adolescents between ages 9 and 14 with ruminating response styles, who were unable to disengage from self-referential negative thoughts, were also characterized by an attentional bias away from happy faces (Hilt & Pollak, 2013).

We found no evidence for gender differences in the relation between facial emotion identification and later onset of depressive disorder or symptoms of anhedonia and sadness, but did find small differences in results depending on whether adolescents with lifetime social or generalized anxiety were in- or excluded. When anxious adolescents were excluded, the predictive value of facial emotion identification for depressive disorder and symptoms of anhedonia slightly strengthened. We are unable to explain the (small) differences in results and can only speculate that social or generalized anxiety disorder, which is correlated with depression and anhedonia, might be reversely associated with facial emotion identification. Including adolescents with social or generalized anxiety disorder may, therefore, have slightly concealed the association between facial emotion identification and risk of depression and anhedonia.

Overall our findings are consistent with the previous studies in which symptom-congruent and content-specific associations between facial emotion processing biases and depression were reported (Bourke, Douglas, & Porter, 2010; Joormann & Gotlib, 2006; Joormann et al., 2007), but inconsistent with many other studies. This is not surprising since, as was already mentioned in the introduction, results of previous studies were quite diverse. Because many different measures

for facial emotion identification and depression were used, it is difficult to compare our results to theirs. Furthermore, most previous studies looked at cross-sectional associations between facial emotion identification and depression; studies investigating facial emotion identification as a potential trait marker for depression are scarce. Our finding that the identification speed pattern across multiple emotions seems to be more relevant than identification speeds of individual emotions cannot be interpreted in relation to prior evidence, because – with the exception of only a few studies focusing on positive-negative bias (Oldehinkel, Hartman, Oort, & Nederhof, 2015; Wright et al., 2009) – this field has remained largely unexplored so far. Hence, awaiting further support our conclusions are tentative.

Our study has several strengths. We are one of the very few who focused on predicting onset of depressive disorder and symptoms rather than on cross-sectional associations. This is the only way to shed light on the possibility of facial emotion identification differences as trait marker of depression or depressive symptoms, which could have implications for treatment and prevention. Our study has a large sample size and a follow-up period of 8 years, and depression and symptoms of anhedonia and sadness were assessed by means of standardized diagnostic interviews. The large sample size enabled us to test all associations regardless of comorbid anxiety and after exclusion of adolescents with lifetime social or generalized anxiety disorder, hereby ensuring internal validity as well as ecological validity. To our best knowledge, we were the first to examine symptoms of anhedonia and sadness separately in relation to facial emotion identification, and also the first to explore the possibility of Happy/Sad facial emotion identification specialization as predictor of depression, anhedonia and sadness.

However, our study is not without limitations. First of all, the inconsistent results reported in earlier studies, as well as the suppressor effects found in our own study, called for consideration of a large number of interrelated hypotheses and various (post hoc) options to find novel patterns. Suppressor effects, i.e., that single predictors were only significant if other predictors were included in the model as well, occur more frequently when predictors are highly correlated and are mainly due to large standard errors of the estimates, suggesting that suppression situations may be less replicable and therefore caution is needed (Tzelgov & Henik, 1991). We performed post hoc analyses in an attempt to explain differences in results between emotions tested separately and when tested together in one model, in the hope to demystify our suppressor findings. Although the post hoc analyses did provide clues to an explanation, the problem of capitalization on chance should be noted. Second, all effect sizes reported in this study are small. Although it can be argued that the effect sizes are reasonable considering the longitudinal nature of our study and the fact that we corrected for early depressive diagnoses and symptoms, it is difficult to estimate the clinical relevance of our findings. Third, symptoms of anhedonia as well as symptoms of sadness were measured by single items from the CIDI screening list. The CIDI is not a survey but a structured diagnostic interview allowing for elaborate explanations and answers, but still the scores are based on one single question concerning anhedonia and

one single question regarding sadness. Finally, the CIDI was assessed retrospectively at age 19. Although specific interview strategies were developed for the CIDI to reduce recall inaccuracies, e.g., decomposing questions and using specific life course events as point of reference (Kessler & Üstün, 2004), the possibility of recall biases cannot be excluded, especially for the earliest diagnoses and symptoms.

Further research is needed on many levels. First of all, our findings need to be confirmed by replication in other samples because of multiple testing and suppressor effects. Second, more research is needed into whether the results of this study can be generalized to older age groups and to clinical populations, since our sample consisted largely of young adults with mild psychiatric problems at the most. Related to this, participants indicated whether they had experienced anhedonia and sadness for 'a period lasting several days or longer', without additional information on the severity of the symptoms and on how many days they had experienced those symptoms. Although there is evidence that subclinical symptoms of anhedonia and sadness predict adult MDD (Pine et al., 1999; Wilcox & Anthony, 2004), more research is needed to determine to what extent mild symptoms of anhedonia and sadness belong to normative adolescent development and to what extent they reflect characteristics that have predictive value for psychopathology. Third, in the facial emotion identification task used in this study participants were asked to judge full intensity facial emotions. It has been argued that, for the sake of ecological validity, more subtle emotions should also be taken into account, for example by using the so-called morphing tasks which show movies of neutral faces gradually changing into full intensity facial emotions. Fourth, our findings call for more focus on intra-individual multi-emotion patterns of facial emotion identification, a so far largely unexplored field of study. Furthermore, more research into possible underlying mechanisms is needed, e.g., reward responsiveness. As depression and anhedonia have already been linked to lower activity levels in reward-related brain areas during different stages of reward processing (Liu, Hairston, Schrier, & Fan, 2011; Pizzagalli et al., 2009; Redlich et al., 2015), a potentially viable direction for future research would be to determine whether emotion identification biases can also be linked to different responses in reward-related brain areas using fMRI methods. Finally, because of the small effect sizes found in our study, training facial emotion identification biases may not, in general, be expected to be efficient if used for treatment and prevention purposes. Training may, however, be effective if limited to adolescents with severe biases. Preliminary positive effects resulted from a randomized controlled trial of training the perception of happiness over sadness in ambiguous facial expressions (Penton-Voak et al., 2012), but more research is needed.

To conclude, from the perspective of contemporary cognitive neuropsychological models of depression according to which emotion processing biases are present before mood starts to deteriorate and mood is only enhanced via changing these emotion processing biases, measuring emotion processing bias and trying to modify these biases is of essence. Our findings point at a rather complex picture in which how fast young adolescents are able to identify happy

facial emotions compared to how fast they identify sad emotions predicts onset of depression and symptoms of anhedonia (but not sadness) within a time frame of 8 years. A possible underlying mechanism could be a less reactive reward system.

# SUPPLEMENTARY MATERIAL

## Supplement 1

Testing procedures of the Amsterdam Neuropsychological Tasks program (ANT) and the 'Identification of Facial Expressions' (IFE) task.

The ANT was assessed in separate rooms at participants' schools or, if no separate rooms were available in participants' schools, at nearby community centers by trained undergraduate psychology students. To ensure that participants understood the instructions and kept in mind that both speed and accuracy of performance were of essence, practice trials were run prior to the test trials and screenshots of the tasks were explained verbally. The ANT started with a test to press a button as soon as a square was presented on the screen, in order to determine participants' baseline speed in simple cognitive decisions, and was followed by six other tasks (for a detailed description of the ANT in TRAILS, see Brunnekreef and colleagues (2007)).

The IFE task, which was the last task of the ANT, was used to measure participants' capacity to identify different facial emotions (De Sonneville, 1999). Six different emotions were tested (happiness, sadness, anger, fear, disgust and surprise), each in a separate subtask by means of 40 trials, 20 trials in which the target emotion was presented and 20 nontarget trials in which a random selection of the other emotions was presented (hence, in total 240 trials). Each subtask consisted of a random sequence of target and non-target trials. Stimuli consisted of digitized high quality color photographs of four adult faces (two men and two women, frontal views) showing distinct expressions of one of the following six emotions: happiness, sadness, anger, fear, disgust and surprise (see Fig. S1 for examples of expressions of the emotions happiness and anger).

**FIGURE S1.** Examples of facial emotion expressions in the 'Identification of Facial Expressions' (IFE) Task. The left face expresses happiness and the right face expresses anger.

Each of the six emotion subtasks started with an instruction and practice session. Participants were shown an example of the target emotion and were instructed to focus on the target emotion and to respond whether the face showed the target emotion or not by clicking a mouse button. Participants had to press the 'yes' button (right button for right-handed participants and left button for left-handed participants) if the picture presented on a computer screen matched the target emotion and 'no' (left button for right-handed participants and right button for left-handed participants) if it did not. After the instruction, eight practice trials followed. Each subtask was preceded by showing a picture of the target emotion and each trial started with a presentation of a fixation cross for 500 ms, after which a stimulus face was presented on the screen until the participant pressed the 'yes' or the 'no' button. The inter-stimulus presentation time was 1000 ms and because of the variable presentation time of the stimuli, inter-trial intervals were also variable. Responses slower than 8000 ms were regarded invalid, as well as responses faster than 250 ms, which were interpreted as accidental responses. If participants did not press either of the answer buttons within 8000 ms or responded faster than 250 ms the trial was replaced by a new similar trial. The valid responses were coded as follows: 'hit' = correct 'yes' response; 'correct rejection' = correct 'no' response; 'miss' = no 'yes' response on target signal; 'false alarm' = 'yes' response on nontarget signal'. All participants were offered all emotion subtasks. Accidentally, participants were unwilling or unable to complete all subtasks, for instance because they did not understand the emotion that was tested.

## Supplement 2

Descriptives for Reaction Times and Error Proportions at Age 11 by Diagnostic Group Between Age 11 and 19

| | Mean (*SD*) | | | |
|---|---|---|---|---|
| **Variables** | **Depression[a] N = 342-344** | **Anhedonia[b] N = 614-617** | **Sadness[c] N = 758-764** | **Healthy[d] N = 797-800** |
| RT Happy[e] | 887 (216) | 881 (203) | 879 (208) | 871 (206) |
| RT Sad[e] | 1197 (284) | 1196 (273) | 1211 (302) | 1207 (278) |
| RT Angry[e] | 1117 (248) | 1111 (257) | 1115 (262) | 1115 (258) |
| RT Fearful[e] | 1100 (266) | 1112 (278) | 1116 (281) | 1101 (273) |
| EP Happy[f] | 3.0 (2.9) | 3.1 (3.3) | 3.2 (3.4) | 3.4 (3.3) |
| EP Sad[f] | 12.1 (8.8) | 12.1 (8.8) | 12.7 (9.1) | 12.8 (9.4) |
| EP Angry[f] | 8.1 (6.0) | 8.1 (5.9) | 8.2 (6.2) | 8.5 (6.0) |
| EP Fearful[f] | 7.2 (6.2) | 7.2 (6.4) | 7.4 (6.5) | 7.6 (6.5) |

[a] CIDI-based DSM-IV diagnosis of major depressive disorder, minor depressive disorder or dysthymia, with age of onset between 11 and 19;
[b] Symptoms of anhedonia for at least several consecutive days between age 11 and 19;
[c] Symptoms of sadness for at least several consecutive days between age 11 and 19;
[d] No depressive disorder or symptoms of anhedonia or sadness;
[e] RT = mean reaction time for correct responses measured in milliseconds, assessed at age 11
[f] EP = mean error proportion, assessed at age 11

## Supplement 3

Results Logistic Regression Analyses of DSM-IV Depression and Symptoms of Anhedonia and Sadness for at Least Several Days Between Age 11 and Age 19 on Facial Emotion Identification Reaction Times at Age 11 – **Without Correcting for Sadness and Anhedonia**

| | | **Anhedonia[a] *without* sadness as covariate N=1732-1773** | | **Sadness[b] *without* anhedonia as covariate N=1748-1793** | |
|---|---|---|---|---|---|
| | | **OR** | **P** | **OR** | **P** |
| Emotions tested separately | RT Happy | 1.04 | .50 | 1.00 | .95 |
| | RT Sad | 0.94 | .24 | 1.02 | .74 |
| | RT Angry | 0.97 | .56 | 0.99 | .80 |
| | RT Fearful | 1.02 | .69 | 1.03 | .53 |
| Models backward selection | RT Happy | 1.17 | .03 | | |
| | RT Sad | 0.84 | .02 | | |
| | RT Angry | | | | |
| | RT Fearful | | | | |
| Post hoc analyses | RT Sad - RT Happy (HS) | 0.88 | .02 | Not tested | Not tested |

[a] Symptoms of anhedonia for at least several consecutive days;
[b] Symptoms of sadness for at least several consecutive days;
All effects were adjusted for error proportions, gender and age at the time of the facial emotion task;
OR = odds ratio; RT = mean reaction time for correct responses; all RTs in this table are standardized (Z-values) with one exception: HS was calculated on unstandardized RT Sad and RT Happy and was standardized afterwards

## Supplement 4

**Sample Without Participants With Social Phobia or Generalized Anxiety Disorder -** Results Logistic Regression Analyses of DSM-IV Depression and Symptoms of Anhedonia and Sadness for at Least Several Days Between Age 11 and Age 19 on Facial Emotion Identification Reaction Times at Age 11

| | | **Depression[a] N=1501-1540** | | **Anhedonia[b] with sadness as covariate N=1434-1469** | | **Sadness[c] with anhedonia as covariate N=1434-1469** | |
|---|---|---|---|---|---|---|---|
| | | **OR** | **P** | **OR** | **P** | **OR** | **P** |
| Emotions tested separately | RT Happy | 1.15 | .06 | 1.06 | .36 | 1.01 | .93 |
| | RT Sad | 1.04 | .64 | 0.91 | .14 | 1.11 | .08 |
| | RT Angry | 1.06 | .42 | 0.97 | .62 | 1.02 | .70 |
| | RT Fearful | 1.06 | .42 | 1.02 | .70 | 1.10 | .11 |
| Multi-emotion models | RT Happy | 1.24 | .03 | 1.28 | .004 | Not tested | Not tested |
| | RT Sad | 0.90 | .30 | 0.77 | .002 | Not tested | Not tested |
| Post hoc analyses | RT Sad - RT Happy (HS) | 0.92 | .25 | 0.82 | .002 | Not tested | Not tested |

[a] CIDI-based DSM-IV diagnosis of major depressive disorder, minor depressive disorder or dysthymia;
[b] Symptoms of anhedonia for at least several consecutive days;
[c] Symptoms of sadness for at least several consecutive days;
All effects were adjusted for error proportions, gender and age at the time of the facial emotion identification task;
OR = odds ratio; RT = mean reaction time for correct responses; all RTs in this table are standardized (Z-values) with one exception: HS was calculated on unstandardized RT Sad and RT Happy and was standardized afterwards

## Supplement 5

Results Logistic Regression Analyses of DSM-IV Depression and Symptoms of Anhedonia and Sadness for at Least Several Days Between Age 11 and Age 19 on Facial Emotion Identification Reaction Times at Age 11, **adjusted for cohort status**

| | | **Depression[a] N=1803-1815** | | **Anhedonia[b] with sadness as covariate N=1716-1727** | | **Sadness[c] with anhedonia as covariate N=1718-1727** | |
|---|---|---|---|---|---|---|---|
| | | **OR** | **P** | **OR** | **P** | **OR** | **P** |
| Emotions tested separately | RT Happy | 1.07 | .29 | 1.04 | .44 | 0.97 | .56 |
| | RT Sad | 0.98 | .69 | 0.93 | .18 | 1.04 | .42 |
| | RT Angry | 1.02 | .81 | 0.99 | .79 | 0.97 | .59 |
| | RT Fearful | 0.97 | .58 | 1.01 | .82 | 1.04 | .48 |
| Multi-emotion models | RT Happy | 1.18 | .06 | 1.21 | .015 | Not tested | Not tested |
| | RT Sad | 0.87 | .12 | 0.81 | .008 | Not tested | Not tested |
| Post hoc analyses | RT Sad - RT Happy (HS) | 0.90 | .11 | 0.86 | .007 | Not tested | Not tested |

[a] CIDI-based DSM-IV diagnosis of major depressive disorder, minor depressive disorder or dysthymia;
[b] Symptoms of anhedonia for at least several consecutive days;
[c] Symptoms of sadness for at least several consecutive days;
All effects were adjusted for error proportions, gender, age at the time of the facial emotion identification task, and cohort status; OR = odds ratio; RT = mean reaction time for correct responses; all RTs in this table are standardized (Z-values) with one exception: HS was calculated on unstandardized RT Sad and RT Happy and was standardized afterwards

# CHAPTER 4

## Reward-related attentional bias at age 16 predicts onset of depression during nine years of follow-up

Published as:
Vrijen, C., Hartman, C. A., & Oldehinkel, A. J. (2019). Reward-Related Attentional Bias at Age 16 Predicts Onset of Depression During Nine Years of Follow-Up. *JAACAP, 58*(3), 329-338. https://doi.org/10.1016/j.jaac.2018.06.009

## ABSTRACT

**Objective.** This study investigated whether low reward responsiveness marks vulnerability for developing depression in a large cohort of never-depressed 16-year-old adolescents who completed a reward task and were subsequently followed for 9 years, during which onset of depression was assessed.

**Method.** Data were collected as part of the TRacking Adolescents' Individual Lives Survey (TRAILS), an ongoing prospective cohort study. Reward responsiveness was assessed by the Spatial Orienting Task at 16 years and depression was assessed at 19 years by the World Health Organization Composite International Diagnostic Interview and at 25 years by the Lifetime Depression Assessment Self-report. Participants who completed the reward task at 16 years, had no previous onset of depression, and were assessed on depression onset at 19 and/or 25 years were included in the present study (*n*=531; 81 became depressed during follow-up).

**Results.** Difficulties in shifting attention from expected non-reward to expected reward and from expected punishment to expected non-punishment at age 16 predicted depression during follow-up. This was only found at an automatic level of information processing.

**Conclusions.** The findings suggest that decreased reward responsiveness at 16 years marks vulnerability for depression. Prevention programs may aim at enhancing at-risk adolescents' responsiveness to cues for potential rewards, particularly in situations in which they are focused on negative experiences.

## INTRODUCTION

There is compelling evidence that depressed individuals and those with a high familial risk for depression are less responsive to rewards than healthy and low-risk individuals (Forbes & Dahl, 2012; Luking et al., 2016; McCabe & Gotlib, 1995; Pizzagalli et al., 2008). This is reflected in decreased activity in reward-related brain areas (Forbes & Dahl, 2012; Luking et al., 2016), decreased hedonic experience of rewards (American Psychiatric Association, 2013), decreased reward learning (Pizzagalli et al., 2008), and decreased attention to rewards (McCabe & Gotlib, 1995). There also are first indications that low reward responsiveness predicts an increase in future depressive symptoms (Forbes et al., 2007; Morgan et al., 2013; Nelson et al., 2016; Rawal et al., 2013; Telzer et al., 2014) and first onset of depressive disorder in adolescents (Bress et al., 2013; Forbes et al., 2007; Nelson et al., 2016; Pan et al., 2017; Rawal et al., 2013; Stringaris et al., 2015). Particularly the prospective association between reward responsiveness and onset of depressive disorder is important to investigate further because of its high clinical relevance and possible implications for prevention. The evidence of such a prospective association is based on only small groups of individuals (ranging from *n*=3 (Forbes et al., 2007) to *n*=44 (Pan et al., 2017)) who were healthy at baseline and depressed at follow-up. The evidence seems largely consistent, with the exception of 2 studies in which opposite effects were reported, that is, increased activity in reward-related brain areas predicted onset of depression (Pan et al., 2017; Telzer et al., 2014).

Our main aim was to investigate whether the idea of low reward responsiveness as a vulnerability marker for subsequent depression would hold up in a large sample of never-depressed 16-year-old adolescents who completed a task that assessed reward-related attentional biases and were subsequently followed for 9 years to assess onset of depressive disorder. We focused on reward responsiveness at the level of attentional processes, because these constitute the first filtering of information and likely contribute to the overall negative cognitive processing biases that characterize depression (Derryberry & Reed, 2002). We investigated how reward expectancies modified attention, and whether low modification of attention by reward expectancies predicted onset of depression. There is evidence that not only reward processing but also punishment processing could be altered in individuals with depression. Depressed individuals have been found to be more sensitive to negative feedback and to show impaired functioning after negative feedback (Elliott et al., 1997; Eshel & Roiser, 2010), and responses to reward and punishment may also interact (Forbes & Dahl, 2012). Therefore, our secondary aim was to investigate how punishment expectancies modify attention, and whether a strong modification of attention by punishment predicts onset of depression. We hypothesized that decreased attentional engagement toward expected reward and non-punishment and increased attentional disengagement from expected reward and non-punishment at 16 years of age would predict onset of depression between 16 and 25 years of age. Automatic and voluntary attentional processes were explored, as was the specificity of findings for depression as opposed to other psychiatric problems (e.g., anxiety).

It is particularly relevant to investigate onset of depression in mid- to late adolescence because this developmental period is marked by a strong increase in incidence of depression (Merikangas et al., 2010), and adolescent onset sets the stage for severe and recurrent depression and impaired academic and interpersonal development in adulthood (Kovacs & Goldston, 1991; Lewinsohn, Rohde, Klein, & Seeley, 1999). Normal adolescent development is characterized by large changes in the reward system (Galvan, 2010) and a peak in reward responsiveness at approximately mid-adolescence (Braams, Duijvenvoorde, Peper, & Crone, 2015). It has been suggested that depression-related differences in reward function might be most pronounced at the time of this peak (Forbes & Dahl, 2012). Elucidating mid-adolescent reward-related attentional biases that predict depression could inform the design of prevention programs to modify these biases already in adolescence and promote positive psychosocial and academic development of at-risk youth.

## METHOD

### Sample and procedure

The data were collected in a subsample of the TRacking Adolescents' Individual Lives Survey (TRAILS), an ongoing prospective cohort study investigating mental health and social development from early adolescence into adulthood. The TRAILS study was approved by the Dutch Central Committee on Research Involving Human Subjects, participants were treated in accordance with the Declaration of Helsinki, and written consent was acquired from all adolescents and their parents.

Starting at approximately 11 years of age, the TRAILS participants have been assessed every 2 to 3 years. They were recruited from primary schools (response rate 90%) in 5 municipalities in the northern region of the Netherlands. Of all eligible children, 2,230 (76%) agreed to participate at T1. Details on the selection procedure and an overview of all measures used have been described elsewhere (De Winter et al., 2005; Oldehinkel, Rosmalen, et al., 2015). For the present study, we used data from the third (T3; mean age 16.3 years, *SD* 0.7), fourth (T4; mean age 19.1 years, *SD* 0.6), and sixth (T6; mean age 25.7, *SD* 0.6) waves (Oldehinkel, Rosmalen, et al., 2015). At T3 (September 2005 to December 2007), 1,816 of the original 2,230 adolescents participated again (retention rate 81%) (Oldehinkel, Rosmalen, et al., 2015), 744 of whom were invited for a series of laboratory tasks in addition to the usual assessments, and 715 (96.1%) agreed to participate. Participants with a high-risk profile were oversampled for the laboratory experiments; 66.0% were characterized by a difficult temperament, lifetime parental psychopathology, or living in a single-parent family. The remaining 34.0% were randomly selected from the TRAILS participants without any of the three risk factors.

We selected all participants of the T3 laboratory tasks (*n*=715) who: (1) completed the reward-related attention bias task, that is, the Spatial Orienting Task (SOT), at T3 with less than 25% outliers (excluded *n*=2); (2) had been subjected to the World Health Organization Composite International Diagnostic Interview (CIDI) at T4 or completed the Lifetime Depression Assessment Self-report (LIDAS) at T6 (excluded *n*=64); (3) did not meet the criteria for bipolar disorder or hypomania, as unipolar and bipolar depressive disorders were expected to be associated with different reward biases (excluded *n*=27); and (4) had not had a depressive disorder (i.e., major depressive disorder or dysthymia) during or before taking the SOT at T3 (excluded *n*=91). This yielded a sample of 531 participants (74.3% of the total cohort participating in T3 laboratory tasks). Figure S1 of the Supplementary Material presents a flowchart of the sampling procedure. For a more detailed description of the selection for the laboratory tasks, see Supplement 1 and Table S1 of the Supplementary Material.

A power analysis for logistic regression (Faul et al., 2009) with a one-sided alpha value set to .05, 531 included participants, a proportion of prospective depressed participants of 0.153, and power set to 0.8 yielded the possibility of finding effects of approximately 35% increased risk of developing depression per *SD* increase in the predictor variable (see Figure S2 of the Supplementary Material).



## Measures

### *Spatial orienting task (SOT)*

The SOT was programmed to be similar to the SOT developed and described by Derryberry and Reed (2002). The task consisted of 4 positive and 4 negative games. During positive games, fast responses resulted in the gain of points; slow responses did not change the score. During negative games, slow responses resulted in the loss of points; fast responses did not change the score. Fast and slow scores were determined relative to participants' own performance (see Supplement 1). Positive games were used to investigate attentional bias to expected reward, and negative games to investigate attentional bias to expected non-punishment.

During each game, two vertical black bars which were displayed against a white background marked the location of the cues and targets, and the score was presented in black at the center of the screen. Participants were instructed to fixate on the score, which was updated after each response, and to avoid moving their eyes. Each trial started with turning off the fixation score for 200 ms and subsequently turning it on again for 250 ms, after which a cue arrow replaced one of the two vertical black bars. The cue arrow served the purpose of orienting participants' attention to one of the two peripheral locations. After a short (250 ms) or long (500 ms) delay, a target, that is, a small vertical gray rectangle, appeared, either centered within the cue arrow (a so-called cued target, see Figure s3a of the Supplementary Material) or centered within the vertical black bar on the other side of the fixation score (an uncued target, see Figure s3b of the Supplementary Material). Participants were informed that a blue-up arrow (easy cue) signaled

that a target appearing in that (cued) location would be easy and a target appearing in the uncued location would be hard. A red-down arrow (hard cue) signaled that a target appearing in the cued location would be hard, and a target appearing in the uncued location would be easy. Participants were also informed that 2/3 of the targets would appear in the location of the cue arrow, and that occasionally no target would appear (catch trials). They were instructed to press the 'b' key on the keyboard as soon as they detected the target and were warned that pressing the key before the target appeared or when no target appeared would result in a loss of 10 points. Five-hundred ms after the 'b' was pressed, or for catch trials 1s after the delay interval, the cue arrow and target were replaced by the two black bars, and a feedback arrow was presented below the centered score. A blue-up arrow indicated a fast response on target trials or a correct non-response on catch trials and a red-down arrow indicated a slow response on target trials or an incorrect response on catch trials. To increase the relevance of the scores and boost the participants' motivation, they were informed that a prize (e.g., a balloon ride) would be awarded to those with the highest scores on the positive games, and that very low scores on the negative games could result in having to start over again until performance was sufficient. For a more detailed description and schematic overview of the SOT, see Supplement 1, Figure S3, and Tables S2, S3, and S4 of the Supplementary Material.

Following Van Hemel-Ruiter and colleagues (2013), attentional bias for reward was operationalized as a relatively faster engagement toward reward and a relatively slower disengagement from reward, that is, (1) faster responses at locations of expected reward than at locations of expected non-reward (faster engagement toward reward) and (2) slower reshifting of attention from expected reward to expected non-reward locations than from expected non-reward to expected reward locations (slower disengagement from reward). Attentional bias for non-punishment was operationalized in a similar way, that is, (1) faster responses at locations of expected non-punishment than at locations of expected punishment (faster engagement toward non-punishment) and (2) slower reshifting of attention from expected non-punishment to expected punishment locations than from expected punishment to expected non-punishment locations (slower disengagement from non-punishment). Separate engagement and disengagement scores were calculated for short (250 ms) and long (500 ms) delays between cues and targets. The short-delay trials tap into relatively automatic and implicit attentional responses and the long-delay trials into more voluntary and explicit attentional responses. Table S5 of the Supplementary Material presents an overview of the calculations of all attentional engagement and disengagement scores used in the statistical analyses. We note that it is not possible to distinguish between difficulties with disengaging from reward and difficulties with shifting toward expected non-reward in the disengagement condition; because there is no neutral condition in the SOT these are 2 sides of the same coin (Derryberry & Reed, 2002).

***Depressive disorder and other psychiatric diagnoses***

The World Health Organization Composite International Diagnostic Interview (CIDI) version 3.0 (Kessler & Üstün, 2004), assessed at T4, and the Lifetime Depression Assessment Self-report (LIDAS; Bot et al., 2017), assessed at T6, were used to determine first onset of a depression after T3, which was operationalized as a lifetime major depressive disorder or dysthymia with age at onset older than at T3. The CIDI and the LIDAS depression diagnoses were determined according to the DSM-IV criteria. For information about the reliability, validity and agreement of the CIDI and the LIDAS, see supplement 1 of the Supplementary Material. For depression according to both the CIDI and LIDAS, the earliest age at onset was used to exclude participants with an age at first onset younger than or equal to age at T3. Lifetime T4 CIDI diagnoses of bipolar disorder or hypomania, and the T6 LIDAS item about ever having been diagnosed with bipolar disorder by a professional were used to exclude participants with lifetime bipolar disorder or hypomania. Other CIDI and LIDAS diagnoses were used for the sensitivity analyses described in the Statistical Analysis section.



***Covariates***

Family socioeconomic position was computed by taking the average score on standardized family income, educational level of the father and mother, and occupational level of the father and mother, assessed at T1. The affective problems scale (13 items) of the Youth Self-Report (Achenbach & Rescorla, 2001) (YSR) was used to assess symptoms of depression at T3.

## Statistical Analysis

Reaction times (RT s) were standardized to mean 0 and SD 1 in order to be able to compare Odds Ratios (*OR*s) across different attentional bias conditions and different diagnostic groups. Using SPSS version 25.0, we performed a series of logistic regression analyses to test the hypotheses that decreased engagement toward expected reward and non-punishment, and increased disengagement from expected reward and non-punishment at 16 years of age would predict onset of depression between 16 and 25 years of age. Because both hypotheses concern effects in a specific direction, 1-sided tests were used (Cho & Abe, 2013; Greenland et al., 2016), that is, for effects in the expected direction 1-sided $p$ = 2-sided $p$ from the SPSS output / 2; for effects in the unexpected direction 1-sided $p$ = 1 - (2-sided $p$ / 2). (The original output is available at https://osf.io/zvw5d/.) First, separate models were tested for reward and non-punishment games, followed by a model including both. All effects were adjusted for gender, age at the time of the SOT, and family socioeconomic position. To correct for multiple tests, we used the classical False Discovery Rate (FDR) method, resulting in an FDR-derived significance threshold of 0.0125 (Benjamini & Hochberg, 1995). (For further details, see supplement 1 and Table S6 of the Supplementary Material.) Results were interpreted as significant only for $p$-values below this threshold.

***Sensitivity analyses***

Because depression is highly comorbid with other psychiatric problems, differences between depressed and non-depressed individuals may be explained by other psychiatric problems. We tested whether the effects found were specific to depression by repeating the analyses after excluding all individuals with a lifetime separation anxiety disorder, agoraphobia, generalized anxiety disorder, obsessive compulsive disorder, panic disorder, social phobia, specific phobia, attention deficit disorder, oppositional defiant disorder, or conduct disorder (sensitivity check 1). Furthermore, although we already excluded participants with clinical depression prior to or during the SOT, our findings could still be driven by subclinical depressive symptoms at the time of the SOT (T3) rather than by a prospective association. Therefore, analyses were repeated while adjusting for scores on the YSR depression scale assessed at T3 (sensitivity check 2). In addition, for 89 participants without an onset of depressive disorder at T4, it was unknown whether they developed depression from T4 to T6. We checked whether the findings still held after excluding this subsample (sensitivity check 3). Then, we repeated the main analyses without adjusting for gender, age and socioeconomic position (sensitivity check 4), and tested each of the eight effects in separate univariate models (sensitivity check 5).

***Open science***

Data and syntax have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/zvw5d/.

# RESULTS

## Descriptive statistics

Adolescents who developed a depressive disorder between 16 and 25 years of age were more likely to be female, have a lower socioeconomic family status, and report more (subclinical) depressive symptoms and other lifetime psychiatric problems at 16 years than their peers who did not develop a depressive disorder (see Table 1).

## Task descriptives

As presented in Table 1, engagement to reward and non-punishment had a positive value for all groups, which means that all groups showed more engagement to reward than to non-reward and more engagement to non-punishment than to punishment. In voluntary trials, difficulty to disengage from reward and non-punishment also showed positive values for all groups, that is, participants in general had more difficulties disengaging from expected reward and non-punishment than from expected non-reward and punishment in trials where they could voluntarily control their attention. On a more automatic level of processing, the depression

**TABLE 1.** Descriptive Statistics of Demographics and Attentional Engagement to and Disengagement From Expected Reward and Non-punishment, Presented Separately for Each Prospective Diagnostic Group

| | Mean (*SD*) / Count (percentage) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Depression (*n*=81)** | | **No depression (*n*=450)** | | **Depression, no other diagnoses (*n*=41)** | | **No psychiatric diagnosis (*n*=305)** | |
| **Demographics and psychiatric problems** | | | | | | | | |
| Socioeconomic family status (*n*=80, *n*=449) | 0.05 | (0.84) | 0.15 | (0.73) | 0.00 | (0.82) | 0.17 | (0.72) |
| Age at time of the SOT | 16.0 | (0.6) | 16.2 | (0.6) | 15.9 | (0.7) | 16.1 | (0.6) |
| Age first onset of depression | 20.0 | (3.2) | - | - | 20.4 | (3.3) | - | - |
| Females | 55 | (68%) | 216 | (48%) | 28 | (68%) | 150 | (49%) |
| Psychiatric diagnoses other than depression[a] | 40 | (49%) | 145 | (32%) | - | - | - | - |
| Depressive symptoms at time of the SOT[b] (*n*=80, *n*=444) | 0.31 | (0.25) | 0.23 | (0.22) | 0.32 | (0.25) | 0.20 | (0.20) |
| **Attentional engagement to and disengagement from expected reward and non-punishment (SOT)** | | | | | | | | |
| Engagement to reward, automatic | 30.28 | (27.94) | 29.54 | (33.15) | 30.29 | (28.81) | 29.57 | (33.67) |
| Difficulty to disengage from reward, automatic | -19.78 | (69.08) | 2.24 | (61.13) | -29.42 | (74.85) | 0.57 | (59.95) |
| Engagement to reward, voluntary control | 40.80 | (41.35) | 36.76 | (50.99) | 50.01 | (36.65) | 37.59 | (50.65) |
| Difficulty to disengage from reward, voluntary control | 8.93 | (52.72) | 5.30 | (51.35) | 11.16 | (45.36) | 2.95 | (52.01) |
| Engagement to non-punishment, automatic | 24.35 | (29.53) | 28.54 | (34.90) | 27.75 | (28.36) | 28.51 | (34.37) |
| Difficulty to disengage from non-punishment, automatic | -20.02 | (71.74) | 1.51 | (61.83) | -19.63 | (77.58) | 1.36 | (61.44) |
| Engagement to non-punishment, voluntary control | 32.16 | (51.17) | 34.95 | (50.47) | 32.84 | (49.01) | 34.36 | (51.31) |
| Difficulty to disengage from non-punishment, voluntary control | 12.03 | (70.63) | 4.11 | (59.56) | 20.76 | (63.75) | 1.54 | (59.42) |

Note: SOT = spatial orienting task

[a] Lifetime separation anxiety disorder, agoraphobia, generalized anxiety disorder, obsessive compulsive disorder, panic disorder, social phobia, specific phobia, attention deficit disorder, oppositional defiant disorder, or conduct disorder

[b] Mean score on the Youth Self-Report affective problems scale: 0 = "not true"; 1 = "somewhat or sometimes true"; 2 = "very or often true"; higher scores reflect more depressive symptoms.

**TABLE 2.** Results Logistic Regression Analyses of Onset of Depression Between Age 16 and Age 25 on Reward Related Attentional Biases at Age 16

| | | Depression[a] (*n*=529) | | Depression without other diagnoses[b] (*n*=346) | |
|---|---|---|---|---|---|
| | | ***OR*** | ***P*** | ***OR*** | ***P*** |
| Reward model | Engagement to reward, automatic | 0.980 | .439 | 0.918 | .319 |
| | Difficulty to disengage from reward, automatic | **0.670** | **<.001** | **0.571** | **<.001** |
| | Engagement to reward, voluntary control | 1.107 | .789 | 1.248 | .888 |
| | Difficulty to disengage from reward, voluntary control | 1.055 | .671 | 1.181 | .824 |
| Non-punishment model | Engagement to non-punishment, automatic | 0.887 | .182 | 0.948 | .383 |
| | Difficulty to disengage from non-punishment, automatic | **0.734** | **.007** | 0.740 | .043 |
| | Engagement to non-punishment, voluntary control | 0.939 | .313 | 0.907 | .292 |
| | Difficulty to disengage from non-punishment, voluntary control | 1.132 | .841 | 1.333 | .945 |
| Full model | Engagement to reward, automatic | 1.009 | .525 | 0.926 | .345 |
| | Difficulty to disengage from reward, automatic | **0.679** | **.001** | **0.591** | **.002** |
| | Engagement to reward, voluntary control | 1.127 | .822 | 1.299 | .921 |
| | Difficulty to disengage from reward, voluntary control | 1.061 | .689 | 1.194 | .839 |
| | Engagement to non-punishment, automatic | 0.867 | .148 | 0.906 | .301 |
| | Difficulty to disengage from non-punishment, automatic | **0.751** | **.011** | 0.780 | .080 |
| | Engagement to non-punishment, voluntary control | 0.931 | .295 | 0.923 | .335 |
| | Difficulty to disengage from non-punishment, voluntary control | 1.141 | .849 | 1.361 | .947 |

Note: All variables were standardized (Z-values) prior to analysis. All effects were adjusted for gender, age at the time of the attentional reward bias task and socioeconomic family status (unadjusted effects are presented in Table S12 of the Supplementary Material). One-sided *p-values* are reported. **Bold** indicates *p-values* below the False Discovery Rate (FDR) derived significance threshold (0.0125). OR = odds ratio.

[a] *DSM-IV* diagnosis of Major Depressive Disorder or Dysthymia between age 16 and 25

[b] Sensitivity check 1: similar to [a], but without other diagnoses (i.e., depressed only versus super healthy)

groups showed negative values on disengagement from reward and non-punishment, that is, less difficulties in disengaging from reward and non-punishment than in disengaging from non-reward and punishment, whereas the control groups showed virtually no difference. For other task related descriptive statistics, see Table S7 and S8 of the Supplementary Material.

## Reward-related attentional biases and onset of depression

Faster disengagement from expected reward and non-punishment during automatic trials predicted onset of depressive disorder ($p <$ FDR-derived significance threshold). See Table 2, left panel (95 % confidence intervals of the odds ratios (ORs) are presented in Table S9 of the Supplementary Material). No other tests reached statistical significance.

## Sensitivity analyses

Excluding participants with comorbid diagnoses did not weaken the effect of faster disengagement from expected reward; if anything the effect became stronger. The effect of faster disengagement from expected non-punishment became slightly weaker and no longer reached statistical significance (sensitivity check 1; see Table 2, right panel). Adjusting for depressive symptoms at T3 (sensitivity check 2; see Table S10 of the Supplementary Material), exclusion of participants without onset of depressive disorder at T4 for whom T6 information was missing (sensitivity check 3; see Table S11 of the Supplementary Material), and repeating the main analyses without adjusting for gender, age and socioeconomic position (sensitivity check 4; see Table S12 of the Supplementary Material) yielded results comparable to the main analyses. Testing each of the eight effects in separate univariate models (sensitivity check 5) again resulted in the same patterns (exact *OR*s and *p*-values available from the authors upon request).

## Additional *post hoc* check

The disengagement scores were computed by means of difference scores, therefore the disengagement effects we found could be explained by the fact that participants who later developed a depression disengaged more easily from locations of expected reward (or non-punishment) to locations of expected non-reward (or punishment), or by the fact that they showed more difficulties in disengaging from locations of expected non-reward (or punishment) to locations of expected reward (or non-punishment), or by both. Figure 1 and Table S8 indicate that the differences between individuals with and without an onset of depression were mainly due to difficulties in shifting from a location of expected non-reward or punishment to a location of expected reward or non-punishment, that is, the main group differences were found in uncued hard automatic response trials (Table S8).

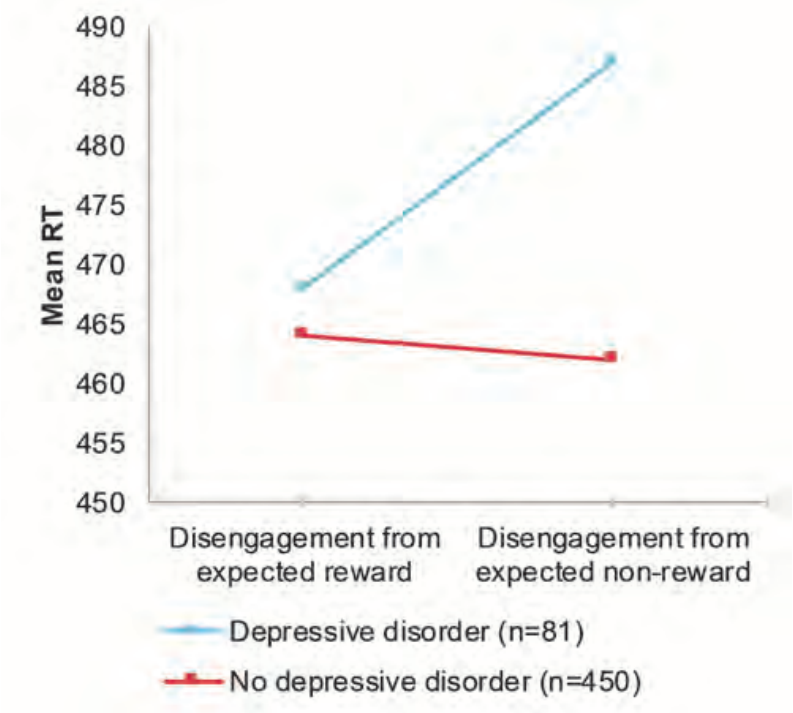



**FIGURE 1.** Mean Reaction Time (RT) of Automatic Disengagement From Locations of Expected Reward to Locations of Expected Non-reward Compared to Automatic Disengagement From Locations of Expected Non-reward to Locations of Expected Reward, Presented Separately for Adolescents With and Without Later Onset of Depression.

Note: Patterns are similar for automatic disengagement from locations of expected non-punishment compared to expected punishment.

## DISCUSSION

This study showed that reduced reward responsiveness predicts future onset of depression. Our results provide evidence in favor of the hypothesis that easier disengagement from expected reward than from expected non-reward and easier disengagement from expected non-punishment than from expected punishment at 16 years of age predicted depressive disorder between 16 and 25 years of age. This was found at the more automatic level and not at the more voluntary level of information processing. Adolescents who would later develop a depression had more difficulties than their never-depressed peers in disengaging from locations of expected negative outcomes (i.e., non-reward and punishment) and subsequently shifting to locations of expected positive outcomes (i.e., reward and non-punishment). Contrary to expectations, we found no evidence that decreased initial engagement toward reward or non-punishment

predicts onset of depressive disorder. Effects were not driven by co-occurring anxiety or other psychiatric problems and therefore seem specific to depression, and they were also not driven by subclinical depressive symptoms at 16 years of age.

Engagement to expected reward in our task should be interpreted as engagement to expected reward when the target actually appears where the reward is expected, that is, in cases where shifting attention from one location to the other is not required. Disengagement from expected non-reward is to be interpreted as a shift toward expected reward after the initial expectation that the target would appear in a different location and would be non-rewarding, that is, when shifting attention from one location to the other is required. The fact that we found a disengagement effect but not an engagement effect suggests that adolescents with vulnerability to depression respond similarly to expected reward, non-reward, non-punishment, or punishment when there is no need to shift attention, but specifically show decreased reward responsiveness in situations in which they are initially focused on negative information.



Our general finding that low reward responsiveness predicts onset of depression is consistent with the preliminary prospective evidence from most previous studies (Bress et al., 2013; Forbes et al., 2007; Nelson et al., 2016; Rawal et al., 2013; Stringaris et al., 2015). Two studies found opposite effects. In one study, increased activity in reward-related brain areas during selfish rewards predicted an increase in depressive symptoms (Telzer et al., 2014); in the other, increased node strength for resting state reward-related brain networks predicted onset of depression (Pan et al., 2017). The latter finding may be less contradictory than it seems, because it has been argued that increased resting state node strength may imply a blunted responsiveness to rewards (Pan et al., 2017). Our findings are also consistent with cognitive neuropsychological models of depression according to which biases toward negative and away from positive information have a central causal role in the development of depression (Roiser et al., 2012), as well as with earlier suggestions that these biases may be driven by a difficulty to disengage from negative stimuli rather than by more initial engagement toward negative stimuli (Gotlib & Joormann, 2010). That we did not find the hypothesized effect of decreased engagement toward expected reward and non-punishment tentatively suggests that vulnerability to future depression is not characterized by problems with engaging in rewarding situations when they present themselves, but more with the incapability to let go of negative situations and redirect attention to situations that may potentially result in reward. This may ultimately result in the overall bias toward negative and away from positive information that characterizes depressed individuals.

Participants who did not develop depression showed no difference in automatic disengagement from expected reward (or non-punishment) versus non-reward (or punishment), whereas one would possibly expect them to show more difficulties in disengaging from expected reward than the other way around (Matton et al., 2017). This lack of an automatic disengagement effect is consistent with previous studies in which the same task was used (Derryberry & Reed,

2002; Matton et al., 2017; van Hemel-Ruiter et al., 2013), but so far no plausible explanation has been given. We propose that the perceived safety of the conditions in which reward and non-punishment were expected and the perceived threat of the conditions in which non-reward and punishment were expected may have moderated participants' attentional scope. Because safe environments have been found to broaden attentional scope, and threatening environments have been found to narrow attentional scope (Friedman & Förster, 2010), individuals may find it easier to redirect their broader focus away from a safe environment (i.e., a location of expected reward/non-punishment) than to redirect their narrow focus away from a threatening environment (i.e., a location of expected non-reward/punishment). This could have masked the difficulties healthy adolescents may have with disengaging from expected reward (or non-punishment), particularly on more automatic levels of information processing, because these are associated with a more narrow attentional scope than voluntary levels (Derryberry & Reed, 2002). In a task containing a neutral condition, not part of the present task, it might be better evaluated if healthy adolescents have more difficulties in disengaging from expected reward than in disengaging from an expected neutral condition.

Reflection that our findings pertain to automatic rather than voluntary levels of information processing is required as well. It has been suggested that depressed individuals are characterized particularly by voluntary, higher-order top-down information processing biases, because previous studies using behavioral reaction time tasks found results on voluntary but not on automatic levels of information processing (Mathews & MacLeod, 2005). It is important to emphasize that our prospective approach of investigating information processing biases in individuals who were healthy at the time of assessment is novel and an important difference from the existing literature. We propose that although voluntary processes may indeed largely explain information processing biases for individuals who already have depression, automatic processes could constitute a vulnerability to depression in not-yet-depressed individuals. Adolescents with an automatic tendency to remain focused on negative situations and a diminished capability to redirect attention to potentially rewarding situations may process disproportionally more negative information (Derryberry & Reed, 2002), which may gradually trigger the more voluntary top-down negative biases that characterize depressed patients. A plausible explanation of why no automatic attentional biases have been found in depressed patients is that they in general perform slower on reaction time tasks, which could mask existing automatic biases. That is, automatic versus voluntary processing tends to be operationalized in the same way for everyone without taking into account individual differences in processing time. Because depressed patients process information slowly, they might not show automatic reward-related differences on a behavioral reaction times task when the reward-related information is presented only briefly.

The findings of the present study suggest that enhancement of reward responsiveness, particularly in situations in which adolescents are focused on negative information, may benefit adolescents who are vulnerable to depression. Whether enhancement of reward responsiveness can actually lower risk of depression can be tested only in an experimental setting, but if it could, the potential gain would be substantial. Assuming that our effect sizes reflect un-confounded causal relations, we estimated that increasing reward-related attentional bias in adolescents with low reward responsiveness by 1 *SD* could prevent 5.3 depression onsets per 100 adolescents treated (see Table S13 for details). Previous attempts to modify attentional biases toward negative information in currently depressed individuals suggest that an increase of 1 *SD* is feasible (Wells & Beevers, 2010; Yang, Ding, Dai, Peng, & Zhang, 2015). However, note that these interventions did not focus on training reward responsiveness specifically in situations in which people are focused on negative information, did not target the specific automatic processes of attention our results apply to (i.e., 250 ms between cue and target), and were aimed at lowering depressive symptoms in already depressed individuals rather than at lowering the risk on first onset of depression.



Our study has notable strengths. We investigated whether reward-related attentional bias predicts onset of depression with a large sample size and a long follow-up period (9 years). Our large sample allowed excluding participants with an onset of depression at the time of or before the assessment of the attention task, which was necessary for disentangling contemporaneous from prospective associations. This led to confirmation of previously reported preliminary evidence (Bress et al., 2013; Forbes et al., 2007; Nelson et al., 2016; Rawal et al., 2013; Stringaris et al., 2015) that biased processing of reward-related information may represent a vulnerability marker for depression, which has important implications for early treatment and prevention. Depressive disorder was assessed by standardized diagnostic interviews at 19 years of age and by a validated depression self-report diagnostic assessment at 25 years. Because depression is commonly characterized by high comorbidity with other psychiatric disorders we started by comparing adolescents with and without a depressive disorder between 16 and 25 years of age regardless of other psychiatric problems, because this best represents depression in the population, and assessed the specificity of these findings for depression by repeating the analyses after excluding adolescents with lifetime psychiatric problems other than depression.

Our study is not without limitations. First, because of the lack of a neutral condition in the SOT, it was impossible to distinguish between difficulties with disengaging from negative information and decreased responsiveness to positive information after an initial focus on negative information. A neutral condition also could have helped to explain the absence of a disengagement effect in healthy controls. Second, the retrospective assessments of the CIDI and the LIDAS at 19 and 25 years of age increased the risk for recall bias, especially for psychiatric problems with a young age at onset. Third, our sample size was sufficiently large to find effect sizes with an odds ratio approximately equal to 0.65, but not for finding more subtle effects

that could also play a role in vulnerability to depression. Fourth, the conclusions of this study are limited to depression onset from mid-adolescence up to early adulthood; they might not generalize to populations with an early or later onset. Fifth, it might have been interesting to compare the reward responsiveness of the group with future depression to the reward responsiveness of a group with past depression, but because of the heterogeneity of the group with past depression (i.e., adolescents could have remitted from depression, or not, and may or may not have experienced several episodes since the first onset of depression), we were not convinced that the results of such a comparison could be interpreted adequately.

To conclude, we found that difficulties in shifting attention from expected non-reward to expected reward and from expected punishment to expected non-punishment at 16 years of age predicted depression during 9 years of follow-up. This was found only at an automatic level of information processing. Our findings suggest that reduced reward responsiveness at 16 years marks vulnerability for depression. Prevention programs may aim at increasing at-risk adolescents' responsiveness to cues for potential rewards, particularly in situations in which they are focused on negative experiences.

# SUPPLEMENTARY MATERIAL

## Supplement 1

### Detailed information about sampling procedure, the Spatial Orienting Task, assessments of depressive disorder, and the multiple test correction method

#### *Overview of the sampling procedures of the T3 laboratory tasks and composition of the final sample*

Participants with a high-risk profile had a higher chance of being selected for the laboratory tasks at T3, and 66.0% had at least one out of three risk factors for developing psychiatric problems, that is, high-risk temperament, lifetime parental psychopathology, or living in a single-parent family. The remaining 34.0% were randomly selected from the TRAILS participants without any of the three risk factors. The following criteria were used for the three risk factors:

1. High-risk temperament criteria: EATQ (Early Adolescent Temperament Questionnaire) Frustration ≥ 90th percentile or EATQ Fear ≥ 90th percentile or EATQ Effortful Control ≤ 10th percentile.
2. Parental psychopathology criterion: at least one parent with severe psychopathology.
3. Single parent criterion: at least one of both biological parents is not part of the family.

See Table S1 of the Supplementary Material for an overview of the composition of the final sample.

#### *Detailed description of the Spatial orienting task (SOT)*

The SOT (Derryberry & Reed, 2002) was the first task of a larger set of laboratory tasks. Extensively trained test assistants were used to optimize standardization of the procedures. The task was assessed on weekdays, in a sound-attenuating room with blinded windows at locations in the participants' town of residence. The task was programmed in E-prime version 1.1 (Psychology Software Tools, Pittsburgh, PA) and was presented on the screen of an Intel Pentium 4 CPU computer with Philips Brilliance 190 P monitor. Participants were seated 50 cm from the screen. The SOT started with a practice positive game and a practice negative game, each consisting of 6 cued, 6 uncued and 2 catch trials. After the practice games the actual task started, which consisted of four positive and four negative games, alternated in sets of two, starting with two positive games. Each game consisted of 32 cued, 16 uncued, and 8 catch trials, presented in random order. Half of the trials of each condition were short-delay trials and the other half long-delay trials. Each game started with a score of zero points. See Table S2 of the Supplementary Material for an overview of the different games and trials.

During each game, two vertical black bars which were displayed against a white background marked the location of the cues and targets, and the score was presented in black at the center of the screen. Participants were instructed to fixate on the score, which was updated after each response, and to avoid moving their eyes. Each trial started with turning off the fixation score for 200 ms and subsequently turning it on again for 250 ms, after which a cue arrow replaced one of the two vertical black bars. The cue arrow served the purpose of orienting participants' attention to one of the two peripheral locations. After a short (250 ms) or long (500 ms) delay, a target, that is, a small vertical gray rectangle, appeared, either centered within the cue arrow (a so-called cued target, see Figure s3a) or centered within the vertical black bar on the other side of the fixation score (an uncued target, see Figure s3b). Participants were informed that a blue-up arrow (easy cue) signaled that a target appearing in that (cued) location would be easy and result in a fast response in 75% of the cases, and a target appearing in the uncued location would be hard and result in a slow response in 75% of the cases. A red-down arrow (hard cue) signaled that a target appearing in the cued location would be hard, that is, too slow in 75% of the cases, and a target appearing in the uncued location would be easy, that is, result in a fast response in 75% of the cases. Participants were also informed that 2/3 of the targets would appear in the location of the cue arrow, and that occasionally no target would appear. They were instructed to press the 'b' key on the keyboard as soon as they detected the target and were warned that pressing the key before the target appeared or when no target appeared would result in a loss of 10 points. Five-hundred ms after the 'b' was pressed, or for catch trials 1s after the delay interval, the cue arrow and target were replaced by the two black bars, and a feedback arrow was presented below the centered score. A blue-up arrow indicated a fast response on target trials or a correct non-response on catch trials and a red-down arrow indicated a slow response on target trials or an incorrect response on catch trials. To increase the relevance of the scores and boost the participants' motivation, they were informed that a prize (e.g., a balloon ride) would be awarded to those with the highest scores on the positive games, and that very low scores on the negative games could result in having to start over again until performance was sufficient. See Table S3 for a schematic overview of the SOT and Table S4 for an example of a short delay (automatic) uncued hard (red) trial during a positive game.

At the end of each positive or negative game, for each individual participant the median reaction time (RT) and standard deviation (*SD*) were computed and used to determine cutoffs for fast and slow responses on the next game of the same type (positive or negative). For the two practice blocks a fixed cutoff of 350 ms was used, because no prior games of the same type were available to calculate personal cutoffs. Following Derryberry and Reed (2002), for easy targets, a response was considered fast if the RT was less than the median plus 0.55 times the *SD* and was otherwise considered slow. For hard targets, a response was considered fast if the RT was less

than the median minus 0.55 times the *SD*, and slow otherwise. Because RTs were expected to be about 25 ms slower after short delays compared to long delays, 12 ms were added to the cut-off for short-delay trials and subtracted from the cut-off for long-delay trials.

Following Derryberry and Reed (2002), SOT trial reaction times (RTs) below 125 ms (probable anticipations) and RTs above 1,000 ms (probable distractions) were treated as outliers and removed before mean RTs were calculated for each of sixteen conditions (positive or negative game * cued or uncued trial * easy or hard trial * short-delay or long-delay trial).

***Detailed description of the CIDI and the LIDAS assessments of depressive disorder***

The World Health Organization Composite International Diagnostic Interview (CIDI, T4) and the Lifetime Depression Assessment Self-report (LIDAS, T6) were used to assess onset of a depression after T3, which was operationalized as a lifetime Major Depressive Disorder or Dysthymia (Dysthymia was only assessed at T4) with an age at onset higher than the age at T3 according to the CIDI or the LIDAS. For both the CIDI and the LIDAS diagnoses we applied the official DSM-IV criteria, including impairment of functioning.



At T4, the CIDI was used to assess psychiatric disorders. The CIDI is a structured diagnostic interview which has been shown to have good reliability and validity in assessing lifetime DSM-IV disorders (Andrews & Peters, 1998; Haro et al., 2006; Kessler et al., 2004). For each of the psychiatric disorders present, questions were asked regarding the age of onset. At T6, participants completed the LIDAS (Bot et al., 2017), a self-report instrument that covers all nine DSM-IV symptoms criteria for Major Depressive Disorder (MDD), supplemented with items on impairment of functioning, age of onset of depression and on other psychiatric disorders diagnosed by a professional. The LIDAS has been shown to have adequate sensitivity and specificity in measuring lifetime MDD as assessed with the CIDI (Bot et al., 2017). In the validation study of the LIDAS three algorithms to determine MDD status were tested. (1) DSM symptoms criterion of at least 5 out of nine symptoms and at least one of two core symptoms; (2) DSM symptoms criterion or self-reported lifetime diagnosis or treatment by a professional, or antidepressant use; (3) similar to algorithm 2 but now controls with other psychiatric disorders are excluded to get a super healthy control group. The specificity ranged from 0.80 (algorithm 3) to 0.86 (algorithm 1) and the sensitivity from 0.66 (algorithm 1) to 0.85 (algorithm 3) (Bot et al., 2017). We used the symptoms criterion, that is, algorithm 1 of the LIDAS validation study, extended with a criterion of impairment in functioning to ensure that the CIDI and the LIDAS criteria closely matched each other and the official DSM criteria for MDD. Please note that in the validation study the impairment of functioning criterion was not used (Bot et al., 2017).

We inspected the agreement between the CIDI and the LIDAS for the 421 participants who completed both assessments in the present study. Forty-eight participants had a depressive disorder according to the LIDAS but not according to the CIDI. Thirty-four of them reported a first onset of depression between the CIDI and the LIDAS and thus for them agreement between

the CIDI and the LIDAS could not be expected. For 358 (93%) of the remaining 387 participants, the CIDI and the LIDAS were in agreement about whether or not they developed a depressive disorder after age 16.

***Multiple test correction method: False Discovery Rate***

To correct for multiple tests, the classical False Discovery Rate (FDR) method (Benjamini & Hochberg, 1995) was used. The FDR method was developed as an instrument to correct for multiple testing while retaining more power than when using the overly conservative Bonferroni correction. The main thought behind this approach is that finding one significant result in 20 tests calls for a more stringent correction than finding 10 significant results in 20 tests. The advantage in power of FDR above more conservative methods increases when more significant results are found and when the total number of tests increases. The FDR method was used in our study because we expected to find multiple significant results.

For the calculation of the classical FDR (Benjamini & Hochberg, 1995), the p-values of the 8 main statistical tests were ranked from low to high. Subsequently, we calculated the FDR-corrected significance thresholds (with alpha set at 0.05), as follows:

$$FDR\ derived\ significance\ threshold = \frac{0.05}{number\ of\ tests / ranking}$$

The resulting FDR-derived thresholds are listed in Table S6 of the Supplementary Material. We determined which of the original p-values were smaller than the accompanying FDR-corrected significance threshold. As can be seen in Table S6, the 2nd p-value was below the FDR-derived significance threshold, but the 3rd p-value not anymore. Since the 2nd p-value was the last one to remain below the FDR-derived significance threshold, the accompanying significance threshold (0.0125) is the significance threshold for all statistical tests performed in the present study.

**TABLE S1.** Number of Participants in the High and Low Risk Profile Groups, Presented Separately for Each Prospective Diagnostic Group

| | Count (percentage) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Depression (*n*=81)** | | **No depression (*n*=450)** | | **Depression, no other diagnoses (*n*=41)** | | **No psychiatric diagnosis (*n*=305)** | |
| Low risk (not A, B, or C) | 24 | 29.6% | 172 | 38.2% | 14 | 34.1% | 130 | 42.6% |
| High-risk temperament (A) | 12 | 14.8% | 69 | 15.3% | 4 | 9.8% | 40 | 13.1% |
| Parental psychopathology (B) | 15 | 18.5% | 69 | 15.3% | 8 | 19.5% | 48 | 15.7% |
| Single-parent family (C) | 11 | 13.6% | 38 | 8.4% | 7 | 17.1% | 31 | 10.2% |
| A and B | 4 | 4.9% | 37 | 8.2% | 3 | 7.3% | 22 | 7.2% |
| A and C | 2 | 2.5% | 13 | 2.9% | 0 | 0% | 7 | 2.3% |
| B and C | 9 | 11.1% | 32 | 7.1% | 3 | 7.3% | 18 | 5.9% |
| A, B, and C | 4 | 4.9% | 20 | 4.4% | 2 | 4.9% | 9 | 3.0% |

**TABLE S2.** Overview of the Games and Trials in the Spatial Orienting Task

| | | Number of trials | | | |
|---|---|---|---|---|---|
| **Block** | **Game** | **Cued** | **Uncued** | **Catch** | **Cut-off for fast responses** |
| 1 | Practice positive game | 6 | 6 | 2 | 350 ms |
| 2 | Practice negative game | 6 | 6 | 2 | 350 ms |
| 3 | Positive game | 32 | 16 | 8 | Based on RT block 1 |
| 4 | Positive game | 32 | 16 | 8 | Based on RT block 2 |
| 5 | Negative game | 32 | 16 | 8 | Based on RT block 3 |
| 6 | Negative game | 32 | 16 | 8 | Based on RT block 4 |
| 7 | Positive game | 32 | 16 | 8 | Based on RT block 5 |
| 8 | Positive game | 32 | 16 | 8 | Based on RT block 6 |
| 9 | Negative game | 32 | 16 | 8 | Based on RT block 7 |
| 10 | Negative game | 32 | 16 | 8 | Based on RT block 8 |

Note: RT = reaction time

**TABLE S3.** Schematic Overview of the Spatial Orienting Task, Derived From Hemel-Ruiter et al. (2013)

| Duration | Description | Sequential screens within one trial | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 200ms | 2 vertical black bars (size 0.16 x 0.64 cm) mark the location of cues and targets | \| | | \| | | | |
| 250ms | Fixation score in between 2 bars (size 0.6 x 0.9 cm per digit) | \| | 0 | \| | | | | |
| 250ms (short-delay trial) or 500ms (long-delay trial) | Cue arrow (size 0.5 x 1.3 cm, shaft width 0.16 cm) replaces left or right bar | Easy (blue) cue – signals high chance at easy condition in which points are gained in positive trials and no points are lost in negative trials | | | Hard (red) cue – signals high chance at hard condition in which no points are gained (in positive trials) or points are lost (in negative trials) | | | |
| | | ↑ | 0 | \| | \| | 0 | ↓ | |
| | | \| | 0 | ↑ | ↓ | 0 | \| | |
| Reaction time (RT) | Target (small vertical gray rectangle, 0.08 x 0.24 cm) appears - press 'b' as fast as possible if you see target, no target: don't press button | 2/3 of targets appear in cued location (easy or hard condition) PRESS B | | | 1/3 of targets appear in uncued location (easy or hard condition) | | | 1/7 of trials no target (catch trial) DON'T PRESS |
| | | Easy condition – much time to react (i.e., own median RT + 0.55 SD) | | | | | | |
| | | ↑ | 0 | \| | ¦ | 0 | ↓ | |
| | | \| | 0 | ↑ | ↓ | 0 | ¦ | |
| | | Hard condition – little time to react (i.e., own median RT – 0.55 SD) | | | | | | |
| | | \| | 0 | ↓ | ¦ | 0 | ↑ | |
| | | ↓ | 0 | \| | ↑ | 0 | ¦ | |
| 500ms after response or after 1s when no response (catch trial) | Black bars are reinstated and arrows signal positive (blue) or negative (red) feedback | \| | 0 ↑ | \| | \| | 0 ↓ | \| | |
| after 250 ms delay | Total score is updated | \| | 0 | \| | | | | |

Note: CM = centimeters; MS = milliseconds; RT = reaction time

**TABLE S4.** Schematic Overview of an Example of a First Automatic (i.e., Short-Delay) Trial during a Positive Game of the Spatial Orienting Task

| Step | Duration | Description | Sequential computer screens | | |
|---|---|---|---|---|---|
| 1 | 200ms | 2 vertical black bars (size 0.16 x 0.64 cm) mark the location of cues and targets | \| | | \| |
| 2 | 250ms | Fixation score in between 2 bars (size 0.6 x 0.9 cm per digit) | \| | **0** | \| |
| 3 | 250ms (short-delay trial) | Cue arrow (size 0.5 x 1.3 cm, shaft width 0.16 cm) replaces left or right bar | \| | **0** | ↓ |
| 4 | Reaction time (RT) | Target (small vertical gray rectangle, 0.08 x 0.24 cm) appears - press 'b' as fast as possible if you see target, no target: don't press button | ¦ | **0** | ↓ |
| 5 | 500ms after response | Black bars are reinstated and the feedback arrow appears | \| | **0** ↑ | \| |
| 6 | after 250 ms delay | Total score is updated | \| | **10** | \| |

Note: Each trial starts with 2 vertical black bars marking the location of cues and targets (**step 1**). In **step 2**, participants are shown their score. Each game starts with a score of 0. In this specific example, a hard cue, that is, a down red arrow, appears at the right side location (**step 3**). This signals a high chance that the target will appear at the cued right side location and that if the target appears at the cued location it will be hard (i.e., fast response required) and there is a high chance that no points will be gained, and that if the target appears in the uncued left bar's location it will be easy (much time to react) and there is a high chance that 10 points will be gained. The target that appears after 250 milliseconds [ms] (**step 4**) is a small vertical gray rectangle in the uncued left bar. This example requires the participant to shift attention from the cued location of expected non-reward to the uncued location of expected reward, that is, disengagement from a location of expected non-reward and subsequent engagement toward the location of expected reward. The blue arrow which appears in **step 5** signals that the response was fast enough to result in a gain of ten points. In **step 6** the score is updated. This is an example of a short delay (automatic) uncued hard (red) positive trial. CM = centimeters

**TABLE S5.** Computation of Attentional Bias Scores for Expected Reward and Non-punishment

| | Positive games | | Negative games | |
|---|---|---|---|---|
| Attentional bias scores | Engagement toward location of expected reward | Difficulty to disengage from location of expected reward (and shift to location of expected non-reward) | Engagement toward location of expected non-punishment | Difficulty to disengage from location of expected non-punishment (and shift to location of expected punishment) |
| Formulas | RT cued hard (red) trials minus RT cued easy (blue) trials | RT uncued easy (blue) trials minus RT uncued hard (red) trials | RT cued hard (red) trials minus RT cued easy (blue) trials | RT uncued easy (blue) trials minus RT uncued hard (red) trials |

Note: All four attentional bias scores were computed for short-delay trials (250 milliseconds [ms]; automatic responses) and long-delay trials (500 ms; voluntary responses) separately, i.e., in total eight attentional bias scores were computed. See Table S4 for an example of a short delay uncued hard (red) trial during a positive game.

**TABLE S6.** False Discovery Rate (FDR) Classical Method, with Alpha Set to 0.05

| Rank | Ascending *p-values* | Hypothesis name | FDR-derived significance thresholds |
|---|---|---|---|
| 1 | 0.001 | Difficulty to disengage from reward, automatic[a] | 0.00625 |
| 2 | 0.007 | Difficulty to disengage from non-punishment, automatic[a] | **0.01250** |
| 3 | 0.182 | Engagement to non-punishment, automatic | 0.01875 |
| 4 | 0.313 | Engagement to non-punishment, voluntary | 0.02500 |
| 5 | 0.439 | Engagement to reward, automatic | 0.03125 |
| 6 | 0.671 | Difficulty to disengage from reward, voluntary | 0.03750 |
| 7 | 0.789 | Engagement to reward, voluntary | 0.04375 |
| 8 | 0.841 | Difficulty to disengage from non-punishment, voluntary | 0.05000 |

Note: **Bold** indicates the FDR-derived significance threshold.
[a] Significant after multiple test correction

**TABLE S7.** Mean Percentage Correct Catch Trials and Mean Percentage Outliers on Target Trials, Presented Separately for Each Prospective Diagnostic Group

| | Mean % (*SD*) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Attentional bias task variables** | **Depression (N=81)** | | **No depression (N=450)** | | **Depression, no other diagnoses (N=41)** | | **No diagnosis (N=305)** | |
| Correct catch trials positive games, easy cues | 59.3 | (19.8) | 60.4 | (22.3) | 61.3 | (19.9) | 59.9 | (23.0) |
| Correct catch trials positive games, hard cues | 65.8 | (19.1) | 65.6 | (21.4) | 67.7 | (17.2) | 65.1 | (22.1) |
| Correct catch trials negative games, easy cues | 53.1 | (24.1) | 52.9 | (23.0) | 54.1 | (24.5) | 53.1 | (23.2) |
| Correct catch trials negative games, hard cues | 56.0 | (22.7) | 57.8 | (22.9) | 59.0 | (23.0) | 58.0 | (23.5) |
| Outliers target trials: RTs <125 ms or >1,000 ms | 5.5 | (3.5) | 5.4 | (3.6) | 5.1 | (3.2) | 5.4 | (3.6) |

Note: Trials with reaction times (RTs) < 125 milliseconds (ms) (probably anticipations) and trials with RTs > 1,000 ms (probable distractions) were considered outliers. Mean percentage correct catch trials reflects the percentage of cases in which participants correctly withheld a response when no target was presented.

**TABLE S8.** Mean Reaction Times in Milliseconds (ms) on Correct Trials, Presented Separately for Each Prospective Diagnostic Group

| | Mean (*SD*) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Spatial orienting task variables** | **Depression (N=81)** | | **No depression (N=450)** | | **Depression, no other diagnoses (N=41)** | | **No diagnosis (N=305)** | |
| Positive game, cued, easy, short delay | 334 | (40) | 333 | (40) | 342 | (40) | 333 | (40) |
| Positive game, cued, hard, short delay | 364 | (43) | 362 | (44) | 373 | (42) | 363 | (45) |
| Positive game, uncued, easy, short delay | 468 | (94) | 464 | (85) | 496 | (95) | 460 | (85) |
| Positive game, uncued, hard, short delay | 487 | (91) | 462 | (84) | 525 | (88) | 460 | (84) |
| Positive game, cued, easy, long delay | 343 | (54) | 339 | (56) | 350 | (50) | 339 | (55) |
| Positive game, cued, hard, long delay | 383 | (65) | 376 | (68) | 400 | (60) | 376 | (68) |
| Positive game, uncued, easy, long delay | 395 | (81) | 375 | (71) | 422 | (77) | 373 | (68) |
| Positive game, uncued, hard, long delay | 386 | (76) | 369 | (67) | 410 | (75) | 370 | (68) |
| Negative game, cued, easy, short delay | 329 | (46) | 327 | (44) | 335 | (48) | 326 | (44) |
| Negative game, cued, hard, short delay | 353 | (41) | 355 | (50) | 363 | (40) | 355 | (50) |
| Negative game, uncued, easy, short delay | 455 | (90) | 451 | (85) | 480 | (94) | 452 | (88) |
| Negative game, uncued, hard, short delay | 476 | (96) | 449 | (87) | 499 | (100) | 450 | (87) |
| Negative game, cued, easy, long delay | 339 | (55) | 328 | (56) | 350 | (55) | 328 | (58) |
| Negative game, cued, hard, long delay | 371 | (64) | 363 | (68) | 383 | (64) | 362 | (69) |
| Negative game, uncued, easy, long delay | 392 | (91) | 373 | (77) | 419 | (89) | 372 | (77) |
| Negative game, uncued, hard, long delay | 380 | (78) | 369 | (73) | 398 | (84) | 371 | (72) |

Note: Short-delay trials (250 ms) assessed more automatic responses and long-delay trials (500 ms) more voluntary ones.

**TABLE S9.** Results Logistic Regression Analyses of Onset of Depression Between Age 16 and Age 25 on Reward Related Attentional Biases at Age 16, Including 95% CIs

| | | Depression[a] (*n*=529) | | Depression without other diagnoses[b] (*n*=346) | |
|---|---|---|---|---|---|
| | | ***OR*** | **95% CI *OR*** | ***OR*** | **95% CI *OR*** |
| Reward model | Engagement to reward, automatic | 0.980 | 0.757 - 1.268 | 0.918 | 0.642 - 1.312 |
| | Difficulty to disengage from reward, automatic | **0.670***** | 0.525 - 0.855 | **0.571***** | 0.402 - 0.810 |
| | Engagement to reward, voluntary control | 1.107 | 0.863 - 1.421 | 1.248 | 0.874 - 1.782 |
| | Difficulty to disengage from reward, voluntary control | 1.055 | 0.834 - 1.334 | 1.181 | 0.832 - 1.677 |
| Non-punishment model | Engagement to non-punishment, automatic | 0.887 | 0.685 - 1.149 | 0.948 | 0.666 - 1.348 |
| | Difficulty to disengage from non-punishment, automatic | **0.734**** | 0.574 - 0.937 | 0.740* | 0.526 - 1.043 |
| | Engagement to non-punishment, voluntary control | 0.939 | 0.729 - 1.209 | 0.907 | 0.641 - 1.285 |
| | Difficulty to disengage from non-punishment, voluntary control | 1.132 | 0.888 - 1.445 | 1.333 | 0.937 - 1.896 |
| Full model | Engagement to reward, automatic | 1.009 | 0.768 - 1.325 | 0.926 | 0.635 - 1.351 |
| | Difficulty to disengage from reward, automatic | **0.679**** | 0.531 - 0.869 | **0.591**** | 0.412 - 0.848 |
| | Engagement to reward, voluntary control | 1.127 | 0.875 - 1.451 | 1.299 | 0.904 - 1.867 |
| | Difficulty to disengage from reward, voluntary control | 1.061 | 0.839 - 1.343 | 1.194 | 0.841 - 1.697 |
| | Engagement to non-punishment, automatic | 0.867 | 0.664 - 1.133 | 0.906 | 0.624 - 1.315 |
| | Difficulty to disengage from non-punishment, automatic | **0.751*** | 0.589 - 0.958 | 0.780 | 0.552 - 1.103 |
| | Engagement to non-punishment, voluntary control | 0.931 | 0.717 - 1.208 | 0.923 | 0.638 - 1.335 |
| | Difficulty to disengage from non-punishment, voluntary control | 1.141 | 0.888 - 1.467 | 1.361 | 0.937 - 1.978 |

Note: Individuals with depression onsets before or at the time of the attentional reward bias task and individuals with lifetime bipolar disorder or hypomania were excluded from the sample. All variables were standardized (Z-values) prior to analysis. All effects were adjusted for gender, age at the time of the attentional reward bias task and socioeconomic family status (unadjusted effects are presented in Table S12). **Bold** indicates *p-values* below the False Discovery Rate (FDR) derived significance threshold of 0.0125. CI = confidence interval; OR = odds ratio.

[a] DSM-IV diagnosis of Major Depressive Disorder or Dysthymia between age 16 and 25

[b] Similar to [a], but without other diagnoses (i.e., depressed only versus super healthy)

*one-sided p-value < .05; **one-sided p-value < .01; ***one-sided p-value < .001

**TABLE S10.** Sensitivity Check 2: Results Logistic Regression Analyses of Onset of Depression Between Age 16 and Age 25 on Reward Related Attentional Biases at Age 16, Adjusted for Depressive Symptoms at Age 16

| | | Depression[a] (*n*=522) | | Depression without other diagnoses[b] (*n*=342) | |
|---|---|---|---|---|---|
| | | ***OR*** | **95% CI *OR*** | ***OR*** | **95% CI *OR*** |
| Reward model | Engagement to reward, automatic | 0.996 | 0.765 - 1.298 | 0.938 | 0.655 - 1.345 |
| | Difficulty to disengage from reward, automatic | **0.674***** | 0.526 - 0.864 | **0.592**** | 0.416 - 0.842 |
| | Engagement to reward, voluntary control | 1.148 | 0.892 - 1.476 | 1.260 | 0.883 - 1.800 |
| | Difficulty to disengage from reward, voluntary control | 1.050 | 0.826 - 1.336 | 1.192 | 0.842 - 1.689 |
| Non-punishment model | Engagement to non-punishment, automatic | 0.885 | 0.682 - 1.149 | 0.904 | 0.633 - 1.292 |
| | Difficulty to disengage from non-punishment, automatic | **0.715**** | 0.558 - 0.917 | **0.660*** | 0.462 - 0.943 |
| | Engagement to non-punishment, voluntary control | 0.964 | 0.747 - 1.243 | 0.925 | 0.653 - 1.311 |
| | Difficulty to disengage from non-punishment, voluntary control | 1.153 | 0.904 - 1.471 | 1.338 | 0.940 - 1.905 |
| Full model | Engagement to reward, automatic | 1.017 | 0.769 - 1.345 | 0.960 | 0.658 - 1.403 |
| | Difficulty to disengage from reward, automatic | **0.685**** | 0.533 - 0.881 | **0.635**** | 0.441 - 0.914 |
| | Engagement to reward, voluntary control | 1.164 | 0.901 - 1.504 | 1.306 | 0.906 - 1.882 |
| | Difficulty to disengage from reward, voluntary control | 1.053 | 0.827 - 1.340 | 1.222 | 0.860 - 1.736 |
| | Engagement to non-punishment, automatic | 0.864 | 0.660 - 1.131 | 0.866 | 0.594 - 1.263 |
| | Difficulty to disengage from non-punishment, automatic | **0.739**** | 0.577 - 0.945 | 0.711* | 0.497 - 1.017 |
| | Engagement to non-punishment, voluntary control | 0.945 | 0.725 - 1.231 | 0.908 | 0.623 - 1.325 |
| | Difficulty to disengage from non-punishment, voluntary control | 1.163 | 0.903 - 1.497 | 1.354 | 0.931 - 1.970 |

Note: Individuals with depression onsets before or at the time of the attentional reward bias task and individuals with lifetime bipolar disorder or hypomania were excluded from the sample; All variables were standardized (Z-values) prior to analysis. All effects were adjusted for gender, age at the time of the attentional reward bias task, socioeconomic family status, and depressive symptoms at age 16. **Bold** indicates *p-values* below the False Discovery Rate (FDR) derived significance threshold of 0.0125. OR = odds ratio; CI = confidence interval.

[a] DSM-IV diagnosis of Major Depressive Disorder or Dysthymia between age 16 and 25

[b] Similar to [a], but without other diagnoses (i.e., depressed only versus super healthy)

*one-sided p-value < .05; **one-sided p-value < .01; ***one-sided p-value < .001

**TABLE S11.** Sensitivity Check 3: Results Logistic Regression Analyses of Onset of Depression Between Age 16 and Age 25 on Reward Related Attentional Biases at Age 16, After Excluding 89 Participants Without Depression Onset According to the CIDI at T4 for Whom no T6 Information Was Available

| | | **Depression[a] (*n*=441)** | | **Depression without other diagnoses[b] (*n*=291)** | |
|---|---|---|---|---|---|
| | | ***OR*** | **95% CI *OR*** | ***OR*** | **95% CI *OR*** |
| Reward model | Engagement to reward, automatic | 0.957 | 0.731 - 1.253 | 0.891 | 0.615 - 1.292 |
| | Difficulty to disengage from reward, automatic | **0.633***** | 0.489 - 0.821 | **0.563***** | 0.392 - 0.809 |
| | Engagement to reward, voluntary control | 1.087 | 0.840 - 1.407 | 1.259 | 0.884 - 1.793 |
| | Difficulty to disengage from reward, voluntary control | 1.140 | 0.887 - 1.465 | 1.209 | 0.849 - 1.722 |
| Non-punishment model | Engagement to non-punishment, automatic | 0.891 | 0.683 - 1.162 | 0.964 | 0.672 - 1.383 |
| | Difficulty to disengage from non-punishment, automatic | **0.728**** | 0.569 - 0.930 | 0.759 | 0.539 - 1.068 |
| | Engagement to non-punishment, voluntary control | 0.917 | 0.706 - 1.190 | 0.859 | 0.602 - 1.226 |
| | Difficulty to disengage from non-punishment, voluntary control | 1.097 | 0.855 - 1.408 | 1.326 | 0.928 - 1.895 |
| Full model | Engagement to reward, automatic | 0.993 | 0.743 - 1.325 | 0.893 | 0.600 - 1.330 |
| | Difficulty to disengage from reward, automatic | **0.638***** | 0.490 - 0.832 | **0.585**** | 0.404 - 0.849 |
| | Engagement to reward, voluntary control | 1.099 | 0.846 - 1.428 | 1.338 | 0.929 - 1.929 |
| | Difficulty to disengage from reward, voluntary control | 1.161 | 0.902 - 1.494 | 1.224 | 0.857 - 1.748 |
| | Engagement to non-punishment, automatic | 0.855 | 0.646 - 1.130 | 0.905 | 0.613 - 1.337 |
| | Difficulty to disengage from non-punishment, automatic | **0.733**** | 0.574 - 0.937 | 0.787 | 0.556 - 1.114 |
| | Engagement to non-punishment, voluntary control | 0.931 | 0.709 - 1.221 | 0.858 | 0.587 - 1.256 |
| | Difficulty to disengage from non-punishment, voluntary control | 1.076 | 0.832 - 1.393 | 1.361 | 0.933 - 1.984 |

Note: Individuals with depression onsets before or at the time of the attentional reward bias task and individuals with lifetime bipolar disorder or hypomania were excluded from the sample. All variables were standardized (Z-values) prior to analysis. All effects were adjusted for gender, age at the time of the attentional reward bias task and socioeconomic family status. **Bold** indicates *p-values* below the False Discovery Rate (FDR) derived significance threshold of 0.0125. CIDI = Composite International Diagnostic Interview; OR = odds ratio; CI = confidence interval.

[a] DSM-IV diagnosis of Major Depressive Disorder or Dysthymia between age 16 and 25

[b] Similar to [a], but without other diagnoses (i.e., depressed only versus super healthy)

*one-sided p-value < .05; **one-sided p-value < .01; ***one-sided p-value < .001

**TABLE S12.** Sensitivity Check 4: Results Logistic Regression Analyses of Onset of Depression Between Age 16 and Age 25 on Reward Related Attentional Biases at Age 16, Unadjusted Effects

| | | Depression[a] (*n*=529) | | Depression without other diagnoses[b] (*n*=346) | |
|---|---|---|---|---|---|
| | | ***OR*** | **95% CI *OR*** | ***OR*** | **95% CI *OR*** |
| Reward model | Engagement to reward, automatic | 0.977 | 0.764 - 1.249 | 0.915 | 0.650 - 1.290 |
| | Difficulty to disengage from reward, automatic | **0.704**** | 0.556 - 0.892 | **0.623**** | 0.448 - 0.867 |
| | Engagement to reward, voluntary control | 1.104 | 0.862 - 1.413 | 1.304 | 0.916 - 1.858 |
| | Difficulty to disengage from reward, voluntary control | 1.075 | 0.852 - 1.357 | 1.142 | 0.811 - 1.609 |
| Non-punishment model | Engagement to non-punishment, automatic | 0.860 | 0.668 - 1.109 | 0.953 | 0.674 - 1.346 |
| | Difficulty to disengage from non-punishment, automatic | **0.709**** | 0.559 - 0.899 | 0.732* | 0.526 - 1.021 |
| | Engagement to non-punishment, voluntary control | 0.979 | 0.763 - 1.255 | 0.952 | 0.678 - 1.335 |
| | Difficulty to disengage from non-punishment, voluntary control | 1.126 | 0.887 - 1.428 | 1.363 | 0.966 - 1.925 |
| Full model | Engagement to reward, automatic | 1.008 | 0.777 - 1.308 | 0.924 | 0.646 - 1.323 |
| | Difficulty to disengage from reward, automatic | **0.711**** | 0.560 - 0.903 | **0.642**** | 0.458 - 0.901 |
| | Engagement to reward, voluntary control | 1.124 | 0.872 - 1.447 | 1.349 | 0.941 - 1.935 |
| | Difficulty to disengage from reward, voluntary control | 1.075 | 0.853 - 1.355 | 1.154 | 0.819 - 1.626 |
| | Engagement to non-punishment, automatic | 0.846 | 0.652 - 1.097 | 0.920 | 0.643 - 1.318 |
| | Difficulty to disengage from non-punishment, automatic | **0.722**** | 0.570 - 0.915 | 0.773 | 0.553 - 1.081 |
| | Engagement to non-punishment, voluntary control | 0.963 | 0.746 - 1.243 | 0.923 | 0.647 - 1.317 |
| | Difficulty to disengage from non-punishment, voluntary control | 1.136 | 0.892 - 1.448 | 1.399 | 0.975 - 2.006 |

Note: Individuals with depression onsets before or at the time of the attentional reward bias task and individuals with lifetime bipolar disorder or hypomania were excluded from the sample. All variables were standardized (Z-values) prior to analysis. Effects were unadjusted for gender, age at the time of the attentional reward bias task and socioeconomic family status. **Bold** indicates *p-values* below the False Discovery Rate (FDR) derived significance threshold of 0.0125. OR = odds ratio; CI = confidence interval.

[a] DSM-IV diagnosis of Major Depressive Disorder or Dysthymia between age 16 and 25

[b] Similar to [a], but without other diagnoses (i.e., depressed only versus super healthy)

*one-sided p-value < .05; **one-sided p-value < .01; ***one-sided p-value < .001

**TABLE S13.** Estimation of the Number of Depression Onsets that May be Prevented by Modification of Reward-Related Bias in Adolescents With a Negative Reward Responsiveness Bias

| | Total sample (*n*=531) | |
|---|---|---|
| **Disengagement from reward** | ***N*** | **Mean predicted probability depression** |
| Dis ≥ 3 SD | 1 | 0.06 |
| 2 SD ≤ Dis < 3 SD | 11 | 0.07 |
| 1 SD ≤ Dis < 2 SD | 63 | 0.10 |
| 0 SD ≤ Dis < 1 SD | 204 | 0.13 |
| -1 SD < Dis < 0 SD | 178 | 0.17 |
| **-2 SD < Dis ≤ -1 SD** | **56** | **0.22** |
| **-3 SD < Dis ≤ -2 SD** | **15** | **0.28** |
| **Dis ≤ -3 SD** | **3** | **0.36** |

The number of depression onsets that might be prevented by enhancing reward-related bias in participants with Dis ≤ -1 SD (indicated in **bold**) with one standard deviation was calculated as: ((3*(36-28)) + (15*(28-22)) + (56*(22-17)))/74 = 5.3%. In words, bias modification may prevent up to 5.3 onsets of depression per 100 treated adolescents. For the most extremely biased group with Dis ≤ -3 SD, bias modification may prevent a max. of 36-28 = 8 onsets per 100 treated adolescents.

Note: A standard deviation (SD) < 0 indicates a faster disengagement from reward than from non-reward, i.e., lower reward responsiveness. This table serves the purpose of providing a first indication of how many depressions may be prevented by 1 SD bias modification in high-risk groups scoring ≤ -1 SD on disengagement from reward. Mean predicted probabilities are based upon a logistic regression analysis with onset of depression between age 16 and age 25 regressed on difficulty to disengage from reward. Effects of other variables were not taken into account. Dis = Disengagement from reward.

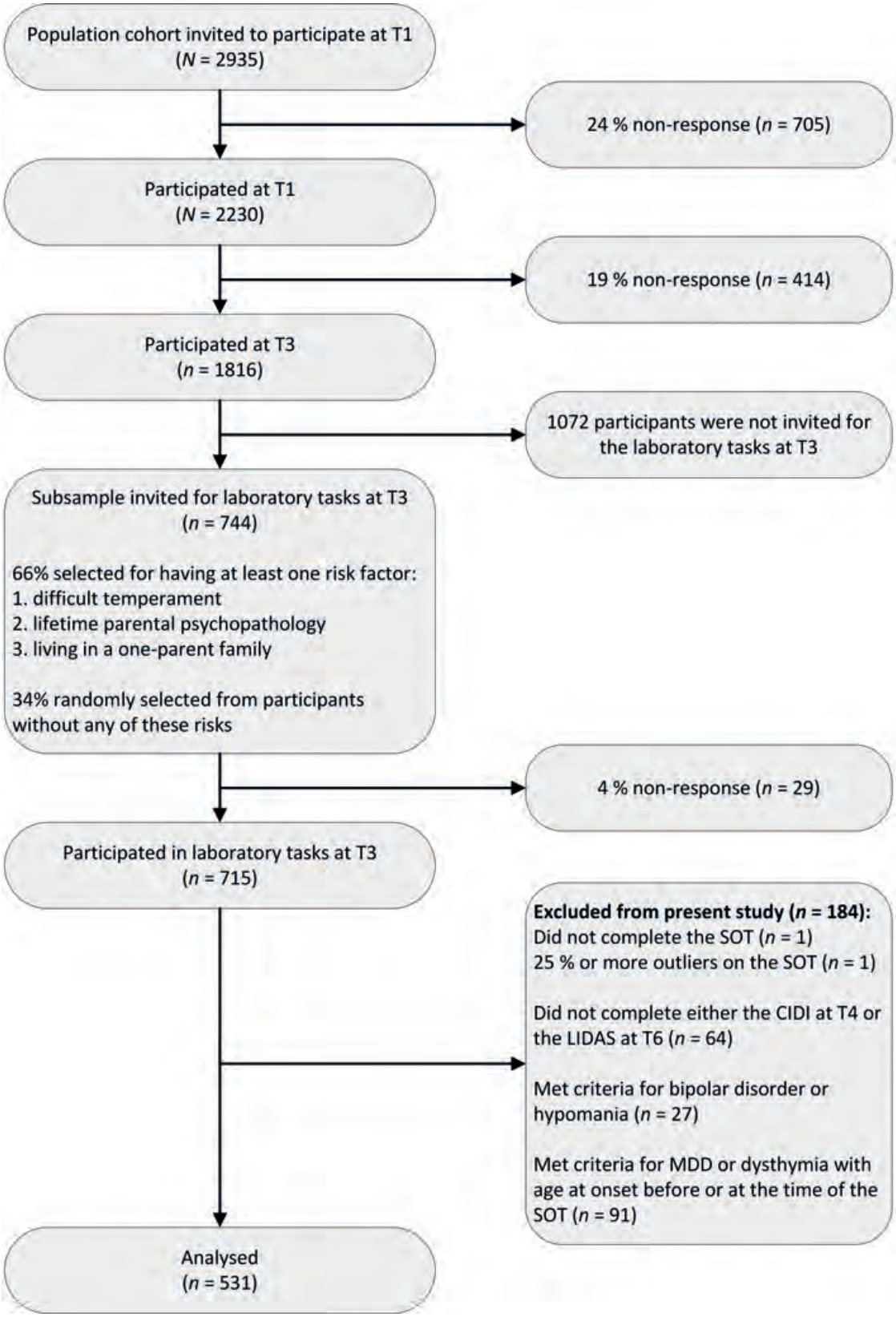



**FIGURE S1.** Flowchart of the Sampling Procedure

Note: CIDI = Composite International Diagnostic Interview; LIDAS = Lifetime Depression Assessment Self-report; MDD = Major Depressive Disorder; SOT = Spatial Orienting Task.

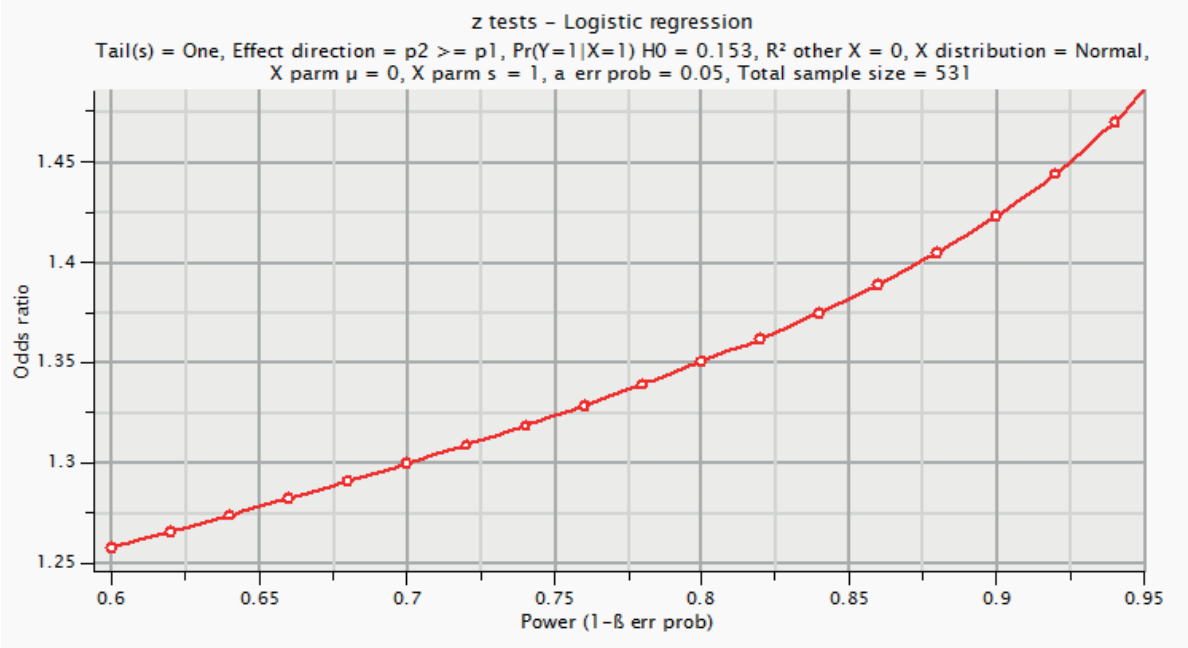



**FIGURE S2.** Power Calculation G*Power Version 3.1.9.2, Based on a Total Sample Size of 531 Individuals Non-Depressed at Baseline, of Which 81 (i.e., 15.3%) Developed a Depressive Disorder During Nine Years of Follow-Up.

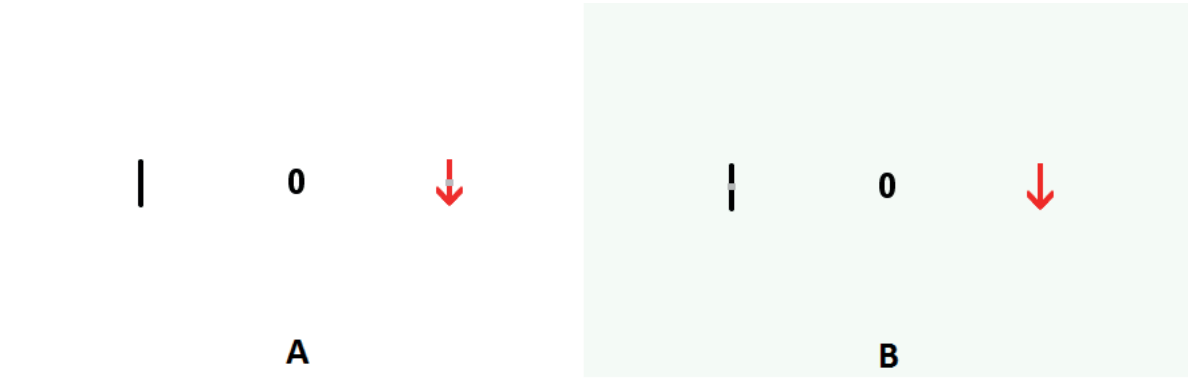



**FIGURE S3.** Panel A is an example of a cued trial (i.e., the gray rectangle target appears at the cued location of the arrow) and panel B is an example of an uncued trial (i.e., the gray rectangle appears at the uncued location of the vertical black bar).

# CHAPTER 5

## Spread the joy: How high and low bias for happy facial emotions translate into different daily life affect dynamics

Published as:
Vrijen, C., Hartman, C. A., Van Roekel, E., de Jonge, P., & Oldehinkel, A. J. (2018). Spread the Joy: How High and Low Bias for Happy Facial Emotions Translate into Different Daily Life Affect Dynamics. *Complexity*, vol. 2018. https://doi.org/10.1155/2018/2674523.

## ABSTRACT

There is evidence that people commonly show a bias toward happy facial emotions during laboratory tasks, that is, they identify other people's happy facial emotions faster than other people's negative facial emotions. However, not everybody shows this bias. Individuals with a vulnerability for depression, for example, show a low happy bias compared to healthy controls. The main aim of this study was to acquire a better understanding of laboratory measures of happy bias by studying how these translate to people's daily life. We investigated whether stable high and low happy bias during a laboratory task were associated with different daily life affect dynamics (i.e., effects from one time interval of 6 hours to the next). We compared the daily life affect dynamics of young adults (age 18-24) with a high bias toward happy facial emotions (*N*=25) to the affect dynamics of young adults with a low bias toward happy emotions (*N*=25). Affect and related measures were assessed three times per day during 30 days. We used multilevel vector autoregressive (VAR) modelling to estimate lag 1 affect networks for the high and low happy bias groups, and used permutation tests to compare the two groups. Compared to their peers with a low happy bias, individuals with a high happy bias more strongly sustained the effects of daily life reward experiences over time. Individuals with a high happy bias may use their reward experiences more optimally in daily life to build resources that promote well-being and mental health. Low reward responsiveness in daily life may be key to why individuals who show a low happy bias during laboratory tasks are vulnerable for depression. This study illustrates the potential benefits of a network approach for unraveling psychological mechanisms.

## INTRODUCTION

It is an interesting phenomenon that some people have a tendency to be relatively fast in identifying other people's happy facial emotions while others are relatively fast in identifying other people's negative facial emotions. Happy bias is an implicit bias of which people are probably unaware themselves, therefore it is commonly assessed with standardized laboratory tasks. There is consistent evidence from studies using these laboratory tasks that people generally show a bias toward happy facial emotions, that is, they commonly identify other people's happy facial emotions faster than other people's negative facial emotions (Leppänen & Hietanen, 2004). However, not everybody shows this bias. Depressed individuals, for example, show a low happy bias compared to healthy controls (Bourke et al., 2010; Joormann & Gotlib, 2006; Surguladze et al., 2004), and there are indications that low happy bias is already present before onset of depression and predicts onset of depression (Joormann et al., 2007; Vrijen, Hartman, & Oldehinkel, 2016). Given these findings and the accumulating evidence that the smallest building blocks of an individual's adaptive and maladaptive affect patterns are found in daily life affect dynamics (Pe et al., 2015; Trull et al., 2015; Wigman et al., 2015), one would expect that high and low happy bias also reflect differences in daily life affect dynamics. However, to date this has not been investigated. In the present study we looked at daily life correlates of laboratory measures of happy bias. We investigated how happy bias during a standardized laboratory task translates to daily life affective dynamics by comparing the daily life affect dynamics (i.e., effects from one time interval of 6 hours to the next) between young adults with a stable high happy bias and young adults with a stable low happy bias. The main aim of this study was to acquire a better understanding of the importance and scope of laboratory measures of happy bias in people's daily life. Our findings are expected to facilitate the interpretation of laboratory measures of happy bias, and, because of our focus on adaptive and maladaptive affect dynamics, may possibly also provide clues to why a low happy bias is associated with an increased risk for depression.

Indications of which daily life affect dynamics promote mental health (i.e., are adaptive) and which ones are associated with depressive problems (i.e., are maladaptive) can be found in both laboratory studies and in studies based on ecological momentary assessments (EMA). Evidence from laboratory tasks suggests that the inability to sustain positive emotions (Horner et al., 2014; McMakin et al., 2009), the inability to sustain activation in neural circuits underlying positive affect and reward over time (Heller et al., 2009), and the incapability to disengage from negative self-referential rumination (Hilt & Pollak, 2013) are associated with depressive symptoms and clinical depression. It was further found that positive affect facilitates recovery from negative emotional experiences (Fredrickson & Levenson, 1998; Tugade & Fredrickson, 2004). EMA studies also indicate that the inability to sustain positive affect over time in daily life is associated with depressive symptoms (e.g., anhedonia), in general as well as in clinically depressed populations (Heininga et al., 2017; Höhn et al., 2013). Additionally, the ability to generate positive affect

from pleasant experiences in daily life predicted fewer symptoms of depression and anxiety in individuals with a history of depression (Wichers et al., 2010) and in individuals who had been exposed to childhood adversity or recent stressful life events (Geschwind et al., 2010). Taken together, this evidence suggests that the following types of affect dynamics are adaptive and promote mental health:

**(1) the ability to sustain positive affect and positive experiences over time** (Heininga et al., 2017; Heller et al., 2009; Höhn et al., 2013; Horner et al., 2014; McMakin et al., 2009), that is, the carry-over of positive affect and positive experiences from one time interval to the next;

**(2) the ability to use positive experiences to generate positive affect and *vice versa*** (Geschwind et al., 2010; Wichers et al., 2010), that is, the carry-over of positive experiences to positive affect and *vice versa* from one time interval to the next;

**(3) the ability to use positive affect and positive experiences to dampen negative affect, negative thoughts (i.e., rumination), and negative experiences** (Fredrickson & Levenson, 1998; Hilt & Pollak, 2013; Tugade & Fredrickson, 2004).

In the present study we investigated whether a high happy bias as compared to a low happy bias during a laboratory task is associated with (1) enhanced responses to positive affect and positive experiences over time in daily life, with more carry-over over time (i.e., from one 6 hour time interval to the next), and more carry-over from one type of positive affect or positive experience to another, and a stronger dampening effect on negative affect, thoughts, and experiences; and (2) diminished responses to negative affect, thoughts, and experiences in daily life, with less carry-over over time (i.e., from one time interval of 6 hours to the next), less carry-over from one type of negative affect, thoughts or experience to another, and weaker dampening effects on positive affect and positive experiences.

We used a network approach to affect dynamics, which entails that psychological symptoms or constructs are represented as interacting components of a complex dynamic system (Borsboom & Cramer, 2013; Bringmann et al., 2013), and that these dynamics define the very nature of the psychological phenomena we study (i.e., mental disorders, well-being) (Bringmann et al., 2016). This approach can be used to investigate cross-sectional associations between symptoms at a specific point in time, but also, as in the present study, to investigate the temporal dynamics of affect over time. These temporal networks consist of 'nodes' (i.e., the variables in the network) and 'edges' (i.e., the directed associations between these nodes from one assessment to the next). See Figure 1 for a fictitious example of a temporal network containing the three nodes 'happiness' (HAP), 'sadness' (SAD), and 'worrying' (WOR). In this network model, as well as in the models we used, each node is predicted by the lag (i.e., *t-1*) of all other variables and itself. SAD at time *t* is, for example, predicted by $HAP_{t-1}$, $WOR_{t-1}$, and $SAD_{t-1}$. Temporal networks can be used to study how different affect components interact as a

dynamic system over time. They provide insightful visualizations of the interplay of the network components and it is also possible to compute centrality indices indicating the importance of each of the components in the network.

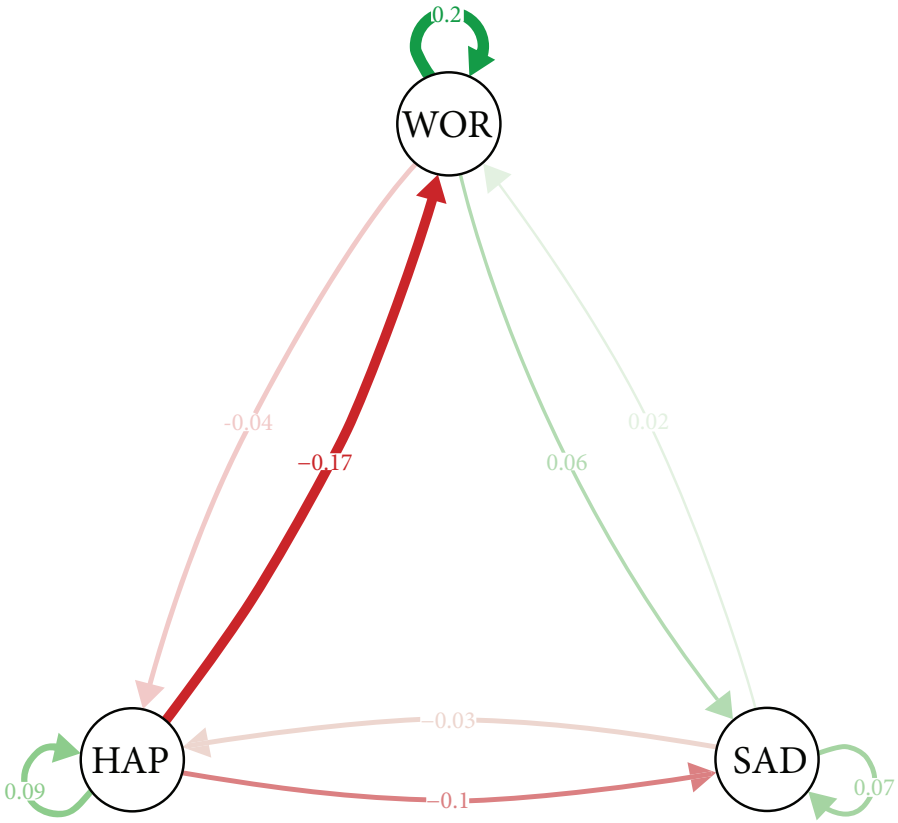



**Figure 1.** Fictitious example of a temporal network containing three nodes. The green edges represent positive directed associations, for example, on average high levels of worrying during one assessment predict high levels of sadness during the next assessment. The red edges represent negative directed associations, for example, on average high levels of happiness during one assessment predict low levels of worrying during the next assessment. The self-loops represent autocorrelations (i.e., the effect of the node on itself from one assessment to the next).

We compared the daily life dynamic affect networks (i.e., effects from one time interval of 6 hours to the next) of two groups of young adults with extreme and stable biases to happy facial emotions during a laboratory task. We selected a high happy bias group, consisting of individuals who were considerably more sensitive to happy emotions than to negative emotions, and a low happy bias group, who showed considerably less bias toward happy emotions, or even a bias toward negative emotions.

The affect dynamics of the high and low happy bias groups were compared on nodes that are associated with reward responsiveness, emotion regulation and depressive symptoms. We selected three nodes that were related to positive affect and positive experiences (for the sake of brevity and readability hereafter referred to as positive nodes): 'feeling joyful', 'pleasant experiences' and 'feeling interested'; and four nodes related to negative affect, negative thoughts, and negative experiences (for the sake of brevity and readability hereafter referred to as negative nodes): 'feeling sad', 'feeling irritated', 'worrying' and 'unpleasant experiences'. The nodes 'feeling interested', 'feeling sad' and 'feeling irritated' closely resemble core symptoms of depression according to the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013), and 'feeling interested' also reflects openness to new experiences and an inclination to actively approach and explore the outside world. The nodes 'feeling joyful' and 'pleasant experiences' are particularly relevant in the light of indications that high transference of positive emotions over time in daily life (Höhn et al., 2013) and the ability to generate boosts of positive affect from pleasant daily life experiences (Geschwind et al., 2010) may protect against affective problems. As opposed to feeling interested, feelings of joy and pleasant experiences are by definition rewarding at the very moment they are experienced. To illustrate the difference, people may cheer because they feel joy or pleasure, but not because they feel interested. More than the other nodes, 'pleasant experiences' and 'unpleasant experiences' not only reflect affective states, but also the type of events individuals are involved in and their ability to seek out rewarding experiences or escape from a cascade of negative events. 'Worrying'[6] was included for its associations with depressive disorder (Nolen-Hoeksema et al., 2008), positive and negative affect (McLaughlin, Borkovec, & Sibrava, 2007), and reduced cognitive control (Beckwé, Deroost, Koster, Lissnyder, & Raedt, 2014). It was explored how these different nodes interact as dynamic systems and if these dynamics differed between the high and the low happy bias group.

We expected that an increase in positive affect and positive experiences, particularly of the directly rewarding positive nodes joy and pleasant experiences, would have larger and longer-lasting effects in the high happy bias group than in the low happy bias group. More specifically, we expected that the high happy bias group would more easily sustain and act upon pleasant experiences and feelings of joy to enhance positive affect and positive experiences and dampen negative affect, negative thoughts and negative experiences. We also expected that pleasant experiences would generalize or carry-over to feelings of joy and the other way around. We thus hypothesized that pleasant experiences and feelings of joy would be stronger predictors in the network of the high happy bias group than in the network of the low happy bias group **(hypothesis 1)**; and that the nodes pleasant experiences and joy would more strongly predict themselves (i.e., pleasant experiences and feelings of joy would be more easily sustained) and

6 The term 'worrying' was used for both negative thoughts about the past, often referred to as 'rumination', and negative thoughts about the future, commonly referred to as 'worrying' (McLaughlin, Borkovec, & Sibrava, 2007; Nolen-Hoeksema, 1991).

each other (i.e., more carry-over between pleasant experiences and feelings of joy) over time in the high happy bias group than in the low happy bias group **(hypothesis 2)**; and that joy and pleasant experiences would more strongly predict the negative affect nodes over time (i.e., a larger dampening effect on negative nodes) in the high happy bias group than in the low happy bias group **(hypothesis 3)**. Further, because of the hypothesized reduced reward responsiveness in the low happy bias group, we expected the negative affect nodes to be stronger predictors in the network of the low happy bias group than in the network of the high happy bias group **(hypothesis 4)**; and that the negative affect nodes would more strongly predict themselves and each other over time in the low happy bias group than in the high happy bias group **(hypothesis 5)**; and more pronounced negative associations between negative nodes and positive nodes over time (i.e., a larger dampening effect on positive nodes) in the low happy bias group than in the high happy bias group **(hypothesis 6)**. Although feeling interested is a positive node, we did not expect it to have a similar role as joy and pleasant experiences as we consider feelings of joy and pleasure as intrinsically rewarding, whereas feeling interested is a more instrumental node, which only potentially leads to reward. More specifically, rather than group differences in the way in which feeling interested influenced other nodes, we expected that joy and pleasant experiences would more strongly predict interest in the high than in the low happy bias group **(hypothesis 7)**.



## METHODS

### Sample

Data were collected as part of the 'No Fun No Glory' (NFNG) study, in which we investigated anhedonia in young adults. The study was approved by the Medical Ethical Committee from the University Medical Center Groningen (no. 2014/508) and registered in the Dutch Clinical Trial Register (NTR5498). Participants were treated in accordance with the Declaration of Helsinki and indicated their informed consent prior to enrollment in the study. The project started with a large online screening survey in the northern part of the Netherlands among 2,937 young adults between 18 and 24 years old. Participants were recruited through advertisements on electronic learning environments of university and higher and intermediate vocational education institutes, pitches during lectures and classes, flyers and advertisements on social media. After subscribing on the study website (www.nofunnoglory.nl), participants received an email with the link to the online survey. The survey contained questionnaires about, for example, pleasure, psychiatric problems and stress, as well as a facial emotion identification task. From the screening survey 69 young adults who suffered from persistent anhedonia and 69 controls were selected for the part of the study in which momentary assessments were completed. For a description of the selection procedure for the anhedonia and control group in the NFNG project, see Section 1 of

the Supplementary Material or Van Roekel and colleagues (2016). From the 138 participants who completed the momentary assessments during the first month, we selected 25 participants with a high happy bias and 25 participants with a low happy bias for the present study.

## Selection high and low happy bias groups

We used extreme bias groups rather than the full happy bias continuum because of both conceptual and methodological considerations. First and foremost, the extreme-group approach more closely fitted our supposition that mainly happy biases in the extremes of the distribution distinguish between adaptive and maladaptive affective mechanisms (Vrijen, Hartman, Lodder, et al., 2016). Second, a particular strength of network analyses is that these can be used to explore group differences in overall patterns of affect dynamics rather than investigating single effects only. Estimating and plotting the networks for each of the groups separately yields more insight into the affect dynamics within these groups than a single network based on the total sample, while it is still possible to test statistically whether specific affect patterns differ between the groups.

We selected participants for the high and low happy bias groups without taking into account whether participants belonged to the anhedonia or the control group. The selection was based on scores on a facial emotion identification morph task participants completed for the first time as part of the online screening survey (T0), and a second time after the first month of momentary assessments (T2); see Figure 2 for a flowchart.

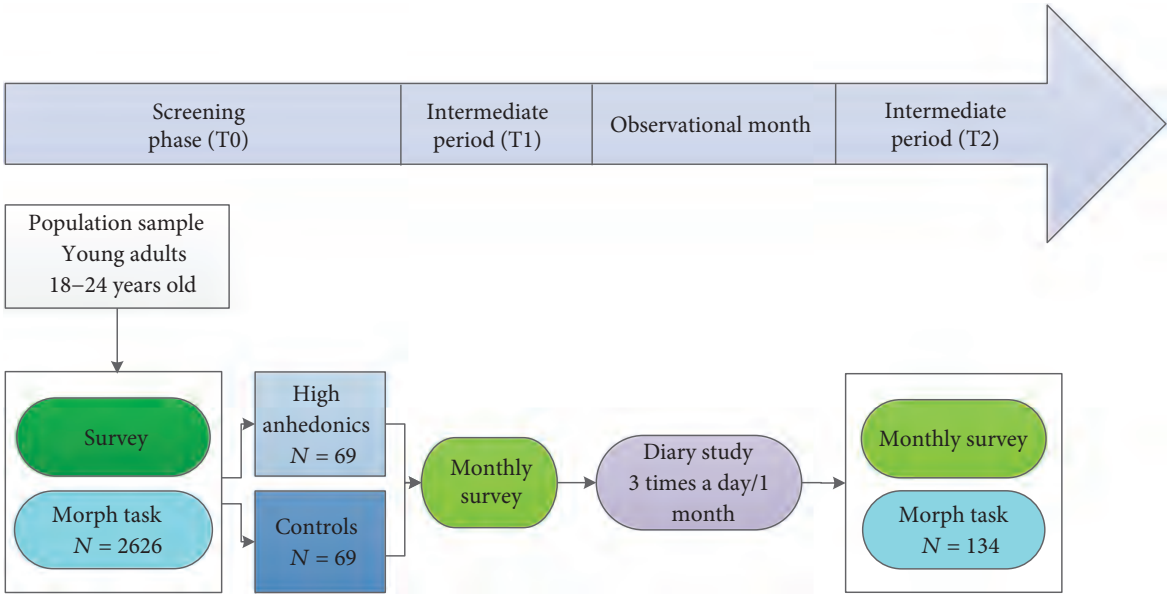



**FIGURE 2.** Flowchart of morph tasks and diary study month (i.e., momentary assessments) used in the current study

We excluded four participants who did not complete the morph task at T2. During the morph task, participants were shown 24 10-second movie clips of neutral faces which slowly changed into one of four emotions: happy, sad, angry or fearful. The participants had to press the spacebar as soon as they identified the emotion the neutral face turned into. For a more detailed description of the morph task, which was a shortened version of a task developed at Radboud University Nijmegen, the Netherlands (Lodder et al., 2015), see Section 1 of the Supplementary Material or Vrijen and colleagues (2016). For each participant the mean reaction time (RT) of correctly identified trials was calculated per emotion, resulting in RT Happy, RT Sad, RT Angry and RT Fearful. We excluded one participant with less than 50% correct answers at T2. Separate happy bias scores were calculated at T0 and T2 by dividing the average of RT Sad, RT Angry and RT Fearful by RT Happy. A higher happy bias means being faster in identifying happy emotions than in sad, angry and fearful emotions.

We were interested in the affect dynamics associated with trait high and low happy bias and compared the average affect dynamics during 30 days of individuals with a stable high happy bias (i.e., stable during these 30 days) to the affect dynamics of individuals with a stable low happy bias. Because stable happy bias and state fluctuations can only be unraveled by using happy bias at two time points, we selected an extreme high stable and an extreme low stable happy bias group based on the ranked happy bias scores at T0 and T2. Happy bias at T0 and Happy bias at T2 were each ranked from low (ranking 133) to high (ranking 1), and selection of the two happy bias groups was based on the summed ranks for T0 and T2. The 25 participants with the highest summed rank were selected for the high happy bias group, and the 25 participants with the lowest summed rank for the low happy bias group. An additional advantage of this approach was that part of the measurement error is also parceled out because a participant is only selected for the high happy bias group if scores on both tasks are high.[7] To ensure that high (or low) scores reflected a high (or low) score relative to the rest of the group on both tasks, we used summed rank scores rather than summed mean scores. In this way scores on both tasks count equally even in the case of general learning effects of the whole group. The middle group, which consisted of participants with moderate or unstable happy bias scores, was excluded from the main analyses. This group was taken into account in part of the *post hoc* sensitivity checks (see Section 4 of the Supplementary Material). See Figure 3 for the individual happy bias scores at T0 and T2 for the high happy bias group, the low happy bias group and the middle group.



7 Please note that we do not mean to suggest that all differences between happy bias at T0 and T2 are due to measurement error. We acknowledge that there may well be state happy bias fluctuations within a person between T0 and T2, but in the present study we are interested in the daily life affect dynamics associated with more stable high and low happy bias.

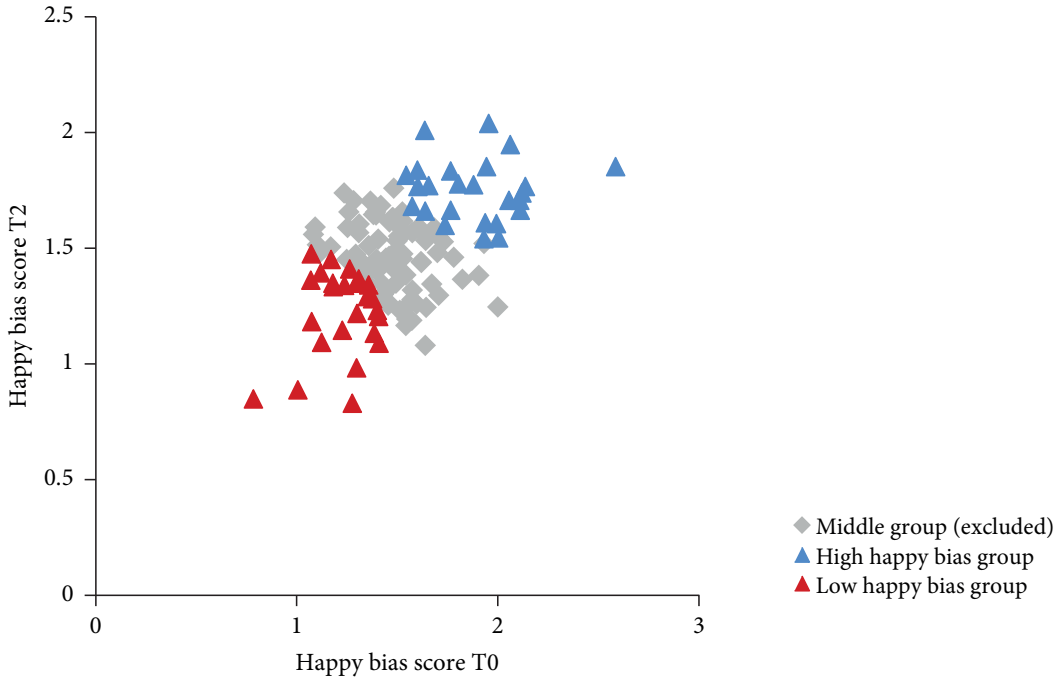



**FIGURE 3.** Happy bias scores T0 and T2 and selection high and low happy bias group

## Ecological momentary assessments

In the online questionnaire participants filled out three times a day, we included items to measure positive affect and positive experiences, negative affect, negative thoughts, and negative experiences, social company, activities, etc. See Van Roekel et al. (2016) for a detailed description of all momentary items that were assessed as part of the No Fun No Glory Study. Starting point as well as times of receiving the questionnaires during the day were personalized to the schedule and preference of the participants. They received a text message on their smartphone with the link to the questionnaire three times a day on fixed times with 6 hour intervals in between (e.g., 10:00 AM, 4:00 PM, 10:00 PM). The questionnaire had to be completed within 2 hours after the first notification. If necessary, reminders were sent after 1 hour and again after 1.5 hours. Completion of the questionnaires took on average 3 minutes.

We included the following items in the present study: Since the last assessment I felt joyful **(JOY)**; Think about the most pleasant event you experienced since the last assessment: how pleasant was this experience for you? **(POS)**; Since the last assessment I felt interested in the things around me **(INT)**; Since the last assessment I felt sad **(SAD)**; Since the last assessment I felt irritated **(IRR)**; Since the last assessment I have been worrying **(WOR)**; Think about the most unpleasant event you experienced since the last assessment: how unpleasant was this experience for you? **(NEG)**. Because we considered JOY, INT, SAD and IRR to be sensitive to overnight recall bias, the morning assessments of these items were phrased more momentarily,

that is, into 'I feel joyful', 'I feel interested in the things around me', 'I feel sad' and 'I feel irritated'. Participants indicated their endorsement to these items by means of a slider on a Visual Analogue Scale (VAS), with "not at all" as its left anchor and "very much" as its right anchor. The position of the slider was transformed into a score between 0 ("not at all") and 100 ("very much").

## Statistical analyses

We provided descriptive statistics for gender, age, education and anhedonia status, calculated mean levels of all variables used in this study and showed them for the high and low happy bias group separately. We used R package psych version 1.6.12 (R Core Team, 2013; Revelle, 2016) to calculate the group and individual Mean Squared Successive Differences (*MSSD*s) for each node, as these indicate the amount of variability from one assessment to the next. We also indicated per node how many participants had an $MSSD < 50$, which was used as a criterion for insufficient variability from one assessment to the next, following Van der Krieke and colleagues (2015). If one group shows a higher average *MSSD* than the other, or contains more participants with an acceptable *MSSD*, it is possible that this results in more power for this group. Therefore, we decided that if more or stronger significant associations were found for this group we would address the possibility that these differences were driven by differences in *MSSD* in the discussion.

We used multilevel vector autoregressive (VAR) modeling in R package mlVAR version 0.4 (Epskamp, Deserno, & Bringmann, 2016; Epskamp, Waldorp, Mõttus, & Borsboom, 2016) to explore the daily life dynamics between JOY, POS, INT, SAD, IRR, WOR, and NEG for the high and low happy bias groups. One of the main advantages of mlVAR was the availability of tools that, in combination with the R packages igraph (Csardi & Nepusz, 2006) and qgraph (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012), allowed not only visualization of networks and centrality indices on a group level, but also visualization of individual variation within groups.

Although exact power calculations are not possible for VAR analyses, a minimum of 50 assessments per person has been recommended for individual VAR analyses (Box, Jenkins, Reinsel, & Ljung, 2015). We performed multilevel VAR analyses for which power is influenced by both the number of assessments and the number of persons. Our analyses were based on three assessments per day for a period of 30 days, that is, 90 assessments per person (with an average of 6 missings per person), and our high and low happy bias groups consisted of 25 participants each. We ensured sufficient power by limiting the number of parameters estimated in mlVAR, that is, we focused mainly on within-subject processes, refrained from investigating the influence of between-subject predictors, and did not estimate correlations between random effects (see below for further details). We have performed a simulation study based on the present study's effect sizes, number of subjects and data points. Eight hundred datasets were simulated in which the individual network models we found were generated as the 'true' models. Fixed and random effects were estimated with the same method and number of nodes as in our main mlVAR analyses. In the present study we used on average 84 time points per subject, and for

this number of time points the simulation study showed high correlations between the true and the estimated fixed and random effects of the simulated datasets (see Figure S1 of the Supplementary Material). This indicates that the method we used is appropriate for our effect sizes, number of subjects and number of time points. Additionally, because all of our hypotheses were based on network patterns rather than on specific effects of a single variable, our findings do not rely on single paths in the network models.

As a first preparatory step we removed linear time trends from the data, because time trends violate the stationarity assumption of VAR analyses and may bias parameter estimation (Rovine & Walls, 2006). We also removed cyclic time of day trends prior to VAR analysis, because mlVAR does not allow controlling for time of day. Linear and cyclic time trends were removed by regressing each variable on time and on dummy variables for afternoon and evening, within each individual. The residuals from these analyses were used as input for the VAR models.

For estimating networks containing both autoregressive and cross-lagged effects it has been recommended to person-mean center all predictors prior to the analyses in order to separate within-subject from between-subject effects (Hamaker & Grasman, 2015; Raudenbush & Bryk, 2002). Because our main interest was to grasp daily life psychological processes which take place within individuals, we separated within-subject from between-subject effects even further by within-person standardization of all network variables prior to the VAR analyses. For comparing the relative strengths of different predictors within and between networks, standardization of the coefficients has been recommended because differences in coefficients may be due to differences in variance (Bulteel, Tuerlinckx, Brose, & Ceulemans, 2016; Schuurman, Ferrer, de Boer-Sonnenschein, & Hamaker, 2016). Using raw coefficients to compute centrality indices has been discouraged as well (Bulteel et al., 2016).

In mlVAR separate lag 1 networks were estimated for the high and low happy bias groups, by means of the lmer function from the Linear Mixed-Effects R package lme4 version 1.1-15 (Bates, Mächler, Bolker, & Walker, 2015). The networks were constructed by performing seven univariate multilevel VAR analyses, one for each dependent variable, and combining the results into a network. In each of the univariate multilevel VAR analyses the dependent variable was predicted by the lag (i.e., *t-1*) of all other variables and itself. This means that, for example, feeling irritated (IRR) at time *t* was predicted by $INT_{t-1}$, $JOY_{t-1}$, $SAD_{t-1}$, $WOR_{t-1}$, $POS_{t-1}$, $NEG_{t-1}$ and $IRR_{t-1}$. The unique direct temporal effects were modeled (Bringmann et al., 2016; Bulteel et al., 2016). Random effects were estimated to account for individual differences. We assumed no correlations between random intercepts and random slopes (orthogonality specification in mlVAR), as the person-mean of each variable was equal to 0 after within-person standardization.

The above-described procedure resulted in a network for the high happy bias and a network for the low happy bias group. For each node of these networks we calculated two centrality indices, outstrength and instrength. The outstrength of a node represents the summed strength, that is, the absolute value of the coefficients, of all outgoing paths from this node at *t-1* to other

nodes at time *t*, and as such reflects how strongly the node predicts other nodes over time. The instrength reflects how strongly a particular node is predicted by other nodes over time, and is computed by the summed strength of all its incoming paths at time *t* from other nodes at *t-1*. In mlVAR the packages igraph version 1.1.2 and qgraph version 1.4.4 were used to plot the networks and to compute and visualize the centrality indices. Autoregressive components were not included in the outstrength and instrength (Epskamp et al., 2012). We compared the group network models and centrality indices of the high and low happy bias group by means of visual inspection. Next, we explored individual differences within the two groups by plotting instrength and outstrength for each person separately, based on the person-specific effects.

In addition to the visual comparisons of the networks and centrality indices, we performed seven permutation tests to test the hypothesized differences between the high and low happy bias groups. Significant results on the permutation test suggest differences between high and low happy bias in the population. The permutation tests compared the observed differences of interest to distributions of possible differences under the null-hypothesis of no differences between the groups. Distributions of possible differences were derived from reshuffling the groups randomly 10,000 times, also called Monte Carlo sampling. For each reshuffle, differences between the two reshuffled groups were estimated with the lmer function of R package lme4, that is, in the same way as the original models had been defined in mlVAR. If an observed difference between the high and low happy bias groups was within 2.5% on either side of the distribution of the 10,000 possible differences, the difference between the high and low happy bias group was considered significant (i.e., $p < .05$). We used an adapted version of the permutation test developed by Viechtbauer (Snippe et al., 2017) to test differences between the high and low happy bias groups which match our hypotheses as described in the Introduction. See Table 1 for a description of the seven hypotheses and their operationalization for the permutation tests.

All of the tested difference scores were based on the fixed effects of the group models. For permutation tests (1) and (4) we used absolute edge weights in order to avoid that expected positive and expected negative associations cancel each other out. All of our hypotheses and therefore also all permutation tests applied to outstrength. We explored possible differences in instrength between the high and low happy bias groups by visual comparison of the networks and centrality plots, and did not use permutation tests because we did not have clear hypotheses in advance.

Finally, we performed multiple sensitivity analyses to explore the robustness of our findings. First, we repeated the mlVAR analyses in Mplus version 8 (Muthén & Muthén, 1998a), which allowed multivariate mlVAR analyses with a Bayesian estimator. Second, although the decision to use extreme groups rather than continuous happy bias measures was driven by valid conceptual and methodological considerations, there were no clear criteria on how extreme the groups should be and therefore the exact number of individuals selected for each group (i.e., 25) was somewhat arbitrary. To assess the robustness of the results based on groups of

**TABLE 1.** Description and Operationalization of the Seven Hypotheses Tested With the Permutation Tests

| | Description | Permutation Test |
|---|---|---|
| **Hypothesis 1** | JOY and POS are stronger predictors in the network of the high happy bias group than in the network of the low happy bias group | **(1)** The total summed absolute edge weight of all outgoing edges from JOY and POS at time *t-1* to all nodes in the network at time *t* (including autoregressive edges) is larger for the high happy bias group than for the low happy bias group |
| **Hypothesis 2** | JOY and POS more strongly predict themselves (i.e., are more easily sustained over time) and each other (i.e., more carry-over between JOY and POS) over time in the high happy bias group than in the low happy bias group | **(2)** The total summed edge weight of all outgoing edges from JOY and POS at time *t-1* to JOY and POS at time *t* (including autoregressive edges) is larger for the high happy bias group than for the low happy bias group |
| **Hypothesis 3** | JOY and POS more strongly predict the negative nodes (i.e., larger dampening effect on negative nodes) over time in the high happy bias group than in the low happy bias group | **(3)** The total summed edge weight of all outgoing edges from JOY and POS at time *t-1* to SAD, IRR, WOR, and NEG at time *t* is larger for the high happy bias group than for the low happy bias group |
| **Hypothesis 4** | The negative nodes are stronger predictors in the network of the low happy bias group than in the network of the high happy bias group | **(4)** The total summed absolute edge weight of all outgoing edges from SAD, IRR, WOR, and NEG at time *t-1* to all nodes in the network at time *t* (including autoregressive edges) is larger for the low happy bias group than for the high happy bias group |
| **Hypothesis 5** | The negative nodes more strongly predict themselves (i.e., are more easily sustained over time) and each other (i.e., more carry-over between the negative nodes) over time in the low happy bias group than in the high happy bias group | **(5)** The total summed edge weight of all outgoing edges from SAD, IRR, WOR, and NEG at time *t-1* to SAD, IRR, WOR, and NEG at time *t* (including autoregressive edges) is larger for the low happy bias group than for the high happy bias group |
| **Hypothesis 6** | More pronounced negative associations between negative nodes and JOY and POS (i.e., larger dampening effect on JOY and POS) over time in the low happy bias group than in the high happy bias group | **(6)** The total summed edge weight of all outgoing edges from SAD, IRR, WOR, and NEG at time *t-1* to JOY and POS at time *t* is larger for the low happy bias group than for the high happy bias group |
| **Hypothesis 7** | JOY and POS more strongly predict INT in the high than in the low happy bias group | **(7)** The total summed edge weight of all outgoing edges from JOY and POS at time *t-1* to INT at time *t* is larger for the high happy bias group than for the low happy bias group |

JOY = feeling joyful; POS = pleasant experiences; INT = feeling interested in the things around me; SAD = feeling sad; IRR = feeling irritated; WOR = worrying; NEG = unpleasant experiences.

25 individuals, we estimated the networks, computed the centrality indices and performed the permutation tests for bias groups of 20, 30, 35 and 40 individuals. We used the same mlVAR methods as in the main analyses. As a third check, to adjust for anhedonia status, we computed subject-specific centrality indices based on the random estimates of the edges of the low and high happy bias networks, and subsequently regressed the subject-specific centrality indices on anhedonia status and happy bias. See section 3-5 in the Supplementary Material for further details.

## Data Availability

Data and syntax have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/4czv3/.

# RESULTS

## Descriptive statistics

### *General demographics*

The high and low happy bias groups were quite comparable in terms of age, gender, and education (see Table 2). Although in the low happy bias group more participants attended university, in both groups all participants were enrolled in higher education. The groups differed considerably in symptoms of anhedonia.



**TABLE 2.** General Demographics and Anhedonia Status

| | **High happy bias group (*n* = 25)** | | **Low happy bias group (*n* = 25)** | |
|---|---|---|---|---|
| | ***Mean* (*SD*)/ Count (%)** | | ***Mean* (*SD*)/ Count (%)** | |
| Age | 21.64 | (1.77) | 20.69 | (2.05) |
| Females | 20 | (80%) | 22 | (88%) |
| University education | 13 | (52%) | 18 | (72%) |
| Higher vocational education | 12 | (48%) | 7 | (28%) |
| Anhedonic[a] | 8 | (32%) | 13 | (52%) |
| Control[a] | 17 | (68%) | 10 | (40%) |
| Switcher[a] | 0 | (0%) | 2 | (8%) |

[a] Participants were classified as anhedonic or control if they met all criteria at T0 and did not change in pleasure levels from one group to the other at either T1 or T2. Otherwise they were classified as switcher.

The descriptive statistics for the facial emotion identification variables are presented in Table 3. The high happy bias group had a mean happy bias score of 1.82, which means that happy facial emotions were identified on average 1.82 times faster than the negative facial emotions sadness, anxiety and fear. The low happy bias group had a mean happy bias score of 1.23, which indicates that this group is on average still faster in identifying happy facial emotions, but the difference between happy and the negative emotions is only small.

**TABLE 3.** Descriptive Statistics of Facial Emotion Identification Scores

| | **High happy bias group (*n* = 25)** | | **Low happy bias group (*n* = 25)** | |
|---|---|---|---|---|
| | ***Mean*** | ***SD*** | ***Mean*** | ***SD*** |
| RT Total | 5522.61 | 574.66 | 5236.69 | 848.37 |
| RT Happy | 3449.08 | 377.55 | 4550.20 | 1034.25 |
| RT Sad | 6797.62 | 801.18 | 6051.65 | 1014.42 |
| RT Angry | 5761.44 | 705.29 | 5189.62 | 916.74 |
| RT Fearful | 6094.69 | 883.37 | 5168.86 | 821.34 |
| Happy bias score | 1.82 | 0.14 | 1.23 | 0.13 |
| Happy bias rank | 36.84 | 18.92 | 218.96 | 22.25 |

RT = Reaction Time; RT Total = mean score on RT Happy, RT Sad, RT Angry and RT Fearful; Happy bias score = mean score on RT Sad, RT Angry and RT Fearful divided by RT Happy; Happy bias rank = summed rank of happy bias score at T0 and T2

Table 4 presents the descriptive statistics of the network variables. On average, the high happy bias group scored higher than the low happy bias group on the positive nodes (JOY, POS and INT) and lower on the negative nodes (SAD, IRR, WOR and NEG). Within-person *SD*s were quite similar across the groups, with the largest differences for IRR and WOR. *MSSD*s of all nodes were larger in the low happy bias group than in the high happy bias group and in both groups for all nodes *MSSD* ≥ 50 for almost all participants, that is, in the high happy bias group 5 participants had an *MSSD* < 50 on SAD, 2 on IRR and 1 on NEG; in the low happy bias group 3 participants had an MSSD < 50 on SAD, 1 on INT, 1 on IRR and 2 on WOR. Both the high and low happy bias group showed low numbers of missings per person on the momentary assessments, for both groups the mean number of missings per person was 6.2 (out of 90), with min. = 1 and max. = 17.

***Descriptive statistics network models***

The network models for the high happy bias group and the low happy bias group are visualized in Figure 4; only the significant edges with *p*-values < .05 are depicted (see Table S1 in the Supplementary Material for the exact coefficients and significance levels of all paths). Green edges represent positive and red edges represent negative associations from one node at time *t-1* to another node at time *t*; the thickness of the edges indicates the strength of the

associations. As we within-person standardized all variables, the edge coefficients represent the change in terms of within-person *SD* in the outcome variable based on one within-person *SD* increase in the predictor variable.

**TABLE 4.** Descriptive Statistics of the Momentary Assessment Items Used as Nodes in the Networks

| | ***Mean*** | | **Average within-person *SD*** | | **Average within-person *MSSD* (*n* with *MSSD* < 50)** | |
|---|---|---|---|---|---|---|
| | **High bias *n* = 2094** | **Low bias *n* = 2095** | **High bias** | **Low bias** | **High bias** | **Low bias** |
| JOY | 60.27 | 56.07 | 12.98 | 12.38 | 237 (0) | 248 (0) |
| POS | 63.35 | 60.06 | 14.76 | 14.02 | 302 (0) | 312 (0) |
| INT | 58.31 | 53.30 | 13.28 | 13.35 | 263 (0) | 275 (1) |
| SAD | 13.96 | 18.49 | 11.16 | 12.34 | 225 (5) | 262 (3) |
| IRR | 15.42 | 19.87 | 13.02 | 15.03 | 302 (2) | 384 (1) |
| WOR | 21.88 | 22.88 | 13.75 | 15.26 | 291 (0) | 310 (2) |
| NEG | 32.82 | 39.27 | 18.75 | 18.79 | 565 (1) | 554 (0) |

JOY = feeling joyful; POS = pleasant experiences; INT = feeling interested in the things around me; SAD = feeling sad; IRR = feeling irritated; WOR = worrying; NEG = unpleasant experiences; *MSSD* = average within-person Mean Squared Successive Difference.
*Note*. These descriptive statistics are based on data from which linear time trends and cyclic time of day trends have already been removed.



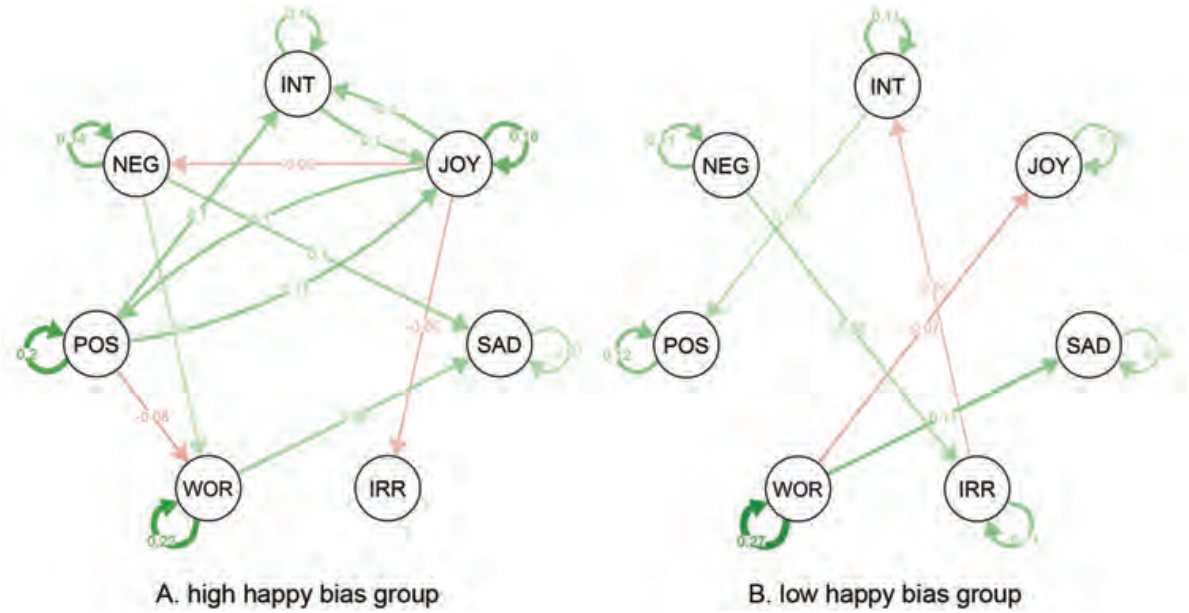



**FIGURE 4.** Significant associations networks high happy bias group (A) and low happy bias group (B). JOY = feeling joyful; POS = pleasant experiences; INT = feeling interested in things around me; SAD = feeling sad; IRR = feeling irritated; WOR = worrying; NEG = unpleasant experiences.

**Associations between positive nodes at time *t-1* and positive nodes at time *t*.** For both groups all autocorrelations of the positive nodes JOY, POS and INT were significant. Autocorrelations of JOY and POS were higher, and JOY, POS and INT were more densely connected to each other in the high happy bias group than in the low happy bias group. The positive edges suggest that an increase in one of the positive nodes is associated with an increase in the others at the next measurement.

**Associations between negative nodes at time *t-1* and negative nodes at time *t*.** For both groups, we found significant autocorrelations of the negative nodes WOR, NEG and SAD, with a higher autocorrelation for WOR in the low happy bias group than in the high happy bias group. The autocorrelation of IRR was only significant in the low happy bias group. The negative nodes SAD, IRR, WOR and NEG showed several positive temporal interrelations in both groups.

**Associations between positive nodes at time *t-1* and negative nodes at time *t*.** For the high happy bias group, a higher score on POS predicted a lower score on WOR, and a higher score on JOY predicted lower scores on IRR and NEG. For the low happy bias group we did not find negative edges from POS and JOY to negative nodes at the next measurement.

**Associations between negative nodes at time *t-1* and positive nodes at time *t*.** A higher score on negative nodes was significantly associated with a lower score on positive nodes at the next measurement for the low happy bias group only, that is, WOR and IRR showed negative associations with JOY and INT.

***Descriptive statistics centrality indices***

**Centrality plots on group level.** Centrality plots for outstrength and instrength are presented in Figure 5, panel A. JOY and POS had the highest outstrength in the high happy bias group and the lowest outstrength in the low happy bias group, indicating that JOY and POS most strongly predicted the other nodes in the high happy bias group and least strongly predicted the other nodes in the low happy bias group. We found the largest differences in instrength between the high and low happy bias group for WOR, JOY and INT, all of which were more strongly predicted by other nodes in the high happy bias group than in the low happy bias group.

**Individual variation in centrality indices.** Individual variation in outstrength and instrength within the two happy bias groups is presented in Figure 5, panel B. In general, instrength showed more individual variation than outstrength. Regardless of the individual variation, the ranges of the high and low happy bias groups hardly overlapped on the outstrength of JOY, indicating that the low and high happy bias groups could be most clearly discriminated on the outstrength of this node.

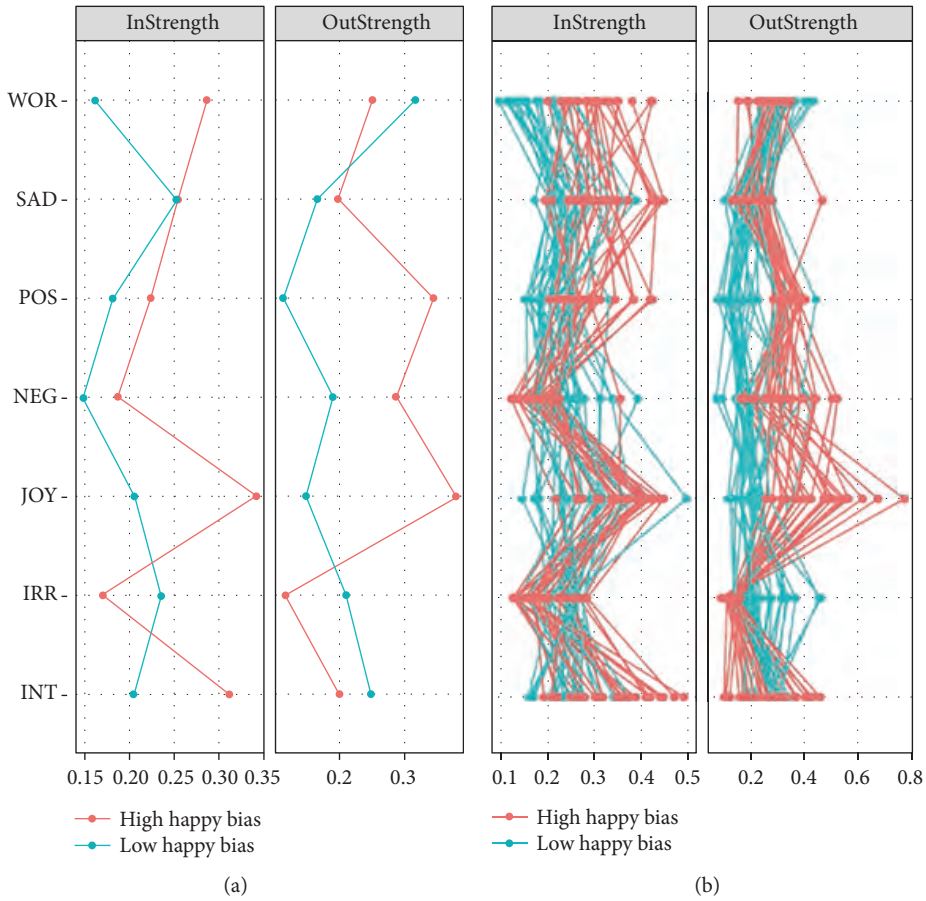



**FIGURE 5.** Centrality indices instrength and outstrength based on the complete network models. In panel **A** the indices are plotted for the high and low happy bias groups and panel **B** illustrates the individual variation within these groups. JOY = feeling joyful; POS = pleasant experiences; INT = feeling interested in things around me; SAD = feeling sad; IRR = feeling irritated; WOR = worrying; NEG = unpleasant experiences.

*Note.* Estimations for panel B are based on multilevel group models (fixed effects) and individual variation within these groups (random effects). No separate individual models were estimated and the lines in panel B should not be interpreted as such.

## Permutation tests of differences between the low and high happy bias group

As a brief reminder, the following seven hypotheses were tested: (1) JOY and POS are stronger predictors in the network of the high happy bias group than in the network of the low happy bias group; (2) JOY and POS more strongly predict themselves and each other over time in the high happy bias group than in the low happy bias group; (3) JOY and POS more strongly predict the negative nodes over time in the high happy bias group than in the low happy bias group; (4) the negative nodes are stronger predictors in the network of the low happy bias group than in the network of the high happy bias group; (5) the negative nodes more strongly predict themselves and each other over time in the low happy bias group than in the high happy bias

group; (6) more pronounced negative associations between negative nodes and positive nodes over time in the low happy bias group than in the high happy bias group; (7) JOY and POS more strongly predict INT in the high than in the low happy bias group. See Table 1 for a more detailed description of the hypotheses and their operationalization for the permutation tests.

Only permutation tests 1 and 2 reached significance at $p < .05$. The observed difference of the total absolute strength of all outgoing edges from JOY and POS between the high and low happy bias groups was 0.61, $p < .01$ (permutation test 1). The observed difference between the two groups of the total strength of all outgoing edges from JOY and POS to JOY and POS was 0.31, $p < .01$ (permutation test 2). We found no significant differences between the two groups for: the total strength of all outgoing edges from JOY and POS to negative nodes (observed difference = -0.24, $p = .13$; permutation test 3); the total absolute strength of all outgoing edges of the negative nodes (observed difference = -0.14, $p = .57$; permutation test 4); the total strength of all outgoing edges from negative nodes to negative nodes (observed difference = -0.15, $p = .36$; permutation test 5); and the total strength of all outgoing edges from negative nodes to JOY and POS (observed difference = 0.14, $p = .18$; permutation test 6); and the total strength of all outgoing edges from JOY and POS to INT (observed difference = 0.12, $p = .06$; permutation test 7).

### Sensitivity analyses

The multivariate Mplus multilevel VAR analyses showed minor differences for several of the network model estimates, but general patterns and main findings were confirmed (for more details see Section 3 of the Supplementary Material). Networks, centrality indices and permutation tests for extreme happy bias groups of 20, 30, 35 and 40 individuals showed the same patterns as the ones based on the original groups of 25 individuals. As expected, group differences became less pronounced as groups became larger and for $N$=40 one of the two permutation tests was no longer significant. For more details, see Section 4 of the Supplementary Material. Finally, controlling for anhedonia status did not change our findings (See section 5 of the Supplementary Material for more details).

## DISCUSSION

Our study is the first to investigate what a bias for happy facial emotions as assessed by a standardized laboratory task pertains to in daily life. We found that feelings of joy and pleasant experiences were stronger predictors in the network of the high happy bias group than in the network of the low happy bias group **(hypothesis 1)**; and that in the high happy bias group joy and pleasant experiences more strongly predicted themselves and each other over time **(hypothesis 2)**. These were robust findings based on both visual inspection of the networks

and centrality indices, and permutation tests. Other group differences were only found by visual comparison but could not be corroborated by the permutation tests: joy and pleasant experiences dampened the negative nodes (i.e., sadness, irritation, worrying, and unpleasant experiences) in the high but not in the low happy bias network **(hypothesis 3)**; the negative nodes dampened joy and pleasant experiences in the low but not in the high happy bias network **(hypothesis 6)**; and joy and pleasant experiences predicted interest in the high but not in the low happy bias network **(hypothesis 7)**. These group differences were present in the study sample, but since the permutation tests were not significant it is not possible to draw inferences about group differences in the population. The fact that the permutation tests did not reveal group differences regarding these hypotheses may be due to large individual differences or small effects. We found no support that negative nodes more strongly predicted the overall affect network or the negative network in the low happy bias group than in the high happy bias group **(hypotheses 4 and 5)**. Although the high and low happy bias groups differed considerably in symptoms of anhedonia, the differences we found between the high and low happy bias groups seem to be driven primarily by happy bias status and could not be (fully) explained by anhedonia status.

Our more specific results may be discussed in terms of the extent to which a certain node predicts other nodes at the next time point (outstrength) or in terms of the extent to which a certain node is predicted by other nodes at the previous time point (instrength). Because our hypotheses applied to outstrength and we did not have clear hypotheses about instrength in advance, our methodological approaches toward outstrength and instrength differed. For outstrength we focused more on hypothesis-testing by means of permutation tests, whereas we used a more exploratory approach based on visual inspection for instrength. Both perspectives will be discussed, starting with the outstrength.

We found that joy and pleasant experiences more strongly predicted affect over time in the high happy bias group than in the low happy bias group. This suggests that individuals with a high happy bias show a bias toward positive affect and positive experiences in their daily lives and may be better capable of sustaining positive affect and positive experiences over a longer period of time and generalizing it to other positive components than individuals with a low happy bias. This is particularly important because these same mechanisms, that is, the inability to sustain positive affect over time (Heininga et al., 2017; Heller et al., 2009; Höhn et al., 2013; Horner et al., 2014; McMakin et al., 2009) and the inability to generate positive effect from pleasant experiences (Geschwind et al., 2010; Wichers et al., 2010), have been associated with depression in previous studies. In the present study we also found indications that the same specific daily life affect dynamics that are associated with a low happy bias are also associated with depressive symptoms. That is, we found that for individuals suffering from anhedonia, which is one of the two core symptoms of depression, joy and pleasant experiences were weaker predictors of affect in the next six hours (see Section 5 of the Supplementary Material). Furthermore, daily

life momentary positive affect during one month has been found to predict life satisfaction and a higher ability to adapt to changing environments after this month (Cohn et al., 2009), which suggests that positive affect in the moment broadens one's attentional scope, and facilitates building valuable cognitive and social resources essential to well-being (Cohn et al., 2009; Fredrickson, 1998). If feelings of joy can be sustained longer and spread to other positive experiences their beneficial influence may be prolonged too. The savoring of positive affect, which includes the anticipation as well as the prolongation of positive affect, has been found to be associated with more life satisfaction and happiness, and with lower levels of neuroticism, depression, and anhedonia (Bryant, 2003).

In previous studies it has also been found that positive affect facilitates recovery from negative experiences (Fredrickson & Levenson, 1998) and that resilient individuals use positive affect to downregulate negative affect (Tugade & Fredrickson, 2004). This is in accordance with our findings that joy and pleasant experiences dampened negative nodes in the high but not in the low happy bias network, and suggests that individuals with a high happy bias may be better equipped to use positive experiences and positive affect to regulate negative affect, thoughts, and experiences than the low happy bias group. However, caution is warranted in interpreting this finding because, although in all different sensitivity analyses joy and pleasant experiences dampened negative nodes in the high but not in the low happy bias group, the permutation test did not reach statistical significance. Therefore no conclusions can be drawn about group differences in the population. It is possible that the permutation test was not significant because of large individual variation in edges from joy and pleasant experiences to negative nodes, but this is only speculation. The regulation of negative nodes by means of positive nodes may also essentially occur on a shorter time-frame, for example, 2 hours. If this is indeed the case then the current method only picked up what was still left of the initial effect several hours later.

With regard to instrength, that is, the extent to which a certain node is predicted by other nodes at the previous time point, we found the largest differences between the high and low happy bias group for worrying, joyfulness and feeling interested, all of which were more strongly predicted by other affect components in the high happy bias group than in the low happy bias group. A possible explanation is that this reflects psychological flexibility, such that for individuals with a high happy bias worrying, joyfulness and feeling interested are more dependent on context. How this might work can be illustrated by comparing the high and low happy bias networks in Figure 4. The level of worrying is influenced by pleasant and unpleasant experiences in individuals with a high happy bias, whereas for individuals with a low happy bias worrying seems to be less dependent on context, have a higher autocorrelation, and consequently tends to lead its own life. As psychological flexibility has been found to be highly important for optimal functioning in many situations, and psychological rigidity has been associated with depression as well as other forms of psychopathology (Kashdan & Rottenberg, 2010; Kuppens, Sheeber, Yap, Whittle, Simmons, & Allen, 2012), the high happy bias group seems to be better off.

Sensitivity analyses were performed to explore the effects of different estimators, statistical packages, group sizes, a multivariate approach, and controlling for anhedonia status. All of these sensitivity analyses but one supported the original main findings completely; for the largest happy bias groups (*N*=40) only partial support was found in the sense that one of the two original main findings was no longer significant. As expected, group differences became less pronounced as groups became larger. Two plausible explanations are, first, that for the happy bias groups of *N*=40 the stability assumption of mlVAR analysis was not met, and, second, that only happy bias in the extremes of the distribution may be associated with the development of adaptive versus maladaptive affective patterns in daily life.

Strengths of our study are, first of all, that we combined the best of two worlds by using a multilevel approach in which within-subject effects were separated from between-subject effects by within-person standardization of all variables prior to the analyses. This enabled us to explore dynamic processes that take place within individuals; at the same time it allowed us to compare the two happy bias groups (Schuurman et al., 2016). Secondly, following recent developments in the field (Klippel et al., 2017; Snippe et al., 2017), in addition to visual comparison of the affect networks and centrality indices, we used permutation methods adapted to our specific hypotheses to test statistically whether the happy bias groups differed in their affect dynamics. A third strength of this study is its high ecological validity, as we assessed affect and related measures three times a day in daily life situations, for a period as long as 30 days, and achieved compliance rates of at least 80%. Additionally, the use of a morph task allowed us to assess the identification of more subtle traces of emotions, which is assumed to give a more ecologically valid perspective than static full intensity facial emotion identification tasks, as in daily life static full intensity facial emotions are quite rare. Finally, we repeated the facial emotion identification task and based our selection of the happy bias groups on individuals' scores on both tasks. This enabled us to select only those participants with a stable happy bias. This was necessary because the daily life affect networks were estimated over a period of 30 days and participants showing large shifts in happy bias from one happy bias group to the other during this period would have added noise to the network models.

Evidently there are also limitations to our study. First, our sample largely consisted of higher educated females, which may limit the generalizability of our findings because gender and level of education may moderate the associations we investigated (Bagozzi, Wong, & Yi, 1999; Demenescu et al., 2014; Houben, Van Den Noortgate, & Kuppens, 2015; McClure, 2000; Nolen-Hoeksema & Aldao, 2011; Strand, Dalgard, Tambs, & Rognerud, 2003). Second, the selection of extreme and stable happy bias groups resulted in small groups of 25 participants, which limits generalizability and resulted in insufficient power to correct for anhedonia status or use multivariate multilevel analysis, which would require the estimation of additional parameters. The disadvantage of the univariate approach is that correlations between the dependent variables and between random effects of the dependent variables were not taken into account.

We presented sensitivity analyses to show the effects of a multivariate approach, different group sizes, and controlling for anhedonia status. However, most of these alternative approaches required multiple conceptual and methodological concessions and can only be interpreted as proxies to our original models. Third, we offered a network approach in which only unique direct temporal effects were studied, and no shared effects (Bringmann et al., 2016; Bulteel et al., 2016). As such, our approach should be regarded as complementary to approaches that take into account shared variance. Fourth, our results were based on assessments that were on average six hours apart. We were unable to grasp dynamic processes that took place within a shorter time frame. Finally, a limitation of our study that applies to all non-experimental study designs is that we cannot make inferences about true causality; our conclusions are confined to 'Granger' causality, that is, if a variable at time *t-1* contains unique information to predict a second variable at time *t*, it is said to Granger cause this second variable (Granger, 1969). We investigated the directed associations between different affect components over time and it is plausible that other factors that were not included in our models explain part of these dynamics and therefore no true causal claims can be inferred from our network models.

Further research is required, first of all to confirm our findings by replication in other samples. Because of the present study's small group sizes the conclusions are tentative awaiting attempts to replicate. Second, the specific conditions in which happy bias influences daily life affect dynamics need to be explored, for example, how extreme the bias needs to be before predicting positive or negative outcomes regarding well-being or mental health. Third, although our findings are promising, it should be noted that the ability to sustain positive emotions has been operationalized in many different ways in previous studies and there are also inconsistencies and unresolved issues, for example with respect to autocorrelation. It has been found that a stronger daily life autocorrelation of positive emotions over time protects against depression (Höhn et al., 2013), but also that strong autocorrelations, for positive as well as negative emotions, predict depression (Houben et al., 2015; Kuppens et al., 2012). Further research is needed to investigate adaptive and maladaptive effects of strong autocorrelations versus psychological flexibility. It seems plausible that strong autocorrelations may indicate resistance to change and thereby limit psychological flexibility, but equally plausible that no carry-over of positive affect and positive experiences over time (weak autocorrelation) may also not be very adaptive. It may be important to consider different time-scales (Koval, Pe, Meers, & Kuppens, 2013), to look at proportions of autocorrelation in relation to cross-lagged paths (relative influence of other nodes) and to distinguish between high autocorrelation with respect to low and high levels of positive and negative affect and experiences. Finally, depression is a heterogeneous construct and specific subtypes of depression may be differentially associated with affect dynamics. A low happy bias could reflect such a subtype and our study suggests

that it can be useful to take happy bias into account when studying affect dynamics. Depressed individuals with a low happy bias may show different affect dynamics compared to depressed individuals with a high happy bias, but this remains to be investigated.

## Conclusions

We compared young adults with a high bias for happy facial emotions during a standardized laboratory task to peers with a low bias for happy facial emotions on their daily life affect dynamics, using a highly personalized approach in which we separated within-subject from between-subject effects. Our most important and robust finding was that joy and pleasant experiences more strongly predicted the affect network of the high happy bias group than that of the low happy bias group. These findings tentatively suggest that individuals with a high happy bias are more responsive to positive, rewarding, experiences and emotions, and more easily sustain them, whereas positive experiences and emotions seem to be more short-lived in the daily life of individuals who lack this happy bias. We propose that high reward responsiveness may be reflected in both a high happy bias during facial emotion identification and the ability to sustain and generalize positive experiences and positive affect in daily life. This may be key to why individuals with a bias toward happy facial emotions are potentially more resilient to developing depression. By using a network approach to compare the daily life affect dynamics of individuals with a high and with a low happy bias we came closer to understanding the daily life mechanisms behind high and low happy bias during a laboratory task. This novel perspective is valuable for interpreting facial emotion processing tasks, as are often assessed in clinical research and practice. The present study illustrates the potential benefits of a network approach for unraveling psychological mechanisms.

# SUPPLEMENTARY MATERIAL

Section 1 of the supplementary material provides detailed descriptions of the No Fun No Glory (NFNG) selection procedures and the facial emotion identification morph task. Section 2 presents the exact coefficients and significance levels of the main analyses, followed by the results of a simulation study that we used to assess the reliability of mlVAR for our specific sample size, number of time points and model specifications. In Sections 3-5 results are presented of the sensitivity analyses we performed to assess the robustness of our findings.

## 1. Procedures

### *Selection of participants for the ecological momentary assessments based on the screening survey*

We based the selection of individuals with persistent anhedonia on a general pleasure item from the Domains of Pleasure Scale (Masselink et al., n.d.). Inclusion criteria were: (1) low levels of pleasure compared to peers, that is, below the 25$^{th}$ percentile; (2) a loss of pleasure, that is, the pleasure level they had been experiencing during the past two weeks was lower than the level they considered normal for themselves; (3) persistent anhedonia, that is, a loss of pleasure persisting for more than two months. Exclusion criteria were: inability to complete an electronic diary three times a day, use of psychotropic medication, professional therapeutic treatment for psychiatric problems and pregnancy. As a tandem skydive was one of the possible interventions offered in the NFNG Project, unwillingness to perform a tandem skydive, factors that prevented safe participation in a tandem skydive (epilepsy, cardiovascular problems, visual or hearing impairments, loose prostheses, a height of more than two meters, a weight of more than 95 kg, incapability to raise legs 90 degrees), and prior experience with skydiving, bungee jumping or base jumping were exclusion criteria as well. This screening process yielded a group of 148 anhedonic individuals, of whom 28 no longer met the inclusion criteria when contacted, 22 refused to participate and 29 could not be included for other reasons, resulting in a group of 69 participants. For each participant included in the anhedonia group, we matched a control participant on age, sex and educational level. Inclusion criteria for the control group were: (1) at least moderately high pleasure levels compared to peers, that is, above the 50$^{th}$ percentile; (2) no loss of pleasure, that is, the pleasure level they had been experiencing during the past two weeks had to be equal to or higher than the level they considered normal for themselves. Exclusion criteria for the control group were identical to those for the anhedonia group. Of the 114 individuals we sent an invitation to participate in the control group one participant no longer met the inclusion criteria, 21 refused to participate and 23 could not be included for other reasons. For a detailed description of the NFNG study, see Van Roekel and colleagues (2016,

2017). Please note that the interventions of the NFNG project started only after the first month of momentary assessments (observation month) and the second facial emotion identification task, therefore they did not interfere with the present study in any way.

***Compensation participants***

Participants received 10 euros for completing the screening survey and additionally participated in a lottery (prizes: 1 travel cheque, 4 iPad minis, and 15 fashion cheques). Participants who were included in the intervention study received a further compensation of 75 euros after completion of the observation month (of which the data were used in the present study), 125 euros after the first intervention month, 200 euros after the second intervention month, and 50 euros for each of the two follow-up measures. To receive these compensations, participants were required to complete at least 80% of the momentary assessments, as well as all monthly questionnaires and blood samples, which were part of the larger NFNG study. A tandem skydive was part of one of the interventions in the NFNG study (data were not used for the present study) and the opportunity to perform a tandem skydive free of charge could also be considered a compensation.

Because the tandem skydive as well as our financial compensation may have attracted a sample with altered reward sensitivity, we checked this by comparing participants who were unwilling to perform a tandem skydive to participants who indicated to be willing or perhaps willing to perform a tandem skydive in the total screening sample ($N$ = 2,937) on reward responsiveness (Van den Berg, Franken, & Muris, 2010). We also compared the reward responsiveness of the participants who were selected for the intervention study and agreed to participate to those who were not selected or did not agree to participate. Reward responsiveness did not significantly differ (i.e., $p > .05$) between participants who were unwilling ($N$ = 376) and those who were willing or 'perhaps' willing ($N$ = 2,561) to perform a skydive, and also not between participants who were selected for the intervention study and agreed to participate ($N$ = 138) and those who were not selected or did not agree to participate ($N$ = 2,799). Therefore, our data hold no evidence that the tandem skydive or the financial compensation participants received in the intervention study attracted a sample with altered reward sensitivity.

***Facial emotion identification task***

We used a morph task developed at Radboud University Nijmegen, the Netherlands (Lodder et al., 2015; Vrijen, Hartman, Lodder, et al., 2016). Stimuli consisted of movie clips that lasted 10 seconds and contained 100 frames depicting the gradual change (i.e., 'morph') from a neutral facial expression to one of four full intensity emotional expressions: happiness, sadness, anger or fear (for examples, see the task description at https://osf.io/9edkh/). The movies had a resolution of 256 x 256 pixels, and were created with FaceMorpher (Luxand Inc., Alexandria, VA, USA) from high quality pictures of six different actors (50% females) from the Radboud Faces Database (Langner et al., 2010). Pictures were cropped with an ovoid frame and converted to gray scale to avoid distracting external cues. Four movies were created of each actor, that is, one for

each emotional expression. The original task contained 48 movie clips, that is, twelve per facial emotion, whereas we used a shortened version of 24 movie clips, that is, six for each emotion. A previous study in a large sample of young adults (Vrijen, Hartman, Lodder, et al., 2016) indicated that the emotion identification patterns and reaction times for the shortened version are highly similar to the ones reported for the original 48-video-clip version of the morph task (Lodder et al., 2015).

The morph task was programmed in Inquisit 4 (Millisecond, Seattle, USA). The task started with the instruction that participants were about to see movies of faces gradually changing from neutral to emotional expressions. Participants were asked to press the space bar as soon as they were able to identify the emotion. After pressing the spacebar the stimulus movie disappeared and participants indicated the emotion they identified by clicking on one of the four emotion labels. After clicking 'next' a fixation cross appeared in the middle of the screen for 500 ms, followed by a new stimulus. The order of the movie clips was randomized for each participant separately. Before the start of the actual task participants were shown a complete 10-s example movie, followed by two practice trials. After the practice trials the instructions were repeated, followed by the actual task consisting of 24 trials.

For each participant the mean reaction time (RT) of correctly identified trials was calculated per emotion, resulting in RT Happy, RT Sad, RT Angry and RT Fearful. RTs were calculated only if participants correctly identified at least four out of six movie clips of a specific emotion, otherwise they were considered unreliable. This resulted in 2 missing values for RT Sad, 1 for RT Angry and 2 for RT Fear at T0; and 1 missing value for RT Sad, 2 for RT Angry and 2 for RT Fear at T2. None of the remaining RT scores reached the maximum value of 10,000 ms, which indicates that the participants always pressed the spacebar before the movie clips stopped. As mentioned in our description of the selection procedure of the high and low happy bias groups, we calculated happy bias scores for each participant by taking the average of their RT scores on the other emotions (RT Sad, RT Angry and RT Fearful) and dividing it by RT Happy. Happy bias was used to select participants for the high and low happy bias group.

## 2. Exact coefficients and significance levels of mlVAR network analyses and results simulation study

**TABLE S1.** Standardized Coefficients for all Paths in the Network

| | JOY | | POS | | INT | | SAD | | IRR | | WOR | | NEG | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **High bias** | **Low bias** | **High bias** | **Low bias** | **High bias** | **Low bias** | **High bias** | **Low bias** | **High bias** | **Low bias** | **High bias** | **Low bias** | **High bias** | **Low bias** |
| $JOY_{t-1}$ | **0.161*** | **0.090*** | **0.101*** | 0.030 | **0.098** | 0.046 | -0.021 | 0.007 | **-0.060** | 0.041 | -0.039 | -0.017 | **-0.060** | 0.007 |
| $POS_{t-1}$ | **0.105*** | 0.018 | **0.200*** | **0.122*** | **0.104*** | 0.034 | -0.028 | -0.023 | -0.007 | -0.016 | **-0.076*** | -0.017 | 0.025 | -0.005 |
| $INT_{t-1}$ | **0.100** | 0.034 | 0.037 | **0.066** | **0.102*** | **0.113*** | 0.014 | -0.046 | -0.011 | -0.036 | 0.016 | -0.041 | -0.021 | -0.025 |
| $SAD_{t-1}$ | 0.051 | 0.001 | -0.002 | 0.010 | 0.016 | 0.007 | **0.070** | **0.079*** | 0.035 | 0.041 | 0.051 | 0.059 | 0.043 | 0.047 |
| $IRR_{t-1}$ | -0.020 | -0.058 | -0.014 | -0.002 | 0.031 | **-0.054** | -0.001 | 0.028 | 0.041 | **0.111*** | -0.045 | -0.023 | 0.007 | 0.045 |
| $WOR_{t-1}$ | -0.032 | **-0.069*** | -0.043 | -0.040 | 0.026 | -0.053 | **0.095*** | **0.112*** | 0.022 | 0.023 | **0.217*** | **0.274*** | 0.033 | 0.021 |
| $NEG_{t-1}$ | -0.034 | -0.025 | 0.026 | -0.034 | -0.038 | 0.010 | **0.096*** | 0.037 | 0.035 | **0.078*** | **0.059** | 0.006 | **0.143*** | **0.109*** |

JOY = feeling joyful; POS = pleasant experiences; INT = feeling interested in the things around me; SAD = feeling sad; IRR = feeling irritated; WOR = worrying; NEG = unpleasant experiences. Bold coefficients were significant at $p < .05$. Bold coefficients marked with an * were significant at $p < .01$ and bold coefficients marked with an * and marked gray were significant at $p < .001$.

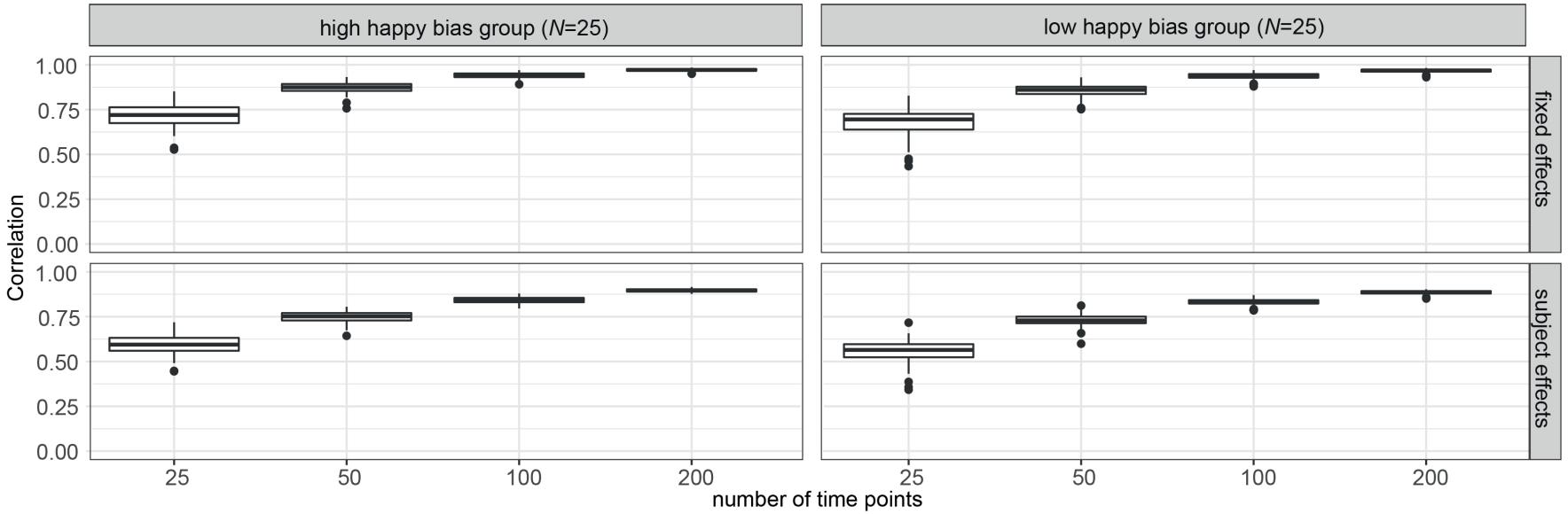



**FIGURE S1.** Results of a simulation study to assess the reliability of mlVAR temporal networks of seven nodes based on 25 subjects per group and a variable number of time points per person (25-200). In total 800 datasets were simulated, that is, 100 for each of the eight conditions (2 happy bias groups * 4 different time point conditions). The subject specific temporal networks found in the present study were generated as the true subject specific network structures. In order to assess how well the estimated networks resemble the true networks, we computed for each dataset the correlations between the true and the estimated subject specific temporal networks and between the true and the estimated fixed temporal networks. Boxplots indicate the distribution of the measures over all of the 100 simulated datasets per condition. In the top panel results are presented for the fixed effects and in the bottom panel for the subject specific effects (i.e., random effects), in the left panel for the high happy bias group and in the right panel for the low happy bias group.

## 3. Sensitivity analyses Mplus, multivariate models

A new version of Mplus (version 8) was launched recently. One of the new features is a package for dynamical structural equation modeling (DSEM) that uses a Bayesian estimator and allows the use of multivariate techniques for samples with a large number of assessments per participants and a small number of participants per group, like ours. We used this new package to perform a sensitivity check. As before, R packages qgraph version 1.4.4 and igraph version 1.1.2 were used to plot the networks and to compute and visualize the centrality indices.

Multivariate lag 1 network models were estimated for the low and high happy bias group separately by using the default specifications of DSEM, that is, Bayesian estimation with non-informative priors based on two independent Markov Chain Monte Carlo chains, and the potential scale reduction (PSR) criterion was used to assess convergence (Gelman & Rubin, 1992; Muthén & Muthén, 1998a). PSR < 1.1 was used as the default convergence criterion, but PSR < 1.05 has been recommended as well. For both bias groups the models converged after 800 iterations, that is, a PSR below 1.1 was reached. Recommendations to increase the number of iterations to check whether the PSR increased again were followed by using 20000 iterations. The PSR slightly increased and showed values > 1.1 after 800 iterations, but after, respectively, 2100 (high happy bias group) and 1800 (low happy bias group) iterations PSR < 1.05 was reached and the PSR remained stable and below 1.05 for the remainder of the 20000 iterations. This long sequence of low PSR values indicates model convergence. The network models and centrality indices reported in Figures S2 and S3 were based on 20000 iterations.



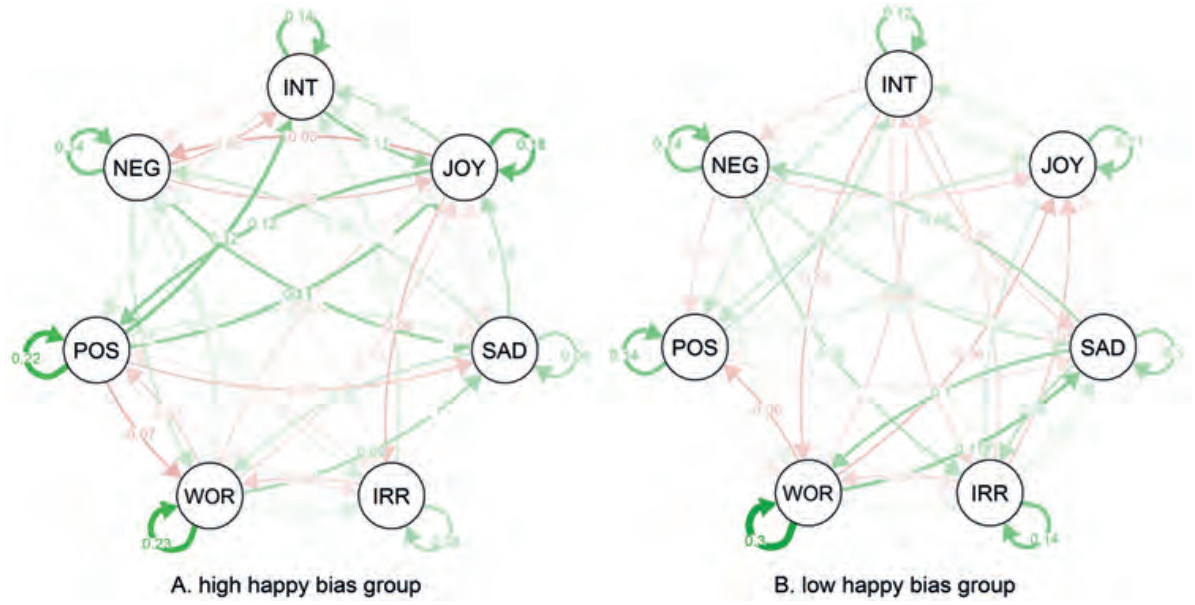



**FIGURE S2.** Complete networks high happy bias group (A) and low happy bias group (B), Mplus multivariate analyses. JOY = feeling joyful; POS = pleasant experiences; INT = feeling interested in things around me; SAD = feeling sad; IRR = feeling irritated; WOR = worrying; NEG = unpleasant experiences.

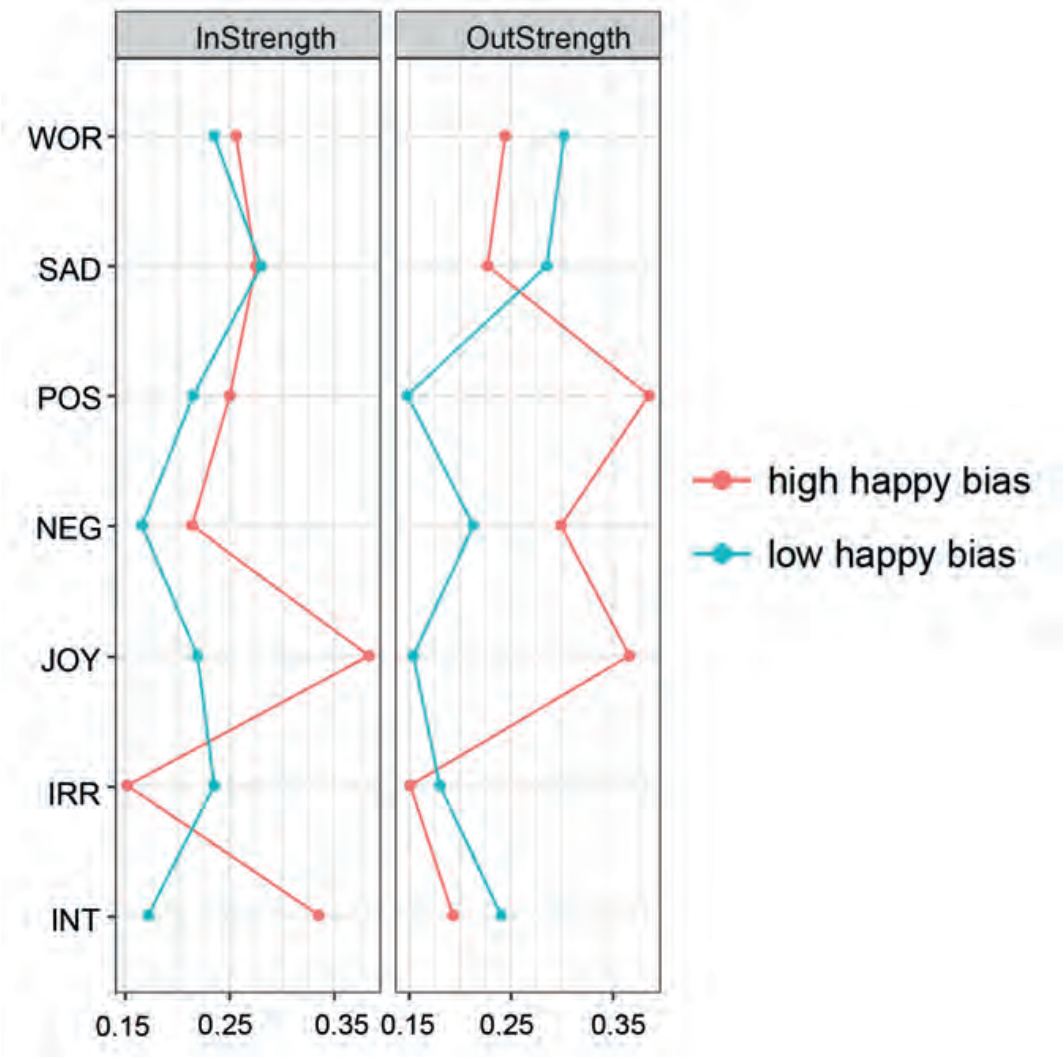



**FIGURE S3.** Centrality indices instrength and outstrength based on the complete network models, Mplus multivariate analyses. JOY = feeling joyful; POS = pleasant experiences; INT = feeling interested in things around me; SAD = feeling sad; IRR = feeling irritated; WOR = worrying; NEG = unpleasant experiences.

Even though this approach differed notably from the original one (e.g., a Bayesian estimator was used and a larger part of the information matrix was estimated by taking into account correlations between the dependent variables), the dynamical patterns are strikingly similar and the results confirm our main findings.

## 4. Sensitivity analyses for different group sizes

In the main analyses we used extreme happy bias groups of 25 participants in each group. Because there were no clear criteria on how extreme the groups should be, the exact number of individuals selected for each group was somewhat arbitrary. To assess the robustness of the results based on groups of 25 individuals we estimated the networks, computed the centrality indices, and performed permutation tests for bias groups of 20, 30, 35 and 40 individuals. We used the same mlVAR methods as in the main analyses.

For the groups of 25 individuals used in our main analyses permutation tests 1 and 2 reached statistical significance (see results section). We found that this was also the case for happy bias groups of *N*=20, *N*=30 and *N*=35 (see Table S2).

**TABLE S2**. Results Permutation Tests for Different Group Sizes

| Group sizes of high and low happy bias group | Permutation test 1 | | Permutation test 2 | |
|---|---|---|---|---|
| | Observed difference | *P* | Observed difference | *P* |
| *N*=20 | 0.52 | 0.015 | 0.28 | 0.011 |
| *N*=25[a] | 0.61 | 0.002 | 0.31 | 0.002 |
| *N*=30 | 0.53 | 0.002 | 0.26 | 0.002 |
| *N*=35 | 0.39 | 0.019 | 0.21 | 0.013 |
| *N*=40 | 0.33 | 0.053 | 0.17 | 0.046 |

*Note.* [a] Main analyses; the others are sensitivity analyses.
**Permutation test 1**: Is the total (i.e., summed) absolute edge weight of all outgoing edges from JOY and POS, including the autoregressive edges, larger in the high happy bias group than in the low happy bias group?
**Permutation test 2**: Is the total edge weight of the outgoing edges from JOY and POS to JOY and POS, including the autoregressive edges, larger in the high happy bias group than in the low happy bias group?

For groups of 40 individuals the observed differences between the two happy bias groups were only half the size of those for the original groups of 25 individuals, and only permutation test 2 reached statistical significance. The results of these sensitivity analyses suggest that differences between the low and high happy bias groups become less pronounced as groups become less extreme. On OSF the network plots and instrength and outstrength plots are available for all groups (*N*=20 - *N*=40; https://osf.io/w823j/).

## 5. Anhedonia as a potential confounder

To assess the plausibility that anhedonia status may drive the results we found for happy bias we first tested whether the general affect dynamics we found to be associated with high versus low happy bias were also associated with anhedonia status. Permutation test 1, which was originally used to test whether the reward-related positive nodes JOY and POS more strongly predicted

affect dynamics in the high happy bias group than in the low happy bias group, was repeated to explore whether a similar difference in affect dynamics could also be found between a control group and an anhedonia group. For the purpose of this sensitivity analysis, an anhedonia status variable was construed in which stability was taken into account. We started with the complete sample of participants who completed the first month of momentary assessments (*N*=138). Participants were assigned to the anhedonia group (*N*=60) only if their pleasure levels were low at T0 (< 25$^{th}$ percentile) and remained low at T1 and T2, that is below the 35$^{th}$ percentile, and they were assigned to the control group (*N*=59) only if their pleasure levels were high at T0 (> 50$^{th}$ percentile) and remained above the 40$^{th}$ percentile. Participants with less stability in anhedonia were treated as missing (*N*=19). We found that JOY and POS more strongly predicted affect dynamics in the control group than in the anhedonia group (observed difference = 0.286, $p$ = 0.047). This suggests that at least partly similar affect dynamics are associated with happy bias and anhedonia status and that it is therefore important to adjust for anhedonia status.

To be able to adjust for anhedonia status, the subject-specific affect network paths (i.e., the random effects) based on the original VAR models were used to calculate the subject-specific total strength of the outgoing edges from JOY and POS. This resulted in a separate value for each participant of how strongly JOY and POS together predicted the other nodes in the network and themselves over time. Subsequently, we ran a linear regression analysis in SPSS version 25 with the total strength of all outgoing edges of JOY and POS per participant as outcome variable and dummy variables of anhedonia status and high versus low happy bias as predictors. Adjusting for anhedonia status still resulted in happy bias predicting the total strength of the outgoing edges from JOY and POS ($\beta$ = .87, $p$ < .001). This suggests that the effects we found for happy bias are not, at least not fully, driven by anhedonia status. For a plot of the individual variation in outstrength and instrength plotted separately for subjects with and without anhedonia, see Figure S4.

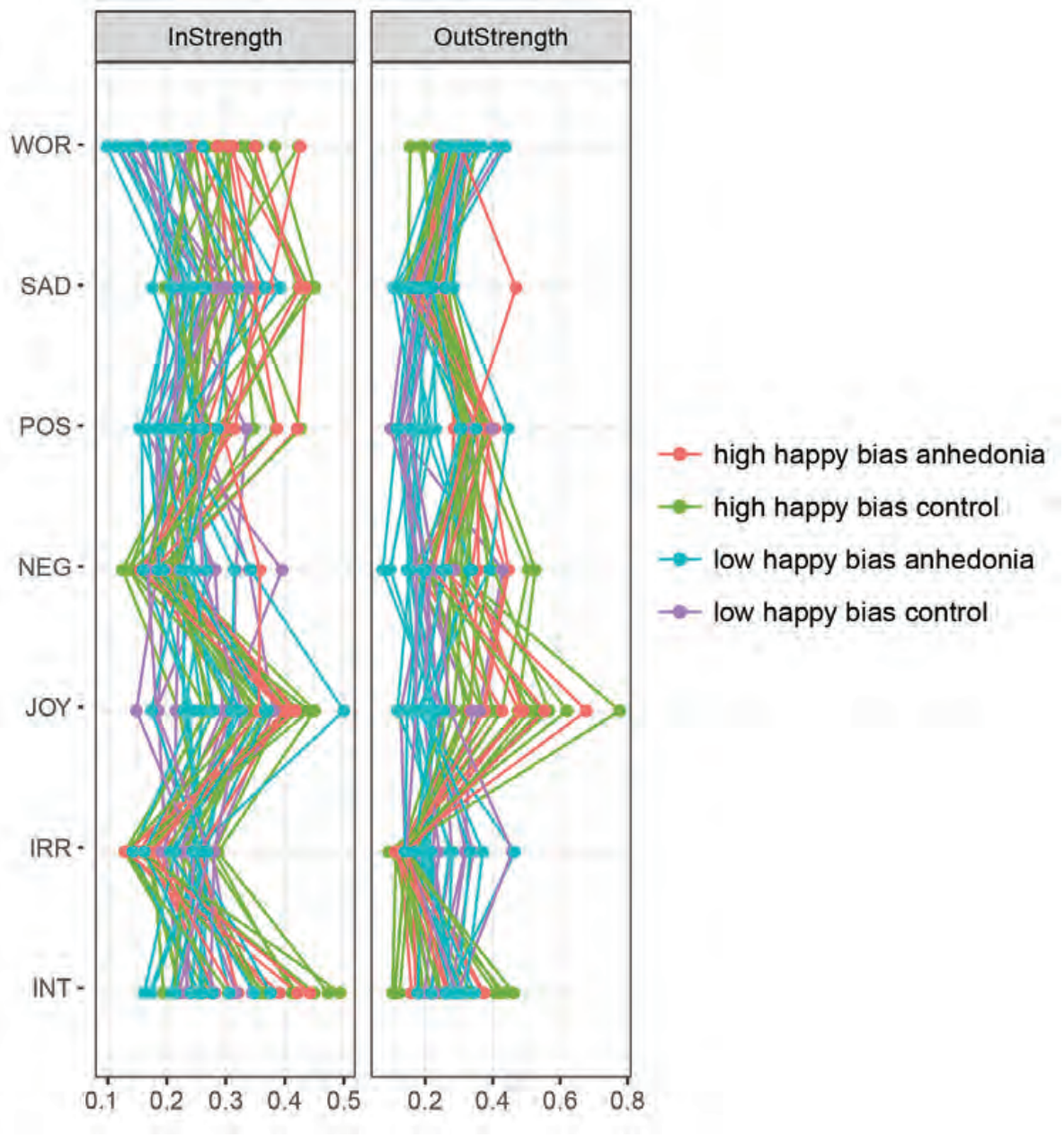


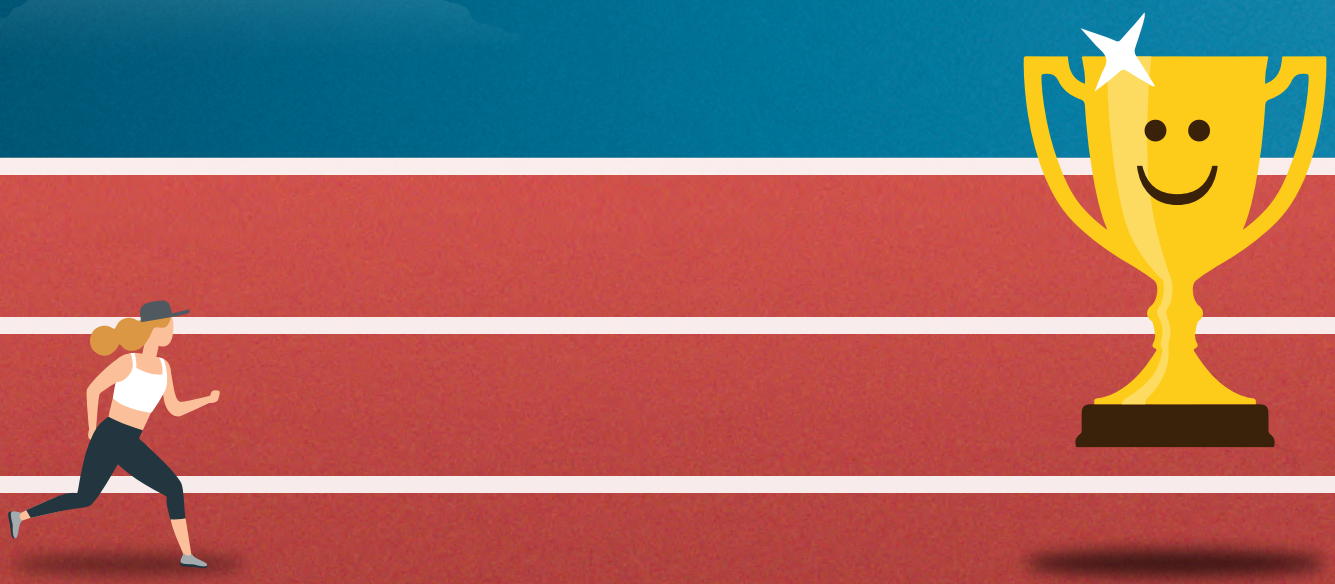

**FIGURE S4.** Individual variation in instrength and outstrength in four subgroups: anhedonic participants with a high happy bias (*n* = 8), anhedonic participants with a low happy bias (*n* = 13), control participants with a high happy bias (*n* = 17) and control participants with a low happy bias (*n* = 10). JOY = feeling joyful; POS = pleasant experiences; INT = feeling interested in things around me; SAD = feeling sad; IRR = feeling irritated; WOR = worrying; NEG = unpleasant experiences.

# CHAPTER 6

## An exploratory randomized controlled trial of personalized lifestyle advice and tandem skydives as a means to reduce anhedonia

Published as:
Van Roekel, E., Vrijen, C., Heininga, V. E., Masselink, M., Bos, E. H., & Oldehinkel, A. J. (2017). An Exploratory Randomized Controlled Trial of Personalized Lifestyle Advice and Tandem Skydives as a Means to Reduce Anhedonia. *Behavior Therapy*, *48*(1), 76–96. https://doi.org/10.1016/j.beth.2016.09.009

The Addendum contains additional analyses not included in the published article



## ABSTRACT

Anhedonia is a major public health concern and has proven particularly difficult to counteract. It has been hypothesized that anhedonia can be deterred by engagement in rewarding social and physical events. The aims of the present study were to examine (1) the effects of personalized lifestyle advice based on observed individual patterns of lifestyle factors and experienced pleasure in anhedonic young adults; and (2) whether a tandem skydive can enhance the motivation to carry out the recommended lifestyle changes. Participants (N=69; $M_{age}$=21.5, SD=2.0; 79.7% female) were selected through an online screening survey among young adults. Inclusion criteria were persistent anhedonia and willingness to perform a tandem skydive. Participants filled out questionnaires on their smartphones for two consecutive months (3 times per day). After the first month, they were randomly assigned to one of three groups: (1) no intervention, (2) lifestyle advice, and (3) lifestyle advice and tandem skydive. The momentary questionnaire data were analyzed using interrupted time series analyses (ITSA) in a multilevel model and monthly pleasure and depression questionnaires by repeated measures ANOVA. No group differences were found in monthly depression and pleasure scores, but the momentary data showed higher positive affect (PA) and pleasure ratings in the month following the intervention in the two intervention groups than in the control group. The tandem skydive did not have any effects above the effects of the lifestyle advice. Our results indicate that providing personalized lifestyle advice to anhedonic young adults can be an effective way to increase PA and pleasure.

## INTRODUCTION

Anhedonia is defined as diminished interest or pleasure in activities that were experienced as pleasurable before (i.e., a loss of pleasure; American Psychiatric Association, 2013). Anhedonia is common among adolescents and young adults and has debilitating consequences, such as a longer and more severe course of depression (Wilcox & Anthony, 2004) and increased risk for suicide (Nock & Kazdin, 2002). Anhedonia can have particularly negative consequences in young adulthood (Gabbay et al., 2015; Kupferberg, Bicks, & Hasler, 2016), as this is a period in which life course decisions are made based on what is satisfying or enjoyable. Despite its detrimental effects, anhedonia has proven particularly difficult to treat (Fusar-Poli et al., 2015; Treadway & Zald, 2011; Wykes, Steel, Everitt, & Tarrier, 2008) and is the symptom that often remains when other symptoms have disappeared (Gutkovich et al., 2010). Further, there is a lack of treatments targeting anhedonia specifically. Hence, there is a need for effective interventions to help anhedonic individuals regain pleasure. In the present study, we aimed to explore whether anhedonia can be deterred by providing anhedonic young adults with personalized lifestyle advice.

Anhedonia is a transdiagnostic symptom that is present in multiple diagnostic categories, such as depression, bipolar disorder, and schizophrenia (American Psychiatric Association, 2013). The few studies that have examined interventions to reduce anhedonia so far have focused particularly on individuals with schizophrenia (Favrod, Giuliani, Ernst, & Bonsack, 2010; Favrod et al., 2015) or depression (Kramer et al., 2014). To our knowledge, two interventions have been developed to target anhedonia in schizophrenic patients: the Anticipatory Pleasure Skills Training (APST; Favrod et al., 2010) and the Positive Emotions Program for Schizophrenia (PEPS; Favrod et al., 2015). APST is a cognitive-sensory intervention aimed at increasing anticipatory pleasure, which consists of 10-25 hours of training. PEPS is a group therapy consisting of eight 1-hour sessions, focused on increasing the anticipation and maintenance of positive emotions. Both interventions have shown positive results in small samples (N=5 for APST; N=31 for PEPS), but have not yet been tested in randomized controlled trials (RCT). For depression, treatment often starts with behavioral activation (BA) treatment (Kanter et al., 2010), to stimulate depressed patients to participate in rewarding activities. In behavioral activation, the therapist and patient decide together which activities are experienced as pleasurable, and the patient is stimulated to set goals to re-engage in these pleasurable activities. BA has been found to be effective in reducing depressive symptoms (Cuijpers et al, 2007) and in reducing depressive symptoms and manic symptoms in bipolar disorder in a first explorative study (Weinstock, Melvin, Munroe, & Miller, 2016). These interventions are all fairly time-consuming, in that they consist of multiple sessions and require a trained psychologist to deliver the interventions. Furthermore, although activity scheduling is part of BA treatments and the PEPS intervention, these treatments lack a thorough and objective analysis of pleasurable experiences. This is important, as anhedonia

is characterized by difficulties recalling pleasurable experiences (Lepage, Sergerie, Pelletier, & Harvey, 2007), and therefore, anhedonic individuals may not be able to recall all activities that they enjoy.

Because of rapid technological developments, it is now possible to overcome these issues by using momentary assessments to provide individuals with insight in their daily life emotions and activities (Myin-Germeys, Klippel, Steinhart, & Reininghaus, 2016; Wichers et al., 2011). A main advantage of momentary assessments is their high ecological validity, as participants report on their mood and activities while they are living their daily lives (Myin-Germeys et al., 2009). Up to now, only one study has used momentary assessments to provide depressed individuals with personalized feedback on the situations in which they experience positive affect (PA) (Kramer et al., 2014). In this study, participants received weekly reports that showed how much PA they experienced in different contexts and during different activities. Importantly, no concrete advice was provided to participants, and the presented feedback was purely descriptive, i.e., the associations were not tested. Although no direct effects were found on momentary ratings of PA (Hartmann et al., 2015), individuals receiving the personalized feedback did decrease more in depressive symptoms on follow-up assessments than individuals who did not receive the feedback (Kramer et al., 2014). The authors did not explore how the personalized feedback affected anhedonia specifically. Whereas these results are very promising, the intervention could be taken one step further, by not only showing participants in which situations and environments they experienced PA, as Kramer et al. (2014) did, but also providing concrete suggestions regarding how they could change their lifestyle. In the present study, we aim to provide anhedonic young adults with personalized advice on which lifestyle factors (e.g., physical exercise) they can change to experience more pleasure, based on observed individual patterns of lifestyle factors and experienced pleasure.

However, even when tailor-made and concrete, lifestyle changes are likely hard to accomplish for anhedonic individuals, as anhedonia is often characterized by a lack of drive to pursue rewarding activities (Treadway, Bossaller, Shelton, & Zald, 2012; Treadway & Zald, 2011). An intense experience that will elicit strong positive responses in all individuals may be necessary to break this vicious circle of low pleasure and motivation, and to reboot the reward system (Grillo, 2016). A free-fall experience such as skydiving may be an acceptable and effective human model to accomplish this goal (Chatterton, Vogelsong, Lu, & Hudgens, 1997), as such an extreme thrill is likely to provide a shock to the reward system and could potentially reset the way individuals respond to their environment. Given the lack of research on this particular topic, it is hard to predict what the precise mechanism is through which a free-fall experience could increase pleasure and motivation. However, the hypothesis that such a shock could provide improvements in mood is very similar to the hypothesis that mood can be improved by means

of electroconvulsive therapy (ECT), which is currently still one of the most effective interventions for treating severe depressed patients and has been found to improve mood and anhedonia (American Psychiatric Association, 2008; Anderson & Fergusson, 2013; Taylor, 2007).

Research in healthy individuals has shown that skydiving provokes strong emotions and physiological responses. Even though it is considered to be safe, falling from a great height evokes a substantial fight-flight response in virtually all individuals (Hare, Wetherell, & Smith, 2013). Physiologically, skydiving results in increases in heart rate, blood pressure, and alpha-amylase and cortisol levels (Chatterton et al., 1997; Hare et al., 2013; Meyer et al., 2015). Psychologically, individuals experience extreme fear before and during the free fall experience (Hare et al., 2013), followed by euphoria afterwards (Meyer et al., 2015). This contrast in emotions can be explained by Solomon's opponent process theory of acquired motivation (Solomon, 1980), which states that humans automatically contrast extreme emotions, such as fear and euphoria.

In addition to these immediate increases in mood in human studies, animal research has provided additional evidence why a free fall experience may boost the reward system (Wang & Tsien, 2011). Mice who experienced a free fall showed increased firing of dopamine neurons in the ventral tegmental area in the brain, which is associated with reward-related motivation. Altogether, these findings indicate that free fall experiences could increase positive affect and motivation, and so set an excellent stage for implementing lifestyle advice that promote more persistent positive feelings.

## The Present Study



The first aim of this exploratory study was to examine whether personalized lifestyle advice, based on observed patterns of pleasure and lifestyle factors, can increase pleasure and positive affect (PA) and decrease depressive symptoms and negative affect (NA) in anhedonic young adults. Secondly, we explored the provoking idea that a free fall may foster the implementation of the lifestyle advice, by testing whether lifestyle advice combined with a tandem skydive had a more positive effect than lifestyle advice only. In order to explore whether a free fall might be effective we chose to investigate its most extreme form, i.e., a tandem skydive, to rule out the possibility that negative findings were due to the experience being not intense enough. This part of the study was a proof-of-concept, to explore the potential effects of skydiving. We investigated both questions in a sample of 69 anhedonic adults, who were randomly assigned to a control group receiving no intervention (N = 22), a group who received lifestyle advice (N = 22), or a group who received both lifestyle advice and a tandem skydive (N = 25). As effects at the group level do not necessarily apply to each individual, we conducted both group-level and individual-level analyses, to further explore whether the intervention may be effective for specific individuals. In contrast to sex differences in depressive symptoms in general and depressed mood specifically (Hankin & Abramson, 2001), levels of anhedonia do not differ between males and females (Bennik, Nederhof, Ormel, & Oldehinkel, 2014; van Roekel et al., 2015). Previous research was inconclusive

about whether the effects of personalized feedback showed sex differences (Kramer et al., 2014). To exclude confounding by possible sex differences, we selected an equal number of males and females in each intervention group.

## METHODS

### Participants

Participants were selected through an online screening survey among 2,937 young adults (*M age* = 21.4 years, *SD* = 1.9, 78% female) from the Northern part of the Netherlands. Inclusion criteria were persistent anhedonia and willingness to perform a skydive. Persistent anhedonia was defined as a pleasure level below the 25$^{th}$ percentile, which was experienced as lower than normal, and lasted for at least two months. Hence, the criteria involved not only a low level of pleasure, but also that this low pleasure level was considered egodystonic and unlikely to be transient. These criteria were assessed by means of three items from the Domains of Pleasure Scale (DOPS; Masselink et al., submitted), which assess (a) level of pleasure in the past two weeks, (b) whether this level represents a change compared to what is considered normal for this individual, and (c) the duration of the loss of pleasure, if any. Exclusion criteria were inability to keep an electronic diary three times a day; current professional treatment for psychiatric problems; use of psychotropic medication; epilepsy; pregnancy; conditions that obstruct participating in a tandem skydive (i.e., loose prostheses; height of more than 2 meters; weight of more than 95 kg; inability to raise one's legs 90 degrees; cardiovascular complaints or problems; and significant visual or hearing impairments); and experience with skydiving, bungee jumping, or base jumping. We aimed for 20 participants in each intervention group (i.e., 60 in total). Because we anticipated some drop-out, we decided to include 4 extra participants in each group (i.e., 72 in total).

For multilevel modeling, a minimum number of 50 Level 2 cases and 20 Level 1 cases is sufficient to accurately estimate regression coefficients and variances (Hox, 2002). As we included over 60 individuals (Level 2) and 180 momentary assessments (Level 1), we considered the power to be adequate to examine our main research question. For the repeated monthly measures, we conducted a post-hoc power analysis in Stata (StataCorp, 2013). Results showed that we needed 21 individuals per group (power = 0.8, alpha = 0.05) to detect a large difference between groups (Cohen's d = 0.8) with a correlation of 0.6 between the repeated assessments. Our smallest group consisted of N = 22 participants; hence it was determined that we had enough power to detect large differences between groups.

## Procedure

Participants were recruited in the northern part of the Netherlands through flyers, electronic learning environments, advertisements on social media, and invitations during lectures and classes. After subscribing on the study website (www.nofunnoglory.nl), participants received an email with the link to the online questionnaire. Participants received a gift card of 10 euro after completion of the questionnaire.

When participants fulfilled the inclusion and exclusion criteria in the screening survey and indicated willingness to participate in further research, they were contacted for the intervention study. They were sent an information letter and informed consent form by email. If they agreed to participate, they had to return the signed informed consent form by mail or email. When participants did not reply, they were sent a reminder or contacted by telephone. Upon receiving the signed informed consent form, participants were invited for an instruction session with one of our four team members, which took place at the University Medical Center Groningen. During this meeting, study staff verified whether participants met study criteria. All study procedures, including study materials, were explained to ensure participant comprehension. Furthermore, an appointment was made for the next meeting (i.e., intervention session), to be held approximately one month later. Please see Figure 1 for a flowchart of the study procedure.

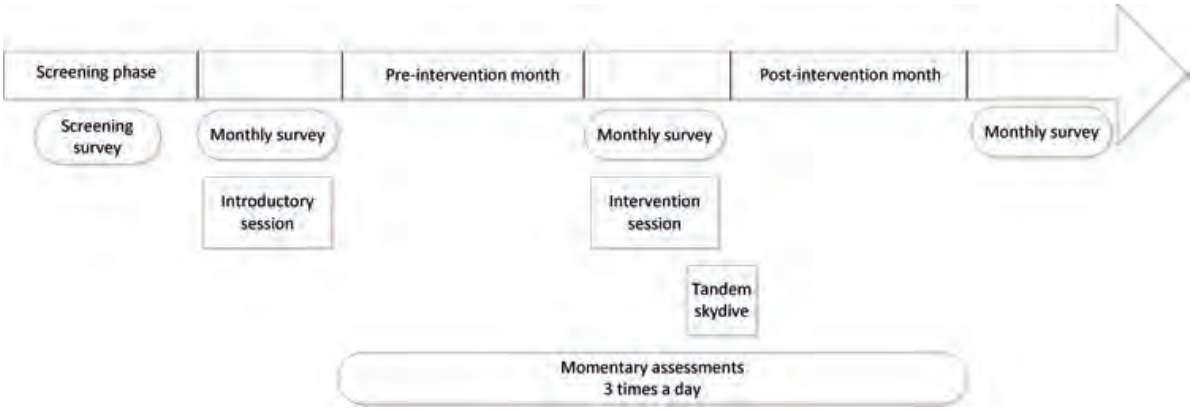



**FIGURE 1.** Flowchart of the study procedures

Participants started with the daily electronic questionnaires within a few days after the instruction session ($M = 1.97$, $SD = 2.47$, range 0-13), depending on when the intervention session could be planned. If scheduling issues dictated that the intervention session had to be planned more than a month after the instruction session, the start of the daily questionnaires was postponed accordingly. The participants received three questionnaires per day, with fixed 6-hour intervals (e.g., 9:00A.M., 3:00P.M., 9:00P.M.). The sampling scheme was determined in consultation with the participant. Each of the daily questionnaires was prompted by a text message containing a link

to the online questionnaire. The questionnaires had to be filled out within two hours after the first notification; reminders were sent after 60 and 90 minutes. Completion of the questionnaire took around three minutes. Participants continued with the daily questionnaires for about three months in total. The first month was the baseline phase, which was used to generate the lifestyle report. Further, data from the baseline phase were compared with data from the second month, the intervention phase, which was used to examine effects of the interventions. Since we wanted to give all participants the opportunity to receive the interventions, each participant was allowed to choose any intervention (that is, no intervention, lifestyle advice, lifestyle advice plus skydive) after the intervention phase. Because the randomization was compromised after the second month only the data assessed before the free choice intervention were included in the present study (i.e., first two months of data). In general, the compliance was excellent, with an average percentage of missing assessments of only 7.6% (range 0.5% - 22.9%).

Intervention allocation was done in four rounds. We used block randomization for males and females separately with a block length of 6 blocks. Hence, the three interventions were equally distributed after each group of 6 males or females. The sequence of the interventions in each block was determined using an online random block generator (www.randomization.com), and this sequence was entered in a blinded Excel file. During each round, the ID numbers of the participants that needed to be assigned to an intervention were entered in this Excel file in a random order (i.e., order was determined based on an online random sequence generator, www.random.org). Participants received their intervention allocation through email 1 week before the intervention. When participants were assigned to the tandem skydive group, they were contacted by telephone to make an appointment for that. If possible, the skydive always took place in the weekend following the intervention session, at a nearby airport. During a 1.5-hour individual meeting with one of the four team members (the intervention session), we discussed the progress with the participants as well as, if a participant belonged to one of the two intervention groups, the lifestyle report and advice (see below for detailed description of the intervention). During the meeting at the end of the intervention month, we discussed the extent to which participants had been able to follow-up on each lifestyle advice.

In addition to the daily electronic questionnaires, participants filled out online questionnaires measuring consummatory pleasure and depressive symptoms on the day of the instruction session (i.e., start of the first month), one day before the intervention session (i.e., end of the first month), and at the end of the intervention month (i.e., end of second month). As participants already knew which intervention they would receive at the second monthly assessment and during the final week of momentary assessments before the intervention session, these assessments were not included in the final analyses.

Participants received a compensation of 75 euros after completion of the first month, 125 euros after the second month, and 250 euros after completion of the third month, providing that they filled out the monthly questionnaires, provided blood samples (i.e., this study was part

of a larger study for which blood samples were included, see Van Roekel et al (2016) for the complete study protocol), and completed more than 80% of the momentary assessments. The present study is registered in the Dutch Clinical Trial Register (NTR5498) and was approved by the Medical Ethical Committee from the University Medical Center Groningen (no. 2014/508).

## Measures

### *Experience Sampling Method (ESM) measures used in the lifestyle reports*

The ESM items used in the lifestyle reports were based on previous ESM studies (e.g., van der Krieke et al., 2015; Van Roekel et al., 2015). The description of all ESM questions that were used to create a personalized lifestyle report for each individual can be found in Appendix A.

### *Pleasure*

We used one item to measure pleasure: 'I experienced pleasure since the last assessment'. Participants were asked to rate the extent to which this statement was applicable by moving a slider along a continuum (i.e., Visual Analogue Scale; VAS) anchored with the words: *not at all* on the left and *very much* on the right. The location of the slider was converted into a score between 0 and 100.

### *Positive Affect (PA)*

PA was measured with 10 items, namely, feeling interested, joyful, determined, calm, lively, enthusiastic, relaxed, cheerful, satisfied, and energetic. These items were rated on a similar VAS scale as described for pleasure (ranging from 0-100). The PA score was calculated by averaging the 10 item scores. Cronbach's alpha was .94 (calculated over all assessments).

### *Negative Affect (NA)*

NA was measured with 8 items: upset, gloomy, sluggish, anxious, bored, irritated, nervous, and listless, rated on a similar VAS scale (ranging from 0-100) as used for pleasure and PA. NA was calculated by averaging the 8 item scores. Cronbach's alpha was .86 (calculated over all assessments).

### *Consummatory pleasure*

Consummatory pleasure was measured in the monthly questionnaires with the DOPS. This scale consists of 21 items, which represent different domains of pleasure experiences: perceptual pleasures (8 items, e.g., I enjoy a good meal, I enjoy pleasant smells), social pleasures (5 items, e.g., I enjoy having close friendships, I enjoy doing things with other people), sexual pleasures (3 items, e.g., I enjoy sex [alone or with someone else]), and pleasure of personal achievements (5 items, e.g., I enjoy getting better at something, I enjoy winning in games or sports). Participants rated on a VAS scale ranging from *not at all* to *very much* (0-100) how much they enjoyed the

described experiences. In order to give each domain equal weight in the overall pleasure score, we first calculated mean scores for each domain, and subsequently averaged these means to create an overall consummatory pleasure score. The consummatory scale of the DOPS showed good convergent validity (Masselink et al., submitted). For the pre- and post-intervention assessments respectively, Cronbach's alphas were.87 and .89 for perceptual pleasures, .89 and .89 for social pleasures, .85 and .91 for sexual pleasures, and .81 and .84 for pleasure of personal achievement.

***Depressive symptoms***

We measured depressive symptoms with the 9-item Patient Health Questionnaire (PHQ-9 (Kroenke, Spitzer, & Williams, 2001)). Participants rated the extent to which they experienced 9 different symptoms in the past two weeks on a scale ranging from *not at all* to *almost every day*. Sample items are "feeling down, depressed or hopeless", "feeling tired or having little energy". As one of the items overlapped with the pleasure scale (i.e., "little interest or pleasure in doing things"), we excluded this item from the scale and created a sum score based on 8 items. The PHQ-9 has obtained good construct and criterion validity in previous research (Kroenke et al., 2001). Cronbach's alpha was .83 for the pre-intervention assessment and .80 for the post-intervention assessment.

***Psychiatric problems***

We measured the prevalence of other psychiatric problems with the Adult Self-Report (ASR) (Achenbach & Rescorla, 2001), which consists of 123 items that are rated on a 3-point scale ranging from *not at all* to *clearly/often*. We used the DSM-IV based scales for depressive problems (14 items; Cronbach's alpha = .77), anxiety problems (7 items; Cronbach's alpha = .76), avoidant personality problems (7 items; Cronbach's alpha = .68), ADHD problems (13 items; Cronbach's alpha = .78), and antisocial problems (20 items; Cronbach's alpha = .67).

***Advice adherence***

The information about advice adherence provided by participants in an interview at the end of the intervention phase was coded into an advice adherence variable, with 0 representing *not at all*, 1 representing *to some extent* and 2 representing *to a large extent*.

## Interventions

***Lifestyle advice***

The lifestyle advice was based on the first 30 days of ESM measurements (i.e., baseline phase; 90 assessments in total). Participants received a lifestyle report, which consisted of four main parts: (1) descriptive information, (2) comparison with norm groups, (3) personal networks, and (4) the advice. The specifics of the variables used in the lifestyle report can be found in Table A1.

All variables were used to construct personal networks, which were based on Autoregressive moving average (ARMA) analyses and automated Vector Autoregressive modeling (VAR) (Brandt & Williams, 2007; Autovar: Emerencia et al., 2015). Pleasure, social context, and activities were additionally used for descriptive purposes, and part of the activities and substance use were used for comparisons with norm groups. See Appendix B for a detailed description of the procedure of the lifestyle advice, the conducted analyses, and the content of the lifestyle report.

Based on the data collected during the baseline phase, potential advice was formulated. Because the personal networks provided information about statistically significant associations between pleasure and lifestyle factors, advice based on these networks were preferred. If no, or an insufficient number of, significant results were found, further advice were based on (a) infrequent lifestyle factors that could not be tested in the analyses, but were experienced as pleasurable according to the descriptive information, or (b) lifestyle factors with a frequency that deviated from the norm group. During the intervention session, the information in the lifestyle reports was explained to and discussed with the participants and they received a paper version of the report to take home, without the lifestyle advice. We checked whether the participants recognized themselves in the presented results and asked them to reflect on the results. Finally, we discussed the potential lifestyle advice with the participant, and decided together with the participants which recommendations were feasible and which were not. For the final advice, we came to an agreement with participants as to how often and when they were going to perform the suggested activities. Each participant received advice about two or three lifestyle factors. After the intervention session, the concept advice was reformulated when needed and final versions were sent to participants through email.

In total, 39 recommendations concerned social activities (e.g., plan more social activities with friends or family), 19 physical activity, 13 physical activity in combination with spending time outside, 20 worrying (e.g., mindfulness exercises), 5 going outside, 10 hobbies, 4 sleep rhythm, 9 time spent on television, internet, or social media, and 2 soft drugs (e.g., cannabis).

***Tandem skydive***

The tandem skydive took place at the certified skydive center 'Eelde-Hoogeveen', in the weekend following the intervention session. In case of bad weather (i.e., rain or hard wind), the skydive was rescheduled, if possible in the same weekend or one week later. The average number of days between the intervention session and the skydive was 5.5 (SD = 5.5, range 1-27 days). After arrival at the skydive center and around 15 minutes before the skydive, the participants received instructions from the skydive instructors. Participants exited a small turbine-powered aircraft (i.e., Cessna 207), at a height of 10,000 feet, safely attached to the tandem skydive instructor. The free fall lasted for around thirty to forty seconds, the total skydive around five minutes.

## Analyses

A randomization check was performed by comparing the intervention groups with regard to educational level (Chi square tests), age, and pre-intervention levels of pleasure and depression (ANOVAs). To examine the effects on the momentarily assessed outcomes (i.e., pleasure, PA, and NA), we used multilevel analyses in Mplus version 7.4 (Muthén & Muthén, 1998b). As we had at least 30 days of data (i.e., 90 assessments) before the intervention (i.e., baseline phase) and at least 30 days of data after the intervention (i.e., intervention phase), we used interrupted time series analyses (ITSA) to explore intervention effects (Huitema & Mckean, 2000). In ITSA, it is possible to explore changes in outcomes between the baseline phase and the intervention phase. Further, an important advantage of ITSA is that it is possible to explore whether changes in the outcome variables are due to trends that were already present in the baseline phase. In the first step, we examined whether the level of pleasure differed between the baseline and intervention phase. We examined the difference in mean level between both phases (i.e., level change) by adding a dummy variable for the phase to the model (i.e., baseline phase versus intervention phase) as a random effect. We modeled the interaction between group (0 for control group, 1 for both intervention groups) and phase to investigate whether the mean scores of the intervention groups had changed more than those of the control group after the intervention.

In the next step, we checked whether this change was gradual or abrupt and whether change was already present in the baseline phase by including time trends in the model, following recommendations by Huitema and McKean (2000). By estimating both the linear change in the baseline phase (i.e., slope), the change in this slope in the intervention phase (i.e., slope change), and the interaction between these time trends and group, we examined whether participants differed in the extent to which they gradually changed in their affect levels and whether this gradual change was already present in the baseline phase. Please note that the estimate for level change in this model no longer represents the change in mean level after the intervention, but the change in mean affect immediately after the intervention was provided. These analyses were repeated for PA and NA. We first compared the control group with both intervention groups, and subsequently examined differences between the lifestyle advice only group and the tandem skydive group. If significant group differences were found (i.e., significant cross-level interactions), we calculated the proportion of variance in the coefficient for level change that was explained by the cross-level interaction following guidelines by Aguinis, Gottfredson, and Culpepper (2013).

In case of significant between-individual differences in level estimates (i.e., significant random variance in the coefficient for the level change), we further examined individual differences in level change by fitting ITSA models for each individual and each outcome variable separately. We used the same design specifications as above (Huitema and McKean, 2000). The individual time-series analyses were controlled for autocorrelation by fitting Auto Regressive Moving Average (ARMA) models to the residuals (Hartmann et al., 1980). Dummy variables for time of

day (i.e., afternoon and evening) were included in each model to adjust for daily cycles. When the model fit was acceptable, we added the dummy variable for level change as a predictor in the model. The estimates for this predictor in the final models were saved. Part of these analyses (i.e., 25%) were repeated by another team member to check consistency of the model-building process. In order to visualize the individual level change estimates, we created forest plots, for each outcome variable separately. To compare whether individual characteristics (i.e., gender, severity of pleasure loss, duration of pleasure loss, advice adherence) were associated with the extent to which individuals improved, we conducted regression analyses or ANOVA (i.e., advice adherence) with the standardized level change estimates as outcome variables.

We further examined the effectiveness of the intervention by conducting repeated measures ANOVAs on the pre- and post-intervention assessments of pleasure and depressive symptoms. This was done in two steps: first we checked whether the two groups who received lifestyle advice differed from the control group. Second, we further compared the tandem skydive group with the lifestyle advice only group.

Effect sizes were calculated for all analyses. For the momentary data, we calculated Cohen's *d* following the formula as suggested by Cohen (1988), dividing the mean group difference by the pooled pre-intervention *SD*. For the mean difference between groups, we included the estimated difference in level effects between the control group and the intervention groups. To calculate the pre-intervention pooled *SD*, we first calculated mean levels of the outcome measure (pleasure, PA, and NA, respectively) in the pre-intervention month per individual. Subsequently, we calculated the *SD* over these within-person mean levels in both intervention groups. In this way, the pooled *SD* reflects the between-person variation in the outcome measures. For the monthly assessments, we used the formula as suggested by Morris (2008), using the pooled pre-intervention *SD*.



## RESULTS

### Participant Flow

In total, 71 participants started with the momentary assessments (see Figure 2 for a flowchart). Two participants dropped out from the study: One due to problems with mobile data subscription and another because she stopped filling out the momentary assessments. One participant did not partake in the skydive due to the possibility of panic attacks related to skydiving. She only received the lifestyle advice but was maintained in the tandem skydive group in the analyses, in line with the intention-to-treat principle. Finally, one person from the lifestyle advice only group was removed from the momentary data analyses as she reported that she had changed her interpretation of the questions with regard to pleasure and PA after the intervention session, which thwarted comparison of data from the baseline and intervention phase. This resulted in a

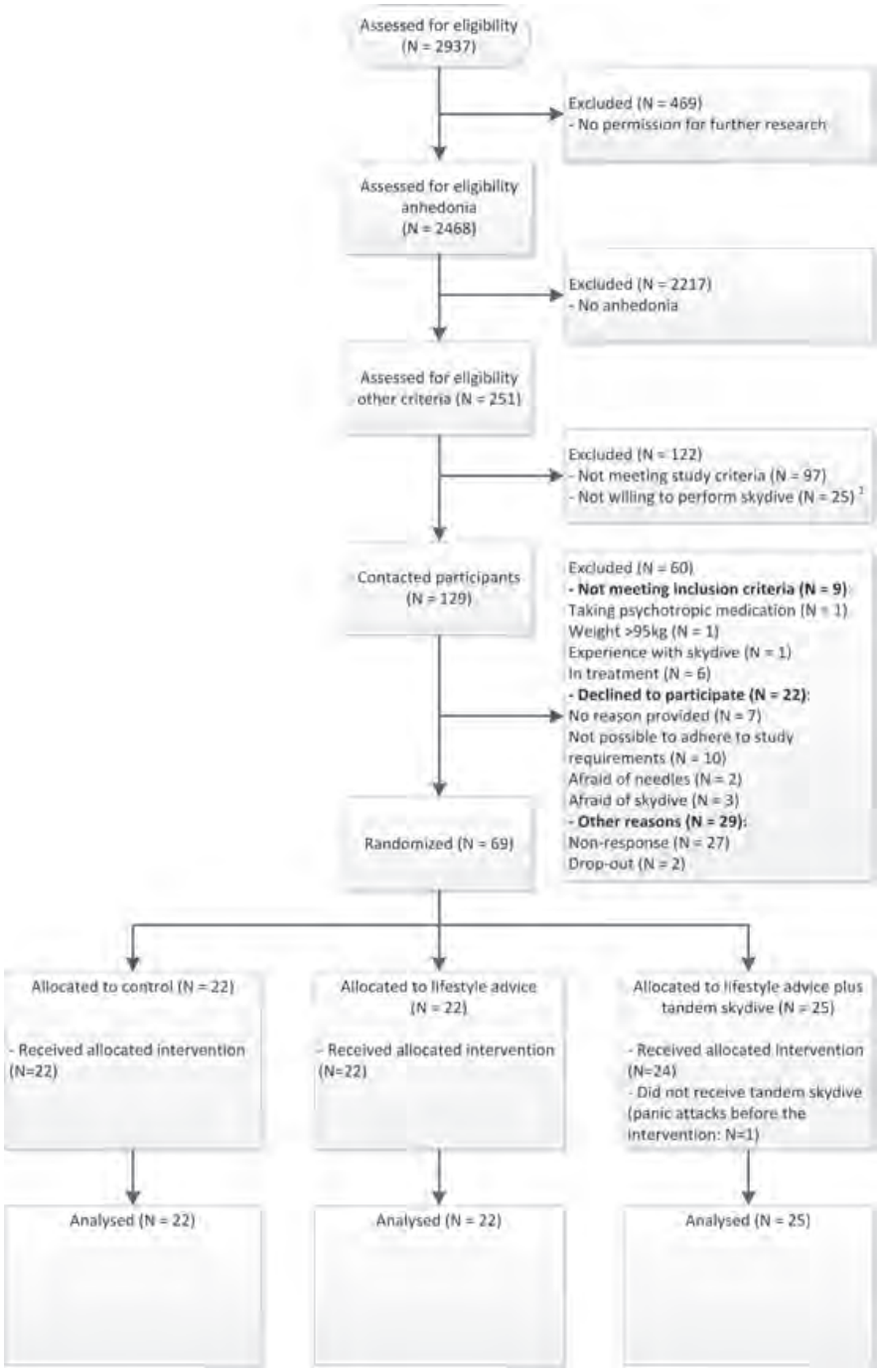



**FIGURE 2.** Flowchart of the participant enrollment.[8] We checked whether the group who was not willing to perform a skydive (N = 25) differed from the group who was willing to perform a skydive (answer yes or maybe), and found no significant differences between these groups on the severity of anhedonia ($t$ = 1.30, $p$ = .13), the level of consummatory pleasure ($t$ = -0.84, $p$ = .40) or depressive symptoms ($t$ = 0.30, $p$ = .77).

8 For one participant in the tandem skydive group, the skydive was delayed for more than 2 weeks (i.e., 27 days). Although we extended the intervention period for this individual so that we had enough assessments after the skydive, this delay could have affected the results. Therefore, we checked whether exclusion of this participant affected our findings. No relevant differences were found for any of the analyses.

final sample of 69 participants for inclusion in the monthly data analyses and 68 participants (22 no intervention, 21 lifestyle advice, 25 lifestyle advice plus tandem skydive) in the momentary data analyses. The number of assessments in the baseline phase ranged between 60 and 135 ($M$ = 84.5, $SD$ = 12.9), the number of assessments in the intervention phase ranged between 69 and 165 assessments ($M$ = 107.7, $SD$ = 15.8).

## Descriptive Statistics

Because we were interested in how knowledge about intervention allocation affected participants in the different intervention groups, we conducted additional analyses in which we compared the first three weeks of momentary assessments with the final week before the intervention session, and the second monthly assessment with the first monthly assessment. Paired samples t-tests showed that aggregated levels of pleasure, PA, and NA did not differ between the first three weeks and the final week of the baseline period in the no intervention group and in the tandem skydive group ($p > .05$). For the lifestyle advice group, mean levels of PA ($t$ [21] = -3.30, $p = .003$) were significantly higher and mean levels of NA ($t$ [21] = 3.72, $p = .001$) significantly lower in the final week, compared to the first three weeks. For the monthly assessments, paired samples t-tests indicated that levels of consummatory pleasure and depressive symptoms did not differ between the pre-intervention assessment and the assessment before the intervention session in in any of the intervention groups ($p > .05$).

Descriptive statistics for each intervention group are presented in Table 1. As educational level may affect the extent to which different activities are enjoyed (e.g., intellectual activities), we checked whether educational levels were equally distributed over the three intervention groups, which was the case ($\chi^2$ (6) = 5.65, $p = .46$). Further, no significant differences were found in age ($F$ [2] = .58, $p$ = .56), baseline scores of pleasure ($F$ [2] = .12, $p$ = .88) or depressive symptoms ($F$ [2] = .27, $p$ = .76) between the three intervention groups. The majority of the participants experienced at least mild depressive symptoms as measured by the PHQ-9 (minimal symptoms: N=6; mild symptoms: N=28; moderate symptoms: N=19; moderately severe symptoms: N=11; severe symptoms: N=5), indicating that we selected a group that experienced significant depressive symptoms in addition to anhedonia. With regard to the prevalence of other psychiatric problems, 7.2% fall in to the clinical range for anxiety problems, 18.8% fall into the clinical range for avoidant personality, and 11.6% into the clinical range for ADHD (see appendix C).

**TABLE 1.** Descriptive Statistics for the Three Intervention Groups

| | No intervention (N = 22) | Lifestyle advice (N = 22) | Lifestyle advice and skydive (N = 25) |
|---|---|---|---|
| Females (%) | 81.8 | 77.3 | 80.0 |
| Age (M [SD]) | 21.4 (2.0) | 21.3 (2.0) | 21.9 (2.0) |
| Current education (%) | | | |
| Intermediate vocational education | 0.0 | 0.0 | 4.0 |
| Higher secondary education | 4.6 | 0.0 | 0.0 |
| Higher vocational education | 45.5 | 36.4 | 24.0 |
| University | 50.0 | 63.6 | 60.0 |
| None | 0.0 | 0.0 | 12.0 [a] |
| Ethnicity (%) | | | |
| Caucasian | 95.5 | 90.9 | 96.0 |
| Latin-American | 4.6 | 0.0 | 0.0 |
| Mixed African American/Caucasian | 0.0 | 4.6 | 4.0 |
| Asian | 0.0 | 4.6 | 0.0 |
| Consummatory pleasure pre-intervention | 58.84 (12.75) | 60.42 (9.62) | 59.72 (9.12) |
| Consummatory pleasure post-intervention | 56.49 (14.72) | 61.41 (9.70) | 58.08 (11.92) |
| Depressive symptoms pre-intervention | 7.91 (3.62) | 8.55 (4.23) | 7.76 (3.62) |
| Depressive symptoms post-intervention | 6.86 (4.06) | 5.73 (3.27) | 6.52 (3.99) |

*Note*. [a] The highest levels of education attained for the three participants who currently did not follow education were applied university, intermediate vocational education, and lower secondary education.

## Group Level Analyses Momentary Data

The results of the group-level analyses of pleasure, PA, and NA can be found in Table 2. In the first models that included only the differences in level between the baseline phase and intervention phase, a significant interaction was found between group and phase for pleasure and PA. The intervention group showed an increase in pleasure and PA after the intervention session, whereas the control group showed no differences in affect level between the two phases. This finding indicates that, on average, participants in the intervention groups experienced higher pleasure and PA in the intervention phase compared to the baseline phase, whereas those in the control group did not. The proportion of random effect variance accounted for by the interaction between phase and group was .09 for pleasure and .14 for PA, showing that respectively 9% and 14% of the variance in the level change estimate for pleasure and PA can be explained by group

differences. For NA, no significant differences between the groups were found, both the control group and the intervention groups significantly decreased in NA in the intervention phase. Effect sizes (i.e., Cohen's d) were 0.45 for pleasure, 0.49 for PA and 0.25 for NA.

In the full model we added the time trends for the baseline and intervention phases to check whether the level change was due to a trend that already started in the baseline phase. As can be seen in Table 2, there were some significant trends in the baseline and intervention phases, but these trends did not differ between groups, indicating that the differences in level estimates between groups were not due to group differences in trends in the baseline phase.

Next, we analyzed differences between the lifestyle advice group and the tandem skydive group (see Table 3). Significant differences were found between the groups, but in opposite direction than expected. The advice only group increased more in PA and decreased more in NA in the intervention phase than the skydive group. No significant differences were found in the full model. These findings indicate that the tandem skydive did not have an additional effect above the effects of the lifestyle advice. [1]

**TABLE 2.** Differences in Estimates of Level and Slope Changes in Pleasure, PA, and NA, Between Control and Intervention Groups

| | **Only level estimate** | | **Complete model** | |
|---|---|---|---|---|
| | **No Intervention (N = 22)** | **Intervention (N = 46)** | **No Intervention (N = 22)** | **Intervention (N = 46)** |
| Pleasure intercept | 53.65 (2.29) | 53.57 (1.26) | 56.43 (2.19) | 54.12 (1.18) |
| Pleasure level estimate | -1.82 (1.34)* | 2.58 (1.09)* | 1.27 (2.18) | 4.83 (1.45) |
| Pleasure slope baseline phase | | | -0.07 (0.03) | -0.01 (0.02) |
| Pleasure slope change | | | 0.06 (0.03) | -0.02 (0.02) |
| PA intercept | 53.65 (2.21) | 53.10 (1.25) | 53.10 (2.12) | 50.45 (1.22) |
| PA level estimate | -0.03 (1.01)*** | 4.75 (0.92)*** | -0.42 (1.90) | 2.94 (1.00) |
| PA slope baseline phase | | | 0.02 (0.03) | 0.06 (0.02) |
| PA slope change | | | -0.02 (0.03) | -0.08 (0.02) |
| NA intercept | 23.15 (1.76) | 24.53 (1.30) | 25.68 (1.99) | 28.27 (1.53) |
| NA level estimate | -2.47 (1.18) | -4.68 (0.77) | -0.20 (1.45) | -1.28 (0.69) |
| NA slope baseline phase | | | -0.06 (0.03) | -0.09 (0.02) |
| NA slope change | | | 0.07 (0.03) | 0.10 (0.03) |

*Note.* PA = Positive Affect, NA = Negative Affect.
* $p < .05$, ** $p < .01$, *** $p < .001$. Please note that the reported significance levels indicate whether the difference between groups (i.e., the interaction) was significant. To calculate the slope in the intervention phase, the coefficients for the slope in the baseline phase and the slope change should be summed.

**TABLE 3.** Differences in Estimates of Level and Slope Changes in PA, Pleasure, and NA, Between Both Intervention Groups

| | Only level estimate | | Complete model | |
|---|---|---|---|---|
| | **Lifestyle advice (N = 22)** | **Lifestyle advice and skydive (N = 24)** | **Lifestyle advice (N = 22)** | **Lifestyle advice and skydive (N = 24)** |
| Pleasure intercept | 53.81 (1.49) | 54.18 (1.89) | 53.95 (1.66) | 52.72 (1.81) |
| Pleasure level estimate | 3.90 (1.71) | 0.12 (1.33) | 4.05 (2.19) | 0.76 (1.63) |
| Pleasure slope baseline phase | | | -0.01 (0.03) | 0.03 (0.02) |
| Pleasure slope change | | | 0.01 (0.03) | -0.08 (0.04) |
| PA intercept | 52.84 (1.92) | 53.93 (1.62) | 49.79 (1.62) | 50.70 (1.82) |
| PA level estimate | 6.35 (1.17)* | 2.61 (1.32)* | 2.42 (1.30) | 1.46 (1.38) |
| PA slope baseline phase | | | 0.07 (0.02) | 0.07 (0.02) |
| PA slope change | | | -0.05 (0.02) | -0.11 (.04) |
| NA intercept | 25.73 (1.96) | 22.85 (1.67) | 29.19 (2.18) | 27.56 (2.09) |
| NA level estimate | -5.90 (1.11)* | -2.90 (0.93)* | -1.35 (0.93) | 0.19 (0.76) |
| NA slope baseline phase | | | -0.08 (0.03) | -0.10 (0.02) |
| NA slope change | | | 0.06 (0.03) | 0.14 (0.04) |

*Note*. PA = Positive Affect, NA = Negative Affect.
* $p < .05$, ** $p < .01$, *** $p < .001$. Please note that the reported significance levels indicate whether the difference between groups (i.e., the interaction) was significant. To calculate the slope in the intervention phase, the coefficients for the slope in the baseline phase and the slope change should be summed.

## Individual Level Analyses Momentary Data

Significant residual random effect variance in the level change estimates was present and hence, individual models were examined. As can be seen in the forest plots (Figures 3, 4, and 5), the majority of participants showed higher mean levels of pleasure and PA and lower mean levels of NA in the intervention phase, compared to the baseline phase, although substantial heterogeneity in the effect sizes were observed and also negative effects were observed, especially in the tandem skydive group. Further, very few differences in level estimates were apparent between the lifestyle advice group and the tandem skydive group.

Next, we examined whether sex, severity of anhedonia, duration of anhedonia, and advice adherence were associated with the extent to which individuals improved. Sex was significantly associated with the level change estimates in PA ($\beta$ = -.32, $p < .05$), in that females showed a lower increase in the level of PA during the intervention phase than males. No associations were found between sex and the level change estimates for pleasure ($\beta$ = -.25, $p$ = .10) or NA ($\beta$ = .14, $p$ = .35). Severity of anhedonia at baseline ($\beta$ = .03, $p$ = .84 for PA; $\beta$ = .07, $p$ = .67 for pleasure; $\beta$ =

-.08, $p$ = .60 for NA), duration of anhedonia at baseline ($\beta$ = -.08, $p$ = .59 for PA; $\beta$ = -.17, $p$ = .27 for pleasure; $\beta$ = .05, $p$ = .74 for NA), and depressive symptoms at baseline ($\beta$ = -.10, $p$ = .35 for PA; $\beta$ = -.16, $p$ = .10 for pleasure; $\beta$ = -.06, $p$ = .41 for NA) were not associated with the level change estimates. Advice adherence was also not related to the level change estimates ($F$ [2] = .72, $p$ = .49 for PA; $F$ [2] = .01, $p$ = .99 for pleasure; $F$ [2] = .63, $p$ = .54 for NA).

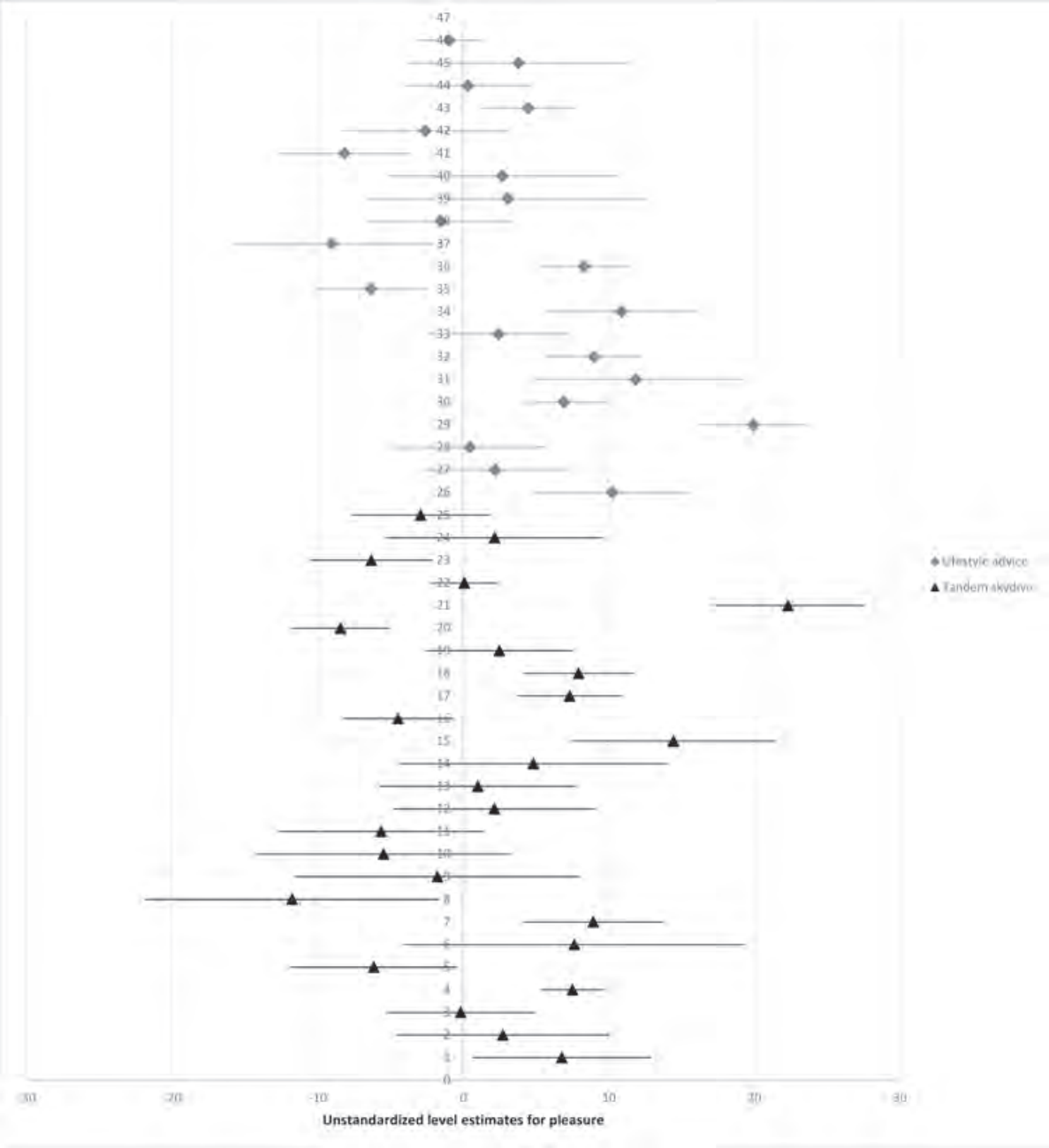



**FIGURE 3.** Forest plot of unstandardized level estimates for pleasure. Numbers on the y-axis represent the different individuals (N = 46).

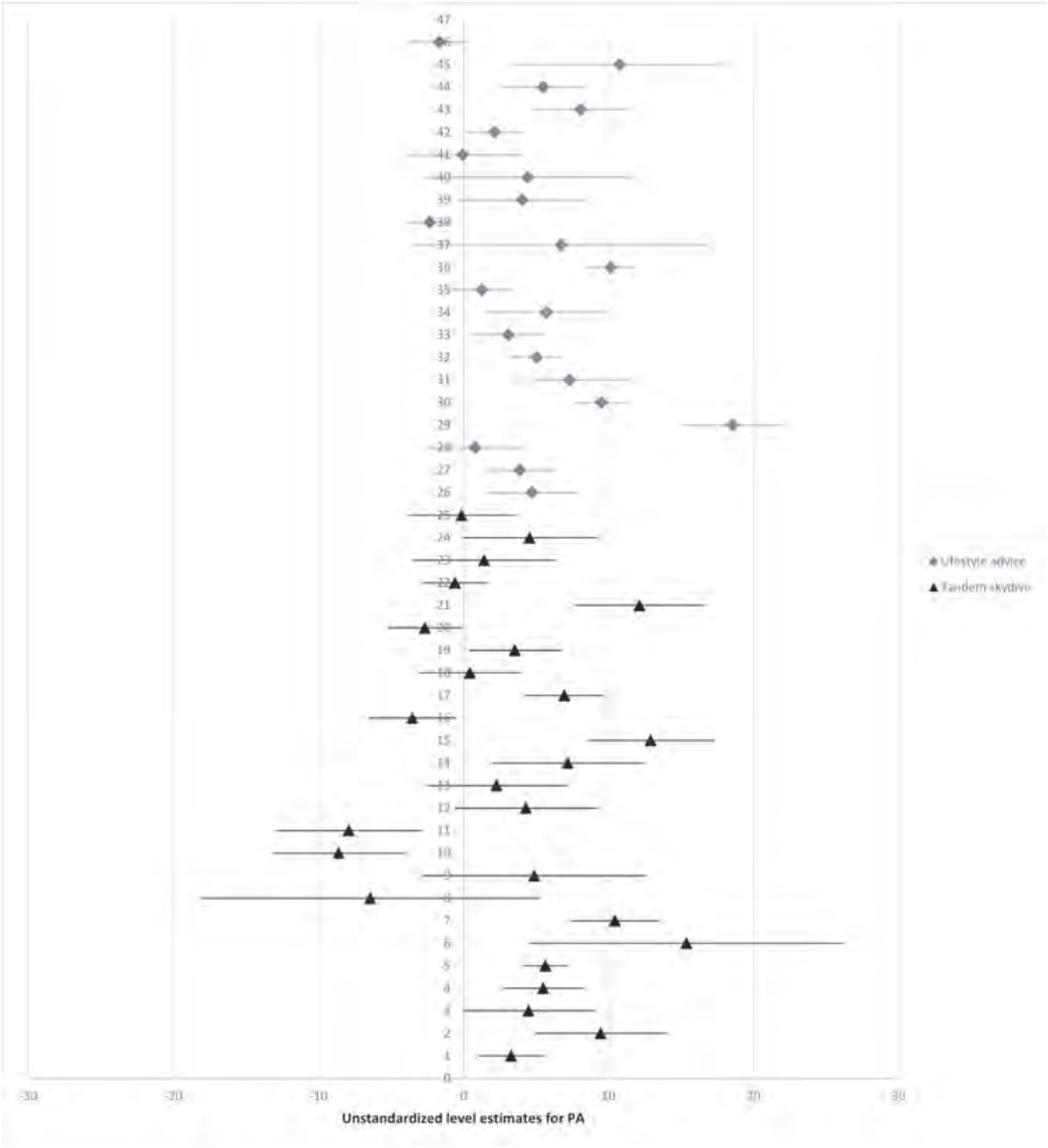



**FIGURE 4.** Forest plot of unstandardized level estimates for PA. Numbers on the y-axis represent the different individuals (N = 46).

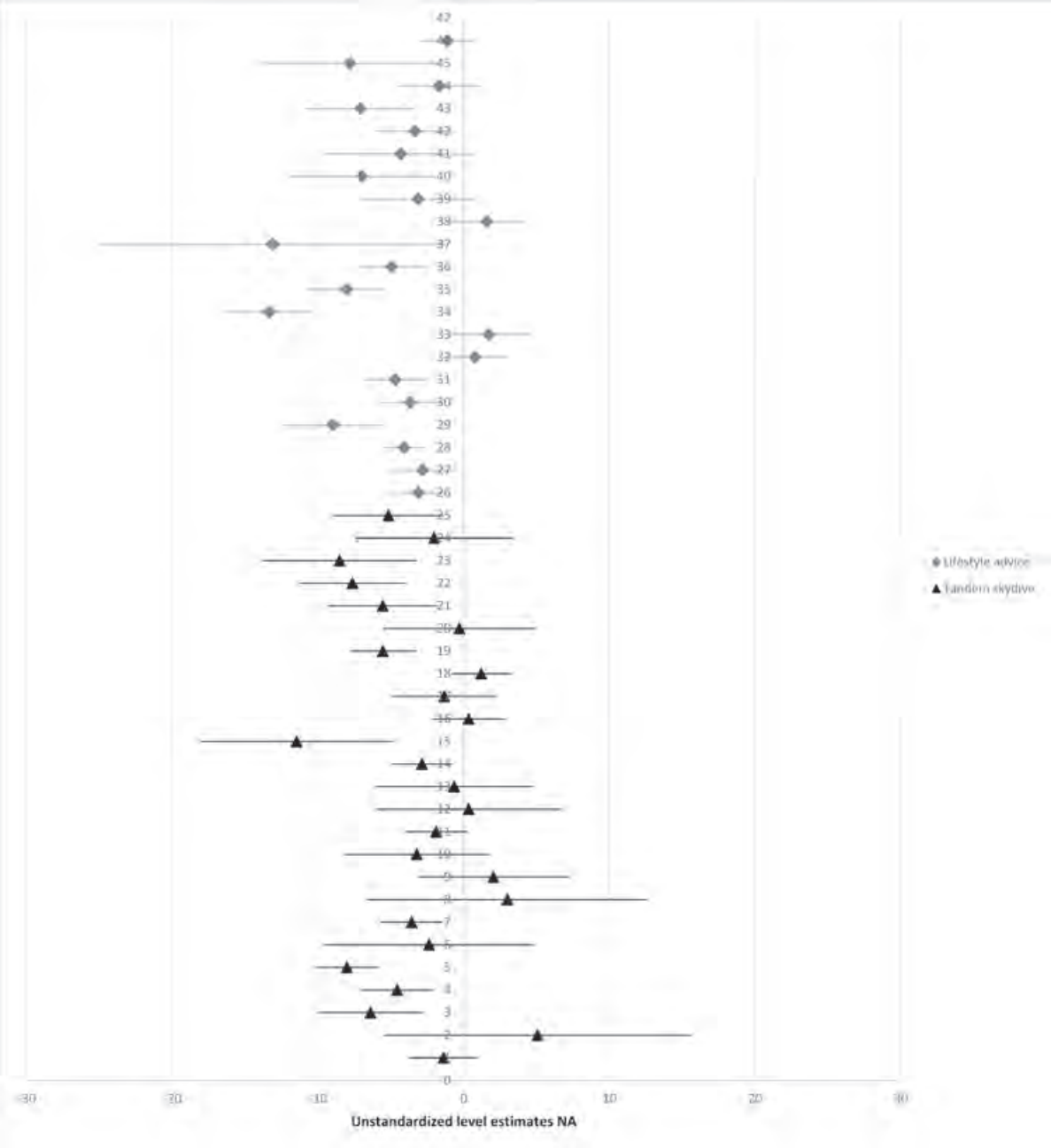



**FIGURE 5.** Forest plot of unstandardized level estimates for NA. Numbers on the y-axis represent the different individuals (N = 46).

## Group Level Analyses Monthly Data

Levels of consummatory pleasure and depressive symptoms for each intervention group at pre- and post-intervention are depicted in Figure 6. A marginally significant group (control versus intervention) by time (pre- and post-intervention) interaction was found for consummatory pleasure ($F$ [2] = 3.38, $p$ = .07), indicating that the intervention group increased more in pleasure

than the control group. No significant interaction was found for depressive symptoms ($F$ [2] = 0.48, $p$ = .49). Effect sizes (Cohen's d) were 0.48 for consummatory pleasure and 0.19 for depressive symptoms. Further, for the advice group versus the tandem skydive group, no significant interaction was found for pleasure ($F$ [2] = 0.32, $p$ = .57; see Figure 6A). For depressive symptoms, a marginally significant group by time interaction was found ($F$ [2] = 3.85, $p$ = .06), in that the lifestyle advice group decreased more in depressive symptoms from pre- to post-intervention than the skydive group (see Figure 6B).

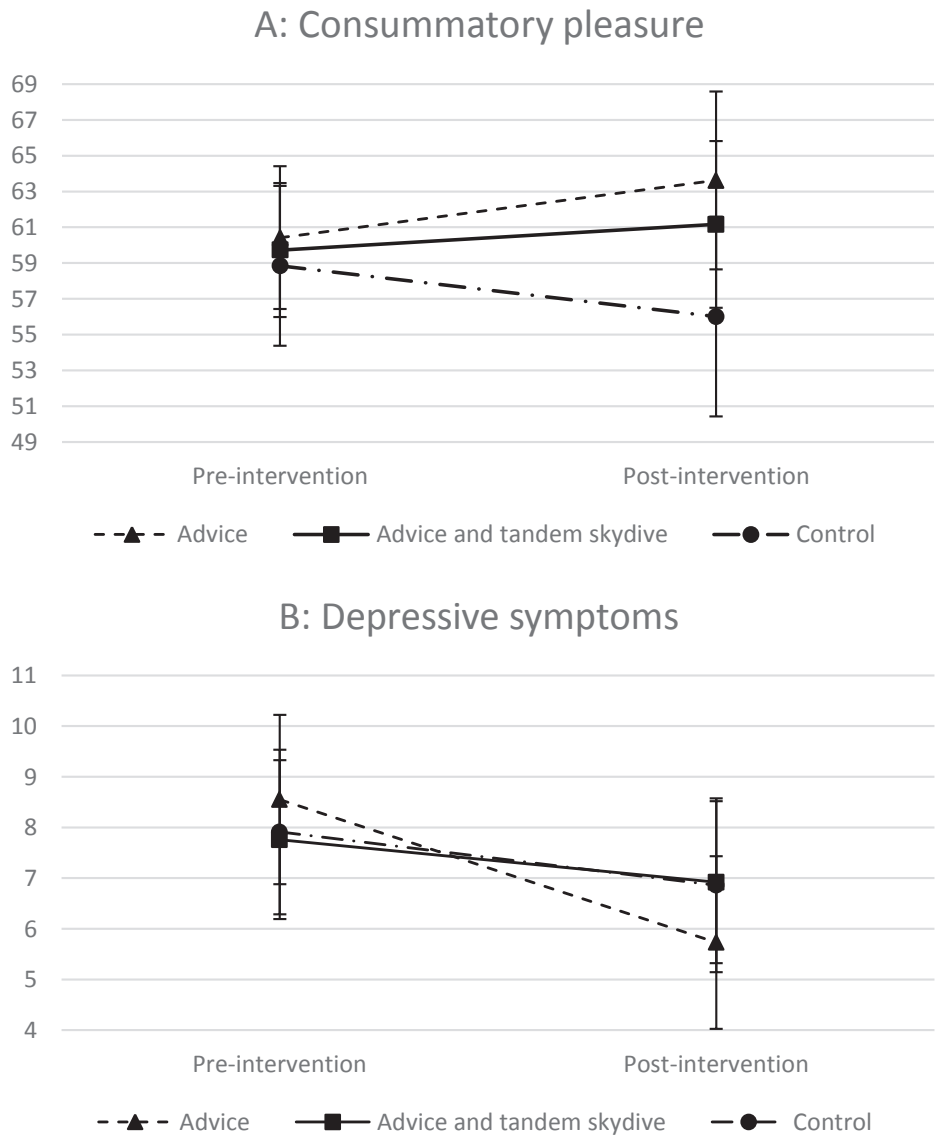



**FIGURE 6.** Comparison between the three intervention groups in monthly assessments of consummatory pleasure as measured by the DOPS (6A) and depressive symptoms as measured by the PHQ-9 (6B).

## DISCUSSION

The main aim of this exploratory study was to examine whether personalized lifestyle advice and tandem skydiving could reduce anhedonia in young adults. Although the small sample size warrants prudence, we showed that the lifestyle advice can increase pleasure and PA in young adults with persistent anhedonia. Despite previous literature suggesting that a free-fall may reboot the reward system and could therewith provide additional benefits, we found no evidence of such an additional effect of the tandem skydive over and above the lifestyle advice.

Our findings indicate that providing anhedonic young adults with lifestyle advice can increase pleasure and PA. However, the effects were not found for all outcome variables and were small to moderate in size, which may be due to two reasons. First, the implementation of the lifestyle advice may take some time to really sink in and become part of an individual's repertoire of behaviors. The only other study providing personalized feedback on PA in depressed patients showed that the differences between groups became significant after 8 weeks (i.e., 2 weeks after the last feedback session) (Kramer et al., 2014) and no significant differences between groups were present during the intervention period (Hartmann et al., 2015). In the present study, we were not able to examine these longer-term effects as our groups were no longer randomized after the intervention phase due to the free-choice intervention. This is a limitation, as we do not know how long the differences between groups in momentary pleasure and PA lasted. Second, participants received their advice in the beginning of the intervention phase and did not receive any reminders or triggers. Using these extra reminders could have increased advice adherence. Third, the forest plots show that there is great heterogeneity in the intervention effects, which dilutes the average group-level effects. This heterogeneity may be explained by individual characteristics, but could also be due to differences in the type of advice provided (e.g., social activity versus physical activity). These findings highlight the need to further explore whether individual characteristics and the type of advice could explain the effectiveness of the lifestyle advice.

In the present study, we provided explicit advice on which behaviors participants could change and expected that this would be more effective than only showing participants in which situations they experienced PA, as was done by Kramer and colleagues (Hartmann et al., 2015; Kramer et al., 2014). The results from our study indicate that concrete advice may indeed be more effective, as we found immediate increases in momentary pleasure and PA, whereas no effects were found on PA in the study in which only feedback was provided (Hartmann et al., 2015). A further explanation for the stronger effects in our study is that our advice was, if possible, based on statistically significant associations with pleasure, whereas the Kramer et al. study only presented descriptive feedback. As our report was based on more robust findings, it may have had a greater impact.

In our analyses, we excluded the measures between the intervention allocation and the intervention session, because participants already knew which intervention they would receive during this period. For momentary levels of PA and NA, we showed that participants in the lifestyle advice only group slightly improved after being informed about the intervention allocation. This indicates that the knowledge that participants would receive help for their problems in the nearby future already improved their PA and NA levels. Interestingly, this was not the case for the tandem skydive group. A possible explanation is that the prospect of a tandem skydive evoked stress and anxiety in some individuals, which counterbalanced any positive effects.

Our findings are conclusive in showing that thrill experiences like a tandem skydive should not be suggested as a possible treatment for anhedonia, as the skydive did not provide the hypothesized motivational boost to kick start the implementation of the lifestyle advice. Possibly, a tandem skydive is only effective in increasing immediate pleasure and motivation, as was found in previous studies (Meyer et al., 2015; Wang & Tsien, 2011), and does not have a persistent effect. Our lack of effects may indicate that a skydive is not effective in decreasing anhedonia. Although we hypothesized that participating in a single skydive could provide a motivational boost to individuals, the effects may have been short-lasting and we cannot exclude the possibility that multiple skydiving experiences are needed to achieve long-lasting effects.

However, previous cross-sectional research has revealed an association between regular skydiving experience and high levels of anhedonia (Franken, Zijlstra, & Muris, 2006), which could indicate that repeated exposure to skydiving leads to increased levels of anhedonia and hence should be avoided. It may also be that individuals who experience anhedonia tend to seek out repeated high-risk activities in order to obtain the 'natural high' associated with these activities. The cross-sectional nature of this study prevents statements about the direction of the effects. Further longitudinal research is needed to explore the direction of effects between skydiving and anhedonia.

It may also be possible that a skydive is effective only for some individuals, as the individual-level analyses suggest. Anecdotal evidence from participants suggests that the skydive may have increased self-esteem and reduced anxiety in some individuals. For example, one of the participants reported that "Whenever I had to do something frightening, I was able to put this into perspective. Because I have done something so extremely scary, everything else seems less frightening". Hence, although a skydive may not be effective in rebooting the reward system and increasing pleasure in general, it might be beneficial for individuals with low self-esteem by boosting their self-confidence. The effectiveness may also depend on how individuals appraised the skydive afterwards, as it did seem to have an effect on the participant described above, who overcame her own fears. Further research should include questionnaires measuring the appraisal of the event and could even use qualitative interview data to explore how the skydive may affect individuals.

The findings from the present study may have clinical implications. Although our effect sizes were small to moderate, it is difficult to interpret these effects in terms of clinical relevance, given that there are no clinical cut-offs with regard to momentary levels of pleasure and PA. Hence, more research is needed in clinical samples in order to explore the clinical relevance of these findings. Pending further research in clinical samples, our intervention could be easily implemented in general health care. For example, providing patients who are on a waiting list to receive treatment with feedback on daily life patterns of mood and behavior during the waiting time could already decrease their symptoms and set the stage for further improvement during the 'actual' treatment. As recent research shows that automated feedback based on momentary assessments is already possible (Emerencia et al., 2015; van der Krieke et al., 2015), such an intervention will not require a lot of time and effort from clinicians. The information derived from the momentary assessments and personalized feedback could provide a valuable starting point for treatment, as it enables tailor-made treatment for each individual. Hence, the most important advantages of personalized feedback are that it is cost-effective, easily accessible, and can be used in addition to existing treatments. It may also be beneficial for patients that do not respond to standard treatments, as the personalized feedback may reveal new potential targets for behavioral change. In addition, automated feedback based on momentary assessments can be used for different kinds of mental health problems, as the variables of interest for the feedback can be adapted. It is important to note that this automated feedback should include a directive component, that is, concrete suggestions on which factors individuals can change, as we showed that this is more effective than only offering feedback without concrete suggestions (Kramer et al., 2014). Our intervention may not be applicable to all individuals with anhedonia. For example, excessive focusing on increasing PA and pleasure might induce mania in individuals for bipolar disorder. Further research is needed in various clinical populations to carefully examine the effects of personalized lifestyle advice.

Finally, we showed that the extent to which individuals improved was highly variable. Surprisingly, variation in improvement was not related to the severity and duration of anhedonia, nor to advice adherence. Although we expected that the extent to which participants followed up on our advice would lead to stronger increases in PA and pleasure, possibly, our broad measure of advice adherence could explain the large variations found. That is, some participants may not have been very adherent in general, but still have followed up on the most effective advice. Further, we only relied on participants' own report of their advice adherence; which may have been socially desirable. Further examination of the actual change in behavior is needed to elucidate these effects (Snippe et al., 2016).

## Strengths and Limitations

The present study had several strengths. It is, to our knowledge, the very first to use momentary data to provide personalized lifestyle advice and evaluate their effects. Because of our unique design consisting of a baseline phase and an intervention phase and the high number of momentary assessments in each phase (i.e., more than 90 per phase), we were able to examine within-person effects, at the group-level and at the individual level. Compliance was excellent, as the average number of completed assessments was 92.4%. Potential explanations for these high compliance rates in comparison to other momentary assessment studies with a long duration (Trull et al., 2008) could be that (1) participants knew that they would receive feedback based on their momentary assessments, which may have enhanced motivation, and (2) participants were instructed that they would only receive the monetary compensation if they filled out at least 80% of the momentary assessments. Our high compliance rate shows that it is feasible to conduct a long-term diary study among anhedonic young adults.

Despite these strengths, several limitations need to be acknowledged. First, although we increased our power by including on average 190 momentary assessments per individual, the number of participants per intervention group was relatively low. This might explain why several of the differences tested did not reach statistical significance, particularly for the monthly collected data. Second, we only selected participants who were willing to conduct a tandem skydive, which may have led to a biased sample. However, this bias, if any, is probably limited, because only 25 (9.3%) of the eligible individuals were excluded because they were not willing to participate in a tandem skydive. Further, we showed that this group did not differ from the group who was willing to perform a skydive in severity of anhedonia, the level of consummatory anhedonia, and depressive symptoms. Third, the majority of our sample was female. Please note that this was due to a high number of females in the survey study, and not due to higher prevalence of anhedonia in females. The small number of males implied that we had little power to test sex differences at the group level. Our findings at the individual level suggest that males benefit more from lifestyle advice than females, but further research in larger samples is needed to replicate these findings. Fourth, one-third of our sample only experienced minimal to mild depressive symptoms and therefore, the generalizability of our findings to individuals with depressive symptoms is limited. However, we showed that the level of depressive symptoms did not affect improvement, which indicates that the lifestyle advice was effective for individuals with different levels of depressive symptoms. Fifth, we cannot exclude the possibility that differences between the control group and the intervention groups are due to the placebo effect of receiving an active intervention. Future research should include a more active control group to examine this alternative explanation. Finally, as mentioned earlier, we were not able to examine long-term effects of the interventions, because we included the free-choice intervention immediately after the intervention period. Further research is needed to explore whether lifestyle advice has a long-term effect on pleasure and PA in anhedonic young adults.

## Conclusions

This study indicates that providing anhedonic young adults with personalized lifestyle advice may increase pleasure and PA. In addition, we showed that it is feasible to collect diary data in anhedonic young adults for a fairly long time without much loss of information. These findings are relevant for clinical practice, as they suggest that it is both promising and feasible to develop momentary interventions that can be implemented as additions to treatment as usual or while patients are on waiting lists for treatment.

# ADDENDUM

In the published version of the paper described in this chapter, the effects of the two interventions (lifestyle advice and tandem skydive) on the main outcome measures PA, NA, pleasure, and depression were investigated. This leaves the question unanswered whether the two interventions also modify positive bias (i.e., spread through the entire model as presented in Fig. 1 of Chapter 1). This is important to establish since a modification of a low positive bias may render an individual less vulnerable to developing depressive symptoms again in the future, and implies that the type of tailored lifestyle advice provided in the No Fun No Glory project may be used to prevent onset of depression. I carried out additional analyses, not presented in the published paper, to test whether lifestyle advice and lifestyle advice combined with the tandem skydive were associated with increases in positive bias.

Positive bias was operationalized in the following two ways: (1) bias toward happy facial emotions during a facial emotion identification morph task (happy bias = average reaction time in identifying sad, angry, and fearful emotions / average reaction time in identifying happy facial emotions; see Chapter 5); (2) Reward responsiveness as measured by the sumscore on the Reward Responsiveness scale (Van den Berg et al., 2010). The first bias is an implicit, relative (compared to negative emotions) bias based on a laboratory task and the second an explicit, absolute bias based on self-report. Repeated measures ANOVAs were conducted on pre- and post-intervention assessments of happy bias and reward responsiveness. Similar to the analyses reported in the published paper, this was done in two steps: first I checked whether the two groups who received lifestyle advice differed from the control group; second, I compared the tandem skydive group with the lifestyle advice only group.

Levels of happy bias and reward responsiveness for each intervention group at pre- and post-intervention are depicted in Figures 7 (happy bias) and 8 (reward responsiveness). No significant group by time interactions were found for happy bias ($F$ [2] = 0.21, $p$ = .65) or reward responsiveness ($F$ [2] = 0.95, $p$ = .33). For the advice group versus the tandem skydive group also no significant interactions were found for happy bias ($F$ [2] = 1.71, $p$ = .20) or reward responsiveness ($F$ [2] = 0.50, $p$ = .49).

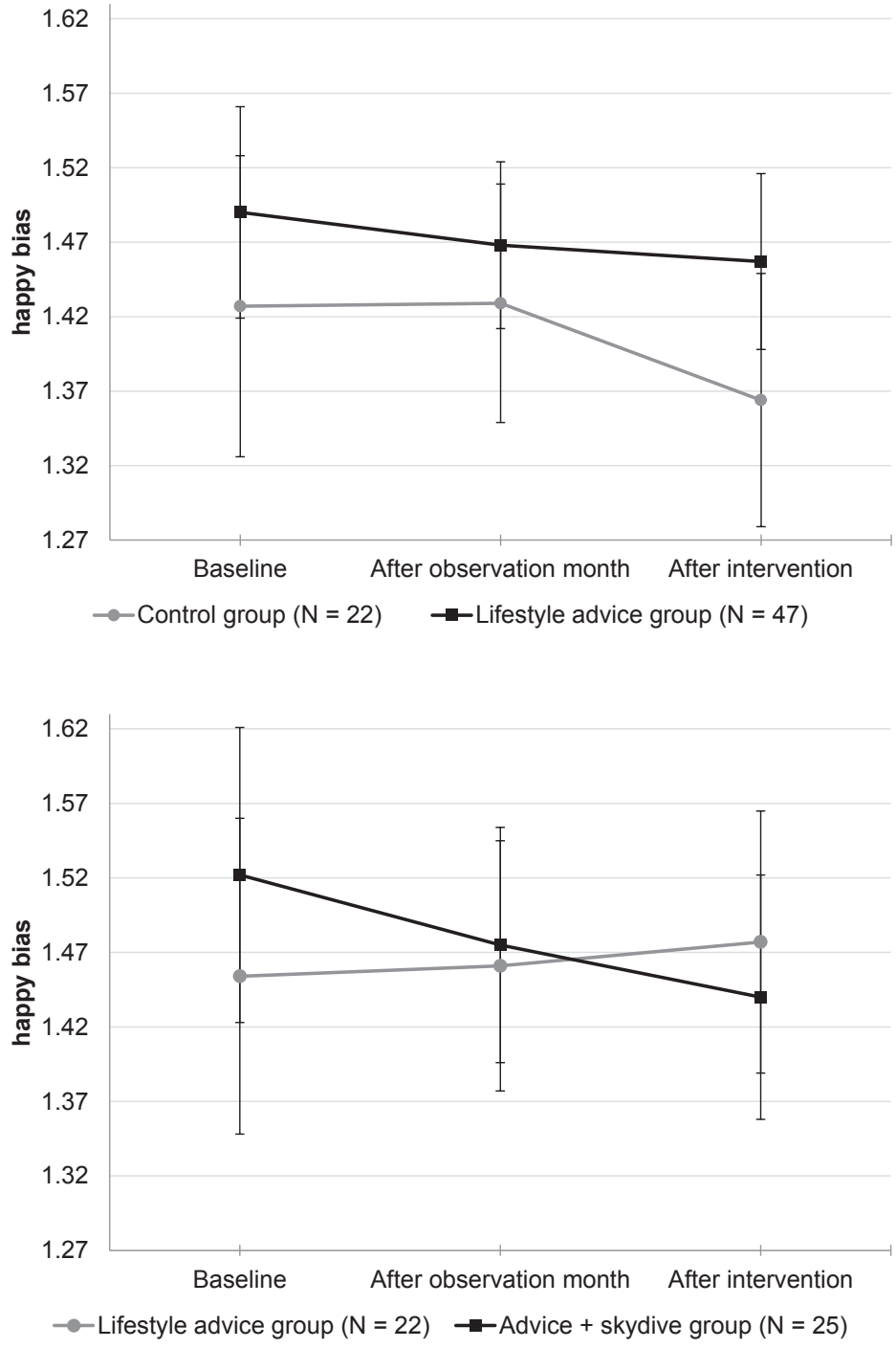



**FIGURE 7.** Comparison between the control group and the group that received lifestyle advice (top panel) and between the group that received lifestyle advice and the group that received lifestyle advice + tandem skydive (bottom panel) on happy bias.

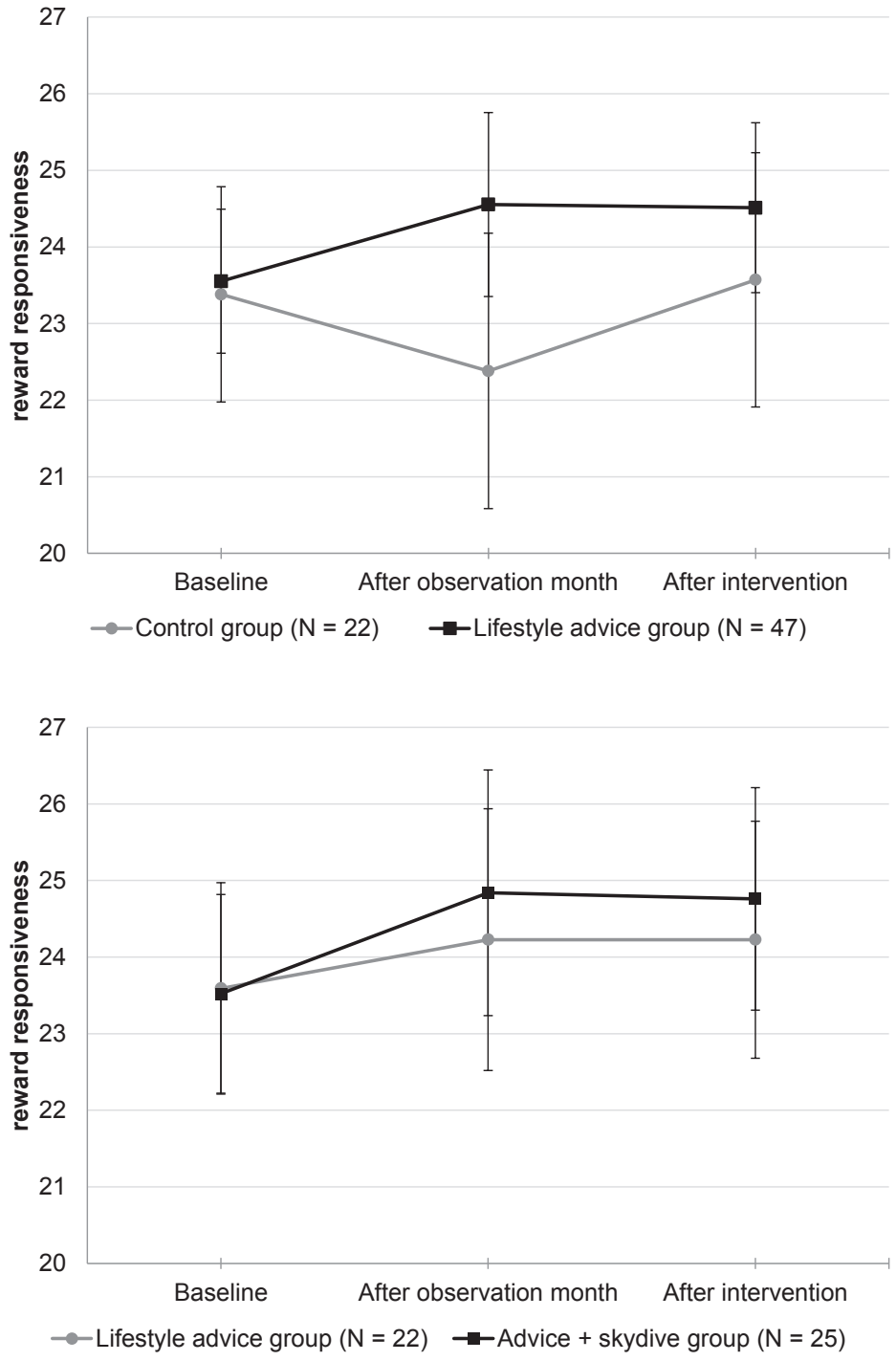



**FIGURE 8.** Comparison between the control group and the group that received lifestyle advice (top panel) and between the group that received lifestyle advice and the group that received lifestyle advice + tandem skydive (bottom panel) on reward responsiveness.

I found no evidence that tailored lifestyle advice alone or in combination with the skydive resulted in a modification of positive bias in anhedonic young adults.

# SUPPLEMENTARY MATERIAL

## Appendix A: Description of Measurement of Model Variables

| Variables | Item | Answer categories |
|---|---|---|
| Pleasure | I have experienced pleasure since the last assessment | VAS 0-100 (not at all – very much) |
| Social context | Since the last assessment, I was (multiple answers possible): | alone; with partner; with family; with friends; with classmates; with acquaintances; with strangers. |
| Activities | I have participated in the following activities since the last assessment (multiple answers possible): | watching TV; listening to music; studying; reading a book or magazine; shopping; using the internet; gaming; work; having sex; household chores; sports; going to a bar or club; going out for dinner; going to the movies, theatre, museum or concert; sleeping; using social media; hobbies; other, namely…. |
| Substance use | I have used the following substances since the last assessment: | nothing; caffeine; nicotine; medication; alcohol; cannabis; stimulants; sedatives; other drugs |
| Stress | How busy have I been since the last assessment | too busy; pleasantly busy; neutral; pleasantly quiet; too quiet |
| Time alone | [When alone has been checked:] Time spend alone: | VAS scale (very little - very much) |
| Social interaction | How much have I been talking to other people? | VAS 0-100 (not at all – very much) |
| Physical activity | I have been physically active since the last assessment | VAS 0-100 (not at all – very much) |
| Time outside | I have been outside since the last assessment | VAS 0-100 (not at all – very much) |
| Worrying | I have been worrying | VAS 0-100 (not at all – very much) |
| Sweet and savory snacks | How many sweet snacks (e.g., cookies, sweets) did I eat?<br>How many savory snacks (e.g., French fries, potato chips) did I eat? | VAS 0-100 (nothing – a lot) |
| Physical discomfort | I have experienced physical discomfort since the last assessments (head ache, diarrhea, heavy legs, etc.) | VAS 0-100 (not at all – very much) |

## Appendix B: Details of the Lifestyle Report

The descriptive information of the lifestyle report included pie charts representing the number of assessments spent in different social contexts and the number of activities (see Figure B1 for an example); a line graph delineating the level of pleasure during the 90 assessments (e.g., Figure B2); two charts showing the level of pleasure experienced in different social contexts and during different activities (e.g., Figure B3); and scatter plots depicting the association between continuous lifestyle factors that were not included in the personal networks and pleasure. For the comparison with norm groups, a bar chart was presented with two bars for each of the variables reported in Table A1 (e.g., Figure B4). The norm data were based on the data from the screening survey among 2937 young adults, which were used to create mean norm scores for males and females separately.

The personal networks were based on two different approaches, because combining categorical and continuous predictors in a single time-series analysis is problematic and not all time-series models can handle large numbers of predictors. For the association between categorical predictors and pleasure (see Table A1), we used Auto Regressive Moving Average (ARMA) models (Brandt & Williams, 2007). By fitting ARMA models to the residuals, we adjusted for autocorrelation in the time series. In all models, we first tested whether we needed to correct for time trends (i.e., linear and quadratic), day of the week, special events (positive or negative), or menstruation for females. The Ljung-Box test was used to explore whether the residuals represented 'white noise'. If this was not the case, we checked whether residual autocorrelation was present by using (partial) autocorrelation functions (ACFs and PACFs) and added ARMA parameters to the regression model until the Ljung-Box test became non-significant. Series that showed heteroscedasticity were stabilized by using the natural log of the scores. The final model selection was based on the Bayesian Information Criterion (BIC) and the Ljung-Box test. When the final model was determined, all categorical predictors were tested in separate models, and significant associations were presented in text in the lifestyle report. Models were implemented using the SPSS22 Forecasting module. The significance level was set to 0.05.

To examine potential reciprocal and time-lagged associations between pleasure and the continuous variables (see Table A1), we used Vector Autoregressive modeling (VAR) (Brandt & Williams, 2007). This procedure was automated by using Autovar, an open source R package that automatically fits and evaluates VAR models (Emerencia et al., 2015). We allowed for a maximum of six variables in each network. Variables were selected based on sufficient variability (i.e., Mean Squared Successive Difference >50; MSSD), and reasonably low skewness (i.e., zskewness < 4). If more than six variables met the criteria, the best distributed variables were chosen (i.e., the least skewed variables of variables with the highest MSSD). If less than six variables met the criteria, only those variables were included in the model. Pleasure was always included, providing that variability and skewness met the criteria mentioned above. We estimated VAR models with one time lag. This yields estimates of the cross-lagged as well as the simultaneous associations

between variables. Trend variables (i.e., linear and quadratic time trend) were included in case of non-stationarity. Dummy variables for time of day (i.e., afternoon, evening) were always included. Models with dummy variables for day of the week and with log-transformed variables were tested. From the models that met the assumptions of stability, homoscedasticity and normality, Autovar chose the best fitting model based on the BIC (i.e., Bayesian Information Criterion). Network plots for simultaneous (i.e., contemporaneous associations) and dynamic (i.e., lag 1 associations) models were created using Data-Driven Documents-3 JavaScript library (http://d3js.org) (see Figures B5 and B6 for examples of simultaneous and dynamic network plots). All analyses were checked by another team member to avoid mistakes.

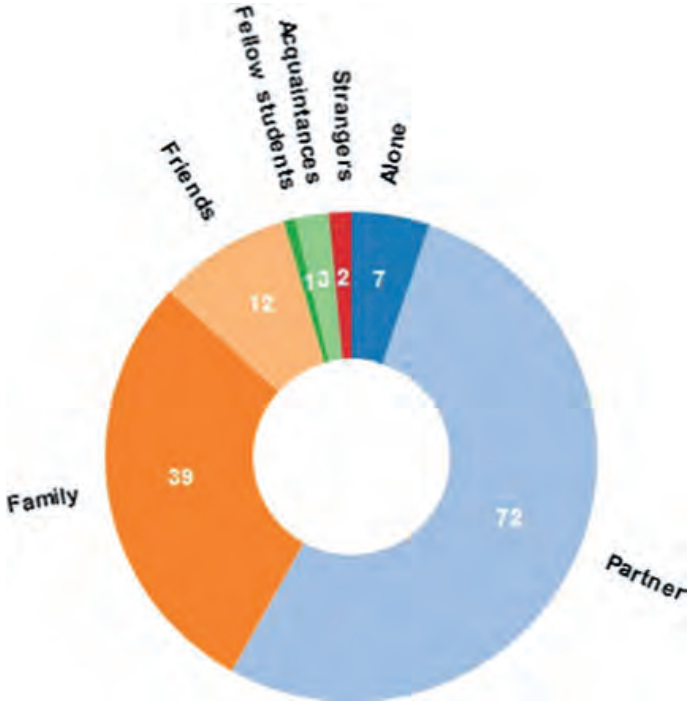



**FIGURE B1.** Pie chart for number of assessments in different types of company.

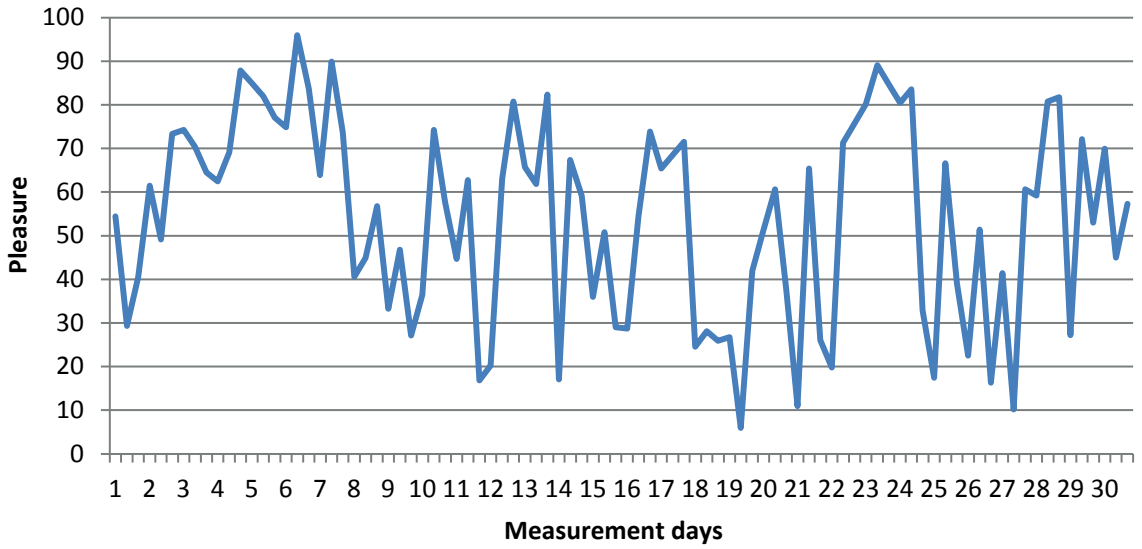



**FIGURE B2.** Line graph depicting level of pleasure during baseline phase

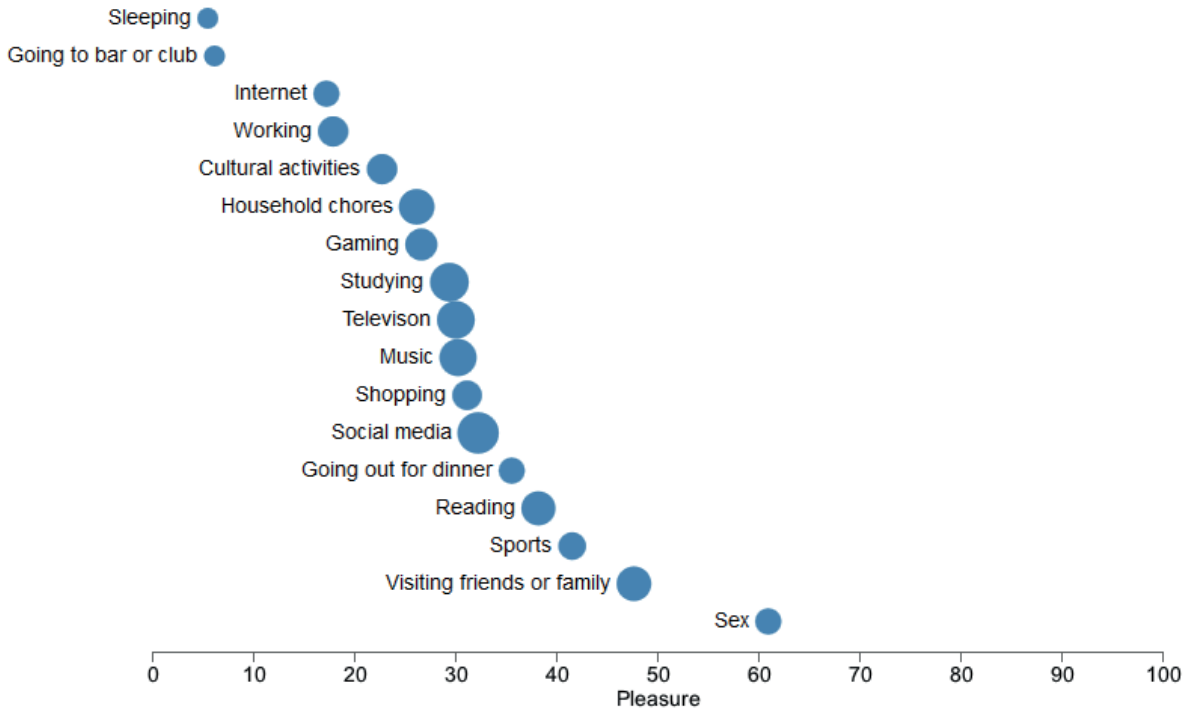



**FIGURE B3.** Activities ranked by level of pleasure

*Note.* Larger circles indicate higher frequency of the activity, smaller circles lower frequency.

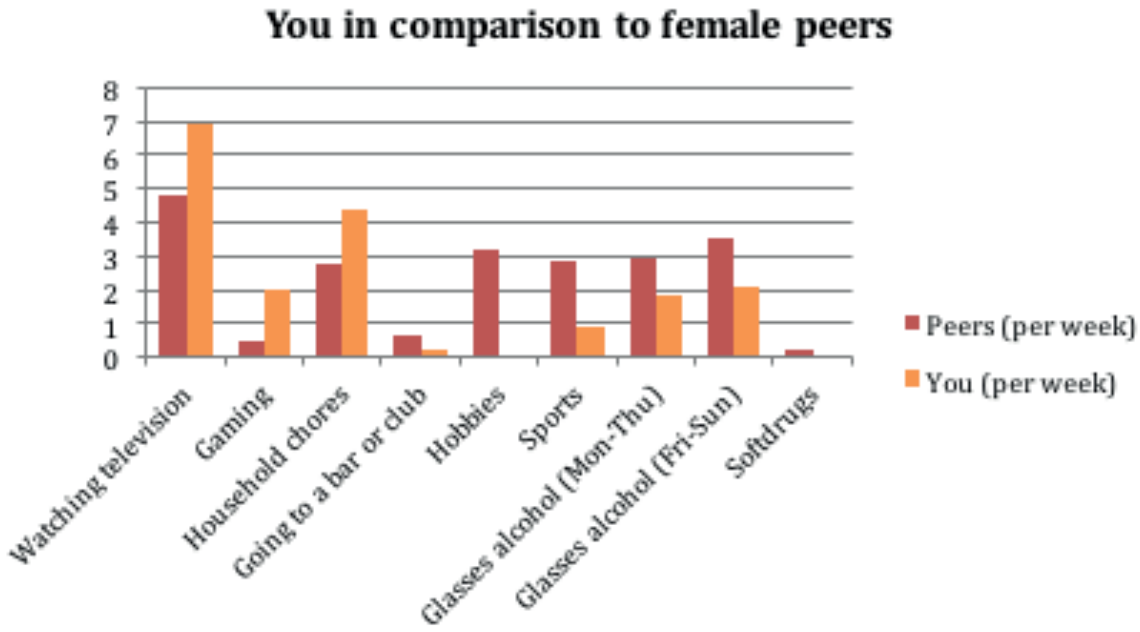



**FIGURE B4.** Frequency of participants' activities compared to norm group

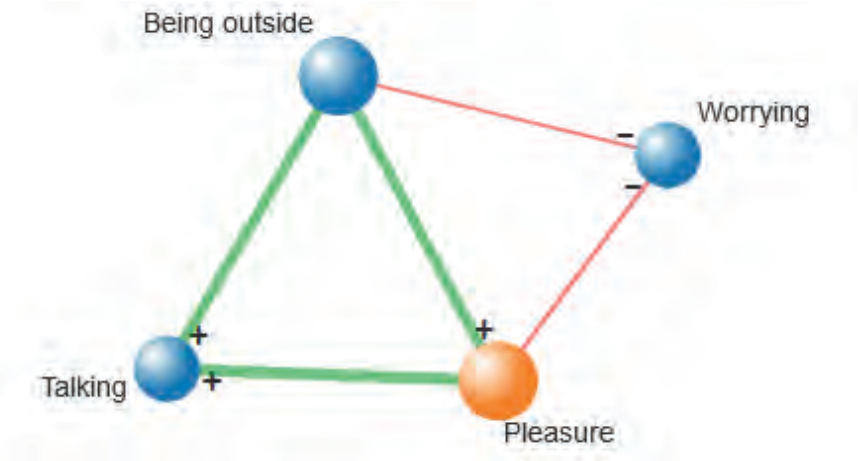



**FIGURE B5.** Personal network for simultaneous associations

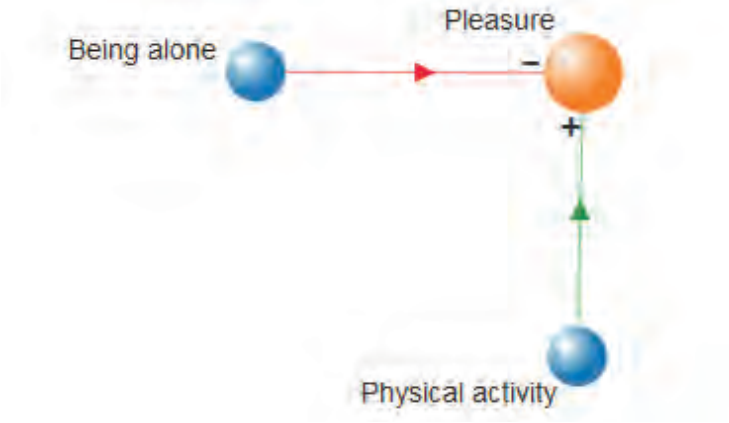



**FIGURE B6.** Personal network for dynamic associations (i.e., lag 1)

## Appendix C: Additional Descriptive Statistics for the Three Intervention Groups

Using the Adult Self-Report questionnaire (ASR; Achenbach & Rescorla, 2001) that was assessed in the screening survey, we calculated cut-off scores for the DSM-based problem scales Depressive problems, Anxiety problems, Avoidant Personality problems, ADHD problems, antisocial personality problems, and total problems. Based on these cut-offs, participant scores were classified as non-clinical, sub-clinical or clinical. Results can be found in Table C1, for each intervention group separately.

**TABLE C1.** Descriptive Statistics for the Three Intervention Groups.

| | No intervention (N = 22) | Lifestyle advice (N = 22) | Lifestyle advice and tandem skydive (N = 25) |
|---|---|---|---|
| Depressive problems | | | |
| Non-clinical | 11 (50%) | 10 (45.5%) | 12 (48%) |
| Sub-clinical | 6 (27.3%) | 6 (27.3%) | 8 (32.0%) |
| Clinical | 5 (22.7%) | 6 (27.3%) | 5 (20.0%) |
| Anxiety problems | | | |
| Non-clinical | 18 (81.8%) | 17 (77.3%) | 21 (84.0%) |
| Sub-clinical | 1 (4.5%) | 3 (13.6%) | 4 (16.0% |
| Clinical | 3 (13.6%) | 2 (9.1%) | 0 (0%) |
| Avoidant personality problems | | | |
| Non-clinical | 17 (77.3%) | 14 (63.6%) | 20 (80.0%) |
| Sub-clinical | 2 (9.1%) | 2 (9.1%) | 1 (4.0%) |
| Clinical | 3 (13.6%) | 6 (27.3%) | 4 (16.0%) |
| ADHD problems | | | |
| Non-clinical | 17 (77.3%) | 18 (81.8%) | 19 (76.0%) |
| Sub-clinical | 2 (9.1%) | 2 (9.1%) | 3 (12.0%) |
| Clinical | 3 (13.6%) | 2 (9.1%) | 3 (12.0%) |
| Antisocial personality problems | | | |
| Non-clinical | 19 (86.4%) | 21 (95.5%) | 23 (92.0%) |
| Sub-clinical | 3 (13.6%) | 1 (4.5%) | 1 (4.0%) |
| Clinical | 0 (0%) | 0 (0%) | 1 (4.0%) |
| Total problems | | | |
| Non-clinical | 13 (59.1%) | 10 (45.5%) | 12 (49%) |
| Sub-clinical | 4 (18.2%) | 7 (31.8%) | 11 (44.0%) |
| Clinical | 5 (22.7%) | 5 (22.7%) | 2 (8%) |
| Depressive symptoms (PHQ-9) | | | |
| Minimal | 3 (13.6%) | 4 (19.0%) | 4 (16.0%) |
| Mild | 11 (50%) | 6 (28.6%) | 10 (40%) |
| Moderate | 5 (22.7%) | 8 (38.1%) | 7 (28.0%) |
| Moderately severe | 3 (13.6%) | 3 (14.3%) | 4 (16.0%) |
| Severe | 0 (0%) | 1 (4.5%) | 0 (0%) |

# CHAPTER 7

## Alpha-amylase reactivity and recovery patterns in anhedonic young adults performing a tandem skydive

Published as:
Vrijen, C.*, Roekel, E. van*, & Oldehinkel, A. J. (2018). Alpha-amylase reactivity and recovery patterns in anhedonic young adults performing a tandem skydive. *PLOS ONE*, *13*(9), e0204556. https://doi.org/10.1371/journal.pone.0204556
*shared first authorship

## ABSTRACT

**Background.** Anhedonia (loss of pleasure) is characterized by low responsiveness to rewards and, by virtue of being one of the two core symptoms of depression, by altered responses to stress. We investigated the effect of an acute stress experience (i.e., a tandem skydive) that was expected to elicit both intense fear and intense euphoria in a sample of anhedonic young adults.

**Objective.** (1) To examine individual differences in alpha-amylase reactivity to and recovery from a tandem skydive in anhedonic young adults; (2) to investigate whether trait depressive and anxiety problems, trait positive affect (PA), i.e., level of pleasure and reward responsiveness, and state anxiety, PA and self-esteem prior to the skydive were associated with alpha-amylase reactivity and recovery patterns; (3) to investigate whether alpha-amylase reactivity and recovery patterns were associated with pre- to post-jump changes in state anxiety, PA, and self-esteem.

**Method.** Participants were 61 individuals with persistent anhedonia (*M*age = 21.38, 78.7% female), who filled out a baseline questionnaire at the start of the study, and momentary questionnaires (3 times per day) before and after the tandem skydive. Alpha-amylase was measured at four time points by means of salivettes (2 before and 2 after the skydive).

**Results.** Alpha-amylase reactivity and recovery patterns were highly similar across individuals, although mean levels varied greatly. No associations were found between any of the trait and state measures and reactivity and recovery. Only state self-esteem was affected by the reactivity and recovery patterns, in that individuals who showed high reactivity and low recovery experienced decreases in self-esteem after the skydive.

**Conclusions.** Alpha-amylase patterns following a tandem skydive in anhedonic individuals are highly similar to patterns previously found in healthy individuals. Although replication is warranted, our findings tentatively suggest that a strong stress response that cannot be downregulated well predicts a decrease in self-esteem.

## INTRODUCTION

Free-fall experiences such as a skydive elicit fight-flight responses in virtually all individuals (Alpers & Adolph, 2008; Price & Bundesen, 2005). This is reflected in strong physiological stress responses (i.e., heart rate, cortisol levels and alpha-amylase levels) that return to normal not too long after completion of the skydive (Chatterton et al., 1997; Hare et al., 2013; Meyer et al., 2015), as well as in psychological responses of intense fear during the free-fall (Alpers & Adolph, 2008; Price & Bundesen, 2005), followed by euphoria after completion of the skydive (Celsi, Rose, & Leigh, 1993; Price & Bundesen, 2005).

There are indications that responses to stress and to reward interact; acute experimental stress has been found to reduce reward-related reactivity (Bogdan & Pizzagalli, 2006; Porcelli, Lewis, & Delgado, 2012). In one study, this effect was found only in individuals with strong stress responses (Berghorst, Bogdan, Frank, & Pizzagalli, 2013). The ability to experience reward has also been found to predict responses to stress and has been suggested to be of particular importance during recovery from stress (Corral-Frías, Nadel, Fellous, & Jacobs, 2016).

Since a tandem skydive evokes both extreme stress and extreme reward, it may be an ultimate natural model to explore the complex relationship between acute stress responsiveness and reward responsiveness. It is particularly relevant to investigate these stress and reward responses in individuals with anhedonia, as anhedonia is characterized by dysfunctions in the reward system (Treadway & Zald, 2011). In the present study, we investigated individual differences in reactivity to and recovery from a tandem skydive in a sample of anhedonic young adults, and tested whether individual variations in trait and state characteristics within this group were associated with reactivity and recovery patterns. (Please note that we use the concepts trait and state to indicate the time-scale of the measurements, that is, from weeks to months (trait) versus daily or momentary experiences (state), without making claims about stability.) In addition, we investigated whether stress reactivity and recovery patterns were associated with changes in state characteristics from pre-skydive to post-skydive.

Anhedonia is characterized by low positive affect (PA), and is one of the two core symptoms of depression. Depression has been associated with worse stress recovery (Burke, Davis, Otte, & Mohr, 2005; Salomon, Clift, Karlsdóttir, & Rottenberg, 2009) and blunted stress reactivity (Burke et al., 2005; Salomon et al., 2009), but also with higher stress reactivity (Tanaka et al., 2012). A study that distinguished between PA and negative affect (NA) found that high arousal NA (i.e., angry, stressed, nervous, worried) predicted alpha-amylase increases, but only in adolescents with high average levels of these emotions; in the same study high arousal PA (feeling strong, active, excited) also predicted increases in alpha-amylase regardless of mean levels (Adam, Hoyt, & Granger, 2011), suggesting that emotional arousal may be more relevant to alpha-amylase reactivity than valence (PA or NA). Another study found no association between pretest state PA and stress reactivity during a social stress test (Oldehinkel et al., 2011). It is possible that these mixed findings are due to different stress reactivity measures, that is, heartrate variability

(Oldehinkel et al., 2011; Salomon et al., 2009), respiratory sinus arrhythmia (Oldehinkel et al., 2011), cortisol (Burke et al., 2005; Oldehinkel et al., 2011) and alpha-amylase (Adam et al., 2011; Tanaka et al., 2012), but they may also be due to the type of stressor and the level of emotional arousal. Increased reactivity has been suggested in particular for novel and uncontrollable experiences (Tanaka et al., 2012) and for high arousal emotions (Adam et al., 2011). As a tandem skydive is a prototypical example of a novel and uncontrollable experience evoking high arousal emotions, we expected that the severity of the affective problems would be positively associated with the reactivity to the skydive.

In addition to the findings described above, there is evidence that trait and state anxiety may affect stress reactivity, but the diversity of results (Allwood, Handwerger, Kivlighan, Granger, & Stroud, 2011; Gerra et al., 2000; Noto, Sato, Kudo, Kurata, & Hirota, 2005; Oldehinkel et al., 2011; Takahashi et al., 2005) precludes expectations regarding the direction of the effects. Furthermore, there are indications that high state self-esteem (Rector & Roger, 1997) is associated with attenuated stress reactivity.

Apart from investigating trait and state characteristics as predictors of stress responses, it is also interesting to explore how stress response patterns in turn predict changes in state characteristics. Despite that several studies have explored state changes in response to experimentally induced stress, hardly any have explicitly examined how stress reactivity and recovery are associated with these changes. One study found that cortisol reactivity to a lab stress task was not related to pre- to post-stress changes in PA in healthy individuals (Buchanan, al'Absi, & Lovallo, 1999); another that heartrate reactivity to a social stress test was associated with pre- to posttest changes in PA and perceived control (Oldehinkel et al., 2011). A third study found that that a fast cortisol recovery was associated with experiencing more PA right after a skydive (Meyer et al., 2015). Because in the latter study only PA assessed after the skydive was taken into account, it is unclear to what extent the association could be explained by pre-skydive PA. In the current study we assessed state anxiety, PA, and self-esteem both before and after the skydive, which enabled us to investigate whether stress reactivity and recovery were associated with changes in these state characteristics.

The main aim of our study was to investigate responses to the acute stress and euphoria evoked by a tandem skydive in young adults suffering from anhedonia who were novice to skydiving. We used salivary alpha-amylase levels as markers for physiological stress reactivity and recovery. Alpha-amylase is an enzyme that is secreted under autonomic regulation and has been found to be highly responsive to acute stress in humans (Nater & Rohleder, 2009; Schumacher, Kirschbaum, Fydrich, & Ströhle, 2013). As such, it has been frequently used as a biomarker for stress. As alpha-amylase not only rapidly increases in response to stress but, as opposed to cortisol, has also been found to decrease in response to relaxation (Takai et al., 2004), it is expected to function as an important marker for both stress reactivity and recovery. We examined (1) individual differences in reactivity to and recovery from a tandem skydive in a

sample of anhedonic young adults; (2) whether trait depressive symptoms, anxiety, and PA (i.e., level of pleasure and reward responsiveness), and state anxiety, PA, and self-esteem prior to the skydive were associated with alpha-amylase reactivity and recovery patterns in response to a skydive; and (3) whether these reactivity and recovery patterns were associated with pre- to post-skydive changes in state anxiety, PA, and self-esteem.

## MATERIALS AND METHODS

### Participants

The present study is part of the larger intervention study "No Fun No Glory", in which anhedonic young adults were given personalized lifestyle advice and exposed to tandem skydives as a possible mean to reduce their symptoms. A detailed description of the study protocol can be found elsewhere (Van Roekel et al., 2016). Participants were recruited from the general population by means of an online screening survey among 2,937 young adults (18-24 years). Inclusion criteria were persistent loss of pleasure and willingness to perform a skydive. It should be noted that, of the total sample that was screened, only 12.8% were unwilling to perform a tandem skydive (N=376) and that these participants did not significantly (i.e., $p > .05$) differ from those who were willing (N=1759; 59.9%), or from those who indicated 'maybe' (N=802; 27.3%), on sex, age, trait PA, depressive symptoms (PHQ-9; Kroenke, Spitzer, & Williams, 2001), and reward responsiveness (RR; Van den Berg et al., 2010). Persistent loss of pleasure was measured with three items of the Domains of Pleasure Scale (DOPS; Masselink et al., n.d.). Participants needed to (1) score below the 25$^{th}$ percentile on level of pleasure, (2) rate this level as less or much less than normal, and (3) report that this loss of pleasure was present for at least two months. Exclusion criteria were inability to keep an electronic diary three times a day; professional treatment for psychiatric problems; use of psychotropic medication; epilepsy; pregnancy; conditions that obstruct participating in a tandem skydive (i.e., loose prostheses; height of more than 2 meters; weight of more than 95 kg; inability to raise one's legs 90 degrees; cardiovascular complaints or problems; and significant visual or hearing impairments); and experience with skydiving, bungee jumping, or base jumping. Hence, all participants were new to skydiving. The sample used in the present study consisted of 61 participants ($M_{age}$ [SD] = 21.38 [1.98], 78.7% female), who were either randomly assigned to the intervention group who received a tandem skydive (N = 24) or initially assigned to another group and chose to participate in the tandem skydive themselves in a later phase of the study (N = 37). These groups did not differ in demographics (i.e., age, gender, BMI), nor in the level of anhedonia, depression, anxiety, or reward responsiveness measured at baseline, before the first (randomized) intervention, and before the second (free-choice) intervention, and also did not differ in alpha-amylase reactivity, recovery or mean level in response to the skydive ($p > .05$).

## Procedure

Participants who were eligible to participate in the intervention study received an information letter and were included in the study after providing their written informed consent. All participants filled out momentary assessments on their smartphone, three times a day with fixed 6h intervals for at least three months. Only data from the week before the skydive and the day of the skydive were used in the present study. Participants filled out monthly questionnaires at the start of the study and after two and three months. The tandem skydive always took place during weekends, at the skydive center Eelde-Hoogeveen. The tandem skydive was performed from a small turbine-powered aircraft (i.e., Cessna 207), at a height of 10,000 feet. Participants were safely attached to the tandem skydive instructor. The duration of the free fall to 5,000 feet was 30 to 40 seconds, the duration of the total skydive was around 5 minutes. Participants were instructed not to eat, drink (except water), smoke or eat chewing-gum in the two hours before the skydive. They had to be present at the skydive center at least half an hour before the skydive. Saliva samples were collected by means of cotton salivettes, which were administered at four time points: 15 minutes after the participants arrived at the skydive center, after they had received the instructions and were ready to board the plane, immediately after they had landed, and 20 minutes after they had landed. Participants were instructed to put the salivette in their mouth without touching it with their hands, and to keep it in their cheek, while occasionally chewing on it to let the cotton be absorbed by saliva for two minutes, as timed by the researcher that was in control of the procedure. The salivettes were immediately stored in a cooling box containing cooling elements. They were transported to Groningen and kept in a fridge for a maximum of two days, and then transported to the laboratory of the University Medical Center Groningen (UMCG). At the lab, the samples were centrifuged at 1000xg at 4°C for 2 minutes, aliquoted and stored at -80°C. The current study was registered in the Dutch Clinical Trial Register (NTR5498) and approved by the Medical Ethical Committee from the University Medical Center Groningen (no. 2014/508).

## Measures

### *Alpha-amylase*

All alpha-amylase samples were analyzed in the same week, on four different days, at the general haematology and chemistry Lab in the UMCG, by means of an enzymatic colorimetric analysis, according to the International Federation of Clinical Chemistry (IFCC) method (Lorentz, 1998), on the Roche Modular P analyzer (Roche Diagnostics GmbH, Mannheim, Germany). Samples were defrosted overnight and centrifuged at 1520xg at room temperature for 3 minutes the next morning. All samples were diluted 1:100 with saline in an automatic onboard dilution step. Samples with alpha-amylase levels that were still too high after the automatic dilution were first diluted with saline 1:10 manually before the automatic dilution of 1:100, resulting in

a total dilution of 1:1000. Results for the manually diluted samples were multiplied by 10. The total coefficients of variation were between 2.5% (for low levels of alpha-amylase) and 5.2% (for high levels of alpha-amylase) with mean inter assay coefficients of variation between 1.7% and 3.2%. We calculated the mean level of the alpha-amylase levels across the four assessments. For reactivity, we calculated the difference between T3 (i.e., first assessment after the skydive) and T2 (last assessment before the skydive) and divided this by the level at T2 to obtain the proportional increase in alpha-amylase compared to each individual's level before the skydive (see Equation 1). For recovery, we calculated the difference between T4 (i.e. second assessment after the skydive) and T3 (first assessment after the skydive) and divided this by the level at T3 to obtain the proportional decrease in alpha-amylase compared to each individual's level right after the skydive (see Equation 2).

$$Reactivity = \frac{(amylase\ T3 - amylase\ T2)}{amylase\ T2} \quad (3)$$

$$Recovery = \frac{(amylase\ T4 - amylase\ T3)}{amylase\ T3} \quad (4)$$

We used these measures of relative change for reactivity and recovery, because we were primarily interested in how much participants' alpha-amylase levels changed from one assessment to the next relative to their own prior alpha-amylase level, and aimed to separate this within-subject effect from a possible confounding (between-subject) effect of mean alpha-amylase level, which is known to be highly variable across individuals.

***Baseline assessments***

Depressive and anxiety problems during the past six months were assessed at baseline with the Adult Self-Report (ASR), which is a standardized questionnaire of behavioral and emotional problems, which has been shown to have good reliability and validity (Achenbach & Rescorla, 2003). We used the ASR depressive problems scale (14 items) and the anxiety problems scale (7 items) that are based on the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV; American Psychiatric Association, 1994). For each item, answer categories are: 0 = 'Not True'; 1 = 'Somewhat or Sometimes True'; 2 = 'Very True or Often True'. In our sample, Cronbach's alpha was .78 for depressive problems and .78 for anxiety problems. Participants filled out the Domains and Dimensions of Pleasure Scale (DDOPS), in which we assessed pleasure during the past two weeks with a VAS ranging from "*I experience little pleasure*" to "*I experience a lot of pleasure*". Reward responsiveness during the past two weeks was measured with the Reward Responsiveness scale (Van den Berg et al., 2010), which consists of 8 items that are scored on a 4-point scale ranging from *strong disagreement* (1) to *strong agreement* (4). Cronbach's alpha was .84.

### *Momentary assessments*

In the momentary assessments participants reported their anxiety, PA and self-esteem. All items were rated on a Visual Analogue Scale (VAS), ranging from *not at all* to *very much*. Participants rated each item by moving a slider along the scale. The position of the slider was transformed into a score between 0-100. Anxiety was a combined measure of the items anxious and nervous (inter-item correlation .86). Positive affect (PA) was measured with six items, which were divided in low arousal PA (i.e., calm and relaxed; inter-item correlation .93), moderate arousal PA (i.e., joyful and cheerful; inter-item correlation .89), and high arousal PA (i.e., enthusiastic and energetic; inter-item correlation .88). We considered low, moderate and high arousal PA separately, because we expected that there may be a difference in how they would affect stress reactivity and recovery. To our best knowledge, this has not yet been tested. Self-esteem was measured with the item 'I was pleased with myself'.

### *Covariates*

Age (Veen et al., 2012), gender (Wingenfeld et al., 2010), BMI (Nater, Rohleder, Schlotz, Ehlert, & Kirschbaum, 2007), smoking (Rohleder & Nater, 2009), alcohol use (Rohleder & Nater, 2009; Veen et al., 2012), physical activity (Rohleder & Nater, 2009), asthma medication (Rohleder & Nater, 2009), and time of day (Marchand, Juster, Lupien, & Durand, 2016; Nater et al., 2007; Out, Granger, Sephton, & Segerstrom, 2013) may have an effect on alpha-amylase. At baseline, participants reported their age, gender, weight, and height. Weight and height were used to calculate BMI. Participants also reported whether they used medication and if so, which type. Reported medications were checked for being prescribed for asthma. Participants reported at each momentary assessment how many cigarettes and how many alcoholic beverages they consumed since the previous assessment. All cigarettes and alcoholic drinks consumed during the week before the skydive were summed into a total smoking and a total alcohol consumption score. Physical activity was measured with the momentary assessment item 'I was physically active', which was rated on a VAS ranging from *not at all* to *very much*. We included the last physical activity measure before the skydive (i.e., either the morning or afternoon assessment, depending on the time of the skydive). For time of day, we included the time of assessment of the second salivette (i.e., immediately before entering the plane). Time of day was calculated as the number of minutes past midnight.

## Strategy of Analyses

### *Alpha-amylase patterns and descriptive statistics*

First, we plotted the group-based mean alpha-amylase levels before (two assessments) and after (two assessments) the tandem skydive, as well as each individual secretion pattern. For the group-based plot as well as for all below-described statistical analyses we used SPSS version 25 (IBM Corporation, 2017); for the individual plots we used R packages dyplr version

0.7.4 (Wickham, Francois, Henry, & Müller, 2017), ggplot2 version 2.2.1 (Wickham, 2009), and cowplot version 0.9.2 (Wilke, 2017), R version 3.4.3 (R Core Team, 2013). Second, we calculated mean levels of and correlations among the trait and state measures, and alpha-amylase mean level, reactivity and recovery. Trait measures were based on the baseline assessments and state measures on the momentary assessments during the day before the skydive and on the day of the skydive. Because previous research found a diurnal pattern in alpha-amylase levels (i.e., levels increased during the day; Adam et al., 2011), we checked whether the timing of the alpha-amylase assessments was associated with mean levels.

***Regression analyses***

We used hierarchical regression models to examine whether trait depressive problems, anxiety problems, and PA (pleasure level and reward responsiveness) were associated with alpha-amylase reactivity and recovery. In addition, we checked whether the association between the predictors and recovery was moderated by reactivity by including the interaction between the predictors and reactivity in the model. All predictors were grand-mean centered before the interaction term was calculated. The same analyses were repeated for state levels of anxiety (i.e., anxious/nervous), PA (i.e., PA low arousal, PA moderate arousal, PA high arousal), and self-esteem, averaged over the day before the skydive. To examine whether reactivity and recovery were associated with a change in state levels from the evening before the skydive to the evening after the skydive, we conducted hierarchical regression analyses with state level at the evening after the skydive as dependent variable (i.e., anxiety, PA low arousal, PA moderate arousal, PA high arousal, self-esteem), and reactivity and recovery as predictors, while controlling for state level at the evening before (step 1). In a second step, we explored whether the effects of recovery on affect were moderated by reactivity by entering the interaction between reactivity and recovery. Again, predictors were grand-mean centered before computing the interaction term.

For all statistically significant findings ($p < .05$), we performed sensitivity analyses to check whether the results changed when adjusted for potential confounders. Since alpha-amylase reactivity and recovery may partly depend upon mean alpha-amylase levels, we also corrected for this variable in our sensitivity analyses.

***Open Science***

Data and syntax have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/4gbew/.

# RESULTS

## Alpha-Amylase Patterns

As can be seen in Figure 1, on average the alpha-amylase levels showed a strong increase between T2 and T3 (reactivity) and a strong decrease between T3 and T4 (recovery). Figure 2 shows that this pattern applied to most individuals. One participant, indicated by the gray cross in Figure 2, showed a suspiciously high alpha-amylase level at T4 combined with very low levels at T1-T3. We decided to discard this individual from further analyses. After removal of this outlier, 57 out of in total 60 participants showed an increase at T3 compared to T2 and an equal number of (partly different) participants showed decreasing alpha-amylase levels at T4 as compared to T3. Mean alpha-amylase levels per participant were highly variable, and ranged between 33 U/mL and 1245 U/mL.

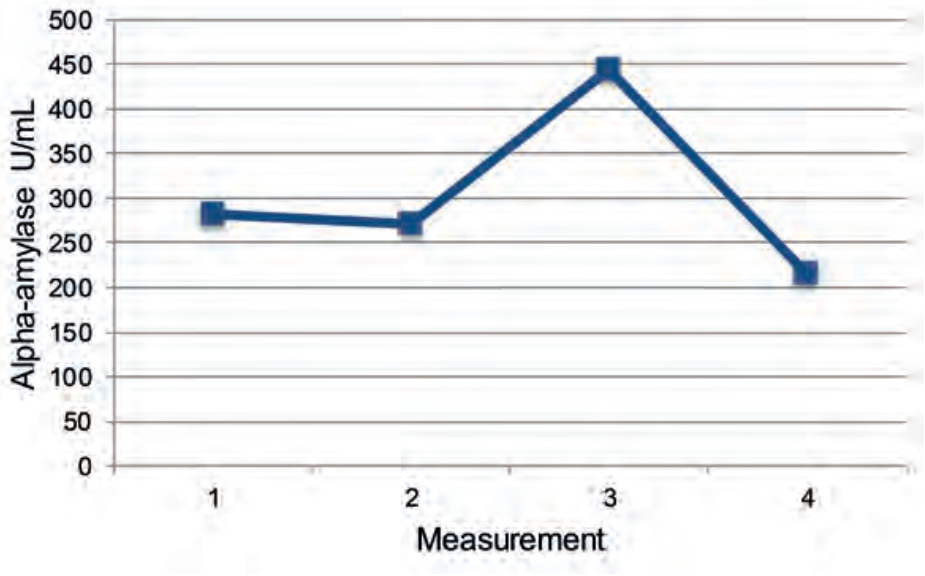



**FIGURE 1.** Plot mean alpha-amylase levels (U/mL) per measurement. The first measurement took place 15 minutes after the participant arrived at the skydive center, the second after they received the instructions and were ready to board the plane, the third immediately after they landed, and the fourth 20 minutes after they landed.

## Descriptive Statistics

Means of and correlations between study variables are depicted in Table 1. The proportional reactivity and recovery scores were moderately correlated, indicating that stronger reactivity is associated with stronger recovery. None of the affect measures were associated with mean

alpha amylase levels, reactivity or recovery. The time at which the skydive took place was not associated with mean alpha-amylase levels ($r$ = .12, p = .37), indicating that mean levels of participants skydiving in the afternoon were not higher than mean levels for participants skydiving in the morning.

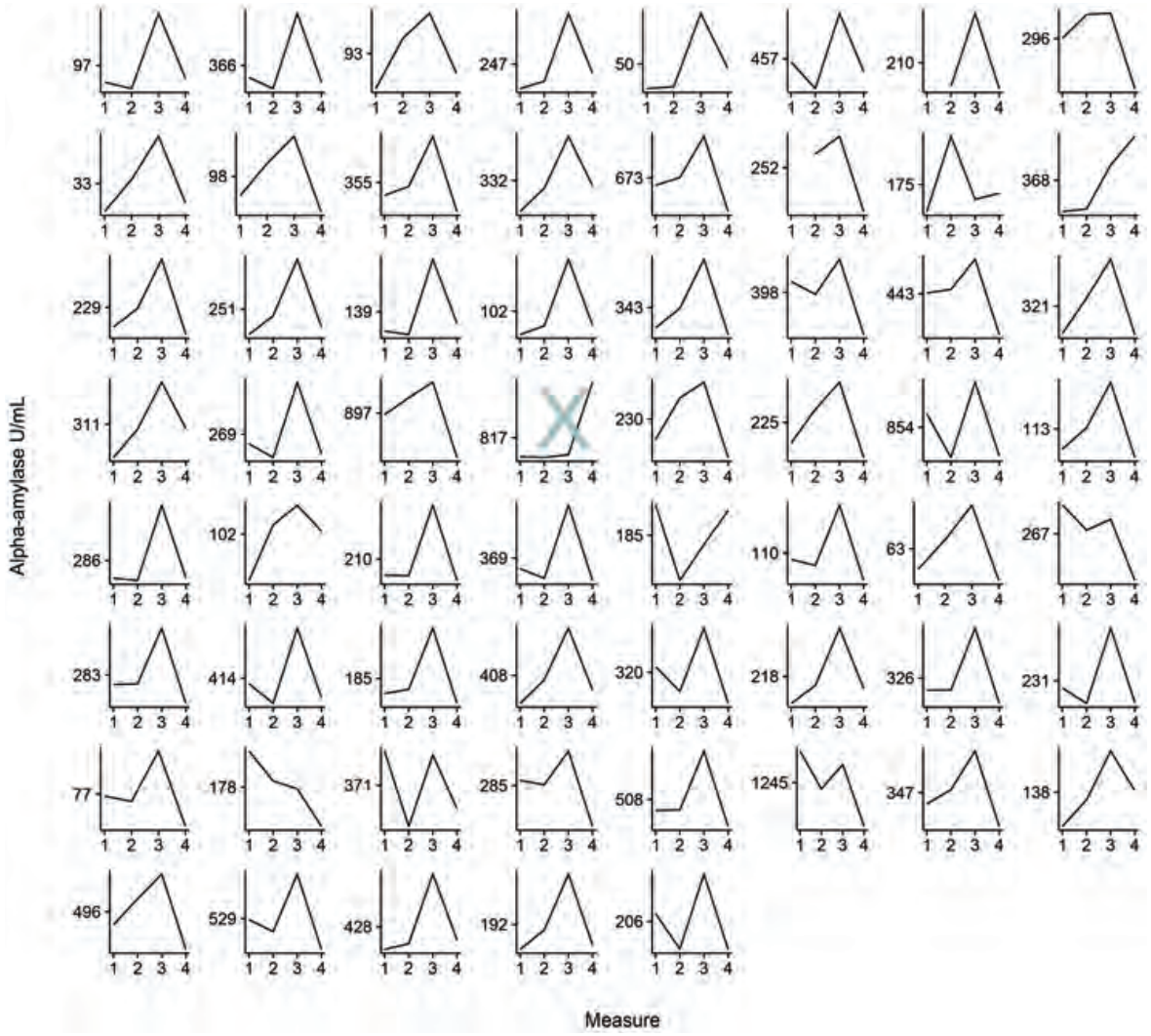



**FIGURE 2.** Individual plots alpha-amylase levels (U/mL) on four measurements surrounding the tandem skydive. The first measurement took place 15 minutes after the participant arrived at the skydive center, the second after they received the instructions and were ready to board the plane, the third immediately after they landed, and the fourth 20 minutes after they landed. The mean score of the four measurements is shown on the y-axis. The gray cross indicates a participant with an outlier on T4. This participant was removed prior to further analyses.

**TABLE 1.** Descriptive Statistics

| | Mean (SD) | Range | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|
| Average alpha-amylase | 303.41 (211.89) | 33.40 – 1244.62 | - | | | | |
| Reactivity | 0.77 (0.71) | -0.19 –3.73 | -.03 | - | | | |
| Recovery | 0.47 (0.21) | -0.25 – 0.78 | .15 | .48** | - | | |
| Baseline depressive problems | 0.79 (0.34) | 0.07 – 1.57 | -.15 | -.02 | -.10 | - | |
| Baseline anxiety problems | 0.86 (0.44) | 0.14 – 1.71 | -.04 | -.11 | -.02 | .60** | - |
| Baseline pleasure | 35.34 (13.93) | 0.53 – 55.74 | .09 | -.01 | -.20 | -.39** | -.13 |
| Baseline reward responsiveness | 25.97 (3.83) | 10 - 32 | -.06 | .07 | -.12 | -.10 | -.05 |
| State anxiety day before skydive | 16.90 (13.38) | 2.00 – 66.95 | -.06 | -.03 | -.12 | .24 | .32* |
| State PA low arousal day before skydive | 60.95 (14.81) | 11.74 – 91.03 | .14 | -.00 | .14 | -.40** | -.44** |
| State PA moderate arousal day before skydive | 58.75 (12.69) | 10.74 – 93.01 | .03 | .09 | .13 | -.36** | -.41** |
| State PA high arousal day before skydive | 55.69 (14.92) | 11.69 – 93.67 | .04 | .08 | -.01 | -.21 | -.33* |
| State self-esteem day before skydive | 52.22 (13.23) | 9.73 – 92.44 | .11 | -.02 | .04 | -.41** | -.38** |
| State anxiety evening before skydive | 17.19 (17.12) | .89 – 96.60 | -.06 | -.06 | -.11 | .21 | .37** |
| State PA low arousal evening before skydive | 60.22 (17.56) | 9.21 – 90.51 | .07 | .03 | .17 | -.42** | -.43** |
| State PA moderate arousal evening before skydive | 60.05 (16.28) | 9.49 – 94.05 | .01 | .14 | .08 | -.41** | -.39** |
| State PA high arousal evening before skydive | 57.20 (16.65) | 11.08 – 93.20 | -.01 | .04 | -.15 | -.30* | -.30* |
| State self-esteem evening before skydive | 52.24 (15.85) | 18.46 – 93.77 | -.05 | -.00 | -.04 | -.36** | -.32* |
| State anxiety evening after skydive | 23.70 (22.85) | 0.71 – 95.98 | -.13 | -.04 | -.06 | .23 | .26 |
| State PA low arousal evening after skydive | 62.89 (16.96) | 5.54 – 92.21 | .11 | -.01 | -.02 | -.32* | -.27* |
| State PA moderate arousal evening after skydive | 68.72 (13.39) | 15.61 – 95.89 | -.02 | -.00 | -.06 | -.31* | -.19 |
| State PA high arousal evening after skydive | 65.85 (14.90) | 17.29 – 95.75 | .03 | -.11 | -.08 | -.12 | -.06 |
| State self-esteem evening after skydive | 63.58 (17.46) | 8.39 – 95.97 | .05 | .03 | .04 | -.30* | -.17 |

*Note*. *** $p < .001$, ** $p < .01$, * $p < .05$. PA = positive affect.

| 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| - | | | | | | | | | | | | | | | |
| .26* | - | | | | | | | | | | | | | | |
| -.09 | .10 | - | | | | | | | | | | | | | |
| .17 | .01 | -.37** | - | | | | | | | | | | | | |
| .16 | .08 | -.14 | .73** | - | | | | | | | | | | | |
| .10 | .15 | .01 | .62** | .87** | - | | | | | | | | | | |
| .29* | .13 | -.24 | .65** | .77** | .74** | - | | | | | | | | | |
| -.04 | .08 | .87** | -.21 | -.10 | .10 | -.19 | - | | | | | | | | |
| .20 | -.05 | -.45** | .87** | .68** | .48** | .55** | -.46** | - | | | | | | | |
| .25 | .04 | -.24 | .56** | .84** | .64** | .66** | -.40** | .74** | - | | | | | | |
| .26 | .15 | -.03 | .48** | .78** | .88** | .64** | -.07 | .52** | .75** | - | | | | | |
| .38** | .18 | -.22 | .42** | .64** | .57** | .85** | -.32* | .52** | .71** | .64** | - | | | | |
| -.24 | .00 | .69** | -.25 | -.11 | .02 | -.15 | .68** | -.39** | -.31* | -.11 | -.27 | - | | | |
| .32* | .19 | -.22 | .69** | .54** | .47** | .35** | -.03 | .72** | .49** | .55** | .29* | -.50** | - | | |
| .25 | .23 | -.00 | .57** | .71** | .74** | .57** | .14 | .48** | .49** | .71** | .42** | .00 | .54** | - | |
| .14 | .22 | .16 | .42** | .55** | .68** | .45** | .26 | .26 | .31* | .58** | .30* | .23 | .32* | .83** | - |
| -.02 | .16 | .02 | .34* | .43** | .46** | .43** | .10 | .23 | .25 | .37** | .26 | .08 | .27 | .74** | .74** |

## Trait and State Measures as Predictors of Alpha-amylase Patterns

First, we explored associations between trait and state measures and reactivity and recovery (Table 2). Trait depressive and anxiety problems were entered in the model simultaneously and both trait PA measures were also entered in one model simultaneously. For the state measures, this was not possible due to multicollinearity issues. No direct associations were found between the trait and state measures and reactivity or recovery.

**TABLE 2.** Regression Analyses for Trait and State predictors on Alpha-Amylase Reactivity and Recovery

| | | Alpha-amylase | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Reactivity | | | | Recovery | | | |
| | | **B** | ***β*** | ***p*** | ***R²*** | **B** | ***β*** | ***P*** | ***R²*** |
| Trait models | Depressive problems | 0.23 | 0.11 | .54 | | -0.09 | -0.14 | .43 | |
| | Anxiety problems | -0.32 | -0.19 | .28 | .02 | 0.03 | 0.06 | .75 | .01 |
| | Pleasure | -0.01 | -0.12 | .39 | | -0.00 | -0.20 | .14 | |
| | Reward responsiveness | 0.02 | 0.09 | .49 | .02 | -0.01 | -0.10 | .47 | .05 |
| State models | Anxiety | -0.00 | -0.04 | .79 | .00 | -0.00 | -0.12 | .39 | .01 |
| | PA low arousal | 0.00 | 0.00 | .98 | .00 | 0.00 | 0.14 | .29 | .02 |
| | PA moderate arousal | 0.01 | 0.10 | .47 | .01 | 0.00 | 0.14 | .30 | .02 |
| | PA high arousal | 0.00 | 0.08 | .54 | .01 | -0.00 | 0.00 | .99 | .00 |
| | Self-esteem | -0.00 | -0.02 | .88 | .00 | 0.00 | 0.04 | .79 | .00 |

*Note.* PA = positive affect. Please note that all state predictors were analyzed in separate regression models, due to high inter-correlations.

Reactivity did not modify the associations between trait and state measures and recovery; all interaction effects were non-significant (depressive problems: $\beta = -0.03$, $p = .80$, anxiety problems: $\beta = 0.02$, $p = .87$, $\Delta R^2$ for full model= .00, level of pleasure: $\beta = 0.02$, $p = .87$, reward responsiveness: $\beta = -0.06$, $p = .67$, $\Delta R^2$ for full model= .00, state anxiety: $\beta = -0.01$, $p = .96$, $\Delta R^2 = .00$, PA low arousal: $\beta = -0.01$, $p = .93$, $\Delta R^2 = .00$, PA moderate arousal: $\beta = 0.01$, $p = .97$, $\Delta R^2 = .00$, PA high arousal: $\beta = -0.01$, $p = .94$, $\Delta R^2 = .00$, and self-esteem: $\beta = 0.15$, $p = .28$, $\Delta R^2 = .02$).

## Alpha Amylase Reactivity and Recovery as Predictors of Changes in Anxiety, PA and Self-esteem

Neither alpha-amylase reactivity nor recovery predicted anxiety, PA and self-esteem reported at the evening after the skydive (Table 3). The interaction between reactivity and recovery was significant for self-esteem ($\beta = 0.46$, $p = .03$, $\Delta R^2 = .10$). As can be seen in Figure 3, when individuals

showed low reactivity, they experienced high levels of self-esteem after the skydive, irrespective of their level of recovery. In contrast, self-esteem was low for individuals who show high levels of reactivity and low levels of recovery. No significant interactions between reactivity and recovery were found for anxiety ($\beta$ =- 0.18, $p$ = .29, $\Delta R^2$= .01), PA low arousal ($\beta$ = 0.10, $p$ = .53, $\Delta R^2$= .00), PA moderate arousal ($\beta$ = 0.34, $p$ = .08, $\Delta R^2$= .05), or PA high arousal ($\beta$ = 0.10, $p$ = .61, $\Delta R^2$= .00).

**TABLE 3.** Regression Analyses for Alpha-Amylase Reactivity and Recovery on State Anxiety, PA, and Self-esteem After the Skydive

| | Anxiety | | | | PA low arousal | | | | PA moderate arousal | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | B | $\beta$ | $p$ | $\Delta R^2$ | B | $\beta$ | $p$ | $\Delta R^2$ | B | $\beta$ | $p$ | $\Delta R^2$ |
| Evening before | 0.78 | 0.69 | .00 | | 0.65 | 0.75 | .00 | | 0.42 | 0.50 | .00 | |
| Reactivity | -0.38 | -0.02 | .90 | | 0.99 | 0.05 | .67 | | 0.15 | 0.01 | .96 | |
| Recovery | 3.87 | 0.04 | .75 | .00 | -13.18 | -0.18 | .15 | .02 | -10.76 | -0.16 | .29 | .02 |

| | PA high arousal | | | | Self-esteem | | | |
|---|---|---|---|---|---|---|---|---|
| | B | $\beta$ | $p$ | $\Delta R^2$ | B | $\beta$ | $p$ | $\Delta R^2$ |
| Evening before | 0.51 | 0.58 | .00 | | 0.30 | 0.26 | .09 | |
| Reactivity | -2.51 | -0.13 | .36 | | 0.89 | 0.04 | .82 | |
| Recovery | -0.19 | -0.00 | .99 | .02 | 1.27 | 0.01 | .93 | .05 |

*Note.* PA = positive affect. Evening before = controlled for level on evening before the skydive. $\Delta R^2$ represents the change in explained variance when reactivity and recovery were added as predictors to the model.

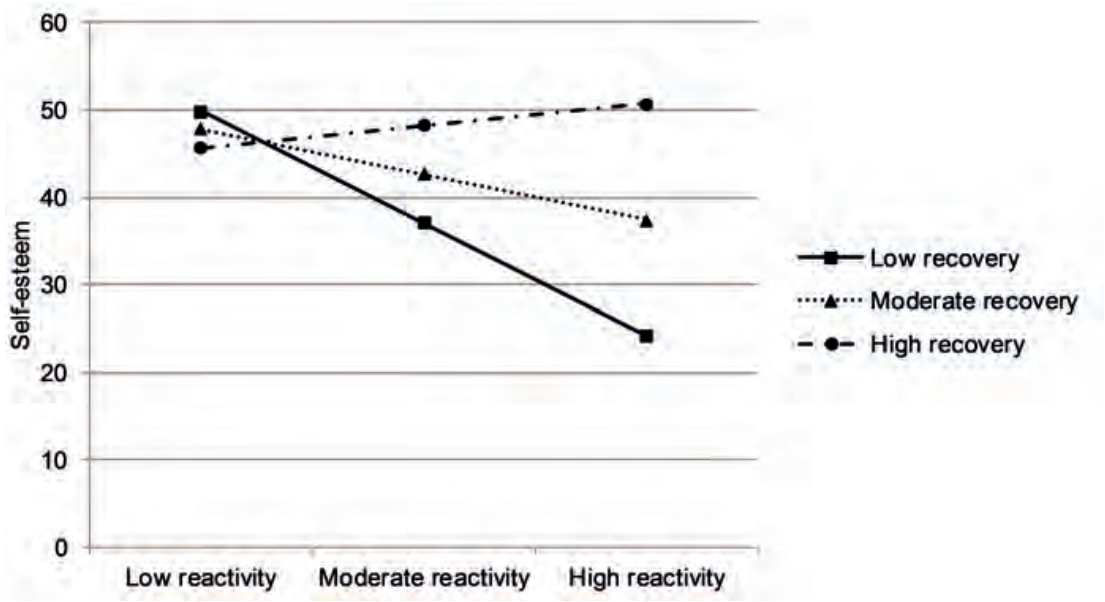



**FIGURE 3.** Interaction between reactivity and recovery on self-esteem

### Sensitivity Analyses

In a final step, we checked for statistically significant findings ($p < .05$) whether controlling for the previously mentioned covariates affected our results (for detailed results see Table S1 of the Supplementary Material). The reported effect that was significant remained significant after entering the covariates.

## DISCUSSION

The main aim of the present study was to investigate reactivity and recovery patterns in alpha-amylase levels evoked by a tandem skydive. Results indicated that, although mean alpha-amylase levels across the measures greatly differed among individuals, the reactivity and recovery patterns were highly similar. We found no associations between trait or state measures and reactivity or recovery. Finally, the association between reactivity and self-esteem after the skydive depended on the level of recovery: individuals who showed high levels of reactivity in combination with low levels of recovery experienced the lowest levels of self-esteem.

While the mean alpha-amylase levels of our participants greatly differed, the reactivity and recovery patterns were highly comparable (see Figure 2); most participants were highly reactive to the skydive (increase in alpha-amylase levels) and showed a quick recovery afterwards (decrease in alpha-amylase levels). Mean levels of alpha-amylase in our study (i.e., 303.41 U/mL) were, as expected, high compared to reported alpha-amylase levels in studies without an acute stressor (i.e., 224.1 U/mL for depressed and 173.9 U/mL for non-depressed individuals; Booij et al., 2015). Similar to a previous skydiving study in non-anhedonic individuals (Chatterton et al., 1997), in our study the highest mean alpha-amylase levels were measured right after landing. However, the mean alpha-amylase levels directly after the skydive were rather low in our sample (444 U/mL), compared to levels found in Chatterton's study (900 U/mL). This may be due to the fact that the participants in Chatterton's study were responsible for opening the parachute during the tandem skydive themselves, which may have added to their stress levels; it may also be partly due to the anhedonic symptoms of our study participants. Chatterton and colleagues reported a 40% decrease in alpha-amylase between the sample collected immediately after landing and the one collected 15 minutes later. This pattern of stress recovery closely matched the pattern found in our anhedonic young adults who showed an average decrease in alpha-amylase of 47% between the sample collected after landing and the sample collected 20 minutes later. Overall, similarities between salivary alpha-amylase reactivity and recovery patterns we report in the present study and those reported in studies on non-anhedonic samples suggest that (sub-clinical) anhedonic symptoms may not largely impact the reactivity to and recovery from ultimate thrill experiences such as a tandem skydive.

Previous research has indicated that trait depression and anxiety may be associated with higher alpha-amylase levels (Booij et al., 2015; Yorbik et al., 2016) and trait depression with increased alpha-amylase reactivity in response to an acute stressor (Tanaka et al., 2012), but we did not find these associations in the present study. These diverging findings might be due to the fact that we examined these associations within an anhedonic sample, whereas most previous studies compared stress responses between healthy individuals and psychiatric populations (Tanaka et al., 2012; Yorbik et al., 2016). We also expected that reward responsiveness would be associated with stress responses to the skydive, particularly that higher reward responsiveness would facilitate stress recovery (Corral-Frías et al., 2016), but we did not find this. Speculatively, perhaps differences in reward responsiveness in daily life situations, as measured with the Reward responsiveness Scale, do not apply to the extreme feelings of euphoria experienced after a skydive. If this is indeed the case slightly less extreme stressors may be more appropriate to investigate if reward responsiveness is related to individual differences in stress recovery. Similarly, levels of depression and anxiety may explain responses to stress in daily life situations, but perhaps not in response to the ultimate stress experience of a skydive. However, further research is needed to explore these potential explanations.

Although the effect was small and should be interpreted with caution, results suggested that the combination of reactivity and recovery rates might affect self-esteem. Individuals who experienced either low reactivity or high recovery showed high post-skydive self-esteem, but individuals who combined high reactivity with low recovery experienced relatively low self-esteem after the skydive. This positive effect of high recovery is partly in line with a previous study in skydivers, which found that greater happiness after the skydive was associated with faster recovery in cortisol (Meyer et al., 2015). Whereas the direction of effects was unclear in this previous study, we found that high recovery was prospectively associated with a positive change in self-esteem. We only found this effect for self-esteem and not for our other mood measures. This could imply that being able to downregulate your stress levels after such an intense stress experience leads to a boost in self-confidence, but not to a boost in mood per se. We acknowledge that this finding was small ($p = .03$), and would not survive multiple testing correction. Hence, it could be a false positive finding and should be interpreted with caution. On the other hand, the effect size ($\beta = 0.46$, explained variance = 10%) was not that small and given our rather small sample size, lack of power may also explain why this finding would not survive correction for multiple testing.



## Strengths and Limitations

The present study had multiple strengths, including the extreme nature of the stressor, which is a novel and uncontrollable experience that elicits a fight-flight response in all individuals. Further, we included intensive momentary measurements before and after the skydive, and hence were able to obtain a detailed picture of affective experiences around the skydive. However, the

results of this study should be interpreted in the context of some limitations. Given that we only included anhedonic individuals, our results may not be generalizable to the general population. Because of the lack of a control group we were also unable to compare stress responses to the skydive between anhedonic and non-anhedonic individuals in the current study, and thus were limited to comparing our findings for anhedonic individuals to findings for healthy individuals reported in other studies. Further, because of the unequal sex distribution (21.7% males) and small sample size, we were not able to examine sex differences in the associations. Another limitation concerns the alpha-amylase sampling. The first (baseline) saliva sample was collected when participants arrived at the skydive center and may have been stressed already. Therefore, we could not examine whether participants recovered to their actual baseline levels. Further, for practical purposes and convenience for the participants we used salivettes to assess alpha-amylase levels, a method that has some limitations (Bosch, Veerman, de Geus, & Proctor, 2011). One of the main limitations is that chewing affects the flow rate of saliva, which may affect the alpha-amylase concentration. Therefore, a passive drooling technique might have been better. Finally, we did not assess how participants experienced the skydive directly and therefore had to rely on how they felt before and after the skydive.

## Conclusion

In sum, we showed that alpha amylase reactivity and recovery patterns in response to a skydive in anhedonic individuals were highly consistent with patterns previously found in healthy individuals. Within an anhedonic sample, trait depressive and anxiety problems, trait PA, and state anxiety and PA were not associated with reactivity and recovery patterns. Finally, although replication in larger samples is warranted, our results tentatively suggest that in individuals with high stress reactivity, the ability to downregulate one's stress levels after a skydive was related to a boost in self-esteem, but not to mood in general.

## SUPPLEMENTARY MATERIAL

**TABLE S1.** Regression Analyses for Alpha-Amylase Reactivity and Recovery on Affect After the Skydive, full model including all covariates.

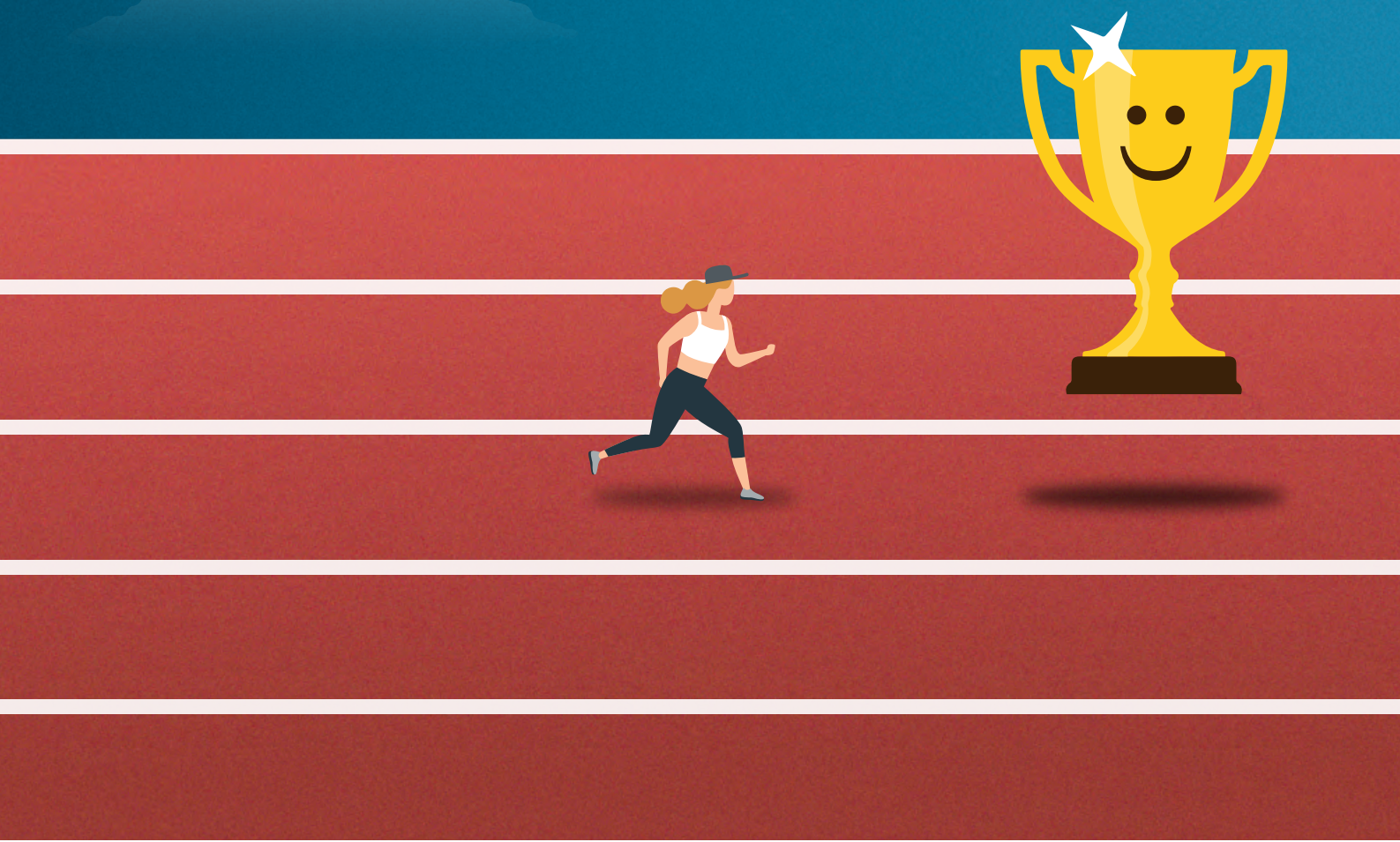
| | Self-esteem | | |
|---|---|---|---|
| | **B** | ***β*** | ***p*** |
| Self-esteem evening before skydive | 0.23 | .20 | .203 |
| Reactivity | -6.32 | -.27 | .263 |
| Recovery | 32.71 | .37 | .106 |
| Reactivity * Recovery | 51.39 | .45 | .045 |
| Mean amylase level | -0.01 | -.13 | .461 |
| Age | 3.02 | .34 | .025 |
| Gender | 4.76 | .11 | .538 |
| BMI | -0.99 | -.17 | .330 |
| Average number of cigarettes (week before skydive) | 0.16 | .13 | .539 |
| Average number of alcoholic beverages (week before skydive) | 0.19 | .09 | .633 |
| Physical activity (assessment before the skydive) | -0.28 | -.20 | .238 |
| Asthma medication | -3.05 | -.05 | .724 |
| Time of day | 0.02 | .16 | .282 |

*Note.* BMI = Body Mass Index

# CHAPTER 8

## Measuring BDNF in saliva using commercial ELISA: Results from a small pilot study

Published as:
Vrijen, C., Schenk, H. M., Hartman, C. A., & Oldehinkel, A. J. (2017). Measuring BDNF in saliva using commercial ELISA: Results from a small pilot study. *Psychiatry Research*, *254*, 340–346. https://doi.org/10.1016/j.psychres.2017.04.034



## ABSTRACT

Brain-derived neurotrophic factor (BDNF) is a protein often studied in psychiatric populations. Commercial ELISA kits have been validated for measuring BDNF in blood plasma and serum, but blood collection is an invasive method which cannot always be used. The aim of this pilot study was to explore the noninvasive alternative of measuring BDNF in saliva. Three different commercial ELISA kits were used to analyze parallel plasma and saliva samples from six healthy adults. In total 33 plasma and 33 saliva samples were analyzed according to manufacturers' standard protocols. BDNF was successfully measured in plasma in two of the three kits, of which the results correlated highly ($r_s$ = .88). BDNF could not be measured reliably in saliva. The results of this pilot study suggest that techniques of commercial BDNF kits may not be ready for noninvasive saliva measurements, which limits the sampling frequency and settings.

## INTRODUCTION

Brain-derived neurotrophic factor (BDNF) is a protein responsible for synaptic plasticity. Differences in BDNF levels have been associated with several psychiatric disorders, such as Major Depressive Disorder, bipolar disorder and schizophrenia (Autry & Monteggia, 2012; Munkholm, Vinberg, & Kessing, 2016; Polyakova et al., 2015; Soares et al., 2016; Toll & Mané, 2015; Wang et al., 2016). In these studies BDNF levels were determined in blood serum or plasma. Blood collection from patients or participants is an invasive method, and poses restrictions to the frequency of sampling and the environments in which sampling can take place. The background of this pilot study is a project in which anhedonia was investigated in young adults and interventions were developed to help them regain pleasure (Van Roekel et al., 2016). Because of its associations with depression, BDNF was considered a relevant protein. Since one of the interventions took place in a situation in which blood collection was not possible, the possibilities of the noninvasive and cost-efficient alternative of measuring BDNF in saliva were explored in a small pilot study. Although not to be interpreted as conclusive evidence, other researchers interested in measuring BDNF in saliva may benefit from the results of this pilot study. Noninvasive alternatives to blood BDNF would be relevant for, among others, research on BDNF fluctuations in psychiatric patient groups, and in the upcoming field of ecological momentary assessment (EMA), which involves frequent assessment of emotions, behaviors and social contexts over time in one's naturalistic environment (Shiffman et al., 2008), and has already been extended to salivary biomarkers cortisol and alpha-amylase (Bitsika, Sharpley, Andronicos, & Agnew, 2015; Booij, Bos, de Jonge, & Oldehinkel, 2016).

It has been demonstrated by means of immunoblotting that BDNF is present in human saliva (Mandel, Ozdener, & Utermohlen, 2009). However, whereas several commercial ELISA kits have been validated for measuring BDNF in plasma and serum, up until now none has been validated for measuring BDNF in saliva. In 2011 Mandel and colleagues assessed the possibilities of quantifying BDNF in saliva using commercial ELISA kits (Chemicon and Promega), but the BDNF levels rarely reached the minimum detection level of the kits (Mandel, Ozdener, & Utermohlen, 2011). Mandel and colleagues suggested that this was most likely due to matrix complexity, and decided not to use the commercial kits, but to develop a sandwich ELISA optimized for measuring BDNF in saliva themselves. In recent years, six studies reported to have successfully quantified BDNF in saliva by using commercial ELISA kits (de Souza et al., 2012, 2014; Ikai et al., 2014; Matsuki et al., 2014; Saruta, Fujino, To, & Tsukinoki, 2012; Tirassa et al., 2012). However, even reported BDNF levels for healthy young adults (Matsuki et al., 2014; Saruta et al., 2012; Tirassa et al., 2012) were very diverse, mean levels ranging from 9 pg/mL to 400 pg/mL. The studies used different ELISA kits and different procedures for collecting, storing and processing; and in most studies BDNF levels below the minimum detection threshold were interpreted as true results rather than discarded as unreliable results. See Table 1 for a comparison of the ELISA kits

**TABLE 1.** A Comparison Between Studies in Which Successful Quantification of BDNF in Saliva Using Commercial ELISA Kits was Reported

| Study | Sample | Saliva collection | ELISA kit & procedure / duplicate? | Centrifugation and freezing procedure |
|---|---|---|---|---|
| de Souza et al. (2012) | 26 adults with Burning Mouth Syndrom (BMS): 25 females ; aged 63.8±12 | (1) Mouth washed with filtered water (2) Saliva formed in mouth during 5 minutes of stimulation spitted in tube | R&D Duoset Not in duplicate | Centrifugation at 3000 rpm for 15 minutes Diluted and then stored at -20°C |
| de Souza et al. (2014) | 30 adults with BMS: 29 females; aged 62.1±12.7 32 controls: 31 females; aged 61.6±12.8 | (1) Mouth washed with filtered water (2) Saliva formed in mouth during 5 minutes of stimulation spitted in tube | R&D Duoset Not in duplicate | Centrifugation at 1509 x *g* for 15 minutes Diluted and then stored at -80°C |
| Ikai et al. (2014) | 50 schizophrenia patients. 25 assigned to yoga (16% males, aged 53.5±9.9) and 25 to control group (17% males, aged 48.2±12.3) | Unclear how saliva was collected | Millipore CYT306, according to manufacturer's instructions | Not reported |
| Tirassa et al. (2012) | 16 young healthy non-smoking female students in first week of their post-menstrual period (aged 21±0.5) | Samples collected at 8.00, 13.00 and 20.00 hrs, after fasting Passive drool in plastic tube | R&D Quantikine Human BDNF immunoassay according to manufacturer's instructions In duplicate | Centrifugation at 10,000 rpm for 10 min Stored at -70°C |
| Matsuki et al. (2014) | 40 healthy fertile female dental hygienist students: in follicular phase (n=24), aged 21.0±3.5 , and in luteal phase (n=16), aged 23.0±3.9 | Samples collected by salivettes between 1 and 4 pm fasted 5 min prior to sampling | Millipore CYT306 according to manufacturer's instructions In duplicate | Centrifuged at 626 × *g* for 15 minutes at 4 °C and stored at –80 °C. Upon thawing, samples were centrifuged once more |
| Saruta et al. (2012) | 50 healthy, non-medicated, non-smoking volunteers (26 males), aged 27±6.4 | Samples were collected by means of salivettes between 9 and 10 am fasted 2 h before sampling, no alcohol 24 h prior | Millipore CYT306 according to manufacturer's instructions In duplicate | Centrifuged at 2000 rpm for 15 min at 4 °C and stored at –80 °C. Upon thawing, samples were centrifuged once more |

| | Dilution / Acidification | Mean levels | Detection range | Results outside detection range |
|---|---|---|---|---|
| | 1:1 dilution acidification: phosphate-buffered saline solution containing protease inhibitors | Before therapy: ± 20 pg/mg<br>After therapy: ± 17 pg/mg | No minimum detection level reported; according to manufacturer: 23.40 pg/mL | Of the 52 reported results, 30 are very close to 0 and 43 are below 23.40<br>Results outside detection range not disqualified |
| | 1:1 dilution acidification: phosphate-buffered saline solution containing protease inhibitors | BMS: 5.27 pg/mg, range: 0-72.14<br>Controls: 10.73 pg/mg, range: 0-129.20 | Study reports a minimum detection level of 10 pg/mL; according to manufacturer: 23.40 pg/mL | Means and ranges suggest many results below minimum detection level<br>Results outside detection range not disqualified |
| | No dilution or acidification | Yoga group: 2.16 pg/mL±3.84 pg/mL<br>Control group: 4.12 pg/mL±8.26 pg/mL | No minimum detection level reported; according to manufacturer: 15 pg/mL | Reported means are far below the minimum detection level<br>Results outside detection range not disqualified |
| | No dilution or acidification | Baseline levels:<br>8.00: ±400 pg/mL<br>13.00: ±360 pg/mL<br>20.00: ±210 pg/mL<br>Post-light therapy:<br>8.00: ±175 pg/mL<br>13.00: ±180 pg/mL<br>20.00: ±170 pg/mL | No minimum detection level reported; According to manufacturer: 62.5 pg/mL | Unclear, since raw results for post-therapy were not reported |
| | No dilution or acidification | Mean follicular phase: 8.5±8.7 pg/mL<br>Mean luteal phase: 13.1±14.3 pg/mL | A minimum detection level of < 4 pg/mL is reported; according to manufacturer: 15 pg/mL | Many results below 15 pg/mL were reported, even below 4 pg/mL<br>Results outside detection range not disqualified |
| | No dilution or acidification | Males: 40.76±4.83 pg/mL<br>Females: 52.64±8.42 pg/mL | A minimum detection level of < 4 pg/mL is reported, but this does not match the manufacturer's information of 15 pg/mL | Unclear, since raw results or ranges were not reported |

and procedures used in the six previous studies, as well as of the mean BDNF levels that were reported. It thus remains unclear whether BDNF in saliva can be quantified reliably by using commercial ELISA kits.

In our study, we aimed to compare salivary BDNF with BDNF in plasma. As opposed to BDNF in serum, which has been reported not to be associated with salivary BDNF (Mandel et al., 2011; Tirassa et al., 2012), correlations between salivary BDNF and BDNF in plasma have not been studied in humans to date. Plasma BDNF levels reflect the momentary circulation of BDNF. Serum BDNF is determined by the sum of plasma BDNF and BDNF released from the platelets in serum (Montag, 2014; Polyakova et al., 2015), which have a life-span of about ten days and can accumulate or store BDNF (Dale, 1997). Therefore, serum BDNF is likely to reflect relatively stable or long-term BDNF levels, and plasma BDNF more momentary or short-term BDNF levels (Bus, 2014; Montag, 2014; Polyakova et al., 2015). Experiments in rats have shown that in situations of acute stress BDNF in the submandibular salivary gland may affect plasma BDNF levels (Tsukinoki et al., 2007).

Developing own methods for measuring BDNF in saliva, as Mandel and colleagues did, is not feasible for most researchers. Salivary BDNF would be accessible for many more researchers if commercial kits could be used for quantification. Because of the inconsistent previous results and ongoing technical developments during the past years, we conducted a pilot study in which three different commercial ELISA kits were used to measure BDNF in plasma and saliva. Main aims of the study were to explore the feasibility of quantifying BDNF in saliva with commercial ELISA kits when following the manufacturers' standard protocols, a comparison between the different ELISA kits, and a comparison between plasma and saliva levels.

## MATERIALS AND METHODS

### Subjects

From six participants, one male (age 43, non-smoking) and five females (age 63, non-smoking; age 28, smoking; age 51, non-smoking; age 37, non-smoking; and age 36, non-smoking), blood and saliva were collected. All participants were healthy and free of medication. From the first three participants (study 1), blood and saliva were collected in the morning on five consecutive days, resulting in a total of five blood and five saliva samples per participant. From the other three participants (study 2) blood and saliva were collected three times a day on two different days, with one resting day in between, resulting in a total of six blood and six saliva samples per participant. Participants were treated in accordance with the Declaration of Helsinki, and written consent was acquired from all participants.

## Fasting procedures

There is evidence that eating and drinking prior to sample collection increases BDNF levels in serum (Bus et al., 2011) and it is generally recommended to avoid eating, drinking and tooth brushing prior to collection of saliva samples (Salimetrics LLC Company, 2012; Wong, 2009). Therefore, it would be best for participants to adhere to a fasting protocol prior to saliva and plasma collection. However, fasting is not possible in all circumstances, particularly if multiple samples per day are collected, and recommended fasting periods differ. We therefore used feasible fasting procedures, adapted to the time of the day, and around noon we tested two different conditions, because around noon it may be most difficult to adhere to a fasting protocol. In study 1, blood and saliva samples were collected after overnight fasting (no eating, drinking or brushing teeth). In study 2, morning blood and saliva samples were collected after overnight fasting and the late afternoon samples after two hours of fasting. On the first day the sample around noon was collected after two hours of fasting, on the second day after 30 minutes of fasting.

## Blood collection

***Study 1***

Blood collection took place at five consecutive days, in the morning at home within half an hour after waking up, using standard sterile techniques. Blood was collected by venipuncture in 10 mL EDTA tubes (BD Biosciences, Franklin Lakes, NJ, USA). Samples were stored in a cooling bag (~8ºC) immediately after collection, transferred to the laboratory, centrifuged at 1650xg at 4°C for 10 minutes. Plasma samples were aliquoted and stored at -80°C within one hour after collection.

***Study 2***

Blood collection took place at the University Medical Center Groningen, three times a day on two different working days: in the morning around 8:30, around noon and in the afternoon around 16:15, using standard sterile techniques. Blood was collected by venipuncture in 4 mL EDTA tubes. Samples were transferred to the laboratory within five minutes, centrifuged at 1400xg at 4°C for 10 minutes. Samples were aliquoted and stored at -80°C within half an hour after collection.

## Saliva collection

***Study 1***

Saliva was collected in the morning right before blood collection. Following the recommendations by Mandel and colleagues we used a passive drooling method instead of cotton-based salivettes to collect saliva, because the use of salivettes may result in decreased levels of salivary BDNF (Mandel et al., 2011). Participants were instructed to tilt their head forward and pool saliva in their mouth. When a sufficient amount of saliva was pooled, participants were asked to drool in a cryovial (Salimetrics, Carlsbad, CA, USA). Samples were stored in a cooling bag (~8ºC) immediately after collection, transferred to the laboratory, centrifuged at 1650xg at 4°C for 10 minutes. Samples were aliquoted and stored at -80°C within one hour after collection.

***Study 2***

Saliva was collected right after blood collection, by passive drooling. Participants were instructed to sit down about 5 minutes after blood collection and drool into a 20 mL glass vial (PerkinElmer, Waltham, MA, USA) until a 1.5 mL marker was reached. Samples were transferred to the laboratory within five minutes, centrifuged at 1400xg at 4°C for 10 minutes. Samples were aliquoted and stored at -80°C within half an hour after collection.

## Procedures ELISA assays

All 33 plasma samples and 33 saliva samples were analyzed in three different commercial ELISA kits: (1) R&D DBD00 (R&D Systems, Minneapolis, MN, USA), (2) LSBio LS-F2402 (LSBio, Seattle WA, USA), and (3) Millipore Chemikine CYT306 (Merck Millipore, Billerica, MA, USA). For sensitivity and detection ranges see Table 2. All kits were validated for human plasma, serum and cell culture supernatants, but not for saliva. Accordingly, the ELISA protocols contained instructions for plasma analyses, but no specific instructions for saliva. Both the R&D kit and the Millipore kit had already been used in previous studies to measure BDNF in saliva (Ikai et al., 2014; Matsuki et al., 2014; Saruta et al., 2012; Tirassa et al., 2012), see also Table 1. The LSBio kit was selected based on the manufacturer's protocol which contained clear instructions and suggested good applicability in our lab.

All plasma analyses were performed according to standard protocol, with the exception that standards and samples were added to the wells in single measures. We did not analyze the samples in duplicate, because we were not interested in the specific individual results but merely in more general patterns of results that could inform us whether BDNF can be quantified in saliva while using commercial ELISA kits according to the manufacturers' protocols. All saliva analyses were performed according to the plasma and serum protocol, with the exceptions that saliva samples were not diluted, and that the saliva standard was based on the diluent that performed best in internal lab tests for construing saliva standards, i.e, a Tris-NaCl-γ-globuline

buffer (0.1 mM Tris/HCl, pH 8.0, 0.1 mM NaCl, 1g/L bovine γ-globulin (Sigma-Aldrich, St. Louis, MO, USA)). For each kit two sets of standards were used, one for plasma and one for saliva. It was unclear what BDNF concentrations to expect in plasma, therefore different dilutions were used for the plasma samples. For the R&D kit a dilution of 1:20 was used and for the LSBio and the Millipore kit a dilution of 1:2. Since low BDNF concentrations in saliva were expected, saliva samples were not diluted. Weighted 4 parameter logistic (4PL) nonlinear regression models were used for fitting the standard curves. Standard curves are presented together with the results in Fig. 1 (see Results).

**TABLE 2.** Sensitivity and Detection Ranges of the ELISA Kits Used in This Study, as Provided by the Manufacturers

| | **Kit 1** | **Kit 2** | **Kit 3** |
|---|---|---|---|
| Manufacturer | R&D | LSBio | Merck Millipore |
| Kit type | Sandwich ELISA | Sandwich ELISA | Sandwich ELISA |
| Serial number | DBD00 | LS-F2402 | CYT306 |
| Sensitivity | 20 pg/mL | 31.2 pg/mL | 15 pg/mL |
| Detection range | 62.5-4000 pg/mL | 31.2-2000 pg/mL | 15-1000 pg/mL |
| Intra-assay variation | <6.2% | <4.9% | 3.7% |
| Inter-assay variation | <11.3% | <7.9% | 8.5% |

## Statistical analyses

Whether the ELISA kits had functioned properly was evaluated by inspecting whether standard curves could be plotted and whether results were within the detection range of the kits. Next, mean BDNF concentrations in plasma and saliva were compared between the three kits, and Spearman correlation coefficients between the different ELISA kits were calculated. Finally, the correlation between BDNF levels in blood and saliva was evaluated.

## Open Science

Data and syntax have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/3kny5/.

## RESULTS

Standard curves could be plotted for BDNF in plasma and in saliva for the R&D and LSBio kits. The Millipore kit could be made operable for neither plasma nor saliva, that is, no standard curve could be plotted. A second attempt with a new Millipore kit yielded the same result. Therefore only the results of the R&D and LSBio kits are provided in this section.

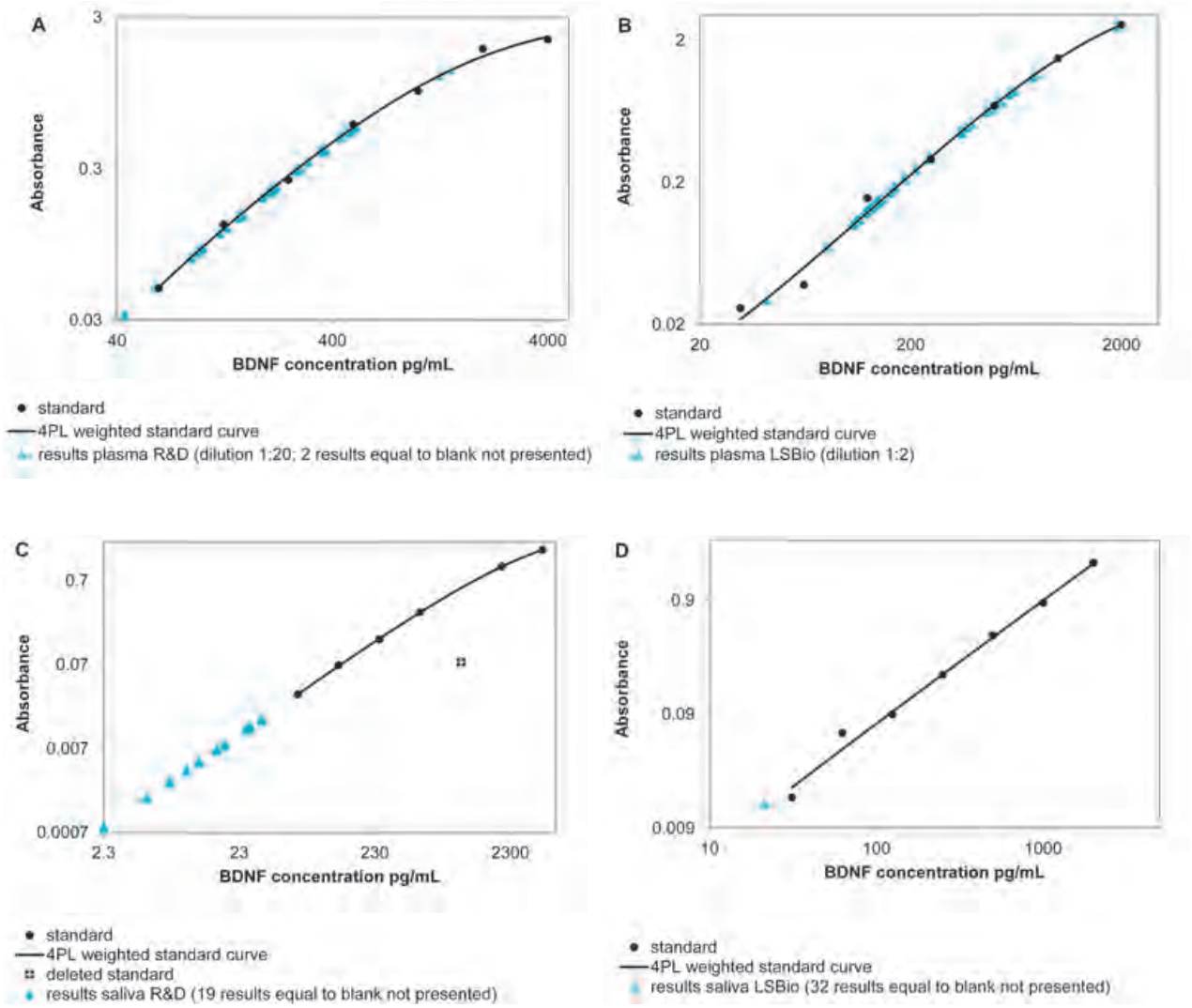



**FIGURE 1.** Standard curves and results of ELISA analyses of BDNF in plasma and saliva.
Results of the plasma and saliva analyses are presented for the R&D and LSBio ELISA kits separately. Plasma results are presented in panels A (R&D) and B (LSBio) and saliva results are presented in panels C (R&D) and D (LSBio). Weighted 4 parameter logistic (4PL) nonlinear regression models were used for fitting the standard curves.

**TABLE 3.** BDNF in Plasma and Saliva Analyzed in two Different ELISA kits

| | Plasma results within detection range* (% ) | Saliva results within detection range* (%) | Saliva results equal to blank | Saliva results above blank but below lowest standard | Mean pg/mL BDNF in plasma *(SD)* | Mean pg/mL BDNF in saliva *(SD)* |
|---|---|---|---|---|---|---|
| R&D kit | 30 (91 %) | 0 (0 %) | 19 (58 %) | 14 (42 %) | 7027 (6065) | - |
| LSBio kit | 33 (100 %) | 0 (0 %) | 32 (97 %) | 1 (3 %) | 795 (876) | - |

Each of the 33 plasma and 33 saliva samples was analyzed in both ELISA kits.
Mean pg/mL BDNF was calculated only for results within the detection range.
*The detection range is the range of the calibrated standard curve, with as its lower limit the first standard and as its upper limit the highest standard.

While almost all plasma results lay within the detection range of the R&D and LSBio kit, no saliva sample fell within the range of the calibrated standard curve and the majority did not even exceed the level of the blank (Table 3 and Fig. 1). BDNF in saliva could not be quantified by either kit. The estimated plasma BDNF levels of the two ELISA kits correlated highly ($r_s = .88$, $p < .001$), the mean concentrations differed by almost a factor ten. Because of the lack of valid saliva results it was not possible to calculate the correlation between BDNF in plasma and in saliva.

## DISCUSSION

By far most of the saliva samples did not exceed blank level, which means that no BDNF concentrations were measured. The saliva samples that did exceed blank level still did not reach the lowest concentration of the calibrated standard curve and can therefore not be considered reliable. Our data thus suggest that it may be impossible to use commercial ELISA kits according to manufacturers' protocols to reliably quantify BDNF in saliva to date. These results are in agreement with Mandel et al. (2011), but not with de Souza et al. (2012, 2014), Ikai et al. (2014), Tirassa et al. (2012), Saruta et al. (2012) and Matsuki et al. (2014), who reported to have successfully quantified BDNF in saliva using commercial ELISA kits. Differences in interpretation of results below the minimum detection level seem to be at least a partial explanation. Whereas de Souza et al. (2012), Ikai et al. (2014) and Matsuki et al. (2014) extrapolated BDNF levels beyond the minimum detection level of the calibrated standard curve, in Mandel et al. (2011) and in our study, BDNF levels beyond the minimum detection level were discarded as unreliable. Saruta et al. (2012) and Matsuki et al. (2014) further reported a minimum detection level of < 4 pg/mL, while the manufacturer's protocol mentions a minimum detection level of 15 pg/mL. Neither Tirassa et al. (2012) nor Saruta et al. (2012) reported on how results below the minimum detection level were handled. It should be noted that the reported mean BDNF levels for healthy young adults varied considerably, even though Saruta et al. (2012) and Matsuki et al. (2014) used the

same ELISA kits. Altogether this suggests that commercial BDNF kits may not yet be sufficiently sensitive to measure BDNF reliably in saliva. This is confirmed by the fact that at this moment no ELISA manufacturers have claimed to have validated their kits for measuring BDNF in saliva.

Results for plasma were, as expected, more promising. Almost all results were within the detection ranges of the kits and the high correlation between the two ELISA kits indicates that a similar construct is measured by both. Absolute BDNF levels showed large differences between kits and cannot be compared, because companies use their own BDNF standards and often do not refer to a more objective standard. From the three ELISA kits we assessed, only the R&D protocol contains a statement about the purity of their standard compared to the World Health Organization (WHO) human BDNF standard and how to convert results. To allow for more meaningful interpretations of absolute BDNF levels, it is essential that more manufacturers provide this type of information.

The goal of our study was to assess whether BDNF in saliva can be measured by following standard protocols of commercial kits, without adapting or optimizing the protocol. This is at the same time a strength and a limitation of this study. It is a strength since it mirrors the approach that would often be followed in psychiatric research, yet it leaves unanswered whether we would have succeeded to measure BDNF in saliva if we had experimented with longer incubation times, the amount of sample material, or made other adaptations to manufacturers' protocol. Furthermore, standards and samples were not analyzed in duplo and therefore, although the analysts who conducted the analyses were well-trained, the possibility of pipetting mistakes cannot be excluded. Another limitation is the small number of subjects participating in our study, i.e., five females and one male. That being said, because of our clear patterns of results for saliva we do not consider it to be likely that our conclusions have been fundamentally influenced by pipetting mistakes and do not expect that including more participants or making minor adaptations to the manufacturers' protocols would change results radically. A final limitation is that the Millipore kit could not be made operable for plasma or saliva. And although this too exemplifies the type of problems encountered when doing this kind of research, it is currently unknown whether this was due to mistakes inside our laboratory or malfunctioning kits.

To conclude, the results of this pilot study tentatively suggest that techniques of commercial BDNF kits are not ready for noninvasive saliva measurements to date, which limits the conditions in which sampling can take place. Because of the benefits of noninvasive and cost effective methods for measuring BDNF, our hopes are that it will be possible to measure BDNF in saliva by using commercial kits in the near future. An alternative approach, which is currently being developed, is to collect blood in less invasive ways (e.g. Hemolink, Tasso Inc, Seattle, WA, USA), which allows the use of commercial ELISA kits that have already been validated for measuring BDNF in plasma and serum. If successful, both approaches are expected to be of great significance to studies in psychiatry.

# CHAPTER 9

General discussion

# INTRODUCTION

The main aim of this dissertation was to study low positive bias as an underlying mechanism of depression. This bias was investigated with different instruments and from different perspectives. Facial emotion identification tasks and a reward task were used to assess a relative bias toward positive stimuli (happy faces, reward and non-punishment) compared to negative stimuli (negative facial emotions, non-reward and punishment). Relatively fast identification of happy facial emotions compared to negative facial expressions (high happy bias) and relatively great attention to reward (high reward responsiveness) both reflect a bias toward positive information. I studied contemporaneous (Chapter 2) and prospective (Chapter 3, 4) associations between a low positive bias and depressive symptoms. The symptom-specificity (Chapter 3) and disorder-specificity of these associations were also investigated (Chapters 2-4). In Chapter 5 I explored whether a high and a low positive bias were associated with different affect dynamics in daily life. From a treatment perspective, I studied the effects of tailored behavioral activation, both on its own and in combination with an extreme free-fall experience, as a means to increase pleasure and boost the motivational system (Chapter 6). I investigated not only whether the two interventions increased pleasure and decreased depressive symptoms, but also whether they modified the positive bias (Chapter 6, *addendum*). In Chapters 7 and 8 two studies are presented on biological markers for depression. Biological responses to a tandem skydive were investigated by means of alpha-amylase (Chapter 7), and Chapter 8 describes a pilot study designed to test whether currently available commercial ELISA kits are sufficiently sensitive to measure brain-derived neurotrophic factor (BDNF) in saliva.

# SUMMARY OF THE MAIN FINDINGS

## Chapter 2

- No significant associations were found between sensitivity to facial emotions during facial emotion identification and symptoms of depression and anxiety;
- Young adults with antisocial behaviors were significantly less sensitive to happy facial emotions than peers without antisocial behaviors, young adults with ADHD symptoms were less sensitive to angry emotions, and those with avoidance problems were less sensitive to both angry and happy emotions;
- The significant effects could not be fully explained by co-occurring psychiatric problems;
- Inspection of the overall pattern of effect sizes regardless of statistical significance reveals generic patterns in that for all psychiatric problem domains the effect sizes for happy and angry facial emotions were larger than the effect sizes for sad and fearful emotions.

## Chapter 3

- A relatively slow speed of identifying happy facial emotions compared to sad facial emotions at age 11 tentatively predicted onset of depressive disorder within 8 years of follow-up;
- Speed of identifying facial emotions predicted symptoms of anhedonia, but not symptoms of sadness (symptom-specificity);
- Results were not driven by co-occurring anxiety disorders (disorder-specificity).

## Chapter 4

- Difficulties in shifting attention from expected non-reward to expected reward and from expected punishment to expected non-punishment at age 16 predicted depression during follow-up;
- This was only found at an automatic level of information processing;
- Levels of initial engagement toward expected reward, non-reward, punishment, and non-punishment did not predict depression;
- Results were not driven by co-occurring psychiatric disorders (disorder-specificity).

## Chapter 5

- Young adults with a high bias toward happy facial emotions showed different daily life affect dynamics compared to their peers with a low happy bias;
- Daily life reward experiences and positive affect during one six hour time interval were stronger predictors of affect and related experiences during the next six hour time interval for the high than for the low happy bias group;
- More specifically, individuals with a high happy bias more strongly sustained their positive experiences and positive affect over time.

## Chapter 6

- Anhedonic young adults who received lifestyle advice tailored to their own subject-specific associations between lifestyle activities and pleasure (based on 3 momentary assessments per day during 30 days) showed a higher increase in momentary positive affect (PA) and pleasure ratings in the month following the intervention than those who did not receive advice;
- Group differences were not found when comparing monthly measures of depression, pleasure, and positive bias (i.e., happy bias and reward responsiveness);
- The tandem skydive did not have any effects on top of the effects of the lifestyle advice.

### Chapter 7

- Alpha-amylase reactivity and recovery patterns were highly similar across anhedonic young adults performing a tandem skydive, although mean alpha-amylase levels varied greatly;
- Alpha-amylase patterns in anhedonic individuals were highly similar to patterns previously found in healthy individuals performing a skydive;
- None of the trait and state measures (depressive and anxiety problems, positive affect, and self-esteem) predicted alpha-amylase reactivity to or recovery from the skydive;
- Individuals who showed high reactivity and low recovery experienced decreases in self-esteem after the skydive. Considering the number of statistical tests, this may well be a false positive finding.

### Chapter 8

- BDNF was successfully measured in plasma in two of the three ELISA kits, of which the results correlated highly (rs =.88);
- BDNF could not be measured reliably in saliva.

## REFLECTION ON THE MAIN FINDINGS

### Low positive bias as a predictor of depression

Below I discuss the results and implications of the studies on low positive bias (low happy bias and low reward bias) as a predictor of depression (Chapters 2-5). Low happy bias and low reward bias were assessed with laboratory tasks and reflect implicit information processing biases of which individuals themselves are probably unaware.

***Contemporaneous associations***

Many previous studies found associations between how fast an individual is able to identify facial emotions and depression. Surprisingly, this was not found in the study described in Chapter 2 of this dissertation. Two possible explanations involve the way sensitivity to happy facial emotions and depression were operationalized. In Chapter 2 the four assessed facial emotions were analyzed separately rather than by means of an aggregated bias score reflecting the relative sensitivity to happy *compared to* negative emotions. This difference may be quite important; perhaps in the context of depression it does not matter much whether someone is slower in identifying happy facial emotions than other people, but depression may be characterized by being slow in identifying happy facial emotions compared to how fast one identifies negative facial emotions. This relative within-subject bias could result in the processing of disproportionally more negative than positive emotional information, which may ultimately lead to negative

feelings and thoughts about oneself and other people, and *vice versa*. From Chapter 3 onward, happy bias was operationalized as within-subject happy bias, that is, how fast an individual is in identifying happy facial emotions compared to negative facial emotions. Note that other studies have found the absolute (rather than only the relative) speed of identifying facial emotions to be associated with depression and anxiety, therefore this first explanation is not sufficient to explain the lack of significant associations reported in Chapter 2. Another explanation may be related to the operationalization of depression. In the study described in Chapter 2, depression was assessed with the Adult Self-Report (ASR). This questionnaire contains separate scales for depressive problems and avoidance problems and the depressive scale contains no avoidance items. Associations were found between sensitivity to happiness and anger in identifying facial emotions and avoidance problems, which are part and parcel of depressive withdrawal behavior, at least in clinically depressed populations. Most other studies were performed in clinical samples and their results may have been driven by increased avoidance behavior and/or decreased approach behavior characteristic to a clinical depression. This may partly explain the lack of significant associations with depression in Chapter 2, although this would need to be confirmed in independent samples. From Chapter 3 onward, depression was assessed either by means of a standardized clinical interview or by means of a self-report questionnaire that closely matches the criteria mentioned in the DSM. This may increase the comparability between the studies presented in this dissertation and previous studies.

***Prospective associations***

Previous studies only provided preliminary evidence that low positive bias predicts future onset of depression and may be a vulnerability marker to develop depression. The studies described in Chapter 3 and 4 of this dissertation were designed to produce evidence to verify or falsify this only incompletely supported hypothesis. I found for two different laboratory tasks and two different developmental stages that low positive bias predicted onset of depression (see Figure 1). Low happy bias at age 11 predicted onset of depression between age 11 and age 19 and low reward bias at age 16 predicted onset of depression between age 16 and age 25. The result for happy bias reflects that adolescents who at age 11 were slower in identifying happy facial emotions compared to sad facial emotions than their peers became depressed more often between age 11 and age 19. The results for reward bias suggests that adolescents who showed difficulties in shifting attention from expected non-reward to expected reward and from expected punishment to expected non-punishment at age 16 became depressed more often than their peers between age 16 and age 25. The reward bias effect was only found for automatic levels of information processing and for disengagement, the implications of which will be discussed in detail later. In this dissertation the prospective associations between positive bias and depression were investigated for the first time in large samples with sufficient power. The findings reported in this dissertation suggest that low positive bias may be a vulnerability marker for depression and a possible target for early interventions to prevent depression.

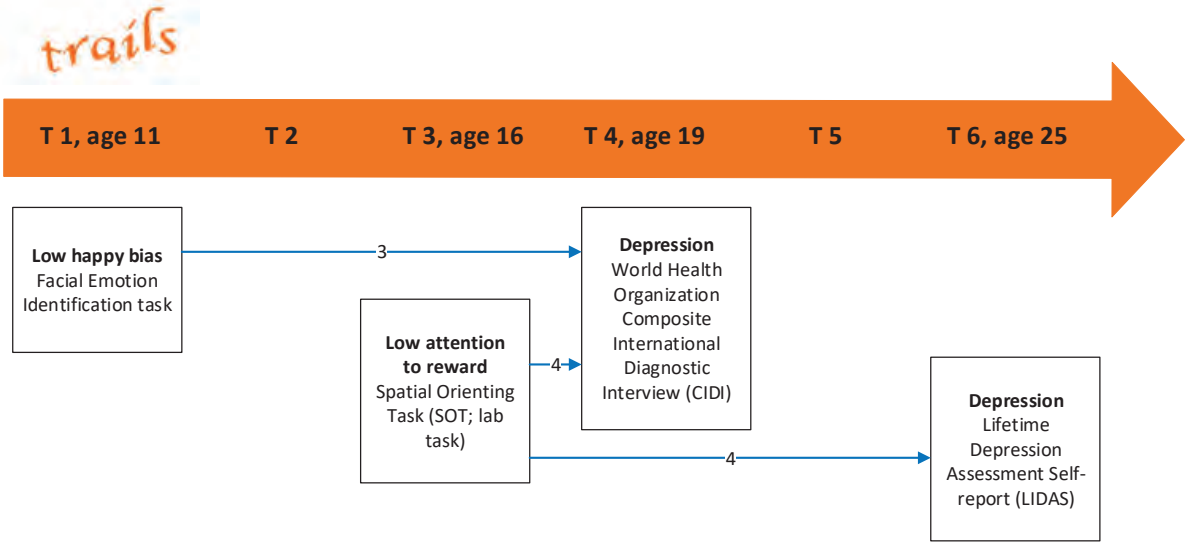



**FIGURE 1.** The blue arrows reflect the prospective associations found in Chapter 3 and 4; the numbers indicate the corresponding chapters. Low happy bias at age 11 predicted onset of depression between age 11 and age 19, and low reward bias at age 16 predicted onset of depression between age 16 and age 25.

### *Symptom- and domain-specificity*

I investigated if results differed for the two core symptoms of depression (symptom-specificity) in a study described in Chapter 3, and found that the association between a low positive bias and depression was driven specifically by the core symptom anhedonia. In Chapter 4 it was not feasible to analyze the two core symptoms separately, but there was already evidence from previous studies that reward responsiveness is primarily related to symptoms of anhedonia (Chase et al., 2010; Luking et al., 2015; Pizzagalli et al., 2008).

The specificity of the results for depression relative to other psychopathology domains or psychiatric disorders (domain-specificity) was investigated in Chapters 2, 3, and 4. Domain-specificity in Chapter 2 was difficult to interpret because no significant findings were found for depression in the first place, although, as outlined above, avoidance may be seen as a symptom of depression. Furthermore, the effects that were found (for antisocial behavior, ADHD symptoms, and avoidance) could not be fully explained by co-occurring psychiatric problems in other domains, suggesting domain-specificity, but they were partly explained by co-occurring psychiatric problems (i.e., adjusting for other psychiatric problems resulted in lower estimates). Moreover, when comparing the unadjusted estimates regardless of significance, similar patterns were found for all five psychiatric problem domains, that is, lower sensitivity for happy and angry facial emotions and no differences in sensitivity for sad and fearful emotions. This suggests generic effects as well. Approach and avoidance mechanisms may play a role in all domains, but the reasons for approaching or avoiding certain stimuli may still be domain-specific. No

clear conclusions can be drawn about domain-specificity based on Chapter 2. Chapters 3 and 4 provide a clearer picture of the domain-specificity of the prospective effects of a low positive bias. The reported associations were domain-specific in that associations were driven by depression and not by anxiety (tested in Chapters 3 and 4) or other psychiatric disorders (only tested in Chapter 4). The approach in Chapters 3 and 4 differed from the approach in Chapter 2; in Chapter 2 other psychiatric disorders were not only used to adjust for other disorders, but also as outcome variables. An additional outcome variable in Chapter 2 was the total score on psychiatric problems, which reflects comorbidity or severity. In Chapters 3 and 4 adjusting for other psychiatric disorders did not result in lower estimates. This suggests that low positive bias predicts onset of depression and not onset of other psychiatric disorders, at least not the same aspect of low positive bias that predicts depression. It is not possible to draw conclusions based on Chapter 3 and 4 about whether a low positive bias may predict total psychiatric problem severity.

***Why would a low positive bias make someone more vulnerable to depression?***

In Chapter 5 I described differences in daily life affect dynamics between individuals with a high and a low happy bias during a laboratory task. A high happy bias as compared to a low happy bias was associated with more carry-over of positive affect and positive experiences over time in daily life (i.e., from one 6 hour time interval to the next), and more carry-over from one type of positive affect or positive experience to another over time. This suggests that, compared to individuals with a high happy bias, individuals with a low happy bias show a lower general bias toward positive experiences and positive affect in their daily lives, and have more difficulties with sustaining positive affect over time and generalizing it to other positive affect related components. This is an important finding, because the affect dynamics of the low happy bias group, that is, the inability to sustain positive affect and generate positive affect from pleasant experiences, have been associated with depression in previous studies (Geschwind et al., 2010; Heller et al., 2009; Höhn et al., 2013; Horner et al., 2014; McMakin et al., 2009; Wichers et al., 2010), and with symptoms of anhedonia in the No Fun No Glory study (Heininga et al., 2017). Furthermore, daily life momentary positive affect has been found to predict life satisfaction and a higher ability to adapt to changing environments (Cohn et al., 2009), and has been suggested to facilitate building valuable cognitive and social resources essential to well-being (Cohn et al., 2009; Fredrickson, 1998). Savoring of positive affect was found to be associated with more life satisfaction and happiness, and with lower levels of affective problems (Bryant, 2003). If positive experiences and positive affect can be sustained longer and spread to other positive components, their beneficial influence of promoting flexibility to changing environments and building cognitive and social resources may be prolonged too.

There is evidence that positive affect facilitates recovery from negative experiences (Fredrickson & Levenson, 1998) and that resilient individuals use positive affect to downregulate negative affect (Tugade & Fredrickson, 2004) whereas vulnerable individuals use this emotion

regulation strategy less. This seems to be in accordance with the findings presented in Chapter 5 that joy and pleasant experiences dampen negative variables (sadness, irritation, worrying, unpleasant experiences) in the high but not in the low happy bias network. However, no significant differences between the high and low happy bias groups were found. It is possible that the permutation test was not significant because of small effect-sizes in combination with large individual variation in associations from joy and pleasant experiences during one 6 hour time interval to negative variables during the next 6 hour time interval, but this is only speculation. The regulation of negative affect by means of positive affect may also essentially occur in a shorter time frame, for example, 2 hours. In this case the current method only picked up what was still left of the initial effect several hours later.

The findings presented in Chapter 5 may suggest that teaching individuals to sustain their positive affect for a longer period of time could be a useful intervention to alleviate or prevent depressive symptoms. However, it is important to note that there is not only evidence that sustaining positive affect is a good thing. It has been found that strong autocorrelations (high inertia), of negative as well as positive emotions, predict depression (Houben et al., 2015; Kuppens et al., 2012). Strong autocorrelation of a particular emotion may signal psychological rigidity, as it implies that the current level of that particular emotion is highly predictive of its own future level, regardless of changes in other emotions or factors included in the model. It has been argued that psychological flexibility is highly important for optimal functioning in many situations, and that psychological rigidity is associated with depression as well as with other forms of psychopathology (Houben et al., 2015; Kashdan & Rottenberg, 2010; Kuppens et al., 2012). Thus, on the one hand, sustaining positive affect over time as shown by a high autocorrelation may be good for well-being and mental health, but on the other hand it may also be an indication of psychological inflexibility, which is bad for mental health. More research is needed on the specific conditions in which high autocorrelation is a good sign and when it is a bad sign. Two conditions that may be worth exploring are the time scale and level to which the high autocorrelation applies. On a short time scale it may, for example, be very important to adapt to the environment flexibly, but on a larger time scale the ability to sustain positive emotions over time may be beneficial for mental health and well-being. A high autocorrelation should also be interpreted differently depending on the level of affect to which the high autocorrelation applies. A high autocorrelation of high levels of positive affect, for example, implies that high levels of positive affect are sustained and a high autocorrelation of low levels of positive affect implies that someone remains trapped in low levels of positive affect.

***Two different operationalizations of positive bias, two different mechanisms?***

I presented low happy bias and low reward bias as two different operationalizations of low positive bias. Although they were used to measure the same concept, evidently, there are also differences between both constructs. Happy bias is based on the identification of other people's facial emotions and may have a more social component, and reward bias was measured in a

socially neutral way with gain of points and a possible prize as rewards and with loss of points and the possibility of having to repeat the task as punishments. Another difference is that the reward task assesses attention and the facial emotion identification tasks measures perception. In the emotion identification task attention and perception cannot be disentangled. It is, for example, impossible to know whether slow perception of a facial emotion is due to reduced attention to that facial emotion. Comparing the two types of tasks, the reward task seems to allow for more specific interpretations than the facial emotion identification tasks. This is not surprising, considering the different traditions in which the tasks were developed. The facial emotion identification tasks have been developed in the neuropsychological tradition where the aim was to establish impairments in emotion recognition, regardless of the specific cognitive processes involved. This in contrast to the reward bias task that has its roots in experimental psychology. It contains different subtasks that can be analyzed separately, for example reward bias and punishment bias, and automatic and more voluntary processing, and in this respect provides more specific information than the facial emotion identification task.

The studies presented in this dissertation, together with previous studies (Bourke et al., 2010; Disner et al., 2011; Luking et al., 2016; McCabe & Gotlib, 1995; Pizzagalli et al., 2008; Roiser et al., 2012; Telzer et al., 2014), suggest that a low positive bias in a wide range of domains and contexts (emotion identification, reward processing, perception, attention, memory, learning, social contexts, non-social contexts), and from different perspectives (maintenance, vulnerability, immediate responses, expectancies, behavior, brain) may be associated with depression. However, it is still possible that one single mechanism is underlying all findings, because for many tasks it is impossible to identify a very specific working mechanism, as illustrated with the example of the facial emotion identification task in which attention and perception cannot be disentangled. Further, although many tasks aim to investigate specific cognitive mechanisms, certain processes may be shared among tasks, making it possible that these shared processed drive the results. One reason why it is often impossible to disentangle different cognitive mechanisms is that the number of different conditions a task can contain is limited. For example, because the task described in Chapter 4 does not contain a neutral condition, switching away from expected non-reward is, by definition, also switching toward expected reward. For the studies described in the present dissertation, it is possible that difficulties with automatic disengagement from expected non-reward to expected reward (Chapter 4) are driving the associations reported in Chapters 2, 3 and 5. It is equally well possible that the facial emotion identification tasks used in these chapters tap into partly different mechanisms, for example, more socially oriented reward processing, or indicate altered processing on a more voluntary level.

***Automatic and voluntary information processing biases***

It has been suggested that depressed individuals are characterized particularly by voluntary, higher-order top-down information processing biases, because previous studies on depression that used behavioral reaction time tasks found results on voluntary but not on automatic levels of information processing (Mathews & MacLeod, 2005). In the studies presented in Chapters 2 and 3 of this dissertation it was not possible to unravel more automatic and more voluntary responses, because in the facial emotion identification task stimuli were shown until participants responded, with a maximum duration of 10,000 ms (Chapter 2) or 8,000 ms (Chapter 3). Mean reaction times varied from 878 ms (happy expressions) to 1,210 ms (sad expressions) in Chapter 3 and from 4,113 ms (happy expressions) to 6,542 ms (sad expressions) in Chapter 2, and these durations are sufficiently large to allow voluntary information processing. In the study presented in Chapter 4 a task was used in which automatic and more voluntary responses were assessed separately. This task can be used to shed light on whether automatic and more voluntary positive biases play a role in the development of depression.

The findings reported in Chapter 4 suggest that difficulties with disengaging from expected non-reward and switching to expected reward on an automatic level of information processing at age 16 predict future onset of depression. No effects are found for more voluntary levels of information processing. Does this contradict earlier studies in which only voluntary biases were found for depressed individuals? It is important to emphasize that the prospective approach of investigating information processing biases in individuals who were healthy at the time of assessment is novel and an important difference from the existing literature. Although voluntary processes may indeed largely explain information processing biases for individuals who already suffer from a depression, it may well be the automatic processes that constitute a vulnerability to depression in not-yet-depressed individuals. Adolescents with an automatic tendency to remain focused on negative situations and a diminished capability to redirect attention to potentially rewarding situations may process more negative information, which may gradually trigger the more voluntary, top-down negative biases that characterize depressed patients. That I found no voluntary biases for vulnerability is thus not in conflict with other studies in which voluntary biases were reported for fully developed depression. But would it not be expected that automatic biases that mark vulnerability for depression are still present in individuals with a full-blown depression? A possible explanation as to why no automatic attentional biases have been found in depressed patients is that they in general perform slower on reaction time tasks, which may mask existing automatic biases. Automatic versus voluntary processing tends to be operationalized in the same way for everyone without taking into account individual differences in processing time. Similarly slow reaction times in depressed patients and controls may therefore reflect more automatic processing in depressed patients and more voluntary processing in healthy controls. Although for behavioral reaction time tasks no associations were found between depression and automatic biases in previous studies, such associations were

reported in fMRI studies (Stuhrmann et al., 2013; Suslow et al., 2010; Victor et al., 2010), which all the more suggests that depressed individuals may indeed show automatic biases, and that some methods may be better suitable for investigating these than others.

***Altered engagement or altered disengagement?***

Because previous studies using behavioral reaction time tasks found results on voluntary but not on automatic levels of information processing (Mathews & MacLeod, 2005), it has been proposed that depression is primarily associated with difficulties in disengaging from negative stimuli and less with an increased engagement toward negative stimuli (Gotlib & Joormann, 2010; Roiser et al., 2012). Apparently, it was considered plausible that disengagement required some sort of voluntary control. This explanation may no longer hold after the aforementioned evidence for automatic biases in fMRI studies (Stuhrmann et al., 2013; Suslow et al., 2010; Victor et al., 2010). In this dissertation the reward task (Chapter 4) allowed for investigating in a more direct way if engagement toward expected reward or disengagement from expected reward predicted depression. The design of the task used in Chapter 4 was based on research by Posner and Peterson, indicating that attention is not a single mechanism, but refers to multiple systems and mechanisms (Derryberry & Reed, 2002; Posner & Petersen, 1990). The posterior attentional system which is relatively reactive, or automatic, consists of three subsystems that together reorient attention from one location to another. Attention is first disengaged from one location, then moved to a new location, and finally needs to be engaged to the new location. The anterior attentional system is responsible for more voluntary, reflective, attentional functions (Derryberry & Reed, 2002; Posner & Petersen, 1990). Posner's model all the more suggests that disengagement can take place on an automatic level of processing and that finding voluntary but not automatic biases in itself does not justify the claim that depression is associated with a decreased disengagement from negative stimuli rather than with an increased engagement toward them.

The findings described in Chapter 4 provide evidence that difficulties in automatic disengagement from non-reward, that is, difficulties in switching from expected non-reward to expected reward on an automatic level of information processing, predict future onset of depression. However, because the task used in Chapter 4 does not contain a neutral condition, switching away from expected non-reward is by definition also switching toward expected reward. With this specific task it cannot be investigated further whether specific problems with switching away from non-reward predict future onset of depression, or problems with switching toward reward, or if they both equally predict onset of depression. Primary engagement to reward or non-reward does not predict onset of depression. Thus, the problem is either with disengagement, or with, what I would call, secondary engagement, that is, engagement to reward after disengaging from non-reward, or with both. Regardless of whether vulnerability for depression is reflected in difficulties with disengaging from non-reward or difficulties with

secondary engagement toward reward, it is worth investigating whether enhancing at-risk adolescents' responsiveness to cues for potential rewards in situations in which they are focused on negative experiences may prevent onset of depression.

***Different time scale, different finding?***

It is interesting to consider the interplay between positive and negative affect and experiences by comparing the findings presented in Chapters 4 and 5. In Chapter 4, I reported that difficulties in switching from expected non-reward to expected reward predict future onset of depression, but that it has remained unclear whether problems with switching away from non-reward predict future onset of depression, or problems with switching toward reward, or if they both equally predict onset of depression. If a low positive bias is primarily characterized by difficulties in letting go of expected non-reward or punishment, it may be expected that increased negative affect at one time point would predict increased negative affect at the next time point, but this was not found in Chapter 5. If the problem lies more in letting go of negative affect specifically in favor of positive affect it may be expected that increased negative affect at one time point would predict decreased positive affect at the next time point, but this was also not found in Chapter 5. It is important to keep in mind that the effects described in Chapter 4 occurred on a shorter time frame than those described in Chapter 5. The Chapter 4 task is used to study attention in a time frame of milliseconds, and the dynamics on a time scale of milliseconds are different from the dynamics studied in Chapter 5, in which affect was assessed each 6 hours and associations thus represented effects that lasted on average 6 hours. The most robust finding reported in Chapter 5 that positive affect is sustained less over time by individuals with a low positive bias than by those with a high positive bias also cannot be directly translated to the time scale of Chapter 4. The discussion about the advantages and disadvantages of inertia versus psychological flexibility for mental health, and particularly the role of time scale, may again be relevant here: which mechanisms play a role may depend on the time scale. On the level of seconds and milliseconds (Chapter 4) it is important to adapt to the environment flexibly, but on a larger time scale (Chapter 5) it may be beneficial to sustain positive emotions over time in order to benefit from them the most in terms of mental health and general well-being. Apart from time scale, other methodological differences discussed elsewhere in this dissertation may evidently account for the different findings reported in Chapters 4 and 5 as well, for example, automatic (Chapter 4) versus voluntary processing (Chapter 5) and implicit (Chapter 4 and the task of Chapter 5) versus explicit (momentary assessments of Chapter 5) assessments.



## Tailored behavioral activation as a treatment for anhedonia

Up to this point the discussion has largely centered around a specific type of positive bias, that is, relative and implicit information processing bias. The positive bias discussed so far is relative because it involves comparing responses to positive information to responses to negative

information (with the exception of Chapter 2). It can be characterized as implicit because both operationalizations of positive bias, happy bias and reward bias, are based on standardized laboratory tasks during which participants are not explicitly asked about their bias but this information was derived from reaction times in response to the tasks. Participants themselves are likely unware to which extend they have a positive bias. Other types of positive biases are absolute biases, which do not depend on a comparison with responses to negative information, and explicit biases, which can for example be assessed with self-report questionnaires. One of the two core symptoms of depression, anhedonia, is characterized by decreased levels of interest and pleasure and may be considered an example of an absolute, explicit low positive bias. The present section is about anhedonia and will be used to discuss the results of the studies described in Chapters 6, 7, and 8, in which I investigated whether behavioral activation may alleviate symptoms of anhedonia, and modify low happy bias (i.e., an implicit and relative low positive bias) and low reward responsiveness (i.e., an explicit and absolute low positive bias). The behavioral activation interventions consist of: (1) tailored lifestyle advice based on an individual's own associations between pleasure and lifestyle factors; or (2) tailored lifestyle advice combined with a tandem skydive.

***Effects of the tailored lifestyle advice on anhedonia***

The findings reported in Chapter 6 suggest that it can be feasible and effective to provide personalized lifestyle advice to young adults suffering from anhedonia to increase their pleasure. This was investigated in a randomized controlled trial in which anhedonic young adults were randomly assigned to a control group receiving no intervention (N = 22), a group who received lifestyle advice (N = 22), or a group who received both lifestyle advice and a tandem skydive (N = 25). Participants reported lifestyle factors and pleasure three times per day during 30 days. The research team of the No Fun No Glory project provided lifestyle advice to participants based on their own associations between pleasure and lifestyle activities during the first period of 30 days, which participants were asked to carry out during a second period of 30 days. We found evidence that this type of tailored behavioral activation can increase daily life levels of pleasure and positive affect. Symptoms of anhedonia have been proven difficult to target and it is promising for clinical practice that tailored behavioral activation can help anhedonic young adults to increase their pleasure. We also showed that it is feasible to assess anhedonic young adults three times per day for more than three months and achieve very good compliance rates. Although there is no way of knowing whether involving fully trained therapists may have resulted in larger effects, the No Fun No Glory study results suggest that researchers without clinical practice, like me and the other researchers working on the project, can provide this type of tailored behavioral activation. In combination with the fast and ongoing technical and statistical advancements, this provides ample opportunities for future research and the development of treatment methods. Much more and diverse data can be used to provide subject-specific advice

than was possible in the early days of behavioral activation when paper daily diaries were used, and there are many promising initiatives to automatize statistical analyses to provide individuals with mental problems with insights in their own daily life patterns (van der Krieke et al., 2017).

I do need to add a few caveats to this optimistic message. First of all, it is important not to overestimate the effect of the tailored lifestyle advice, because effect-sizes were only small to moderate. Furthermore, results were mixed with respect to different outcome measures. Whereas significant increases in pleasure and positive affect (PA) were found when comparing momentary measurements, no significant increases were found when comparing the monthly pleasure measures which were reported over a period of two weeks. Also no effects were found on momentary negative affect (NA) or monthly depressive symptoms reported over the past 2 weeks. Possible explanations for the differences in effects on momentary and monthly PA and pleasure measures are, first, statistical power is less when comparing 2 monthly measures for 2 groups than when comparing 2*90 measures for 2 groups. Because the smallest group contained only 22 individuals, analyses on the monthly measures this could have had low power. Another possibility involves recall bias. Anecdotal evidence from the participants made clear that several of them had difficulties with estimating their average pleasure over a longer period of time, while they were able to estimate it over the past 6 hours. They seemed to remember lower levels of pleasure when they had to report pleasure over a longer period of time, possibly because of negative cognitive reappraisal, for example by weighting negative experiences more than positive ones. Of course I can only speculate that this may be part of the reason why significant effects were found for momentary positive affect and pleasure measures but not for monthly ones. Effects on momentary pleasure and PA but not on NA may be explained by the fact that pleasure was directly targeted by the lifestyle advice. Although NA might be expected to decrease as a result too, alternative scenarios are plausible as well. One option is that increasing pleasure may sometimes interfere with useful tasks that need to be done and results in feelings of guilt or in stress. This was taken into account when discussing the concept advice with participants, for example by discussing whether they had time to meet more with friends next to their exam schedule, but this may not have prevented negative thoughts.

In the published paper on which Chapter 6 was based only effects on the main outcome measures PA, NA, pleasure, and depression were investigated. This still left the question unanswered whether behavioral activation also modifies positive bias (i.e., spreads through the entire model as presented in Figure 1 of Chapter 1). To investigate this I tested whether behavioral activation increased positive bias operationalized as (1) implicit, relative happy bias based on a laboratory task; (2) explicit, absolute reward responsiveness based on self-report. I found no empirical evidence that tailored behavioral activation modified either type of positive bias (see *addendum*, Chapter 6). It was only possible for me to test the effect on positive bias during the month in which the advice was carried out, because after this month participants were free to choose an intervention and there was no longer a control group to compare to.

***Effects of the tandem skydive on anhedonia***

Contrary to what had been hypothesized, I found no evidence that the tandem skydive had an effect in addition to the effect of the lifestyle advice. The hypothesis that a skydive would give a boost to the reward system was based on the assumption that a skydive is such an extreme experience that it evokes universal physiological and psychological responses. One possible explanation of the lack of an effect of the tandem skydive is that anhedonic individuals, or at least some of them, may have shown blunted physiological responses to the skydive. However, the alpha-amylase reactivity and recovery patterns in response to the skydive showed strong similarities between anhedonic participants and were comparable to patterns reported for non-anhedonic individuals. The mean level of the stress peak was lower than in another study in healthy individuals (Chatterton et al., 1997), but it is unclear whether this was due to the anhedonic symptoms of the participants in the No Fun No Glory study, or due to differences in circumstances, for example, the fact that in the other study individuals were responsible for opening the parachute themselves, which could have added to the stress level. Another explanation for the lack of effect from the skydive is that the individuals with the most severe motivational problems may have benefitted the most from the skydive, yet this specific group was the most likely to not participate in the No Fun No Glory study, as they would not have been motivated to complete the screening survey in the first place. By far most of the participants who took part of the interventions were students who were attending classes and had social encounters and interactions with other people. They experienced persistent loss of pleasure, and there were large differences in self-reported motivational problems between the control group and the anhedonic group, but the anhedonic group may not have suffered from the most severe motivational deficits and behavioral withdrawal. Other explanations for the lack of an effect might be that the motivational boost of the tandem skydive was only short-lived, or that the experience was too stressful altogether, with feelings and memories of fear preventing the motivational boost to happen, or that after a tandem skydive other daily life activities seemed rather dull and the motivation to engage in these activities actually decreased rather than increased.

The original plan was not only to study alpha-amylase in saliva on the day of the tandem skydive, but also another biomarker associated with depression and susceptibility to stress, namely brain-derived neurotrophic factor (BDNF). BDNF is a protein that regulates neuronal plasticity and has been suggested as a biomarker for successful treatment of depression (Gourgouvelis, Yielder, Clarke, Behbahani, & Murphy, 2018; Polyakova et al., 2015). Acute physiological stressors such as exercise have been associated with increases in BDNF (Gourgouvelis et al., 2018; Heyman et al., 2012). It would have been interesting to study individual differences in BDNF reactivity and recovery in response to the acute stress experience of the tandem skydive and whether specific reactivity and recovery patterns were associated with decreases in symptoms of anhedonia. However, the results of the pilot study presented in Chapter 8 suggest that BDNF cannot be

measured reliably in saliva with the currently available commercial ELISA kits. Therefore, the intended BDNF analyses were cancelled and I have no information on BDNF in response to the skydive and whether this may explain why the tandem skydive in addition to the lifestyle advice did not result in a stronger increase in pleasure than the lifestyle advice alone.

## STRENGTHS AND LIMITATIONS

### Strengths

The studies described in this dissertation have several strengths. For the studies focused on associations between positive bias and depression large samples (TRAILS and No Fun No Glory) were used. Conclusions about the prospective associations between positive bias and depression were based upon different developmental periods and different tasks in which positive bias was assessed. For the studies described in Chapters 2 and 5 a so-called 'morph' task was used in which movie clips were shown of neutral faces which gradually changed into full intensity facial emotions. This made it possible to assess the identification of more subtle traces of emotions, which is assumed to give a more ecologically valid perspective than the often-used static full intensity facial emotion tasks (Joormann and Gotlib, 2006; Niedenthal et al., 2000). In everyday life full intensity facial emotions are rare, but we encounter subtle traces of facial emotions all the time. The benefit of the emotion identification morph task we used is that it enabled us to tap into individual differences in perceiving these frequently occurring everyday life social situations. The ecological momentary assessment data from the No Fun No Glory study made it possible to investigate the underlying mechanisms of positive bias in daily life, by looking at affect dynamics on a micro-level. The use of a multilevel approach in which within-subject effects were separated from between-subject effects by within-person standardization of all variables prior to the analyses resulted in a combination of the best of two worlds. It enabled exploring dynamic processes that take place within individuals; at the same time it allowed comparing two groups.

Additionally, two promising interventions to improve symptoms of anhedonia were developed and tested in young adults. The combination of monthly surveys, tasks, biomarkers, and momentary assessments completed three times a day for a long period of time resulted in valuable information. The No Fun No Glory study design with a large number of momentary assessments (about 90) in the baseline as well as in the intervention phase, with virtually no drop-out and excellent compliance rates, allowed investigating within-person effects both at a group level and at an individual level. The combination with biomarkers made it possible to explore physiological stress reactivity and recovery of anhedonic young adults in response to a tandem skydive. The findings suggest that they respond rather similarly to a skydive as healthy young adults, thereby discarding blunted stress responses to the skydive in anhedonic individuals as a likely explanation of the lack of an effect of the tandem skydive. Finally, taking the cognitive

model presented in the introductory chapter as the starting point of this dissertation led to new research questions about whether behavioral activation would in the end not only improve depressive symptoms but also increase positive bias. This was important to establish since a reduction of the biases may render the individual less vulnerable to developing depressive symptoms again in the future. If behavioral activation had resulted in an increase in positive bias, this would have suggested that this method may be used to prevent first onset of depression.

## Limitations

### *Tasks used to assess positive bias*

In the reflections on the main findings presented in this dissertation most limitations of the reward task and the facial emotion identification tasks were already discussed. Therefore, the most important limitations are only briefly repeated here:

1. In Chapter 2, associations are described between sensitivity to happy facial emotions and depressive symptoms. No conclusions can be drawn about associations between relative within-subject happy bias (i.e., relative sensitivity to happy facial emotions compared to negative facial emotions) and depressive symptoms, because this was not investigated.
2. Because the reward task (Chapter 4) contained no neutral condition, it was impossible to distinguish between difficulties with disengaging from negative information and reduced responsiveness to positive information after an initial focus on negative information. Other criticisms of spatial tasks that assess engagement and disengagement (Clarke, MacLeod, & Guastella, 2013; Grafton & MacLeod, 2014) also apply to the reward task: (1) the attentional focus on the fixation point (i.e., the score in the reward task) at the start of each trial is neither necessitated nor verified; (2) there may be individual differences in switching attention from the central fixation point to the distal reward cues and this may have a confounding influence on the task and make it impossible to separate fully between engagement and disengagement; (3) difficulty to disengage from negative cues may be intertwined with a freezing (or narrowing, as discussed in Chapter 4) response to negative cues, without the possibility to unravel them.
3. Particularly with respect to the facial emotion identification tasks (Chapter 2, 3, and 5) it is difficult to identify the specific underlying processes that are addressed. This makes interpretation of the results more difficult. Attention and perception of facial emotions cannot be unraveled, neither can automatic and voluntary processing.

***Two sides of the same coin: sustaining high levels positive affect and being stuck in low levels of positive affect***

4. Chapter 5 describes a study in which network analyses were used to compare the daily life affect dynamics between a high and a low positive bias group. In this study evidence was found that increased positive experiences and positive affect during one time interval predicted increased positive experiences and affect during the next time interval. It should be kept in mind, however, that a positive association between positive affect at two successive time points not only implies that increased positive affect at one time point predicts increased positive affect at the next time point, but also that decreased positive affect predicts decreased positive affect at the next time point. Thus, the same finding may mean that high levels of positive affect are sustained or that low levels of positive affect are sustained. Mean levels of positive affect were higher for the group that showed the highest autocorrelation of positive affect, which suggests that the high autocorrelation should be interpreted more toward sustaining high levels of positive affect, but it was not possible to test this statistically with the method described in Chapter 5.

***Ecological momentary assessments***

5. The 6 hour time interval was selected because of the practical consideration that it would be quite a burden to participants to complete daily questionnaires multiple times per day for a period of 3.5 months. Three assessments per day was considered the maximum possible in the light of the desired high compliance rates. It would certainly have been interesting to study activities and affect dynamics on shorter time scales. Theoretically, it is important not to choose too long time intervals between momentary assessments because important information may not be measured. With three assessments per day any dynamics taking place on a shorter time scale were ignored.
6. Fixed rather than semi-random time intervals were used for the momentary assessments in the No Fun No Glory study. A disadvantage of fixed time intervals is that participants know when to expect the questionnaires and may adapt their activities accordingly. For example, if participants expect a new questionnaire in 15 minutes, they may postpone doing groceries or going to the gym until after completing the questionnaire. This is problematic because momentary assessments are used to study people's daily life as it is and the ecological validity is compromised if the study interferes with this. This problem was partly countered by ensuring that most of the awake time of the participants was covered by the momentary assessments. That is, participants completed 3 momentary assessments per day with 6 hour time intervals in between and (most) questions were answered retrospectively about activities and mood during the past 6 hours. This does not solve the problem that participants may have postponed activities when expecting a

questionnaire, but it does ensure that possibly postponed activities, as well as the mood participants experienced during these activities, did still end up in the answers they provided.

7. Answering questions about the past 6 hours introduces recall bias. Perception and recall may be influenced by the emotional and motivational state participants are in while perceiving and reporting (Kihlstrom et al., 2000; Shiffman et al., 2008). This is in fact a problem of each and every study in which questions are not limited to subjects' current state, and it is unclear how large the recall bias is in the studies described in this dissertation or whether results would have been different without this bias. In some cases recalled mood may actually be a stronger predictor than mood at the time it was experienced (Shiffman et al., 2008), but it must be acknowledged that recalled mood does not necessarily, in fact more likely not, reflect mood at the moment it was experienced.

***Lifestyle advice and skydive***

8. Participants were not blinded for group allocation and there was no active control group receiving an alternative intervention. Participants knew whether they received an intervention or were allocated to the control group, this may have led them to expect a positive or negative outcome, and may have affected the outcome. As a consequence, the possibility that the found effects are placebo effects cannot be excluded.
9. The lifestyle advice cannot be seen separately from the conversation with the participants during which the concept advice was discussed and finalized. Therefore, effects we found may have been the result of general 'therapeutic' mechanisms, particularly because the participants who received a lifestyle advice in general had a longer meeting with the researchers than participants who had been assigned to the control condition.
10. Participants received no reminders of the lifestyle advice during the month in which they were supposed to carry it out. Weekly reminders could have prompted them to comply to the advice more frequently and may have improved the effects of the advice on depressive symptoms.
11. Young adults with the most severe motivational problems probably did not participate in the No Fun No Glory study. This means that it is unclear if the behavioral activation would have been feasible and would have benefitted these individuals too. One might also argue that particularly this group was expected to benefit most from the tandem skydive, but this remains speculation.
12. Only anhedonic participants performed a tandem skydive in the No Fun No Glory study. Therefore, it was not possible to compare physiological stress responses to the skydive between anhedonic and non-anhedonic individuals within our own study. I was limited to comparing our findings for anhedonic individuals to findings for healthy individuals reported in other studies that used partly different methods.

## FUTURE RESEARCH

Several findings in this dissertation call for replication in other samples. The prospective associations between positive bias and depression were investigated only in the TRAILS sample; it is therefore important to try to replicate them in a different sample, preferably with a wider range of tasks and more recently developed tasks assessing positive bias. Particularly the study presented in Chapter 4 is important to replicate, because this is the first study in which an automatic low positive bias, measured with a behavioral laboratory task, predicts depression. In Chapter 5, I presented a novel approach to positive bias in daily life by comparing a high and a low happy bias group on their daily life affect dynamics. Repeating this study in a larger sample is required to proceed beyond the level of tentative suggestions.

If the prospective findings are replicated it may be viable to explore whether modification of a low positive bias can be used as an early intervention in adolescence to prevent depression. It should be noted, however, that there is no convincing evidence that attempts to modify biases in general successfully alleviate currently experienced depressive symptoms, let alone prevent them. A recent meta-analysis in which quality of the studies and publication bias were taken into account found a small effect of bias modification attempts on current depressive symptoms, but this effect was reported to be most likely largely due to outliers and publication bias rather than a reflection of the existence of a real effect (Cristea, Kok, & Cuijpers, 2015). A second meta-analysis was performed in which only successful bias modifications were included and the effect of successful bias modification on emotional vulnerability was investigated (Grafton et al., 2017). The findings of the latter study suggest that successful bias modification decreases emotional vulnerability. From the discrepancy between the two meta-analyses it is at least clear that of the many different types of procedures that have been used to achieve bias modifications, many have failed. Thus, more research on how successful bias modification can be achieved and what type of bias modification is most successful is needed before we can even start thinking about using it as a preventive strategy.

Positive bias was a central construct in this dissertation. Both operationalizations of positive bias (happy bias and reward bias) are examples of implicit measures or indirect assessments of positive bias. Positive bias may also be assessed using explicit measures, for example questionnaires that directly assess positive bias. Implicit measures have lower face validity than explicit measures but have the advantage of being less sensitive to recall and response bias (Everaert, Podina, & Koster, 2017; Gotlib & Joormann, 2010). Moreover, it is evident that biases in behavioral or psychophysiological responses individuals themselves are unaware of can only be assessed with implicit measures. In a recent meta-analytic study on the association between depression and interpretation biases for ambiguous emotional information, significant findings were only reported for explicit and not for implicit measures of interpretation bias (Everaert et al., 2017). Additionally, higher effect-sizes were reported for relative interpretation bias scores (i.e., scores on positive interpretations compared to scores on negative interpretations) than for

absolute interpretation bias scores (i.e., isolated scores on positive or negative interpretations). The studies using explicit measures often used relative interpretation bias scores and studies using implicit measures mostly tested absolute scores. Within the pool of studies using implicit measures, studies using the relative scores and studies using the absolute scores were not analyzed separately by Everaert and colleagues (2017). The results reported in this dissertation suggest that this may be an interesting direction for further meta-analytic research, and also to compare effects of implicit relative bias scores to effects of explicit relative bias scores.

It is important to know more about how positive bias is shaped throughout the development from childhood to adolescence. In this dissertation I reported that a low positive bias may increase the risk of depression. In the cohort study that was used to test this hypothesis positive bias was measured at age 11 and age 16, but there was no information available before age 11. It thus remains unknown whether a low positive bias before age 11 also predicts depression and how positive bias develops over time. There is evidence that particular types of positive bias, such as incentive motivation and responsiveness to rewards, show a peak in adolescence compared to childhood and adulthood (Braams et al., 2015; Luciana, Wahlstrom, Porter, & Collins, 2012; Luking et al., 2016). For a fictitious illustration of such a peak in normal development, see the black curve plotted in Figure 2. It is unknown whether the absence of a peak in positive bias predicts depression, and whether the absence of such a peak is only noticeable in adolescence, or can already be predicted from a low positive bias in childhood. The gray curves plotted in Figure 2 represent different possibilities of deviations of normal reward development that could predict depression. It is also unclear if and how positive bias is transferred from parents to children. Positive bias in parents may shape the environments they select for their children and their parenting style, or positive bias may be transferred through genes associated with positive bias. These are promising directions for future research in depression.

Positive bias may not only be relevant in the context of depression or other psychiatric problems, but also in many other contexts, because how we respond to positive and negative information affects every aspect of our lives. It is, for example, likely that a positive bias influences how social contexts are perceived and encountered, and is itself also shaped by social experiences. The longitudinal study of positive bias in the context of broader social development may yield important findings on how positive bias shapes our perceptions of other people and their intentions, how we can benefit, or suffer, from social relationships, and how social experiences shape positive bias.

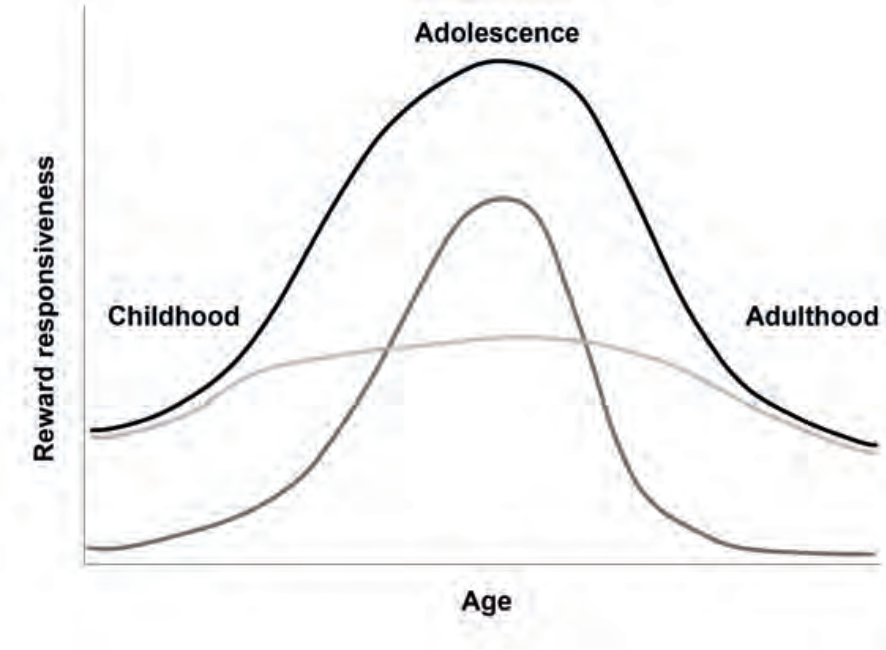



**FIGURE 2.** Fictitious plots of the development of reward responsiveness. The black curve represents normative development of reward responsiveness and shows a clear peak around mid-adolescence. The gray curves represent two types of deviations from normative development that may predict depression: the light gray curve represents a missing peak in adolescence; the dark gray curve an overall decreased but normatively shaped reward responsiveness from childhood onward.

With the model presented in Chapter 1 in mind, it would be interesting to investigate whether a combination of behavioral activation and positive bias modification may be used to alleviate depressive symptoms or prevent onset of depression. Targeting multiple components of the model simultaneously may result in more rigorous and longer-lasting effects. The promising findings presented in this dissertation suggest that it can be feasible and effective to provide tailored lifestyle advice to increase pleasurable activities (Chapter 6), and that it is worth investigating whether learning to savor positive affect and use positive experiences to generate positive affect may benefit individuals with a low positive bias (Chapter 5), and whether adolescents at risk for depression may benefit from learning how to disengage from negative situations and refocus their attention to positive situations (Chapter 4). A novel and exciting direction for future research would be to provide tailored lifestyle advice combining more general advice and advice provided in the moment. Initiatives to provide feedback in the moment are currently developed at Tilburg University by Dr. Eeske van Roekel and I am looking forward to the first results. The proposed feedback in the moment that I have in mind would take place in daily life with actual feedback in daily life situations, for example feedback to sustain positive affect given to participants on their smartphone at the moment they report high positive affect or pleasant experiences (based on Chapter 5), or advice to actively seek for

positive stimuli at the moment they report a negative mood or worrying (based on Chapter 4). This approach can be regarded as a type of bias modification, but a different type of bias modification than most attempts in which specific laboratory tasks are used to modify biases.

## REFERENCES


Achenbach, T. M. (1991). *Manual for the Youth Self-Report and 1991 profile*. Burlington: VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research center for children, youth and families, Burlington, VT.

Achenbach, T. M., & Rescorla, L. A. (2003). *Manual for the ASEBA adult forms & profiles*. Burlington, VT: University of Vermont.

Adam, E. K., Hoyt, L. T., & Granger, D. A. (2011). Diurnal alpha amylase patterns in adolescents: Associations with puberty and momentary mood states. *Biological Psychology*, *88*(2–3), 170–173. https://doi.org/10.1016/j.biopsycho.2011.07.007

Adams, R. B., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion (Washington, D.C.)*, *5*(1), 3–11. https://doi.org/10.1037/1528-3542.5.1.3

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-Practice Recommendations for Estimating Cross-Level Interaction Effects Using Multilevel Modeling. *Journal of Management*, *39*(6), 1490–1528. https://doi.org/10.1177/0149206313478188

Allwood, M. A., Handwerger, K., Kivlighan, K. T., Granger, D. A., & Stroud, L. R. (2011). Direct and moderating links of salivary alpha-amylase and cortisol stress-reactivity to youth behavioral and emotional adjustment. *Biological Psychology*, *88*(1), 57–64. https://doi.org/10.1016/j.biopsycho.2011.06.008

Alpers, G. W., & Adolph, D. (2008). Exposure to heights in a theme park: Fear, dizziness, and body sway. *Journal of Anxiety Disorders*, *22*(4), 591–601. https://doi.org/10.1016/j.janxdis.2007.05.008

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4th ed.)*. Washington, DC.

American Psychiatric Association. (2008). *The Practice of Electroconvulsive Therapy: Recommendations for Treatment, Training, and Privileging (A Task Force Report of the American Psychiatric Association)*. American Psychiatric Pub.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Anderson, I. M., & Fergusson, G. M. (2013). Mechanism of action of ECT. In *The ECT Handbook* (pp. 1–7). RCPsych Publications.

Anderson, I. M., Shippen, C., Juhasz, G., Chase, D., Thomas, E., Downey, D., … Deakin, J. F. W. (2011). State-dependent alteration in face emotion recognition in depression. *The British Journal of Psychiatry: The Journal of Mental Science*, *198*(4), 302–308. https://doi.org/10.1192/bjp.bp.110.078139

Andrews, G., & Peters, L. (1998). The psychometric properties of the Composite International Diagnostic Interview. *Social Psychiatry and Psychiatric Epidemiology*, *33*(2), 80–88. https://doi.org/10.1007/s001270050026

Archer, J., Hay, D. C., & Young, A. W. (1992). Face processing in psychiatric conditions. *British Journal of Clinical Psychology*, *31*(1), 45–61. https://doi.org/10.1111/j.2044-8260.1992.tb00967.x


*

Aspan, N., Bozsik, C., Gadoros, J., Nagy, P., Inantsy-Pap, J., Vida, P., & Halasz, J. (2014). Emotion recognition pattern in adolescent boys with Attention-Deficit/Hyperactivity Disorder. *BioMed Research International*, *2014*, e761340. https://doi.org/10.1155/2014/761340

Autry, A. E., & Monteggia, L. M. (2012). Brain-Derived Neurotrophic Factor and Neuropsychiatric Disorders. *Pharmacological Reviews*, *64*(2), 238–258. https://doi.org/10.1124/pr.111.005108

Bagozzi, R. P., Wong, N., & Yi, Y. (1999). The Role of Culture and Gender in the Relationship between Positive and Negative Affect. *Cognition and Emotion*, *13*(6), 641–672. https://doi.org/10.1080/026999399379023

Basso, M. R., Schefft, B. K., Ris, M. D., & Dember, W. N. (1996). Mood and global-local visual processing. *Journal of the International Neuropsychological Society*, *2*(3), 249–255. https://doi.org/10.1017/S1355617700001193

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67 (1)*, 1–48. https://doi.org/doi:10.18637/jss.v067.i01

Beck, A. T. (1967a). *Depression: Causes and treatment*. Philadelphia: University of Philadelphia Press.

Beck, A. T. (1967b). *Depression: Clinical, Experimental, and Theoretical Aspects*. University of Pennsylvania Press.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Beck, A. T., & Haigh, E. A. P. (2014). Advances in Cognitive Theory and Therapy: The Generic Cognitive Model. *Annual Review of Clinical Psychology*, *10*(1), 1–24. https://doi.org/10.1146/annurev-clinpsy-032813-153734

Beckwé, M., Deroost, N., Koster, E. H. W., Lissnyder, E. D., & Raedt, R. D. (2014). Worrying and rumination are both associated with reduced cognitive control. *Psychological Research*, *78*(5), 651–660. https://doi.org/10.1007/s00426-013-0517-5

Bediou, B., Brunelin, J., d'Amato, T., Fecteau, S., Saoud, M., Hénaff, M.-A., & Krolak-Salmon, P. (2012). A comparison of facial emotion processing in neurological and psychiatric conditions. *Frontiers in Psychology*, *3*, 98. https://doi.org/10.3389/fpsyg.2012.00098

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Bennik, E. C., Nederhof, E., Ormel, J., & Oldehinkel, A. J. (2014). Anhedonia and depressed mood in adolescence: course, stability, and reciprocal relation in the TRAILS study. *European Child & Adolescent Psychiatry*, *23*(7), 579–586. https://doi.org/10.1007/s00787-013-0481-z

Berghorst, L. H., Bogdan, R., Frank, M. J., & Pizzagalli, D. A. (2013). Acute stress selectively reduces reward sensitivity. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00133

Bhagwagar, Z., Cowen, P. J., Goodwin, G. M., & Harmer, C. J. (2004). Normalization of enhanced fear recognition by acute SSRI treatment in subjects with a previous history of depression. *American Journal of Psychiatry*, *161*(1), 166–168. https://doi.org/10.1176/appi.ajp.161.1.166

Bistricky, S. L., Atchley, R. A., Ingram, R., & O'Hare, A. (2014). Biased processing of sad faces: an ERP marker candidate for depression susceptibility. *Cognition & Emotion*, *28*(3), 470–492.

Bistricky, S. L., Ingram, R. E., & Atchley, R. A. (2011). Facial affect processing and depression susceptibility: Cognitive biases and cognitive neuroscience. *Psychological Bulletin*, *137*(6), 998–1028. https://doi.org/10.1037/a0025348

Bitsika, V., Sharpley, C. F., Andronicos, N. M., & Agnew, L. L. (2015). Hypothalamus–pituitary–adrenal axis daily fluctuation, anxiety and age interact to predict cortisol concentrations in boys with an autism spectrum disorder. *Physiology & Behavior*, *138*, 200–207. https://doi.org/10.1016/j.physbeh.2014.11.010

Blair, R. J. R., Mitchell, D. G. V., Peschardt, K. S., Colledge, E., Leonard, R. A., Shine, J. H., … Perrett, D. I. (2004). Reduced sensitivity to others' fearful expressions in psychopathic individuals. *Personality and Individual Differences*, *37*(6), 1111–1122. https://doi.org/10.1016/j.paid.2003.10.008

Blair, R. J. R. (2013). The neurobiology of psychopathic traits in youths. *Nature Reviews Neuroscience*, *14*(11), 786–799. https://doi.org/10.1038/nrn3577

Bogdan, R., & Pizzagalli, D. A. (2006). Acute Stress Reduces Reward Responsiveness: Implications for Depression. *Biological Psychiatry*, *60*(10), 1147–1154. https://doi.org/10.1016/j.biopsych.2006.03.037

Bongers, I. L., Koot, H. M., van der Ende, J., & Verhulst, F. C. (2004). Developmental Trajectories of Externalizing Behaviors in Childhood and Adolescence. *Child Development*, *75*(5), 1523–1537.

Booij, S. H., Bos, E. H., Bouwmans, M. E. J., Faassen, M. van, Kema, I. P., Oldehinkel, A. J., & Jonge, P. de. (2015). Cortisol and α-Amylase Secretion Patterns between and within Depressed and Non-Depressed Individuals. *PLOS ONE*, *10*(7), e0131002. https://doi.org/10.1371/journal.pone.0131002

Booij, S. H., Bos, E. H., de Jonge, P., & Oldehinkel, A. J. (2016). The temporal dynamics of cortisol and affective states in depressed and non-depressed individuals. *Psychoneuroendocrinology*, *69*, 16–25. https://doi.org/10.1016/j.psyneuen.2016.03.012

Borsboom, D., & Cramer, A. O. J. (2013). Network Analysis: An Integrative Approach to the Structure of Psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Bosch, J. A., Veerman, E. C. I., de Geus, E. J., & Proctor, G. B. (2011). α-Amylase as a reliable and convenient measure of sympathetic activity: don't start salivating just yet! *Psychoneuroendocrinology*, *36*(4), 449–453. https://doi.org/10.1016/j.psyneuen.2010.12.019

Bot, M., Middeldorp, C. M., Geus, E. J. C. de, Lau, H. M., Sinke, M., Nieuwenhuizen, B. van, … Penninx, B. W. J. H. (2017). Validity of LIDAS (LIfetime Depression Assessment Self-report): a self-report online assessment of lifetime major depressive disorder. *Psychological Medicine*, *47*(2), 279–289. https://doi.org/10.1017/S0033291716002312

Bourke, C., Douglas, K., & Porter, R. (2010). Processing of facial emotion expression in major depression: a review. *The Australian and New Zealand Journal of Psychiatry*, *44*(8), 681–696. https://doi.org/10.3109/00048674.2010.496359

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*(2), 129–148. https://doi.org/10.1037/0003-066X.36.2.129

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.

*

Braams, B. R., Duijvenvoorde, A. C. K. van, Peper, J. S., & Crone, E. A. (2015). Longitudinal Changes in Adolescent Risk-Taking: A Comprehensive Study of Neural Responses to Rewards, Pubertal Development, and Risk-Taking Behavior. *Journal of Neuroscience*, *35*(18), 7226–7238. https://doi.org/10.1523/JNEUROSCI.4764-14.2015

Bradley, B. P., Mogg, K., White, J., Groom, C., & De Bono, J. (1999). Attentional bias for emotional faces in generalized anxiety disorder. *British Journal of Clinical Psychology*, *38*(3), 267–278. https://doi.org/10.1348/014466599162845

Brandt, P. T., & Williams, J. T. (2007). *Multiple Time Series Models*. SAGE.

Brent, D. A., Perper, J. A., Moritz, G., Allman, C., Friend, A., Roth, C., … Baugher, M. (1993). Psychiatric risk factors for adolescent suicide: a case-control study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*(3), 521–529. https://doi.org/10.1097/00004583-199305000-00006

Bress, J. N., Foti, D., Kotov, R., Klein, D. N., & Hajcak, G. (2013). Blunted neural response to rewards prospectively predicts depression in adolescent girls. *Psychophysiology*, *50*(1), 74–81. https://doi.org/10.1111/j.1469-8986.2012.01485.x

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., … Kuppens, P. (2016). Assessing Temporal Emotion Dynamics Using Networks. *Assessment*, *23*(4), 425–435. https://doi.org/10.1177/1073191116645909

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., … Tuerlinckx, F. (2013). A Network Approach to Psychopathology: New Insights into Clinical Longitudinal Data. *PLOS ONE*, *8*(4), e60188. https://doi.org/10.1371/journal.pone.0060188

Brunnekreef, J. A., De Sonneville, L. M. J., Althaus, M., Minderaa, R. B., Oldehinkel, A. J., Verhulst, F. C., & Ormel, J. (2007). Information processing profiles of internalizing and externalizing behavior problems: evidence from a population-based sample of preadolescents. *Journal of Child Psychology and Psychiatry*, *48*(2), 185–193. https://doi.org/10.1111/j.1469-7610.2006.01695.x

Bryant, F. (2003). Savoring Beliefs Inventory (SBI): A scale for measuring beliefs about savouring. *Journal of Mental Health*, *12*(2), 175–196. https://doi.org/10.1080/0963823031000103489

Buchanan, T. W., al'Absi, M., & Lovallo, W. R. (1999). Cortisol fluctuates with increases and decreases in negative affect. *Psychoneuroendocrinology*, *24*(2), 227–241. https://doi.org/10.1016/S0306-4530(98)00078-X

Bui, E., Anderson, E., Goetter, E. M., Campbell, A. A., Fischer, L. E., Feldman Barrett, L., & Simon, N. M. (2015). Heightened sensitivity to emotional expressions in generalised anxiety disorder, compared to social anxiety disorder, and controls. *Cognition & Emotion*, 1–8. https://doi.org/10.1080/02699931.2015.1087973

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016). Using Raw VAR Regression Coefficients to Build Networks can be Misleading. *Multivariate Behavioral Research*, *51*(2–3), 330–344. https://doi.org/10.1080/00273171.2016.1150151

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

Bus, B. A. A. (2014). *Reassessing the neurotrophic hypothesis of depression : what can be learned from peripheral measures?* 's-Hertogenbosch : BOXPress. Retrieved from http://repository.ubn.ru.nl/handle/2066/129677

Bus, B. A. A., Molendijk, M. L., Penninx, B. J. W. H., Buitelaar, J. K., Kenis, G., Prickaerts, J., … Voshaar, R. C. O. (2011). Determinants of serum brain-derived neurotrophic factor. *Psychoneuroendocrinology*, *36*(2), 228–239. https://doi.org/10.1016/j.psyneuen.2010.07.013

Calvo, M. G., & Marrero, H. (2009). Visual search of emotional faces: The role of affective content and featural distinctiveness. *Cognition and Emotion*, *23*(4), 782–806. https://doi.org/10.1080/02699930802151654

Carver, C. S. (2006). Approach, Avoidance, and the Self-Regulation of Affect and Action. *Motivation and Emotion*, *30*(2), 105–110. https://doi.org/10.1007/s11031-006-9044-7

Celsi, R. L., Rose, R. L., & Leigh, T. W. (1993). An Exploration of High-Risk Leisure Consumption Through Skydiving. *Journal of Consumer Research*, *20*(1), 1–23.

Chase, H. W., Frank, M. J., Michael, A., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2010). Approach and avoidance learning in patients with major depression and healthy controls: relation to anhedonia. *Psychological Medicine*, *40*(3), 433–440. https://doi.org/10.1017/S0033291709990468

Chatterton, R. T., Vogelsong, K. M., Lu, Y., & Hudgens, G. A. (1997). Hormonal Responses to Psychological Stress in Men Preparing for Skydiving. *The Journal of Clinical Endocrinology & Metabolism*, *82*(8), 2503–2509. https://doi.org/10.1210/jcem.82.8.4133

Chawla, N., & Ostafin, B. (2007). Experiential avoidance as a functional dimensional approach to psychopathology: An empirical review. *Journal of Clinical Psychology*, *63*(9), 871–890. https://doi.org/10.1002/jclp.20400

Cheverud, J. M. (2001). A simple correction for multiple comparisons in interval mapping genome scans. *Heredity*, *87*(1), 52–58. https://doi.org/10.1046/j.1365-2540.2001.00901.x

Cho, H.-C., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate? *Journal of Business Research*, *66*(9), 1261–1266. https://doi.org/10.1016/j.jbusres.2012.02.023

Clark, L., Chamberlain, S. R., & Sahakian, B. J. (2009). Neurocognitive Mechanisms in Depression: Implications for Treatment. *Annual Review of Neuroscience*, *32*(1), 57–74. https://doi.org/10.1146/annurev.neuro.31.060407.125618

Clarke, P. J. F., MacLeod, C., & Guastella, A. J. (2013). Assessing the role of spatial engagement and disengagement of attention in anxiety-linked attentional bias: a critique of current paradigms and suggestions for future research directions. *Anxiety, Stress, & Coping*, *26*(1), 1–19. https://doi.org/10.1080/10615806.2011.638054

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2 edition). Hillsdale, N.J: Routledge.

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 (Washington, D.C.)*, *9*(3), 361–368. https://doi.org/10.1037/a0015952

Collins, W. A., & Laursen, B. (2004). Changing Relationships, Changing Youth: Interpersonal Contexts of Adolescent Development. *The Journal of Early Adolescence*, *24*(1), 55–62. https://doi.org/10.1177/0272431603260882

*

Corr, P. J. (2001). Testing problems in J. A. Gray's personality theory: a commentary on Matthews and Gilliland (1999). *Personality and Individual Differences*, *30*(2), 333–352. https://doi.org/10.1016/S0191-8869(00)00028-3

Corral-Frías, N. S., Nadel, L., Fellous, J.-M., & Jacobs, W. J. (2016). Behavioral and self-reported sensitivity to reward are linked to stress-related differences in positive affect. *Psychoneuroendocrinology*, *66*, 205–213. https://doi.org/10.1016/j.psyneuen.2016.01.012

Costello, E. J., Erkanli, A., & Angold, A. (2006). Is there an epidemic of child or adolescent depression? *Journal of Child Psychology and Psychiatry*, *47*(12), 1263–1271. https://doi.org/10.1111/j.1469-7610.2006.01682.x

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *The British Journal of Psychiatry*, *206*(1), 7–16.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Systems 1695*. Retrieved from http://igraph.org

Cubillo, A., Halari, R., Smith, A., Taylor, E., & Rubia, K. (2012). A review of fronto-striatal and fronto-cortical brain abnormalities in children and adults with Attention Deficit Hyperactivity Disorder (ADHD) and new evidence for dysfunction in adults with ADHD during motivation and attention. *Cortex*, *48*(2), 194–215. https://doi.org/10.1016/j.cortex.2011.04.007

Cuijpers, P., van Straten, A., & Warmerdam, L. (2007). Behavioral activation treatments of depression: A meta-analysis. *Clinical Psychology Review*, *27*(3), 318–326. https://doi.org/10.1016/j.cpr.2006.11.001

Dale, G. L. (1997). Platelet kinetics. *Current Opinion in Hematology 1997*, *4*(5), 330–334.

Dalili, M. N., Penton-Voak, I. S., Harmer, C. J., & Munafò, M. R. (2014). Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychological Medicine*, *FirstView*, 1–10. https://doi.org/10.1017/S0033291714002591

De Sonneville, L. M. J. (1999). Amsterdam Neuropsychological Tasks: A computer-aided assessment program. In B. P. L. M. den Brinker, P. J. Beek, A. N. Brand, F. J. Maarse, & L. J. M. Mulder, *Cognitive Ergonomics, Clinical Assessment and Computer-assisted Learning: Computers in Psychology* (pp. 187–203). Lisse, The Netherlands: Swets & Zeitlinger.

de Souza, F. T. A., Amaral, T. M. P., dos Santos, T. P. M., Abdo, E. N., Aguiar, M. C. F., Teixeira, A. L., … Silva, T. A. (2012). Burning Mouth Syndrome: A Therapeutic Approach Involving Mechanical Salivary Stimulation. *Headache: The Journal of Head & Face Pain*, *52*(6), 1026–1034. https://doi.org/10.1111/j.1526-4610.2011.02037.x

de Souza, F. T. A., Kummer, A., Silva, M. L. V., Amaral, T. M. P., Abdo, E. N., Abreu, M. H. N. G., … Teixeira, A. L. (2014). The Association of Openness Personality Trait with Stress-Related Salivary Biomarkers in Burning Mouth Syndrome. *Neuroimmunomodulation*, *22*(4), 250–255. https://doi.org/10.1159/000367714

De Winter, A. F., Oldehinkel, A. J., Veenstra, R., Brunnekreef, J. A., Verhulst, F. C., & Ormel, J. (2005). Evaluation of non-response bias in mental health determinants and outcomes in a large sample of pre-adolescents. *European Journal of Epidemiology*, *20*(2), 173–181. https://doi.org/10.1007/s10654-004-4948-6

Demenescu, L. R., Stan, A., Kortekaas, R., van der Wee, N. J. A., Veltman, D. J., & Aleman, A. (2014). On the connection between level of education and the neural circuitry of emotion perception. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00866

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology*, *111*(2), 225–236.

Disner, S. G., Beevers, C. G., Haigh, E. A. P., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*, *12*(8), 467–477. https://doi.org/10.1038/nrn3027

Ekers, D., Richards, D., & Gilbody, S. (2008). A meta-analysis of randomized trials of behavioural treatment of depression. *Psychological Medicine*, *38*(5), 611–623. https://doi.org/10.1017/S0033291707001614

Elliott, R., Sahakian, B. J., Herrod, J. J., Robbins, T. W., & Paykel, E. S. (1997). Abnormal response to negative feedback in unipolar depression: evidence for a diagnosis specific impairment. *Journal of Neurology, Neurosurgery, and Psychiatry*, *63*(1), 74–82.

Elliott, R., Sahakian, B. J., Mckay, A. P., Herrod, J. J., Robbins, T. W., & Paykel, E. S. (1996). Neuropsychological Impairments in Unipolar Depression: The Influence of Perceived Failure on Subsequent Performance. *Psychological Medicine*, *26*(5), 975–990.

Emerencia, A., van der Krieke, L., Bos, E., de Jonge, P., Petkov, N., & Aiello, M. (2015). Automating vector autoregression on electronic patient diary data. *IEEE Journal of Biomedical and Health Informatics*, *PP*(99), 1–1. https://doi.org/10.1109/JBHI.2015.2402280

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, *048*(i04). Retrieved from http://econpapers.repec.org/article/jssjstsof/v_3a048_3ai04.htm

Epskamp, S., Deserno, M. K., & Bringmann, L. F. (2016). mlVAR: Multi-Level Vector Autoregression. R package version 0.4. https://CRAN.R-project.org/package=mlVAR

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2016). Discovering psychological dynamics in time-series data. *ArXiv Preprint ID 1609.04156v2*, 1–31.

Erceg-Hurn, D. M., & Mirosevich, V. M. (2008). Modern robust statistical methods: An easy way to maximize the accuracy and power of your research. *American Psychologist*, *63*(7), 591–601. https://doi.org/10.1037/0003-066X.63.7.591

Ernst, M., & Fudge, J. L. (2009). A developmental neurobiological model of motivated behavior: Anatomy, connectivity and ontogeny of the triadic nodes. *Neuroscience & Biobehavioral Reviews*, *33*(3), 367–382. https://doi.org/10.1016/j.neubiorev.2008.10.009

Eshel, N., & Roiser, J. P. (2010). Reward and Punishment Processing in Depression. *Biological Psychiatry*, *68*(2), 118–124. https://doi.org/10.1016/j.biopsych.2010.01.027

Evans, S., Fleming, S. M., Dolan, R. J., & Averbeck, B. B. (2011). Effects of emotional preferences on value-based decision making are mediated by mentalizing not reward networks. *Journal of Cognitive Neuroscience*, *23*(9), 2197–2210. https://doi.org/10.1162/jocn.2010.21584

Everaert, J., Podina, I. R., & Koster, E. H. W. (2017). A comprehensive meta-analysis of interpretation biases in depression. *Clinical Psychology Review*, *58*, 33–48. https://doi.org/10.1016/j.cpr.2017.09.005

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Favrod, J., Giuliani, F., Ernst, F., & Bonsack, C. (2010). Anticipatory Pleasure Skills Training: A New Intervention to Reduce Anhedonia in Schizophrenia. *Perspectives in Psychiatric Care*, *46*(3), 171–181. https://doi.org/10.1111/j.1744-6163.2010.00255.x

Favrod, J., Nguyen, A., Fankhauser, C., Ismailaj, A., Hasler, J.-D., Ringuet, A., … Bonsack, C. (2015). Positive Emotions Program for Schizophrenia (PEPS): a pilot intervention to reduce anhedonia and apathy. *BMC Psychiatry*, *15*, 231. https://doi.org/10.1186/s12888-015-0610-y

Fils, J. M., Penick, E. C., Nickel, E. J., Othmer, E., DeSouza, C., Gabrielli, W. F., & Hunter, E. E. (2010). Minor Versus Major Depression: A Comparative Clinical Study. *Primary Care Companion to The Journal of Clinical Psychiatry*, *12*(1). https://doi.org/10.4088/PCC.08m00752blu

Fombonne, E., Simmons, H., Ford, T., Meltzer, H., & Goodman, R. (2001). Prevalence of Pervasive Developmental Disorders in the British Nationwide Survey of Child Mental Health. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(7), 820–827. https://doi.org/10.1097/00004583-200107000-00017

Forbes, E. E., Christopher May, J., Siegle, G. J., Ladouceur, C. D., Ryan, N. D., Carter, C. S., … Dahl, R. E. (2006). Reward-related decision-making in pediatric major depressive disorder: an fMRI study. *Journal of Child Psychology and Psychiatry*, *47*(10), 1031–1040. https://doi.org/10.1111/j.1469-7610.2006.01673.x

Forbes, E. E., & Dahl, R. E. (2005). Neural systems of positive affect: Relevance to understanding child and adolescent depression? *Development and Psychopathology*, *17*(3), 827–850. https://doi.org/10.1017/S095457940505039X

Forbes, E. E., & Dahl, R. E. (2012). Research Review: Altered reward function in adolescent depression: what, when and how? *Journal of Child Psychology and Psychiatry*, *53*(1), 3–15. https://doi.org/10.1111/j.1469-7610.2011.02477.x

Forbes, E. E., Hariri, A. R., Martin, S. L., Silk, J. S., Moyles, D. L., Fisher, P. M., … Dahl, R. E. (2009). Altered Striatal Activation Predicting Real-World Positive Affect in Adolescent Major Depressive Disorder. *American Journal of Psychiatry*, *166*(1), 64–73. https://doi.org/10.1176/appi.ajp.2008.07081336

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*(5), 633–639. https://doi.org/10.1016/j.biopsych.2006.05.026

Franken, I. H. A., Zijlstra, C., & Muris, P. (2006). Are nonpharmacological induced rewards related to anhedonia? A study among skydivers. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *30*(2), 297–300. https://doi.org/10.1016/j.pnpbp.2005.10.011

Fredrickson, B. L. (1998). What Good Are Positive Emotions? *Review of General Psychology : Journal of Division 1, of the American Psychological Association*, *2*(3), 300–319. https://doi.org/10.1037/1089-2680.2.3.300

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology. The broaden-and-build theory of positive emotions. *The American Psychologist*, *56*(3), 218–226.

Fredrickson, B. L. (2004). The Broaden-and-Build Theory of Positive Emotions. *Philosophical Transactions: Biological Sciences*, *359*(1449), 1367–1377.

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. https://doi.org/10.1080/026999398379718

Freitas-Ferrari, M. C., Hallak, J. E. C., Trzesniak, C., Filho, A. S., Machado-de-Sousa, J. P., Chagas, M. H. N., … Crippa, J. A. S. (2010). Neuroimaging in social anxiety disorder: A systematic review of the literature. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *34*(4), 565–580. https://doi.org/10.1016/j.pnpbp.2010.02.028

Friedman, R. S., & Förster, J. (2010). Implicit Affective Cues and Attentional Tuning: An Integrative Review. *Psychological Bulletin*, *136*(5), 875–893. https://doi.org/10.1037/a0020495

Frijda, N. H. (2007). *The Laws of Emotion*. New York: Psychology Press.

Fu, C. H. Y., Williams, S. C. R., Brammer, M. J., Suckling, J., Kim, J., Cleare, A. J., … Bullmore, E. T. (2007). Neural Responses to Happy Facial Expressions in Major Depression Following Antidepressant Treatment. *American Journal of Psychiatry*, *164*(4), 599–607. https://doi.org/10.1176/ajp.2007.164.4.599

Furl, N., Gallagher, S., & Averbeck, B. B. (2012). A selective emotional decision-making bias elicited by facial expressions. *PLOS ONE*, *7*(3), e33461. https://doi.org/10.1371/journal.pone.0033461

Fusar-Poli, P., Papanastasiou, E., Stahl, D., Rocchetti, M., Carpenter, W., Shergill, S., & McGuire, P. (2015). Treatments of Negative Symptoms in Schizophrenia: Meta-Analysis of 168 Randomized Placebo-Controlled Trials. *Schizophrenia Bulletin*, *41*(4), 892–899. https://doi.org/10.1093/schbul/sbu170

Gabbay, V., Johnson, A. R., Alonso, C. M., Evans, L. K., Babb, J. S., & Klein, R. G. (2015). Anhedonia, but not Irritability, Is Associated with Illness Severity Outcomes in Adolescent Major Depression. *Journal of Child and Adolescent Psychopharmacology*, *25*(3), 194–200. https://doi.org/10.1089/cap.2014.0105

Gaebel, W., & Wölwer, W. (1992). Facial expression and emotional face recognition in schizophrenia and depression. *European Archives of Psychiatry and Clinical Neuroscience*, *242*(1), 46–52. https://doi.org/10.1007/BF02190342

Galvan, A. (2010). Adolescent Development of the Reward System. *Frontiers in Human Neuroscience*, *4*. https://doi.org/10.3389/neuro.09.006.2010

Garralda, M. E., & Bailey, D. (1988). Child and family factors associated with referral to child psychiatrists. *The British Journal of Psychiatry*, *153*(1), 81–89. https://doi.org/10.1192/bjp.153.1.81

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Gerra, G., Zaimovic, A., Zambelli, U., Timpano, M., Reali, N., Bernasconi, S., & Brambilla, F. (2000). Neuroendocrine Responses to Psychological Stress in Adolescents with Anxiety Disorder. *Neuropsychobiology*, *42*(2), 82–92. https://doi.org/10.1159/000026677

Gershon, J. (2002). A Meta-Analytic Review of Gender Differences in ADHD. *Journal of Attention Disorders*, *5*(3), 143–154. https://doi.org/10.1177/108705470200500302

*

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), 241–247. https://doi.org/10.1016/j.euroneuro.2010.11.004

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*(2), 129–138. https://doi.org/10.1111/j.1600-0447.2009.01525.x

Glenn, A. L., & Yang, Y. (2012). The Potential Role of the Striatum in Antisocial Behavior and Psychopathy. *Biological Psychiatry*, *72*(10), 817–822. https://doi.org/10.1016/j.biopsych.2012.04.027

Gore, F. M., Bloem, P. J., Patton, G. C., Ferguson, J., Joseph, V., Coffey, C., … Mathers, C. D. (2011). Global burden of disease in young people aged 10–24 years: a systematic analysis. *The Lancet*, *377*(9783), 2093–2102. https://doi.org/10.1016/S0140-6736(11)60512-6

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. *Archives of General Psychiatry*, *67*(4), 380–387. https://doi.org/10.1001/archgenpsychiatry.2010.13

Gotlib, I. H., & Joormann, J. (2010). Cognition and Depression: Current Status and Future Directions. *Annual Review of Clinical Psychology*, *6*(1), 285–312. https://doi.org/10.1146/annurev.clinpsy.121208.131305

Gotlib, I. H., Kasch, K. L., Traill, S., Joormann, J., Arnow, B. A., & Johnson, S. L. (2004). Coherence and specificity of information-processing biases in depression and social phobia. *Journal of Abnormal Psychology*, *113*(3), 386–398. https://doi.org/10.1037/0021-843X.113.3.386

Gotlib, I. H., Krasnoperova, E., Yue, D. N., & Joormann, J. (2004). Attentional Biases for Negative Interpersonal Stimuli in Clinical Depression. *Journal of Abnormal Psychology*, *113*(1), 127–135. https://doi.org/10.1037/0021-843X.113.1.121

Gourgouvelis, J., Yielder, P., Clarke, S. T., Behbahani, H., & Murphy, B. A. (2018). Exercise Leads to Better Clinical Outcomes in Those Receiving Medication Plus Cognitive Behavioral Therapy for Major Depressive Disorder. *Frontiers in Psychiatry*, *9*. https://doi.org/10.3389/fpsyt.2018.00037

Grafton, B., & MacLeod, C. (2014). Enhanced probing of attentional bias: The independence of anxiety-linked selectivity in attentional engagement with and disengagement from negative information. *Cognition and Emotion*, *28*(7), 1287–1302. https://doi.org/10.1080/02699931.2014.881326

Grafton, B., MacLeod, C., Rudaizky, D., Holmes, E. A., Salemink, E., Fox, E., & Notebaert, L. (2017). Confusing procedures with process when appraising the impact of cognitive bias modification on emotional vulnerability. *The British Journal of Psychiatry*, *211*(5), 266–271. https://doi.org/10.1192/bjp.bp.115.176123

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424–438. https://doi.org/10.2307/1912791

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology*, *31*(4), 337–350. https://doi.org/10.1007/s10654-016-0149-3

Grillo, L. (2016). A Possible Role of Anhedonia as Common Substrate for Depression and Anxiety. *Depression Research and Treatment*, *2016*, e1598130. https://doi.org/10.1155/2016/1598130

Gutkovich, Z., Morrissey, R. F., Espaillat, R. K., Dicker, R., Gutkovich, Z., Morrissey, R. F., … Dicker, R. (2010). Anhedonia and Pessimism in Hospitalized Depressed Adolescents. *Depression Research and Treatment*, e795173. https://doi.org/10.1155/2011/795173

Hamaker, E. L., & Grasman, R. P. P. P. (2015). To center or not to center? Investigating inertia with a multilevel autoregressive model. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01492

Hankin, B. L., & Abramson, L. Y. (2001). Development of gender differences in depression: An elaborated cognitive vulnerability-transactional stress theory. *Psychological Bulletin*, *127*, 773–796. https://doi.org/10.1037//0033-2909.127.6.773

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

Harmer, C. J., Goodwin, G. M., & Cowen, P. J. (2009). Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. *The British Journal of Psychiatry*, *195*(2), 102–108. https://doi.org/10.1192/bjp.bp.108.051193

Harmer, C. J., Shelley, N. C., Cowen, P. J., & Goodwin, G. M. (2004). Increased Positive Versus Negative Affective Perception and Memory in Healthy Volunteers Following Selective Serotonin and Norepinephrine Reuptake Inhibition. *American Journal of Psychiatry*, *161*(7), 1256–1263. https://doi.org/10.1176/appi.ajp.161.7.1256

Haro, J. M., Arbabzadeh-Bouchez, S., Brugha, T. S., De Girolamo, G., Guyer, M. E., Jin, R., … Kessler, R. C. (2006). Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health Surveys. *International Journal of Methods in Psychiatric Research*, *15*(4), 167–180. https://doi.org/10.1002/mpr.196

Hartmann, D. P., Gottman, J. M., Jones, R. R., Gardner, W., Kazdin, A. E., & Vaught, R. S. (1980). Interrupted Time-Series Analysis and Its Application to Behavioral Data. *Journal of Applied Behavior Analysis*, *13*(4), 543–559. https://doi.org/10.1901/jaba.1980.13-543

Hartmann, J. A., Wichers, M., Menne-Lothmann, C., Kramer, I., Viechtbauer, W., Peeters, F., … Simons, C. J. P. (2015). Experience Sampling-based personalized feedback and positive affect: A randomized controlled trial in depressed patients. *PLoS ONE*, *10*(6), e0128095. https://doi.org/10.1371/journal.pone.0128095

Heininga, V. E., Van Roekel, E., Ahles, J. J., Oldehinkel, A. J., & Mezulis, A. H. (2017). Positive affective functioning in anhedonic individuals' daily life: Anything but flat and blunted. *Journal of Affective Disorders*, *218*(Supplement C), 437–445. https://doi.org/10.1016/j.jad.2017.04.029

Heller, A. S., Johnstone, T., Shackman, A. J., Light, S. N., Peterson, M. J., Kolden, G. G., … McEwen, B. S. (2009). Reduced Capacity to Sustain Positive Emotion in Major Depression Reflects Diminished Maintenance of Fronto-Striatal Brain Activation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(52), 22445–22445.

*

Heyman, E., Gamelin, F.-X., Goekint, M., Piscitelli, F., Roelands, B., Leclair, E., … Meeusen, R. (2012). Intense exercise increases circulating endocannabinoid and BDNF levels in humans—Possible implications for reward and depression. *Psychoneuroendocrinology*, *37*(6), 844–851. https://doi.org/10.1016/j.psyneuen.2011.09.017

Hilt, L. M., & Pollak, S. D. (2013). Characterizing the Ruminative Process in Young Adolescents. *Journal of Clinical Child & Adolescent Psychology*, *42*(4), 519–530. https://doi.org/10.1080/15374416.2013.764825

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. https://doi.org/10.1371/journal.pone.0075655

Horner, M. S., Siegle, G. J., Schwartz, R. M., Price, R. B., Haggerty, A. E., Collier, A., & Friedman, E. S. (2014). C'mon Get Happy: Reduced Magnitude and Duration of Response During a Positive-Affect Induction in Depression. *Depression and Anxiety*, *31*(11), 952–960. https://doi.org/10.1002/da.22244

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*, *141*(4), 901–930. https://doi.org/10.1037/a0038822

Hox, J. J. (2002). *Multilevel Analysis: Techniques and Applications*. Lawrence Erlbaum Associates.

Huisman, M., Oldehinkel, A. J., De Winter, A., Minderaa, R. B., Bildt, A. de, Huizink, A. C., … Ormel, J. (2008). Cohort Profile: The Dutch 'TRacking Adolescents' Individual Lives' Survey'; TRAILS. *International Journal of Epidemiology*, *37*(6), 1227–1235. https://doi.org/10.1093/ije/dym273

Huitema, B. E., & Mckean, J. W. (2000). Design Specification Issues in Time-Series Intervention Models. *Educational and Psychological Measurement*, *60*(1), 38–58. https://doi.org/10.1177/00131640021970358

IBM Corporation. (2017). IBM SPSS Statistics for Windows. Version 25. Armonk NY: IBM Corporation.

Ikai, S., Suzuki, T., Uchida, H., Saruta, J., Tsukinoki, K., Fujii, Y., & Mimura, M. (2014). Effects of Weekly One-Hour Hatha Yoga Therapy on Resilience and Stress Levels in Patients with Schizophrenia-Spectrum Disorders: An Eight-Week Randomized Controlled Trial. *The Journal of Alternative and Complementary Medicine*, *20*(11), 823–830. https://doi.org/10.1089/acm.2014.0205

Ingram, R. E. (1984). Toward an information-processing analysis of depression. *Cognitive Therapy and Research*, *8*(5), 443–477. https://doi.org/10.1007/BF01173284

Isaac, L. (2012). Facing the future: face-emotion processing deficits as a potential biomarker for various psychiatric and neurological disorders. *Cognitive Science*, 171. https://doi.org/10.3389/fpsyg.2012.00171

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., … Prince, S. E. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology*, *64*(2), 295–304. https://doi.org/10.1037/0022-006X.64.2.295

Jacobson, N. S., Martell, C. R., & Dimidjian, S. (2001). Behavioral Activation Treatment for Depression: Returning to Contextual Roots. *Clinical Psychology: Science and Practice*, *8*(3), 255–270. https://doi.org/10.1093/clipsy.8.3.255

Joormann, J., & Gotlib, I. H. (2006). Is this happiness I see? Biases in the identification of emotional facial expressions in depression and social phobia. *Journal of Abnormal Psychology*, *115*(4), 705–714. https://doi.org/10.1037/0021-843X.115.4.705

Joormann, J., Talbot, L., & Gotlib, I. H. (2007). Biased processing of emotional information in girls at risk for depression. *Journal of Abnormal Psychology*, *116*(1), 135–143. https://doi.org/10.1037/0021-843X.116.1.135

Judd L. L., Akiskal H. S., Zeller P.J., & et al. (2000). Psychosocial disability during the long-term course of unipolar major depressive disorder. *Archives of General Psychiatry*, *57*(4), 375–380. https://doi.org/10.1001/archpsyc.57.4.375

Kahler, C. W., McHugh, R. K., Leventhal, A. M., Colby, S. M., Gwaltney, C. J., & Monti, P. M. (2012). High hostility among smokers predicts slower recognition of positive facial emotion. *Personality and Individual Differences*, *52*(3), 444–448. https://doi.org/10.1016/j.paid.2011.11.009

Kanter, J. W., Manos, R. C., Bowe, W. M., Baruch, D. E., Busch, A. M., & Rusch, L. C. (2010). What is behavioral activation?: A review of the empirical literature. *Clinical Psychology Review*, *30*(6), 608–620. https://doi.org/10.1016/j.cpr.2010.04.001

Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and Exploration: Facilitating Positive Subjective Experiences and Personal Growth Opportunities. *Journal of Personality Assessment*, *82*(3), 291–305. https://doi.org/10.1207/s15327752jpa8203_05

Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review*, *30*(7), 865–878. https://doi.org/10.1016/j.cpr.2010.03.001

Kendler, K. S. (2012). The dappled nature of causes of psychiatric illness: replacing the organic-functional/hardware-software dichotomy with empirically based pluralism. *Molecular Psychiatry*, *17*(4), 377–388. https://doi.org/10.1038/mp.2011.182

Kessler, R. C., Abelson, J., Demler, O., Escobar, J. I., Gibbon, M., Guyer, M. E., … Zheng, H. (2004). Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMH-CIDI). *International Journal of Methods in Psychiatric Research*, *13*(2), 122–139. https://doi.org/10.1002/mpr.169

Kessler, R. C., Merikangas, K. R., Berglund, P., Eaton, W. W., Koretz, D. S., & Walters, E. E. (2003). Mild disorders should not be eliminated from the DSM-V. *Archives of General Psychiatry*, *60*(11), 1117–1122. https://doi.org/10.1001/archpsyc.60.11.1117

Kessler, R. C., & Üstün, T. B. (2004). The World Mental Health (WMH) Survey Initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research*, *13*(2), 93–121. https://doi.org/10.1002/mpr.168

Kessler R. C., Chiu W., Demler O., & Walters E. E. (2005). Prevalence, severity, and comorbidity of 12-month dsm-iv disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, *62*(6), 617–627. https://doi.org/10.1001/archpsyc.62.6.617

Kihlstrom, J. F., Eich, E., Sandbrand, D., & Tobias, B. A. (2000). Emotion and memory: Implications for self-report. In *The science of self-report: Implications for research and practice.* (pp. 81–99). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

*

Klippel, A., Viechtbauer, W., Reininghaus, U., Wigman, J., van Borkulo, C., Myin-Germeys, I., & Wichers, M. (2017). The Cascade of Stress: A Network Approach to Explore Differential Dynamics in Populations Varying in Risk for Psychosis. *Schizophrenia Bulletin*. https://doi.org/10.1093/schbul/sbx037

Knutson, B., Bhanji, J. P., Cooney, R. E., Atlas, L. Y., & Gotlib, I. H. (2008). Neural Responses to Monetary Incentives in Major Depression. *Biological Psychiatry*, *63*(7), 686–692. https://doi.org/10.1016/j.biopsych.2007.07.023

Kohler, C. G., Hoffman, L. J., Eastman, L. B., Healey, K., & Moberg, P. J. (2011). Facial emotion perception in depression and bipolar disorder: A quantitative review. *Psychiatry Research*, *188*(3), 303–309. https://doi.org/10.1016/j.psychres.2011.04.019

Kovacs, M., & Goldston, D. (1991). Cognitive and social cognitive development of depressed children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *30*(3), 388–392. https://doi.org/10.1097/00004583-199105000-00006

Koval, P., Pe, M. L., Meers, K., & Kuppens, P. (2013). Affect dynamics in relation to depressive symptoms: Variable, unstable or inert? *Emotion*, *13*(6), 1132–1141. https://doi.org/10.1037/a0033579

Kramer, I., Simons, C. J. P., Hartmann, J. A., Menne-Lothmann, C., Viechtbauer, W., Peeters, F., … 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. https://doi.org/10.1002/wps.20090

Kret, M. E., & Ploeger, A. (2015). Emotion processing deficits: A liability spectrum providing insight into comorbidity of mental disorders. *Neuroscience & Biobehavioral Reviews*, *52*, 153–171. https://doi.org/10.1016/j.neubiorev.2015.02.011

Kringelbach, M. L., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, *72*(5), 341–372. https://doi.org/10.1016/j.pneurobio.2004.03.006

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kupferberg, A., Bicks, L., & Hasler, G. (2016). Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews*, *69*, 313–332. https://doi.org/10.1016/j.neubiorev.2016.07.002

Kuppens, P., Sheeber, L. B., Yap, M. B. H., Whittle, S., Simmons, J. G., & Allen, N. B. (2012). Emotional inertia prospectively predicts the onset of depressive disorder in adolescence. *Emotion (Washington, D.C.)*, *12*(2), 283–289. https://doi.org/10.1037/a0025046

Lamers, F., van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., van Balkom, A. J. L. M., … Penninx, B. W. J. H. (2011). Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *The Journal of Clinical Psychiatry*, *72*(3), 341–348. https://doi.org/10.4088/JCP.10m06176blu

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & Knippenberg, A. van. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Lejuez, C. W., Hopko, D. R., LePage, J. P., Hopko, S. D., & McNeil, D. W. (2001). A brief behavioral activation treatment for depression. *Cognitive and Behavioral Practice*, *8*(2), 164–175. https://doi.org/10.1016/S1077-7229(01)80022-5

LeMoult, J., Joormann, J., Sherdell, L., Wright, Y., & Gotlib, I. H. (2009). Identification of Emotional Facial Expressions Following Recovery From Depression. *Journal of Abnormal Psychology*, *118*(4), 828–833. https://doi.org/10.1037/a0016944

Lepage, M., Sergerie, K., Pelletier, M., & Harvey, P.-O. (2007). Episodic Memory Bias and the Symptoms of Schizophrenia. *The Canadian Journal of Psychiatry*, *52*(11), 702–709. https://doi.org/10.1177/070674370705201104

Leppänen, J. M. (2006). Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. *Current Opinion in Psychiatry*, *19*(1), 34–39. https://doi.org/10.1097/01.yco.0000191500.46411.00

Leppänen, J. M., & Hietanen, J. K. (2004). Positive facial expressions are recognized faster than negative facial expressions, but why? *Psychological Research*, *69*(1–2), 22–29. https://doi.org/10.1007/s00426-003-0157-2

Leventhal, A. M. (2008). Sadness, depression, and avoidance behavior. *Behavior Modification*, *32*(6), 759–779. https://doi.org/10.1177/0145445508317167

Lewinsohn, P. M., Clarke, G. N., Seeley, J. R., & Rohde, P. (1994). Major Depression in Community Adolescents: Age at Onset, Episode Duration, and Time to Recurrence. *Journal of the American Academy of Child & Adolescent Psychiatry*, *33*(6), 809–818. https://doi.org/10.1097/00004583-199407000-00006

Lewinsohn, P. M., Rohde, P., Klein, D. N., & Seeley, J. R. (1999). Natural Course of Adolescent Major Depressive Disorder: I. Continuity Into Young Adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*(1), 56–63. https://doi.org/10.1097/00004583-199901000-00020

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.

Lewinsohn, P. M., Sullivan, J. M., & Grosscup, S. J. (1980). Changing reinforcing events: An approach to the treatment of depression. *Psychotherapy: Theory, Research & Practice*, *17*(3), 322–334. https://doi.org/10.1037/h0085929

Li, J., & Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, *95*(3), 221–227. https://doi.org/10.1038/sj.hdy.6800717

Liu, X., Hairston, J., Schrier, M., & Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *35*(5), 1219–1236. https://doi.org/10.1016/j.neubiorev.2010.12.012

Lodder, G. M. A., Scholte, R. H. J., Goossens, L., Engels, R. C. M. E., & Verhagen, M. (2015). Loneliness and the social monitoring system: Emotion recognition and eye gaze in a real-life conversation. *British Journal of Psychology*, *107*, 135–153. https://doi.org/10.1111/bjop.12131

Lopez-Duran, N. L., Kuhlman, K. R., George, C., & Kovacs, M. (2013). Facial emotion expression recognition by children at familial risk for depression: high-risk boys are oversensitive to sadness. *Journal of Child Psychology and Psychiatry*, *54*(5), 565–574. https://doi.org/10.1111/jcpp.12005

*

Lorentz, K. (1998). Approved recommendation on IFCC methods for the measurement of catalytic concentration of enzymes. Part 9. IFCC method for alpha-amylase (1,4-alpha-D-glucan 4-glucanohydrolase, EC 3.2.1.1). International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). Committee on Enzymes. *Clinical Chemistry and Laboratory Medicine*, *36*(3), 185–203. https://doi.org/10.1515/CCLM.1998.034

Luciana, M., Wahlstrom, D., Porter, J. N., & Collins, P. F. (2012). Dopaminergic modulation of incentive motivation in adolescence: age-related changes in signaling, individual differences, and implications for the development of self-regulation. *Developmental Psychology*, *48*(3), 844–861. https://doi.org/10.1037/a0027432

Luking, K. R., Pagliaccio, D., Luby, J. L., & Barch, D. M. (2015). Child Gain Approach and Loss Avoidance Behavior: Relationships With Depression Risk, Negative Mood, and Anhedonia. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(8), 643–651. https://doi.org/10.1016/j.jaac.2015.05.010

Luking, K. R., Pagliaccio, D., Luby, J. L., & Barch, D. M. (2016). Reward Processing and Risk for Depression Across Development. *Trends in Cognitive Sciences*, *20*(6), 456–468. https://doi.org/10.1016/j.tics.2016.04.002

Mandel, A. L., Ozdener, H., & Utermohlen, V. (2009). Identification of pro- and mature brain-derived neurotrophic factor in human saliva. *Archives of Oral Biology*, *54*(7), 689–695. https://doi.org/10.1016/j.archoralbio.2009.04.005

Mandel, A. L., Ozdener, H., & Utermohlen, V. (2011). Brain-derived Neurotrophic Factor in Human Saliva: ELISA Optimization and Biological Correlates. *Journal of Immunoassay & Immunochemistry*, *32*(1), 18–30. https://doi.org/10.1080/15321819.2011.538625

Marchand, A., Juster, R.-P., Lupien, S. J., & Durand, P. (2016). Psychosocial determinants of diurnal alpha-amylase among healthy Quebec workers. *Psychoneuroendocrinology*, *66*, 65–74. https://doi.org/10.1016/j.psyneuen.2016.01.005

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion*, *5*(1), 119–124. https://doi.org/10.1037/1528-3542.5.1.119

Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *32*(3), 454–465. https://doi.org/10.1016/j.neubiorev.2007.08.003

Masselink, M., Van Roekel, E., Heininga, V. E., Vrijen, C., Nederhof, E., & Oldehinkel, A. J. (n.d.). The Domains of Pleasure Scale (DOPS): A novel questionnaire to measure anhedonia, Manuscript submitted for publication.

Masunami, T., Okazaki, S., & Maekawa, H. (2009). Decision-making patterns and sensitivity to reward and punishment in children with attention-deficit hyperactivity disorder. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *72*(3), 283–288.

Mathews, A., & MacLeod, C. (2005). Cognitive Vulnerability to Emotional Disorders. *Annual Review of Clinical Psychology*, *1*(1), 167–195. https://doi.org/10.1146/annurev.clinpsy.1.102803.143916

Matsuki, C., To, M., Kondo, Y., Sugiyama, H., Yamamoto, Y., Shimizu, T., … Tsukinoki, K. (2014). Associations between brain-derived neurotrophic factor and estradiol in women's saliva. *Neuroendocrinology Letters*, *35*(3).

Matton, A., de Jong, P., Goossens, L., Jonker, N., Van Malderen, E., Vervaet, M., … Braet, C. (2017). Sensitivity for Cues Predicting Reward and Punishment in Young Women with Eating Disorders. *European Eating Disorders Review*, *25*(6), 501–511. https://doi.org/10.1002/erv.2541

Mazzucchelli, T., Rees, C., & Kane, R. (2009). Behavioral Activation Treatments for Depression in Adults: A Meta-analysis and Review. *Clinical Psychology: Science and Practice*, *16*(4), 383–411. https://doi.org/10.1111/j.1468-2850.2009.01178.x

McCabe, S., & Gotlib, I. (1995). Selective attention and clinical depression: performance on a deployment-of-attention task. *Journal of Abnormal Psychology*, *104*(1), 241–245. https://doi.org/10.1037//0021-843X.104.1.241

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin*, *126*(3), 424–453. https://doi.org/10.1037/0033-2909.126.3.424

McLaughlin, K. A., Borkovec, T. D., & Sibrava, N. J. (2007). The Effects of Worry and Rumination on Affect States and Cognitive Activity. *Behavior Therapy*, *38*(1), 23–38. https://doi.org/10.1016/j.beth.2006.03.003

McLeod, J. D., & Fettes, D. L. (2007). Trajectories of failure: The educational careers of children with mental health problems. *AJS; American Journal of Sociology*, *113*(3), 653–701. https://doi.org/10.1086/521849

McMakin, D. L., Santiago, C. D., & Shirk, S. R. (2009). The time course of positive and negative emotion in dysphoria. *The Journal of Positive Psychology*, *4*(2), 182–192. https://doi.org/10.1080/17439760802650600

Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., … Swendsen, J. (2010). Lifetime Prevalence of Mental Disorders in U.S. Adolescents: Results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(10), 980–989. https://doi.org/10.1016/j.jaac.2010.05.017

Meyer, B., Pilkonis, P. A., & Beevers, C. G. (2004). What's in a (neutral) face? Personality disorders, attachment styles, and the appraisal of ambiguous social cues. *Journal of Personality Disorders*, *18*(4), 320–336. https://doi.org/10.1521/pedi.2004.18.4.320

Meyer, V. J., Lee, Y., Böttger, C., Leonbacher, U., Allison, A. L., & Shirtcliff, E. A. (2015). Experience, cortisol reactivity, and the coordination of emotional responses to skydiving. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00138

Millan, M. J. (2006). Multi-target strategies for the improved treatment of depressive states: Conceptual foundations and neuronal substrates, drug discovery and therapeutic application. *Pharmacology & Therapeutics*, *110*(2), 135–370. https://doi.org/10.1016/j.pharmthera.2005.11.006

Montag, C. (2014). The Brain Derived Neurotrophic Factor and Personality. *Advances in Biology*, *2014*, e719723. https://doi.org/10.1155/2014/719723

*

Montagne, B., Schutters, S., Westenberg, H. G. M., Honk, J. van, Kessels, R. P. C., & Haan, E. H. F. de. (2006). Reduced sensitivity in the recognition of anger and disgust in social anxiety disorder. *Cognitive Neuropsychiatry*, *11*(4), 389–401. https://doi.org/10.1080/13546800444000254

Morgan, J. K., Olino, T. M., McMakin, D. L., Ryan, N. D., & Forbes, E. E. (2013). Neural response to reward as a predictor of increases in depressive symptoms in adolescence. *Neurobiology of Disease*, *52*, 66–74. https://doi.org/10.1016/j.nbd.2012.03.039

Morris, S. B. (2008). Estimating Effect Sizes From Pretest-Posttest-Control Group Designs. *Organizational Research Methods*, *11*(2), 364–386. https://doi.org/10.1177/1094428106291059

Munkholm, K., Vinberg, M., & Kessing, L. V. (2016). Peripheral blood brain-derived neurotrophic factor in bipolar disorder: a comprehensive systematic review and meta-analysis. *Molecular Psychiatry*, *21*(2), 216–228. https://doi.org/10.1038/mp.2015.54

Muthén, L. K., & Muthén, B. O. (1998a). *Mplus User's Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (1998b). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.

Myin-Germeys, I., Klippel, A., Steinhart, H., & Reininghaus, U. (2016). Ecological momentary interventions in psychiatry. *Current Opinion in Psychiatry*, *29*(4), 258–263. https://doi.org/10.1097/YCO.0000000000000255

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*, 1533–1547. https://doi.org/10.1017/S0033291708004947

Nater, U. M., & Rohleder, N. (2009). Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: Current state of research. *Psychoneuroendocrinology*, *34*(4), 486–496. https://doi.org/10.1016/j.psyneuen.2009.01.014

Nater, U. M., Rohleder, N., Schlotz, W., Ehlert, U., & Kirschbaum, C. (2007). Determinants of the diurnal course of salivary alpha-amylase. *Psychoneuroendocrinology*, *32*(4), 392–401. https://doi.org/10.1016/j.psyneuen.2007.02.007

Nelson, B. D., Perlman, G., Klein, D. N., Kotov, R., & Hajcak, G. (2016). Blunted Neural Response to Rewards as a Prospective Predictor of the Development of Depression in Adolescent Girls. *American Journal of Psychiatry*, *173*(12), 1223–1230. https://doi.org/10.1176/appi.ajp.2016.15121524

Niedenthal, P. M., Halberstadt, J. B., Margolin, J., & Innes-Ker, Å. H. (2000). Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology*, *30*(2), 211–222. https://doi.org/10.1002/(SICI)1099-0992(200003/04)30:2<211::AID-EJSP988>3.0.CO;2-3

Nock, M. K., & Kazdin, A. E. (2002). Examination of affective, cognitive, and behavioral factors and suicide-related outcomes in children and young adolescents. *Journal of Clinical Child & Adolescent Psychology*, *31*(1), 48–58. https://doi.org/10.1207/S15374424JCCP3101_07

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology*, *100*(4), 569–582. https://doi.org/10.1037/0021-843X.100.4.569

Nolen-Hoeksema, S., & Aldao, A. (2011). Gender and age differences in emotion regulation strategies and their relationship to depressive symptoms. *Personality and Individual Differences*, *51*(6), 704–708. https://doi.org/10.1016/j.paid.2011.06.012

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking Rumination. *Perspectives on Psychological Science*, *3*(5), 400–424. https://doi.org/10.1111/j.1745-6924.2008.00088.x

Noto, Y., Sato, T., Kudo, M., Kurata, K., & Hirota, K. (2005). The relationship between salivary biomarkers and state-trait anxiety inventory score under mental arithmetic stress: a pilot study. *Anesthesia and Analgesia*, *101*(6), 1873–1876. https://doi.org/10.1213/01.ANE.0000184196.60838.8D

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, *4*(1), 95–102. https://doi.org/10.1038/82959

Oldehinkel, A. J., Hartman, C. A., Oort, F. V. A. van, & Nederhof, E. (2015). Emotion recognition specialization and context-dependent risk of anxiety and depression in adolescents. *Brain and Behavior*, *5*(2), 1–10. https://doi.org/10.1002/brb3.299

Oldehinkel, A. J., Ormel, J., Bosch, N. M., Bouma, E. M. C., Van Roon, A. M., Rosmalen, J. G. M., & Riese, H. (2011). Stressed out? Associations between perceived and physiological stress responses in adolescents: The TRAILS study. *Psychophysiology*, *48*(4), 441–452. https://doi.org/10.1111/j.1469-8986.2010.01118.x

Oldehinkel, A. J., Rosmalen, J. G., Buitelaar, J. K., Hoek, H. W., Ormel, J., Raven, D., … Hartman, C. A. (2015). Cohort Profile Update: The TRacking Adolescents' Individual Lives Survey (TRAILS). *International Journal of Epidemiology*, *44*(1), 76-76n. https://doi.org/10.1093/ije/dyu225

Ormel, J., Raven, D., van Oort, F., Hartman, C. A., Reijneveld, S. A., Veenstra, R., … Oldehinkel, A. J. (2015). Mental health in Dutch adolescents: a TRAILS report on prevalence, severity, age of onset, continuity and co-morbidity of DSM disorders. *Psychological Medicine*, *45*(02), 345–360. https://doi.org/10.1017/S0033291714001469

Ottenbreit, N. D., & Dobson, K. S. (2004). Avoidance and depression: the construction of the Cognitive–Behavioral Avoidance Scale. *Behaviour Research and Therapy*, *42*(3), 293–313. https://doi.org/10.1016/S0005-7967(03)00140-2

Ottenbreit, N. D., Dobson, K. S., & Quigley, L. (2014). An examination of avoidance in major depression in comparison to social anxiety disorder. *Behaviour Research and Therapy*, *56*, 82–90. https://doi.org/10.1016/j.brat.2014.03.005

Out, D., Granger, D. A., Sephton, S. E., & Segerstrom, S. C. (2013). Disentangling sources of individual differences in diurnal salivary α-amylase: Reliability, stability and sensitivity to context. *Psychoneuroendocrinology*, *38*(3), 367–375. https://doi.org/10.1016/j.psyneuen.2012.06.013

Pan, P. M., Sato, J. R., Salum, G. A., Rohde, L. A., Gadelha, A., Zugman, A., … Stringaris, A. (2017). Ventral Striatum Functional Connectivity as a Predictor of Adolescent Depressive Disorder in a Longitudinal Community-Based Sample. *American Journal of Psychiatry*, *174*(11), 1112–1119. https://doi.org/10.1176/appi.ajp.2017.17040430

Pe, M. L., Kircanski, K., Thompson, R. J., Bringmann, L. F., Tuerlinckx, F., Mestdagh, M., … Gotlib, I. H. (2015). Emotion-Network Density in Major Depressive Disorder. *Clinical Psychological Science*, *3*(2), 292–300. https://doi.org/10.1177/2167702614540645

Pelc, K., Kornreich, C., Foisy, M.-L., & Dan, B. (2006). Recognition of Emotional Facial Expressions in Attention-Deficit Hyperactivity Disorder. *Pediatric Neurology*, *35*(2), 93–97. https://doi.org/10.1016/j.pediatrneurol.2006.01.014

Penton-Voak, I. S., Bate, H., Lewis, G., & Munafò, M. R. (2012). Effects of emotion perception training on mood in undergraduate students: randomised controlled trial. *The British Journal of Psychiatry*, *201*(1), 71–72. https://doi.org/10.1192/bjp.bp.111.107086

Penton-Voak, I. S., Thomas, J., Gage, S. H., McMurran, M., McDonald, S., & Munafò, M. R. (2013). Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. *Psychological Science*, *24*(5), 688–697. https://doi.org/10.1177/0956797612459657

Pine, D. S., Cohen, E., Cohen, P., & Brook, J. (1999). Adolescent depressive symptoms as predictors of adult depression: moodiness or mood disorder? *The American Journal of Psychiatry*, *156*(1), 133–135. https://doi.org/10.1176/ajp.156.1.133

Pizzagalli, D. A., Holmes, A. J., Dillon, D. G., Goetz, E. L., Birk, J. L., Bogdan, R., … Fava, M. (2009). Reduced Caudate and Nucleus Accumbens Response to Rewards in Unmedicated Subjects with Major Depressive Disorder. *The American Journal of Psychiatry*, *166*(6), 702–710. https://doi.org/10.1176/appi.ajp.2008.08081201

Pizzagalli, D. A., Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2008). Reduced Hedonic Capacity in Major Depressive Disorder: Evidence from a Probabilistic Reward Task. *Journal of Psychiatric Research*, *43*(1), 76–87. https://doi.org/10.1016/j.jpsychires.2008.03.001

Pizzagalli, D. A., Jahn, A. L., & O'Shea, J. P. (2005). Toward an Objective Characterization of an Anhedonic Phenotype. *Biological Psychiatry*, *57*(4), 319–327. https://doi.org/10.1016/j.biopsych.2004.11.026

Plana, I., Lavoie, M.-A., Battaglia, M., & Achim, A. M. (2014). A meta-analysis and scoping review of social cognition performance in social phobia, posttraumatic stress disorder and other anxiety disorders. *Journal of Anxiety Disorders*, *28*(2), 169–177. https://doi.org/10.1016/j.janxdis.2013.09.005

Polyakova, M., Stuke, K., Schuemberg, K., Mueller, K., Schoenknecht, P., & Schroeter, M. L. (2015). BDNF as a biomarker for successful treatment of mood disorders: A systematic & quantitative meta-analysis. *Journal of Affective Disorders*, *174*, 432–440. https://doi.org/10.1016/j.jad.2014.11.044

Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2016). Attentional Bias for Positive Emotional Stimuli: A Meta-Analytic Investigation. *Psychological Bulletin*, *142*(1), 79–106.

Porcelli, A. J., Lewis, A. H., & Delgado, M. R. (2012). Acute stress influences neural circuits of reward processing. *Decision Neuroscience*, *6*, 157. https://doi.org/10.3389/fnins.2012.00157

Posner, M. I., & Petersen, S. E. (1990). The Attention System of the Human Brain. *Annual Review of Neuroscience*, *13*(1), 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Price, I. R., & Bundesen, C. (2005). Emotional changes in skydivers in relation to experience. *Personality and Individual Differences*, *38*(5), 1203–1211. https://doi.org/10.1016/j.paid.2004.08.003

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. https://doi.org/10.1037/0033-2909.114.3.510

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. SAGE.

Rawal, A., Collishaw, S., Thapar, A., & Rice, F. (2013). 'The risks of playing it safe': a prospective longitudinal study of response to reward in the adolescent offspring of depressed parents. *Psychological Medicine*, *43*(1), 27–38. https://doi.org/10.1017/S0033291712001158

Rector, N. A., & Roger, D. (1997). The stress buffering effects of self-esteem. *Personality and Individual Differences*, *23*(5), 799–808. https://doi.org/10.1016/S0191-8869(97)00095-0

Redlich, R., Dohm, K., Grotegerd, D., Opel, N., Zwitserlood, P., Heindel, W., … Dannlowski, U. (2015). Reward Processing in Unipolar and Bipolar Depression: A Functional MRI Study. *Neuropsychopharmacology*, *40*(11), 2623–2631. https://doi.org/10.1038/npp.2015.110

Revelle, W. (2016). psych: Procedures for Psychological, Psychometric, and Personality Research (Version 1.6.12). Retrieved from https://cran.r-project.org/web/packages/psych/index.html

Rinck, M. (2013). Training approach-avoidance of smiling faces affects emotional vulnerability in socially anxious individuals. *Frontiers in Human Neuroscience*, *7*, 481. https://doi.org/10.3389/fnhum.2013.00481

Rohleder, N., & Nater, U. M. (2009). Determinants of salivary α-amylase in humans and methodological considerations. *Psychoneuroendocrinology*, *34*(4), 469–485. https://doi.org/10.1016/j.psyneuen.2008.12.004

Roiser, J. P., Elliott, R., & Sahakian, B. J. (2012). Cognitive Mechanisms of Treatment in Depression. *Neuropsychopharmacology*, *37*(1), 117–136. https://doi.org/10.1038/npp.2011.183

Rommelse, N. N. J., Geurts, H. M., Franke, B., Buitelaar, J. K., & Hartman, C. A. (2011). A review on cognitive and brain endophenotypes that may be common in autism spectrum disorder and attention-deficit/hyperactivity disorder and facilitate the search for pleiotropic genes. *Neuroscience and Biobehavioral Reviews*, *35*(6), 1363–1396. https://doi.org/10.1016/j.neubiorev.2011.02.015

Rosenthal, M. Z., Kim, K., Herr, N. R., Smoski, M. J., Cheavens, J. S., Lynch, T. R., & Kosson, D. S. (2011). Speed and accuracy of facial expression classification in avoidant personality disorder: a preliminary study. *Personality Disorders*, *2*(4), 327–334. https://doi.org/10.1037/a0023672

Rottenberg, J., Gross, J. J., & Gotlib, I. H. (2005). Emotion context insensitivity in major depressive disorder. *Journal of Abnormal Psychology*, *114*(4), 627–639. https://doi.org/10.1037/0021-843X.114.4.627

Rovine, M. J., & Walls, T. A. (2006). Multilevel Autoregressive Modeling of Interindividual Differences in the Stability of a Process. In T. A. Walls & J. L. Schafer, *Models for Intensive Longitudinal Data*. Oxford University Press.

Salimetrics LLC Company. (2012). Saliva Collection and Handling Advice. 3rd Edition. Retrieved January 14, 2017, from https://www.salimetrics.com/assets/documents/Saliva_Collection_Handbook.pdf

Salomon, K., Clift, A., Karlsdóttir, M., & Rottenberg, J. (2009). Major Depressive Disorder is Associated with Attenuated Cardiovascular Reactivity and Impaired Recovery among Those Free of Cardiovascular Disease. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *28*(2), 157–165. https://doi.org/10.1037/a0013001

Salovey, P., & Mayer, J. D. (1990). Emotional Intelligence. *Imagination, Cognition and Personality*, *9*(3), 185–211. https://doi.org/10.2190/DUGG-P24E-52WK-6CDG

Saruta, J., Fujino, K., To, M., & Tsukinoki, K. (2012). Expression and Localization of Brain-Derived Neurotrophic Factor (BDNF) mRNA and Protein in Human Submandibular Gland. *Acta Histochemica et Cytochemica*, *45*(4), 211–218. https://doi.org/10.1267/ahc.12005

*

Sato, W., Kochiyama, T., Uono, S., & Yoshikawa, S. (2010). Amygdala integrates emotional expression and gaze direction in response to dynamic facial expressions. *NeuroImage*, *50*(4), 1658–1665. https://doi.org/10.1016/j.neuroimage.2010.01.049

Schlund, M. W., Hudgins, C. D., Magee, S., & Dymond, S. (2013). Neuroimaging the temporal dynamics of human avoidance to sustained threat. *Behavioural Brain Research*, *257*, 148–155. https://doi.org/10.1016/j.bbr.2013.09.042

Schönenberg, M., Schneidt, A., Wiedemann, E., & Jusyte, A. (2015). Processing of dynamic affective information in adults with ADHD. *Journal of Attention Disorders*, 1087054715577992. https://doi.org/10.1177/1087054715577992

Schumacher, S., Kirschbaum, C., Fydrich, T., & Ströhle, A. (2013). Is salivary alpha-amylase an indicator of autonomic nervous system dysregulations in mental disorders?—A review of preliminary findings and the interactions with cortisol. *Psychoneuroendocrinology*, *38*(6), 729–743. https://doi.org/10.1016/j.psyneuen.2013.02.003

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods*, *21*(2), 206–221. https://doi.org/10.1037/met0000062

Seidel, E.-M., Habel, U., Kirschner, M., Gur, R. C., & Derntl, B. (2010). The impact of facial emotional expressions on behavioral tendencies in women and men. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(2), 500–507. https://doi.org/10.1037/a0018169

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual Review of Clinical Psychology*, *4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Shin, L. M., & Liberzon, I. (2009). The Neurocircuitry of Fear, Stress, and Anxiety Disorders. *Neuropsychopharmacology*, *35*(1), 169–191. https://doi.org/10.1038/npp.2009.83

Shinohara, K., Honyashiki, M., Imai, H., Hunot, V., Caldwell, D. M., Davies, P., … Churchill, R. (2013). Behavioural therapies versus other psychological therapies for depression. *The Cochrane Database of Systematic Reviews*, (10), CD008696. https://doi.org/10.1002/14651858.CD008696.pub2

Singh, S. D., Ellis, C. R., Winton, A. S. W., Singh, N. N., Leung, J. P., & Oswald, D. P. (1998). Recognition of facial expressions of emotion by children with Attention-Deficit Hyperactivity Disorder. *Behavior Modification*, *22*(2), 128–142. https://doi.org/10.1177/01454455980222002

Sinzig, J., Morsch, D., & Lehmkuhl, G. (2008). Do hyperactivity, impulsivity and inattention have an impact on the ability of facial affect recognition in children with autism and ADHD? *European Child & Adolescent Psychiatry*, *17*(2), 63–72. https://doi.org/10.1007/s00787-007-0637-9

Snippe, E., Simons, C. J., Hartmann, J. A., Menne-Lothmann, C., Kramer, I., Booij, S. H., … Wichers, M. (2016). Change in daily life behaviors and depression: Within-person and between-person associations. *Health Psychology*, *35*(5), 433–441. https://doi.org/10.1037/hea0000312

Snippe, E., Viechtbauer, W., Geschwind, N., Klippel, A., de Jonge, P., & Wichers, M. (2017). The Impact of Treatments for Depression on the Dynamic Network Structure of Mental States: Two Randomized Controlled Trials. *Scientific Reports*, *7*. https://doi.org/10.1038/srep46523

Soares, A. T., Andreazza, A. C., Rej, S., Rajji, T. K., Gildengers, A. G., Lafer, B., … Mulsant, B. H. (2016). Decreased Brain-Derived Neurotrophic Factor in Older Adults with Bipolar Disorder. *The American Journal of Geriatric Psychiatry*, *24*(8), 596–601. https://doi.org/10.1016/j.jagp.2016.02.052

Solomon, R. L. (1980). The opponent-process theory of acquired motivation: The costs of pleasure and the benefits of pain. *American Psychologist*, *35*(8), 691–712. https://doi.org/10.1037/0003-066X.35.8.691

StataCorp. (2013). *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.

Stevens, J. P. (2002). *Applied Multivariate Statistics for the Social Sciences* (4th ed.). New Jersey: Lawrence Erlbaum Associates. Retrieved from http://socialsciencesebooksdownload.blogspot.nl/2013/03/applied-multivariate-statistics-for.html

Stirling, L. J., Eley, T. C., & Clark, D. M. (2006). Preliminary evidence for an association between social anxiety symptoms and avoidance of negative faces in school-age children. *Journal of Clinical Child & Adolescent Psychology*, *35*(3), 431–439. https://doi.org/10.1207/s15374424jccp3503_9

Strand, B. H., Dalgard, O. S., Tambs, K., & Rognerud, M. (2003). Measuring the mental health status of the Norwegian population: A comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36). *Nordic Journal of Psychiatry*, *57*(2), 113–118. https://doi.org/10.1080/08039480310000932

Stringaris, A., Vidal-Ribas Belil, P., Artiges, E., Lemaitre, H., Gollier-Briant, F., Wolke, S., … Paillère-Martinot, M.-L. (2015). The Brain's Response to Reward Anticipation and Depression in Adolescence: Dimensionality, Specificity, and Longitudinal Predictions in a Community-Based Sample. *American Journal of Psychiatry*, *172*(12), 1215–1223. https://doi.org/10.1176/appi.ajp.2015.14101298

Stuhrmann, A., Dohm, K., Kugel, H., Zwanzger, P., Redlich, R., Grotegerd, D., … Dannlowski, U. (2013). Mood-congruent amygdala responses to subliminally presented facial expressions in major depression: associations with anhedonia. *Journal of Psychiatry & Neuroscience: JPN*, *38*(4), 249–258. https://doi.org/10.1503/jpn.120060

Stuhrmann, A., Suslow, T., & Dannlowski, U. (2011). Facial emotion processing in major depression: a systematic review of neuroimaging findings. *Biology of Mood & Anxiety Disorders*, *1*, 10. https://doi.org/10.1186/2045-5380-1-10

Surguladze, S. A., Young, A. W., Senior, C., Brébion, G., Travis, M. J., & Phillips, M. L. (2004). Recognition Accuracy and Response Bias to Happy and Sad Facial Expressions in Patients With Major Depression. *Neuropsychology*, *18*(2), 212–218. https://doi.org/10.1037/0894-4105.18.2.212

Suslow, T., Konrad, C., Kugel, H., Rumstadt, D., Zwitserlood, P., Schöning, S., … Dannlowski, U. (2010). Automatic Mood-Congruent Amygdala Responses to Masked Facial Expressions in Major Depression. *Biological Psychiatry*, *67*(2), 155–160. https://doi.org/10.1016/j.biopsych.2009.07.023

Tajima-Pozo, K., Yus, M., Ruiz-Manrique, G., Lewczuk, A., Arrazola, J., & Montañes-Rada, F. (2016). Amygdala Abnormalities in Adults With ADHD. *Journal of Attention Disorders*. https://doi.org/10.1177/1087054716629213

Takahashi, T., Ikeda, K., Ishikawa, M., Kitamura, N., Tsukasaki, T., Nakama, D., & Kameda, T. (2005). Anxiety, reactivity, and social stress-induced cortisol elevation in humans. *Neuro Endocrinology Letters*, *26*(4), 351–354.

*

Takai, N., Yamaguchi, M., Aragaki, T., Eto, K., Uchihashi, K., & Nishikawa, Y. (2004). Effect of psychological stress on the salivary cortisol and amylase levels in healthy young adults. *Archives of Oral Biology*, *49*(12), 963–968. https://doi.org/10.1016/j.archoralbio.2004.06.007

Tanaka, Y., Ishitobi, Y., Maruyama, Y., Kawano, A., Ando, T., Okamoto, S., … Akiyoshi, J. (2012). Salivary alpha-amylase and cortisol responsiveness following electrical stimulation stress in major depressive disorder patients. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *36*(2), 220–224. https://doi.org/10.1016/j.pnpbp.2011.10.005

Taylor, S. (2007). Electroconvulsive Therapy: A Review of History, Patient Selection, Technique, and Medication Management: *Southern Medical Journal*, *100*(5), 494–498. https://doi.org/10.1097/SMJ.0b013e318038fce0

Telzer, E. H., Fuligni, A. J., Lieberman, M. D., & Galván, A. (2014). Neural sensitivity to eudaimonic and hedonic rewards differentially predict adolescent depressive symptoms over time. *Proceedings of the National Academy of Sciences*, *111*(18), 6600–6605. https://doi.org/10.1073/pnas.1323014111

Tirassa, P., Iannitelli, A., Sornelli, F., Cirulli, F., Mazza, M., Calza, A., … Pacitti, F. (2012). Daily serum and salivary BDNF levels correlate with morning-evening personality type in women and are affected by light therapy. *Rivista Di Psichiatria*, *47*(6), 527–534. https://doi.org/10.1708/1178.13059

Toll, A., & Mané, A. (2015). Brain-derived neurotrophic factor levels in first episode of psychosis: A systematic review. http://dx.doi.org/10.5498/wjp.v5.i1.154

Tomasi, D., & Volkow, N. D. (2012). Abnormal Functional Connectivity in Children with Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, *71*(5), 443–450. https://doi.org/10.1016/j.biopsych.2011.11.003

Treadway, M. T., Bossaller, N., Shelton, R. C., & Zald, D. H. (2012). Effort-Based Decision-Making in Major Depressive Disorder: A Translational Model of Motivational Anhedonia. *Journal of Abnormal Psychology*, *121*(3), 553–558. https://doi.org/10.1037/a0028813

Treadway, M. T., & Zald, D. H. (2011). Reconsidering anhedonia in depression: Lessons from translational neuroscience. *Neuroscience & Biobehavioral Reviews*, *35*(3), 537–555. https://doi.org/10.1016/j.neubiorev.2010.06.006

Trull, T. J., Lane, S. P., Koval, P., & Ebner-Priemer, U. W. (2015). Affective Dynamics in Psychopathology. *Emotion Review*, *7*(4), 355–361. https://doi.org/10.1177/1754073915590617

Trull, T. J., Solhan, M. B., Tragesser, S. L., Jahng, S., Wood, P. K., Piasecki, T. M., & Watson, D. (2008). Affective instability: Measuring a core feature of borderline personality disorder with ecological momentary assessment. *Journal of Abnormal Psychology*, *117*(3), 647–661. https://doi.org/10.1037/a0012532

Tsukinoki, K., Saruta, J., Muto, N., Sasaguri, K., Sato, S., Tan-Ishii, N., & Watanabe, Y. (2007). Submandibular Glands Contribute to Increases in Plasma BDNF Levels. *Journal of Dental Research*, *86*(3), 260–264. https://doi.org/10.1177/154405910708600312

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. https://doi.org/10.1037/0022-3514.86.2.320

Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, socioeconomic status, and birth cohort differences on the children's depression inventory: a meta-analysis. *Journal of Abnormal Psychology*, *111*(4), 578–588.

Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications, and applications. *Psychological Bulletin*, *109*(3), 524–536. https://doi.org/10.1037/0033-2909.109.3.524

Van den Berg, I., Franken, I. H. A., & Muris, P. (2010). A New Scale for Measuring Reward Responsiveness. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00239

Van der Krieke, L., Jeronimus, B. F., Blaauw, F. J., Wanders, R. B. K., Emerencia, A. C., Schenk, H. M., … Jonge, P. D. (2015). HowNutsAreTheDutch (HoeGekIsNL): A crowdsourcing study of mental symptoms and strengths. *International Journal of Methods in Psychiatric Research*. https://doi.org/10.1002/mpr.1495

Van der Krieke, L., Blaauw, F. J., Emerencia, A. C., Schenk, H. M., Slaets, J. P. J., Bos, E. H., … Jeronimus, B. F. (2017). Temporal Dynamics of Health and Well-Being: A Crowdsourcing Approach to Momentary Assessments and Automated Generation of Personalized Feedback. *Psychosomatic Medicine*, *79*(2), 213–223. https://doi.org/10.1097/PSY.0000000000000378

Van Hemel-Ruiter, M. E., de Jong, P. J., Oldehinkel, A. J., & Ostafin, B. D. (2013). Reward-related attentional biases and adolescent substance use: The TRAILS study. *Psychology of Addictive Behaviors*, *27*(1), 142–150. https://doi.org/10.1037/a0028271

Van Meel, C. S., Heslenfeld, D. J., Oosterlaan, J., Luman, M., & Sergeant, J. A. (2011). ERPs associated with monitoring and evaluation of monetary reward and punishment in children with ADHD. *Journal of Child Psychology and Psychiatry*, *52*(9), 942–953. https://doi.org/10.1111/j.1469-7610.2010.02352.x

Van Roekel, E., Bennik, E. C., Bastiaansen, J. A., Verhagen, M., Ormel, J., Engels, R. C. M. E., & Oldehinkel, A. J. (2015). Depressive symptoms and the experience of pleasure in daily life: An exploration of associations in early and late adolescence. *Journal of Abnormal Child Psychology*, 1–11. https://doi.org/10.1007/s10802-015-0090-z

Van Roekel, E., Masselink, M., Vrijen, C., Heininga, V. E., Bak, T., Nederhof, E., & Oldehinkel, A. J. (2016). Study protocol for a randomized controlled trial to explore the effects of personalized lifestyle advices and tandem skydives on pleasure in anhedonic young adults. *BMC Psychiatry*, *16*, 182. https://doi.org/10.1186/s12888-016-0880-z

Van Roekel, E., Vrijen, C., Heininga, V. E., Masselink, M., Bos, E. H., & Oldehinkel, A. J. (2017). An Exploratory Randomized Controlled Trial of Personalized Lifestyle Advice and Tandem Skydives as a Means to Reduce Anhedonia. *Behavior Therapy*, *48*(1), 76–96. https://doi.org/10.1016/j.beth.2016.09.009

Veen, G., Giltay, E. J., Vreeburg, S. A., Licht, C. M. M., Cobbaert, C. M., Zitman, F. G., & Penninx, B. W. J. H. (2012). Determinants of salivary evening alpha-amylase in a large sample free of psychopathology. *International Journal of Psychophysiology*, *84*(1), 33–38. https://doi.org/10.1016/j.ijpsycho.2012.01.005

Veldman, K., Bültmann, U., Stewart, R. E., Ormel, J., Verhulst, F. C., & Reijneveld, S. A. (2014). Mental health problems and educational attainment in adolescence: 9-year follow-up of the TRAILS study. *PLOS ONE*, *9*(7), e101751. https://doi.org/10.1371/journal.pone.0101751

*

Victor, T. A., Furey, M. L., Fromm, S. J., Ohman, A., & Drevets, W. C. (2010). Relationship between amygdala responses to masked faces and mood state and treatment in major depressive disorder. *Archives of General Psychiatry*, *67*(11), 1128–1138. https://doi.org/10.1001/archgenpsychiatry.2010.144

Vrijen, C., Hartman, C. A., Lodder, G. M. A., Verhagen, M., de Jonge, P., & Oldehinkel, A. J. (2016). Lower Sensitivity to Happy and Angry Facial Emotions in Young Adults with Psychiatric Problems. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01797

Vrijen, C., Hartman, C. A., & Oldehinkel, A. J. (2016). Slow identification of facial happiness in early adolescence predicts onset of depression during 8 years of follow-up. *European Child & Adolescent Psychiatry*, 1–12. https://doi.org/10.1007/s00787-016-0846-1

Wang, D. V., & Tsien, J. Z. (2011). Convergent Processing of Both Positive and Negative Motivational Signals by the VTA Dopamine Neuronal Populations. *PLoS ONE*, *6*(2), e17047. https://doi.org/10.1371/journal.pone.0017047

Wang, T.-Y., Lee, S.-Y., Chen, S.-L., Chang, Y.-H., Wang, L.-J., Chen, P. S., … Lu, R.-B. (2016). Comparing clinical responses and the biomarkers of BDNF and cytokines between subthreshold bipolar disorder and bipolar II disorder. *Scientific Reports*, *6*, 27431. https://doi.org/10.1038/srep27431

Weinstock, L. M., Melvin, C., Munroe, M. K., & Miller, I. W. (2016). Adjunctive Behavioral Activation for the Treatment of Bipolar Depression: A Proof of Concept Trial. *Journal of Psychiatric Practice*, *22*(2), 149–158. https://doi.org/10.1097/PRA.0000000000000142

Wells, T. T., & Beevers, C. G. (2010). Biased attention and dysphoria: Manipulating selective attention reduces subsequent depressive symptoms. *Cognition and Emotion*, *24*(4), 719–728. https://doi.org/10.1080/02699930802652388

Whelan, R. (2008). Effective Analysis of Reaction Time Data. *Psychological Record*, *58*(3), 475–482.

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*(1), 191–195. https://doi.org/10.1016/j.jad.2009.11.010

Wichers, M., Simons, C. J. P., 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. https://doi.org/10.1111/j.1600-0447.2011.01749.x

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis. R version package 2.2.1. http://ggplot2.org*. New York: Springer-Verlag.

Wickham, H., Francois, R., Henry, L., & Müller, K. (2017). dplyr: A Grammar of Data Manipulation. R package version 0.7.4. Retrieved from https://CRAN.R-project.org/package=dplyr

Wigman, J. T. W., van Os, J., Borsboom, D., Wardenaar, K.J., Epskamp, S., Klippel, A., … Wichers, M. (2015). Exploring the underlying structure of mental disorders: cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. *Psychological Medicine*, *45*(11), 2375–2387. http://dx.doi.org/10.1017/S0033291715000331

Wilcox, H. C., & Anthony, J. C. (2004). Child and adolescent clinical features as forerunners of adult-onset major depressive disorder: retrospective evidence from an epidemiological sample. *Journal of Affective Disorders*, *82*, 9–20. https://doi.org/10.1016/j.jad.2003.10.007

Wilcox, R. R. (2012). *Introduction to Robust Estimation and Hypothesis Testing*. Academic Press.

Wilke, C. O. (2017). cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2". R package version 0.9.2. Retrieved from https://CRAN.R-project.org/package=cowplot

Williams, M., Moss, S., Bradshaw, J., & Mattingley, J. (2005). Look at me, I'm smiling: Visual search for threatening and nonthreatening facial expressions. *Visual Cognition*, *12*(1), 29–50. https://doi.org/10.1080/13506280444000193

Wingenfeld, K., Schulz, M., Damkroeger, A., Philippsen, C., Rose, M., & Driessen, M. (2010). The diurnal course of salivary alpha-amylase in nurses: an investigation of potential confounders and associations with stress. *Biological Psychology*, *85*(1), 179–181. https://doi.org/10.1016/j.biopsycho.2010.04.005

Wong, D. T. (2009). *Salivary Diagnostics*. John Wiley & Sons.

Wright, S. L., Langenecker, S. A., Deldin, P. J., Rapport, L. J., Nielson, K. A., Kade, A. A., … Zubieta, J.-K. (2009). Gender specific disruptions in emotion processing in younger adults with depression. *Depression and Anxiety*, *26*(2), 182–189. https://doi.org/10.1002/da.20502

Wykes, T., Steel, C., Everitt, B., & Tarrier, N. (2008). Cognitive Behavior Therapy for Schizophrenia: Effect Sizes, Clinical Models, and Methodological Rigor. *Schizophrenia Bulletin*, *34*(3), 523–537. https://doi.org/10.1093/schbul/sbm114

Yang, W., Ding, Z., Dai, T., Peng, F., & Zhang, J. X. (2015). Attention Bias Modification training in individuals with depressive symptoms: A randomized controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry*, *49*(Part A), 101–111. https://doi.org/10.1016/j.jbtep.2014.08.005

Yorbik, O., Mutlu, C., Ozturk, O., Altinay, D. K., Tanju, I. A., & Kurt, I. (2016). Salivary alpha amylase levels in youths with anxiety disorders. *Psychiatry Research*, *235*, 148–153. https://doi.org/10.1016/j.psychres.2015.11.021

Zeiss, A. M., Lewinsohn, P. M., & Muñoz, R. F. (1979). Nonspecific improvement effects in depression using interpersonal skills training, pleasant activity schedules, or cognitive training. *Journal of Consulting and Clinical Psychology*, *47*(3), 427–439.

Zuckerbrot, R. A., & Jensen, P. S. (2006). Improving recognition of adolescent depression in primary care. *Archives of Pediatrics & Adolescent Medicine*, *160*(7), 694–704.

*

# NEDERLANDSE SAMENVATTING (DUTCH SUMMARY)

**Blije gezichten en andere beloningen**

Weinig aandacht voor positieve informatie en veel aandacht voor negatieve informatie als onderliggend mechanisme van depressie

## Achtergrond

Doorgaans laten mensen sterke reacties zien op positieve informatie. Denk bijvoorbeeld aan een lach op het gezicht van iemand die je op straat tegenkomt, of het winnen van een prijs. Maar niet iedereen heeft dit in dezelfde mate; sommige mensen reageren minder sterk op positieve informatie en juist sterker op negatieve informatie. Deze mensen hebben een zogenaamde lage positieve bias. In mijn proefschrift onderzocht ik het verband tussen een lage positieve bias en depressie. Depressie is een veelvoorkomende psychische aandoening, die vaak begint in de adolescentie of jongvolwassenheid, erg belastend is voor degenen die er aan lijden en moeilijk te behandelen is. Zie Box 1 voor extra informatie over depressie. Om beter te begrijpen hoe een depressie ontstaat en mogelijk voorkomen kan worden is het belangrijk te onderzoeken welke mechanismen een rol spelen bij het tot stand komen, de instandhouding en de behandeling van depressie. Ik heb verschillende instrumenten gebruikt om een lage positieve bias als onderliggend mechanisme van depressie te onderzoeken. Hierbij heb ik gegevens die elke twee à drie jaar verzameld waren gecombineerd met fijnmaziger gegevens die drie keer per dag werden verzameld. Ik heb gebruik gemaakt van zelfrapportages, gestructureerde klinische interviews, gegevens afkomstig uit gestandaardiseerde gedragstaken en biologische markers.

## Hoe kun je positieve bias meten?

Veel mensen zijn zich er niet van bewust dat ze anders reageren op positieve dan op negatieve informatie. Je kunt het ze daarom niet simpelweg vragen. De mate van positieve bias wordt meestal gemeten met taken die niet vereisen dat mensen zich bewust zijn van hun bias. In de studies die ik beschrijf in mijn proefschrift werd positieve bias gemeten met verschillende gestandaardiseerde gedragstaken: twee gezichtsemotieherkenningstaken en een beloningstaak. Tijdens de emotieherkenningstaak die gebruikt is in hoofdstuk 3 kregen deelnemers foto's te zien van gezichten die blij, boos, verdrietig of bang keken. De deelnemers moesten zo snel mogelijk op een knop drukken als ze wisten welke gezichtsuitdrukking ze zagen. In hoofdstukken 2 en 5 werd gebruik gemaakt van een ander type gezichtsemotieherkenningstaak; tijdens deze taak zagen deelnemers filmpjes van neutraal kijkende gezichten die geleidelijk in blije, boze, verdrietige of bange gezichten veranderden. Deelnemers moesten wederom zo snel mogelijk

op een knop drukken als ze een gezichtsuitdrukking herkenden. In hoofdstuk 4 heb ik een beloningstaak gebruikt om te meten in hoeverre aandacht wordt beïnvloed door verwachtingen over beloningen. Met deze taak werd de invloed van verwachtingen over beloningen op zowel snelle automatische aandachtsprocessen als langzamere, meer bewuste aandachtsprocessen gemeten. Voor elk van deze drie gebruikte taken heb ik op basis van de reactietijden van de deelnemers hun bias voor positieve informatie berekend. Het relatief langzaam herkennen van blije gezichtsemoties vergeleken met negatieve gezichtsemoties en een relatieve ongevoeligheid voor beloningen zijn beide voorbeelden van een lage positieve bias.

**Box 1: Extra informatie over depressie**

**Wanneer krijgt iemand de diagnose depressie?**
Depressie (*Major Depressive Disorder*, MDD), is een veelvoorkomende psychiatrische stoornis. Volgens het handboek van psychiatrische stoornissen de DSM-5 moet voor een diagnose depressie tenminste twee weken lang vrijwel dagelijks sprake zijn van tenminste vijf van de volgende negen symptomen:

(1) Sombere stemming; (2) Verlies van plezier of interesse in bijna alle activiteiten; (3) Aanzienlijk gewichtsverlies of gewichtstoename; (4) Slapeloosheid of veelvuldig slapen; (5) Psychomotorische gejaagdheid of geremdheid; (6) Vermoeidheid of verlies van energie; (7) Gevoelens van waardeloosheid of buitensporige schuldgevoelens; (8) Verminderd vermogen om na te denken, zich te concentreren of beslissingen te nemen; (9) Herhaalde gedachten over de dood of gedachten aan zelfmoord.

Sombere stemming en verlies van plezier zijn zogenaamde kernsymptomen van depressie. Om de diagnose depressie te krijgen moet iemand in ieder geval één van deze twee kernsymptomen ervaren. De diagnose depressie wordt alleen gegeven als de symptomen leiden tot aanzienlijke beperkingen in iemands dagelijkse functioneren.

## Voor dit proefschrift gebruikte gegevens

Voor hoofdstukken 3 en 4 heb ik gebruik gemaakt van gegevens die zijn verzameld binnen de grote longitudinale studie TRAILS – dit staat voor *TRacking Adolescents' Individual Lives Survey* (www.trails.nl). Binnen TRAILS wordt mentale gezondheid en sociale ontwikkeling bestudeerd van de vroege adolescentie tot in de volwassenheid. Deelnemers worden iedere twee tot drie jaar gemeten, vanaf hun elfde jaar. Hoofdstuk 3 is gebaseerd op gegevens van 1840 deelnemers en voor de taak die gebruikt is in hoofdstuk 4 waren gegevens van 531 deelnemers beschikbaar.

Voor hoofdstukken 2, 5, 6 en 7 heb ik gegevens gebruikt die we hebben verzameld binnen het *No Fun No Glory* onderzoek, een studie naar biologische, psychologische en sociale mechanismen van verlies van plezier in de adolescentie en vroege volwassenheid (www.nofunnoglory.nl). Binnen dit onderzoek hebben we door middel van vragenlijsten en een gezichtsemotieherkenningstaak gegevens verzameld van 2577 jongeren (screening). Vervolgens hebben we hieruit 69 deelnemers met aanhoudend verlies van plezier geselecteerd voor een intensieve interventiestudie, waarin ze drie maanden lang drie keer per dag dagboekjes invulden. Ook 69 deelnemers zonder verlies van plezier hebben de dagboekjes een maand lang ingevuld. Voor hoofdstuk 2 zijn de screeningsgegevens gebruikt en voor hoofdstuk 5, 6 en 7 de gegevens van de interventiestudie.

## Wat al bekend was en welke vragen ik wilde beantwoorden in mijn proefschrift

Voordat ik aan mijn proefschrift begon was er al bewijs dat mensen die depressief zijn minder sterk op positieve informatie reageren en juist sterker op negatieve informatie. Maar er was ook nog veel onbekend. Hieronder bespreek ik de vragen die nog onvoldoende beantwoord waren door andere studies, hoe ik ze heb geprobeerd te beantwoorden in mijn proefschrift en wat de uitkomsten waren.

### 1. Gaat een lage positieve bias vooraf aan de ontwikkeling van depressieve klachten?

Eerdere studies hebben niet alleen laten zien dat mensen met een depressie een minder sterke positieve bias hebben dan mensen zonder depressieve klachten, maar ook eerste aanwijzingen opgeleverd dat een dergelijk gebrek aan positieve bias al aanwezig is voorafgaand aan de depressieve klachten. Deze eerste aanwijzingen waren gebaseerd op kleine studies en er was verder onderzoek nodig om te onderzoeken of een lage positieve bias daadwerkelijk voorafgaat aan depressie. Een lage positieve bias zou een kwetsbaarheid voor depressie kunnen betekenen en een oorzaak kunnen zijn bij het ontstaan van depressie.

Deze vraag heb ik onderzocht in hoofdstukken 3 en 4. Ik vond dat jongeren die op elfjarige leeftijd langzamer zijn in het herkennen van blije gezichtsemoties een grotere kans hebben op een depressie in de daaropvolgende acht jaar. Het bleek hierbij met name relevant te zijn hoe snel jongeren blije gezichtsemoties herkenden ten opzichte van hoe snel ze verdrietige gezichtsemoties herkenden. Ook vond ik dat zestienjarigen die ongevoelig waren voor beloningen een grotere kans hadden om in de daaropvolgende negen jaar depressief te worden. De kans op depressie hing bij deze jongeren vooral samen met de moeite die zij tijdens de beloningstaak hadden om hun aandacht te verplaatsen van een negatieve situatie (beloning onwaarschijnlijk) naar een positieve situatie (beloning waarschijnlijk). Het effect werd alleen gevonden als de jongeren heel snel op de informatie over de beloning moesten reageren; niet

als ze meer tijd kregen om na te denken over hun reactie. Het relatief langzaam herkennen van de blije gezichtsemoties van andere mensen vergeleken met de snelheid waarmee negatieve gezichtsemoties worden herkend (hoofdstuk 3) en een relatieve ongevoeligheid voor beloningen (hoofdstuk 4) zijn beide kenmerken van een lage positieve bias. Kortom, in beide hoofdstukken vond ik dat een lage positieve bias voorafgaat aan depressie. Mogelijk reflecteert een lage positieve bias een kwetsbaarheid voor depressie.

## 2. Hoe specifiek is een lage positieve bias voor depressieve klachten?

Het was op basis van eerdere studies nog niet duidelijk hoe specifiek een lage positieve bias is voor depressie. Omdat mensen met depressieve klachten vaak ook andere psychische problemen hebben is het mogelijk dat het eerder gevonden verband tussen lage positieve bias en depressie veroorzaakt werd door psychische klachten die vaak samen gaan met een depressie, zoals angstklachten.

Deze vraag heb ik onderzocht in hoofdstukken 2, 3 en 4. In hoofdstuk 2 onderzocht ik de samenhang tussen gezichtsemotieherkenning en vijf psychische aandoeningen: depressie, angst, vermijdingsproblematiek, ADHD (aandachtstekortstoornis met hyperactiviteit) en antisociaal gedrag. Jongeren met antisociale gedragsproblemen en vermijdingsproblematiek waren minder gevoelig voor blije gezichtsemoties. ADHD symptomen en vermijdingsproblematiek hingen samen met een lagere gevoeligheid voor boze gezichtsemoties. Ik vond geen samenhang tussen gevoeligheid voor gezichtsemoties en depressie en angst en kon daarom niet onderzoeken of dit verband specifiek was voor depressie. In hoofdstukken 3 en 4 vond ik het al eerder beschreven prospectieve verband tussen lage positieve bias en depressie. Hier heb ik de specificiteit wel getoetst. De gevonden verbanden bleken specifiek te zijn voor depressie en konden niet beter verklaard worden door gelijktijdig aanwezige andere psychische klachten.

## 3. Is een lage positieve bias specifiek voor één van beide kernsymptomen van depressie?

De kernsymptomen van depressie zijn verlies van plezier en depressieve stemming. Om een diagnose van depressie te krijgen moet een persoon, naast een aantal andere criteria, tenminste één van deze twee kernsymptomen hebben. Er zijn aanwijzingen dat de twee kernsymptomen van depressie deels verschillende oorzaken hebben. Daarom is het denkbaar dat een lage positieve bias samenhangt met het ene kernsymptoom maar niet met het andere kernsymptoom.

In hoofdstuk 3 onderzocht ik of het gevonden prospectieve verband tussen een lage positieve bias en depressie specifiek is voor verlies van plezier of voor depressieve stemming. Ik vond dat een lage positieve bias alleen het kernsymptoom verlies van plezier voorspelde en niet het kernsymptoom depressieve stemming.

## 4. Wat betekent een lage positieve bias in het dagelijks leven?

Er was nog niet veel bekend over wat het nu precies betekent als mensen een lage positieve bias laten zien tijdens gedragstaken. Hangt dit ook samen met bepaalde gedragingen en emoties in het dagelijks leven? Als we beter begrijpen hoe het dagelijks leven van mensen met een hoge positieve bias verschilt van dat van mensen met een lage positieve bias kunnen we meer inzicht krijgen in waarom individuen met een lage positieve bias vatbaarder zijn voor depressie.

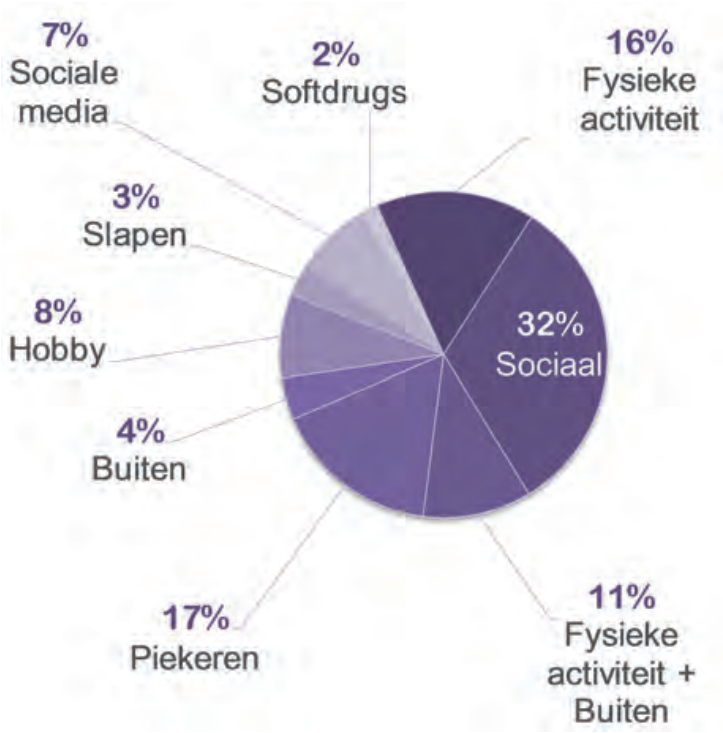
In hoofdstuk 5 onderzocht ik of bias voor blije gezichten tijdens een laboratoriumtaak samenhing met de dynamiek tussen positieve en negatieve gebeurtenissen, gedachten en emoties in het dagelijks leven van jongvolwassenen. Denk bijvoorbeeld aan vragen als hoe een positieve gebeurtenis gedurende de ochtend de hoeveelheid positieve emoties in de middag beïnvloedt; en of dit verschilt tussen mensen die tijdens de laboratoriumtaak een lage en een hoge positieve bias laten zien. Gebeurtenissen, gedachten en emoties werden gedurende een periode van dertig dagen drie keer per dag uitgevraagd. Vergeleken met leeftijdsgenoten met een lage bias voor blije gezichten, lieten deelnemers met een hoge bias voor blije gezichten een sterker effect zien van positieve emoties en positieve ervaringen op zes uur later gerapporteerde positieve emoties en ervaringen. Zij hielden langer vast aan hun positieve ervaringen. Dit kan mogelijk verklaren waarom mensen met een lage bias voor blije gezichten meer kans hebben om depressief te worden; mensen met een hoge positieve bias zijn wellicht beter in staat hun positieve ervaringen optimaal te gebruiken in het dagelijks leven. Er zijn aanwijzingen uit eerder onderzoek dat mensen die een positieve stemming langer vast kunt houden meer open staan voor sociale contacten en het leren van nieuwe dingen. Dit zou er voor kunnen zorgen dat ze minder snel in een negatieve spiraal terechtkomen.

## 5. Hoe kunnen we informatie over de wisselwerking tussen leefstijlfactoren en emoties in het dagelijks leven gebruiken om depressieve symptomen te verminderen en positieve bias te verhogen?

Kunnen we jongvolwassenen die last hebben van verlies van plezier helpen hun plezier terug te krijgen door ze een op de persoon toegesneden leefstijladvies te geven gebaseerd op wat voor hen specifiek samenhangt met plezier in het dagelijks leven? En zijn we in staat om op deze manier ook de lage positieve bias te verhogen? Dit laatste zou kunnen betekenen dat niet alleen depressieve klachten maar ook vatbaarheid voor depressieve klachten verminderd zou kunnen worden. Binnen het *No Fun No Glory* project hebben we ons voor het beantwoorden van deze vragen laten inspireren door een beproefde behandelmethode voor depressie, namelijk gedragsactivatie. Dit is een veelgebruikte behandelmethode waarbij patiënten gedurende een bepaalde periode dagelijks hun activiteiten en stemming bijhouden om op deze manier inzicht te krijgen in het verband tussen beide en aangespoord worden tot het doen van meer activiteiten die samenhangen met een positieve stemming. Deze aanpak lijkt te werken als behandeling voor depressie, maar het was nog onduidelijk of de aanpak ook specifiek werkte voor verlies van plezier en een lage positieve bias. Ook was nog onduidelijk of één keer per dag

activiteiten en stemming rapporteren genoeg is; meerdere keren per dag rapporteren kan meer inzicht geven in de dynamiek tussen activiteiten en stemming. Met de huidige technieken en statistische methodes is vaker rapporteren ook goed mogelijk.

**Box 2: De gegeven leefstijladviezen**




We hebben het vaakst advies gegeven om meer sociale activiteiten te ondernemen, gevolgd door meer te bewegen, bijvoorbeeld te sporten of te wandelen. Deelnemers bij wie piekeren samenhing met minder plezier hebben we geadviseerd mindfulness oefeningen te doen of iedere dag een pieker-halfuurtje in te plannen. Ook kreeg een aantal deelnemers het advies om, in plaats van televisie te kijken of actief te zijn op sociale media, met hun hobby bezig te zijn of af te spreken met vrienden.

In hoofdstuk 6 heb ik op basis van gegevens van 69 jongvolwassenen onderzocht wat de effecten zijn van gepersonaliseerd leefstijladvies op verlies van plezier. Deelnemers rapporteerden 30 dagen lang drie keer per dag leefstijlfactoren en stemming. Ons onderzoeksteam heeft deelnemers persoonlijk leefstijladvies gegeven dat gebaseerd was op de voor elke deelnemer specifieke verbanden tussen leefstijlfactoren en plezier gedurende deze 30 dagen. Zie Box 2 voor de verschillende soorten advies die gegeven zijn. Deelnemers werd gevraagd het leefstijladvies uit te voeren gedurende een nieuwe periode van 30 dagen waarin ze wederom drie keer per dag leefstijlfactoren en plezier rapporteerden. Dit onderzoek heeft aanwijzingen opgeleverd dat

met behulp van dit type op de persoon toegesneden gedragsactivatie een toename van plezier en positief affect in het dagelijks leven bereikt kan worden. Hoewel de effecten klein waren is dit toch een belangrijke bevinding vanwege het feit dat verlies van plezier een symptoom is dat lastig te bestrijden is. Mogelijk kan een dergelijke aanpak ook werken in de klinische praktijk; het lijkt zinvol om hier verder onderzoek naar te doen en de aanpak verder te verbeteren. Een belangrijke uitkomst is daarnaast dat het haalbaar bleek om jongvolwassenen met verlies van plezier drie keer per dag gedurende langere tijd vragenlijstjes te laten invullen met lage uitval en hoge invulpercentages. Ik vond geen aanwijzingen dat onze adviezen ook hielpen om de positieve bias van jongvolwassenen met verlies van plezier te verhogen.

## 6. Kan een ultiem spannende vrije-val ervaring helpen om de motivatie te verhogen van jongvolwassenen die verlies van plezier ervaren?

Uit eerder onderzoek is gebleken dat verlies van plezier vaak samen gaat met verlies van motivatie voor dagelijkse bezigheden. Een lage motivatie zou er voor kunnen zorgen dat de leefstijladviezen die we in het *No Fun No Glory* project gaven aan mensen met verlies van plezier niet zouden worden opgevolgd. Daarom wilden we onderzoeken of het mogelijk was die motivatie te stimuleren. Het was bekend dat een parachutesprong sterke fysiologische en psychologische effecten heeft en onderzoek met muizen suggereert dat een vrije-val ervaring zou kunnen resulteren in verhoogde motivatie. Wij hebben jongvolwassenen een tandem parachutesprong laten doen om hun motivatie te verhogen voor het opvolgen van de leefstijladviezen. Onderdeel van dit onderzoek was met de biologische markers alfa-amylase en BDNF (brain-derived neurotrophic factor) te onderzoeken of jongvolwassenen met verlies van plezier de verwachte reacties op stress rondom de parachutesprong lieten zien en welke individuele verschillen hierin bestonden. Alfa-amylase is een enzym dat sterk reageert op acute stress en BDNF is een eiwit dat zorgt voor verbindingen in ons brein die nodig zijn voor de signaaloverdracht tussen zenuwcellen en een rol speelt in de regulatie van reacties op stress.

In hoofdstuk 6 onderzocht ik of een tandem parachutesprong de motivatie van jongvolwassenen met verlies van plezier zou kunnen verhogen om de leefstijladviezen beter uit te voeren en of dit voor hen zou leiden tot hogere plezierniveaus en een hogere positieve bias. Ik vond geen effect. De parachutesprong gecombineerd met het leefstijladvies zorgde niet voor hogere plezierniveaus dan het leefstijladvies alleen en ook niet voor een hogere positieve bias. Een mogelijke verklaring hiervoor zou kunnen zijn dat mensen met verlies van plezier minder sterke fysiologische reacties laten zien op een parachutesprong en de sprong daarom niet effectief was. Dit lijkt echter op basis van hoofdstuk 7 geen plausibele verklaring. In hoofdstuk 7 onderzocht ik binnen de groep van 61 jongvolwassenen met verlies van plezier die de tandem parachutesprong hadden gedaan de alfa-amylase toename als reactie op de parachutesprong (stress reactiviteit) en de afname van alfa-amylase na de parachutesprong (herstel van de stress). Alfa-amylase werd gemeten met salivettes – dit zijn tamponvormige watjes die je in je mond moet houden totdat ze zijn volgezogen met speeksel – op vier verschillende tijdstippen (twee

keer voor en twee keer na de parachutesprong). Vrijwel alle deelnemers lieten vergelijkbare alfa-amylase reactiviteit en herstel patronen zien, hoewel gemiddelde alfa-amylase niveaus wel erg varieerden. Deze studie had geen controlegroep, maar de in deze studie gevonden alfa-amylase patronen in mensen met verlies van plezier lijken grotendeels gelijk aan de patronen die andere studies vonden voor gezonde individuen. Mensen met verlies van plezier lijken dus niet heel anders te reageren op een parachutesprong dan mensen zonder verlies van plezier.

Als gezegd was er ook het plan om onderzoek te doen naar BDNF-niveaus rondom de parachutesprong. Om te weten of BDNF betrouwbaar kon worden gemeten in speeksel moest er eerst een pilot studie worden gedaan. Voor onderzoek naar BDNF worden vaak zogenaamde ELISA-kits gebruikt. De commerciële ELISA-kits die op de markt zijn zijn gevalideerd voor het bepalen van BDNF in bloed (serum en plasma), maar niet in speeksel. Voor hoofdstuk 8 heb ik drie verschillende commercieel beschikbare ELISA-kits gebruikt om speeksel en bloed (plasma) te analyseren dat op verschillende momenten was afgenomen bij zes gezonde volwassenen. In totaal zijn 33 plasma- en 33 speekselmonsters geanalyseerd waarbij telkens de standaardprotocollen van de fabrikanten werden gevolgd. Met de drie commercieel beschikbare ELISA-kits die we hebben getest bleek BDNF niet betrouwbaar gemeten te kunnen worden in speeksel. Hoewel er studies zijn die wel uitkomsten van BDNF in speeksel gemeten met dergelijke kits hebben gerapporteerd, lijkt het er op dat in een groot aantal van deze studies ruis is geïnterpreteerd als een valide meting van BDNF. Omdat het tijdens de pilot niet lukte BDNF te meten in speeksel, werden de voorgenomen BDNF-analyses geannuleerd.

## Sterke kanten en beperkingen

Het is een sterke kant van dit proefschrift dat de studies over verbanden tussen positieve bias en depressie gebaseerd zijn op grote groepen deelnemers. Een ander sterk punt is dat de bevinding dat jongeren met een lage positieve bias een groter risico hebben op het ontwikkelen van een depressie gebaseerd is op verschillende soorten taken waarmee positieve bias gemeten is en dat dit verband gevonden werd voor verschillende leeftijdsgroepen. Verder zijn binnen het grotere *No Fun No Glory* project waar dit proefschrift onderdeel van uitmaakt twee veelbelovende interventies ontwikkeld en getest. De combinatie van maandelijkse vragenlijsten, taken en biologische markers met de vragenlijsten (dagboekjes) die drie keer per dag werden ingevuld heeft, mede omdat we vrijwel geen uitval van deelnemers hadden en hoge invulpercentages, geresulteerd in rijke en waardevolle gegevens op basis waarvan ik in mijn proefschrift een lage positieve bias vanuit verschillende invalshoeken heb kunnen belichten.

De studies die ik presenteer in mijn proefschrift kennen uiteraard ook beperkingen. Ik zal hier twee belangrijke beperkingen toelichten. De eerste beperkende factor is dat het voor de taken die ik heb gebruikt voor het meten van de mate van positieve bias lastig is om een heel specifiek mechanisme aan te wijzen dat verantwoordelijk is voor de bias. Zo is het bijvoorbeeld op basis van de gebruikte emotieherkenningstaken niet mogelijk om te achterhalen of een

langzame herkenning van blije gezichtsemoties komt door een verminderde waarneming van blije gezichtsemoties of door een verminderde aandacht voor blije gezichtsemoties. Hoewel aandacht niet specifiek wordt gemeten in de gezichtsemotieherkenningstaken hebben alle gebruikte taken met elkaar gemeen dat aandacht een rol kan spelen. Een ander voorbeeld van waarom het lastig is op basis van de gebruikte taken onderscheid te maken tussen verschillende cognitieve mechanismen is dat de in hoofdstuk 4 beschreven taak geen neutrale conditie heeft. Ik vond dat jongeren die op zestienjarige leeftijd meer moeite hadden om hun aandacht te verplaatsen van een negatieve situatie naar een positieve situatie vaker een depressie ontwikkelden in de daaropvolgende negen jaar. Door het ontbreken van een neutrale conditie in de taak is het echter onduidelijk of jongeren met een verhoogd risico op depressie met name problemen hadden met het loslaten van negatieve situaties of het verplaatsen van aandacht naar positieve situaties.

De tweede beperking die ik wil bespreken gaat over de leefstijlinterventie. Deelnemers aan de *No Fun No Glory* interventiestudie wisten zelf uiteraard of ze leefstijladvies kregen of niet en het kan zijn dat deelnemers alleen op basis van de groep waarin ze waren ingedeeld al bepaalde verwachtingen hadden over de uitkomst. Ik kan niet uitsluiten dat deze verwachtingen voor een deel gezorgd hebben dat de groep die een leefstijladvies kreeg een grotere toename in plezier rapporteerde dan de groep die dit advies niet kreeg. Dit wordt ook wel een placebo-effect genoemd. Daarnaast kan het leefstijladvies dat werd gegeven niet los worden gezien van de gesprekken waarin het advies met de deelnemers werd besproken. Het is mogelijk dat de toename in plezier voor een deel toe te schrijven is aan algemene 'therapeutische' kenmerken in plaats van aan het opvolgen van de leefstijladviezen.

## Aanbevelingen voor toekomstig onderzoek

Misschien wel de belangrijkste aanbeveling voor toekomstig onderzoek is het repliceren van een aantal bevindingen uit dit proefschrift. Als gerepliceerd kan worden dat een lage positieve bias tijdens de adolescentie een voorspeller is voor depressie later in het leven is het belangrijk te gaan onderzoeken wat de voorspellers zijn van een lage positieve bias, of deze bias te veranderen valt en of dit vervolgens de kans op depressie verlaagt.

Het is belangrijk te onderzoeken hoe voorkeur voor positieve informatie gevormd wordt gedurende de ontwikkeling van kind tot adolescent. Er zijn aanwijzingen dat bepaalde soorten positieve bias, zoals beloningsgevoeligheid, een piek laten zien in de adolescentie vergeleken met de kindertijd en de volwassenheid. Het is nog onduidelijk of het ontbreken van een dergelijke piek een voorspeller is voor depressie en of het ontbreken daarvan alleen zichtbaar is gedurende de adolescentie of ook al voorspeld kan worden door een lage beloningsgevoeligheid tijdens de kindertijd. Het is verder ook nog onduidelijk of en hoe een positieve bias wordt doorgegeven van ouders op kinderen. Kennis over de voorspellers van een lage positieve bias kan inzicht geven in mogelijkheden voor preventie.

Een vernieuwende manier om de voorkeur voor positieve informatie van jongeren actief te beïnvloeden is een combinatie van algemeen leefstijladvies en advies in het moment. Jongeren zouden bijvoorbeeld directe feedback in situaties in hun dagelijks leven kunnen krijgen via hun smartphone. Te denken valt aan instructies om hun positieve stemming vast te houden op momenten dat ze een positieve stemming rapporteren, of advies om positieve activiteiten te zoeken op momenten waarop ze een negatieve stemming rapporteren of zich overmatig piekeren. Dit zou hun positieve bias kunnen helpen verhogen. De voorgestelde methode verschilt van de meer gangbare methodes waarbij gedragstaken gebruikt worden om bias te beïnvloeden in een laboratoriumsituatie.

# DANKWOORD

Er zijn veel mogelijke werelden waarin dit proefschrift nooit zou zijn voltooid en ik wil hierbij iedereen heel erg bedanken die heeft bijgedragen aan de totstandkoming van dit proefschrift.

Ik wil beginnen mijn geweldige promotoren te bedanken. Dit proefschrift was er niet geweest als mijn eerste promotor het niet had aangedurfd mij met mijn destijds wat beperkte statistische kennis aan te nemen als promovendus op het No Fun No Glory project.

*Tineke*, ik vind het nog steeds heel bijzonder dat je mij destijds hebt aangenomen en wil je bedanken voor de geweldige en inspirerende begeleiding die je al die jaren hebt gegeven en voor de fijne werksfeer waarin altijd veel ruimte was voor het ontwikkelen van eigen ideeën. Ik had het gevoel dat ik alles aan je kon voorleggen en ook best eens minder slimme vragen mocht stellen. De snelheid waarmee je altijd feedback gaf, zelfs als je het heel erg druk had, is werkelijk indrukwekkend en ik heb bewondering voor de manier waarop je mijn stukken altijd beter maakte. Je hebt niet alleen het schrijven van mijn proefschrift mogelijk gemaakt, maar je bent daarnaast ook in een hele belangrijke mate verantwoordelijk voor het plezier dat ik heb gehad tijdens mijn promotietraject!

*Catharina*, mijn afspraken met jou en Tineke waren altijd heel leuk, interessant en inspirerend. Daarnaast had ik het gevoel dat ik op ieder moment je kamer kon binnenvallen als ik met een conceptueel of statistisch probleem worstelde. Ik realiseer me dat dat niet altijd goed uit kwam en vind het geweldig dat je er toch steeds de tijd voor nam om met me mee te denken. Met jou brainstormen over een probleem zorgde er altijd voor dat ik weer een stapje verder kwam. Je hebt je passie voor onderzoek doen zeker ook op mij overgedragen. Bedankt voor alles!

*Peter*, ik wil je heel erg bedanken voor de fijne samenwerking. Als copromotor heb ik wat minder intensief met jou samengewerkt, maar ik vond de gesprekken die we hadden altijd erg prettig en je net andere insteek was zeker een waardevolle aanvulling.

Dan natuurlijk de rest van ons No Fun No Glory team….

*Eeske, Vera en Maurits*, jullie waren geweldig om mee samen te werken. Ik heb genoten van al onze discussies, maar ook van het feit dat we goed konden lachen samen, en natuurlijk van die doodenge parachutesprong die we hebben gedaan met het team en onze strijd om de beste 1 april grap (ik zou de komende jaren maar een beetje alert blijven als ik jullie was...). Ik hoop ergens in de toekomst weer met jullie samen te kunnen werken!

*

*Eeske en Tina*, jullie waren mijn kamergenoten toen ik begon en *Jennifer* de laatste jaren. Bedankt voor alle inhoudelijke feedback en leuke momenten. En vooral voor *Tina* en *Jennifer*, ik weet hoe graag jullie de bende op mijn bureau hadden willen ordenen in mooie stapeltjes, maar bedankt dat jullie dit niet hebben gedaan. *Tina*, ik vind het fantastisch om nu als postdoc te mogen werken op jouw ERC project en ben heel blij dat ik op deze manier mijn wetenschappelijke carrière kan voortzetten.

Verder wil ik natuurlijk ook iedereen van het ICPE bedanken voor de samenwerking, de fijne sfeer op de afdeling en de borrels. *Dennis, Margo en Esther* wil ik nog extra bedanken voor de fijne samenwerking en alle supersnelle hulp met data en administratieve zaken. En *Tom*, je hebt al weer een tijdje een nieuwe baan maar in de periode dat jij deel uitmaakte van ons team heeft het grootste deel van de dataverzameling plaatsgevonden en was je werkelijk onmisbaar. *Esther* (Nederhof), jij was in de beginfase ook betrokken bij het No Fun No Glory project en ik wil je bedanken voor de fijne samenwerking in deze periode. *Elske*, jij hebt ons de beginselen van VAR en ARIMA bijgebracht. Je filmpjes waren heel erg nuttig en ook wil ik je bedanken voor je bereidheid om met ons mee te denken over alle te volgen stappen van analyse tot leefstijladvies. *Ando* en *Rob*, bedankt voor jullie bijdrage aan de analyses en de visuele weergave van de resultaten voor de deelnemers die een leefstijladvies kregen. *Evelien*, heel erg fijn dat ik altijd kon aankloppen met statistische vragen. Ik heb het heel erg naar mijn zin gehad bij het ICPE en vind het dan ook erg leuk dat ik tijdens mijn postdoc bij Pedagogiek met TRAILS Next data ga werken en op deze manier een connectie met het ICPE houd.

*Laura Bringmann*, *Sacha Epskamp* en *Wolfgang Viechtbauer* wil ik bedanken voor de stimulerende methodologische discussies over netwerkanalyses, mlVAR en permutatietoetsen. Jullie input is van grote waarde geweest voor hoofdstuk 5.

Ik wil alle participanten bedanken die deel hebben genomen aan de TRAILS en No Fun No Glory onderzoeken. Zonder jullie inzet was dit proefschrift er niet geweest! Ik heb verschrikkelijk veel van jullie geleerd. Deelnemers aan de tandem parachutesprong, jullie zagen er zo veel dapperder uit dan ikzelf toen het uiteindelijk ook mijn beurt was.

*Tandemmasters* van Skydive Hoogeveen, bedankt voor al jullie geduld met onze speekselafnameprocedures en jullie interesse in ons project.

Ik wil iedereen van de UMCG labs met wie ik heb samengewerkt bedanken. In het bijzonder *Roelof* voor de goede coördinatie en *Jan* (Koerts) voor alle uitleg en hulp bij het interpreteren van resultaten.

Op 1 augustus ben ik begonnen als postdoc bij Pedagogiek en ik wil mijn nieuwe collega's bedanken dat ze zo aardig en sociaal zijn en daarmee mijn overgang naar de nieuwe werkplek zo makkelijk hebben gemaakt. Ik kijk er naar uit om de komende jaren met jullie samen te werken!

Tenslotte wil ik mijn familie en vrienden bedanken.

Mijn ouders hebben uiteraard de belangrijkste voorwaarden geschapen voor de totstandkoming van dit proefschrift. Dat valt natuurlijk eigenlijk helemaal niet in woorden uit te drukken. *Lieve mam en Henk, pap en Bas*, bedankt voor alles wat jullie me hebben geleerd en voor jullie liefde en steun! Ik weet dat jullie er altijd voor me zijn en dat is heel veel waard.

*Woodie*, je bent het geweldigste broertje dat ik me kan voorstellen. Ik kan altijd op je rekenen en ik kan ook verschrikkelijk lachen met jou. Door de afstand zien we elkaar de laatste tijd wat weinig, maar mijn goede voornemen is om vaker langs te komen in Zagreb. Ik wil jou en *Stefi*, de kleine *Karlo en Filip*, en natuurlijk ook mijn vrienden bedanken voor alle steun en leuke momenten door de jaren heen.

*Barbara, Janien en Monique*, jullie ken ik al heel lang en ik ben blij dat jullie mijn vrienden zijn. Ik geniet altijd erg van onze etentjes, herinneringen ophalen aan vroeger, sauna bezoekjes, en zo af en toe eens naar de zon in de winter.

*Lieve Wouter*, dit proefschrift was er misschien niet geweest als jij niet zo graag in Groningen had willen blijven wonen dat ik ging nadenken over carrièremogelijkheden buiten de Randstad en me toen realiseerde dat ik eigenlijk erg graag terug de wetenschap in wilde. Maar nog veel belangrijker dan dat is je onophoudelijke steun, en dat je me laat realiseren dat werken niet alles is. Het leven met jou is heel leuk!

*

## RESEARCH INSTITUTE SHARE

This thesis is published within the **Research Institute SHARE** (Science in Healthy Ageing and healthcaRE) of the University Medical Center Groningen / University of Groningen. Further information regarding the institute and its research can be obtained from the internet site: http://www.share.umcg.nl/

Recent SHARE theses can be found in the list below ((co-) supervisors are between brackets).

**2019**

**Joustra ML**
Similar but different; implications for the One versus Many Functional Somatic Syndromes Discussion
*(prof JGM Rosmalen, prof SJL Bakker, dr KAM Janssens)*

**Nandi T**
Neural control of balance in increasingly difficult standing tasks
*(prof T Hortobagyi, prof BE Fisher, dr CJM Lamoth, dr GJ Salem)*

**Fontanella F**
Antenatal diagnosis and management of fetal megacystis and lower urinary tract obstruction
*(prof CM Bilardo, prof D Oepkes)*

**2018**

**Dammen L van**
Women's health and wellbeing: the roles of early life adversity, stress and lifestyle
*(prof A Hoek, prof TJ Rosenboom, dr H Groen, dr SR de Rooij)*

**Dorland-Pel HF**
Work functioning in cancer patients: looking beyond return to work
*(prof U Bültmann, prof AV Ranchor, dr FI Abma, dr CAM Roelen)*

**Simoons M**
Somatic monitoring of patients with mood and anxiety disorders; problem definition, implementation and further explorations
*(prof EN van Roon, prof RA Schoevers, dr H Mulder, dr HG Ruhé)*

*

**Bouma AJ**
The barrier-belief approach; a new perspective of changing behavior in primary care
*(prof RL Diercks, prof A Dijkstra, prof CP van Wilgen)*

**Looijmans A**
Lifestyle interventions in patients with a severe mental illness; addressing self-management and living environment to improve health
*(prof RA Schoevers, prof RP Stolk, dr F Jörg, dr E Corpeleijn)*

**Phí XA**
Breast cancer screening in women at elevated risk; comparative evaluation of screening modalities to inform practice
*(prof GH de Bock, prof N Houssami, dr MJW Greuter)*

**Kooiman TJM**
The use of self-tracking technology for health
*(prof CP van der Schans, dr M de Groot, dr A Kooy)*

**Schulz T**
Adjustment to kidney transplantation; predictors of perceived health and psychological distress
*(prof AV Ranchor, prof RJ Ploeg, prof JJ Homan-van der Heide, dr J Niesing)*

**Verburg PE**
Pregnancy outcome in South Australia; population and cohort studies
*(prof JJHM Erwich, prof GA Dekker, prof CT Roberts)*

**Krops LA**
Physical activity in hard-to-reach physically disabled people; development, implementation and effectiveness of a community-based intervention
*(prof JHB Geertzen, prof PU Dijkstra, dr R Dekker)*

**Golenia L**
A new perspective on the development of motor variability during middle childhood
*(prof E Otten, dr MM Schoemaker, dr RM Bongers)*

**Bebber J van**
Computerized adaptive testing in primary care: CATja
*(prof RR Meijer, dr JTW Wardenaar-Wigman, dr A Wunderink)*

**Drenth JC**
Motor function, paratonia and glycation cross-linked in older people; motor function decline and paratonia and their relation with Advanced Glycation End-products
*(prof CP van der Schans, prof SU Zuidema, prof I Bautmans, dr JSM Hobbelen)*

**Selivanova A**
Health-state valuation using discrete choice models
*(prof E Buskens, dr PFM Krabbe)*

For more 2018 and earlier theses visit the SHARE website.

# SHORT BIOGRAPHY AND LIST OF PUBLICATIONS

## Short biography

Charlotte Vrijen was born on 11 March 1978 in Rolde, the Netherlands. In 1996 she completed her secondary education and started studying philosophy at the University of Groningen. In 2000 she spent a semester abroad at University College London. She graduated in 2001 and worked from 2001-2005 as a PhD student on Prof. Lodi Nauta's VIDI project 'The Critique of Language in the Philosophy of the Later Middle Ages, Renaissance and the Twentieth Century. Patterns of Conceptual Similarity in Philosophical Argument'. She wrote her thesis on the philosopher Gilbert Ryle (1900-1976) and studied both his published and his unpublished writings, partly in Oxford libraries. Near the end of her PhD Charlotte realized that although she very much enjoyed thinking about philosophical dilemmas, she missed a direct practical purpose of what she was doing. She switched careers and worked for the Dutch government for 7 years. In 2013 she returned to science by starting a second PhD in Prof. Tineke Oldehinkel's VICI project 'No Fun No Glory' at the Interdisciplinary Center Psychopathology and Emotion regulation (ICPE) at the University Medical Center Groningen. From August 2018 onward, Charlotte has been working as a postdoctoral researcher on Dr. Tina Kretschmer's ERC-funded project 'Ghosts from the past: Consequences of Adolescent Peer Experiences across social contexts and generations' (CAPE) at the Faculty of Behavioural and Social Sciences at the University of Groningen. Within this project she investigates genetic and epigenetic mechanisms of intergenerational transmission of peer experiences.

## Peer-reviewed publications

### *2019*

**Vrijen, C.**, Hartman, C.A., & Oldehinkel, A.J. (2019). Reward-Related Attentional Bias at Age 16 Predicts Onset of Depression During Nine Years of Follow-Up. *Journal of the American Academy of Child & Adolescent Psychiatry, 58*(3), 329-338. https://doi.org/10.1016/j.jaac.2018.06.009

### *2018*

**Vrijen, C.**, Hartman, C.A., van Roekel, E., de Jonge, P., & Oldehinkel, A.J. (2018). Spread the Joy: How High and Low Bias for Happy Facial Emotions Translate into Different Daily Life Affect Dynamics. *Complexity*. https://doi.org/10.1155/2018/2674523

van Roekel, E., Heininga, V.E., **Vrijen, C.**, Snippe, E., Oldehinkel, A.J. (2018). Reciprocal associations between positive emotions and motivation in daily life: Network analyses in anhedonic individuals and healthy controls. *Emotion*. http://dx.doi.org/10.1037/emo0000424

*

### *2017*

van Roekel, E., **Vrijen, C.**, Heininga, V.E., Masselink, M., Bos, E.H., & Oldehinkel, A.J. (2017). An Exploratory Randomized Controlled Trial of Personalized Lifestyle Advice and Tandem Skydives as a Means to Reduce Anhedonia. *Behavior Therapy*, *48*(1), 76-96. DOI: 10.1016/j.beth.2016.09.009

**Vrijen, C.**, Schenk, H.M., Hartman, C.A., & Oldehinkel, A.J. (2017). Measuring BDNF in saliva using commercial ELISA: Results from a small pilot study. *Psychiatry Research*, *254*(Supplement C), 340–346. https://doi.org/10.1016/j.psychres.2017.04.034

### *2016*

van Roekel, E., Masselink, M., **Vrijen, C.**, Heininga, V.E., Bak, T., Nederhof, E., & Oldehinkel, A.J. (2016). Study protocol for a randomized controlled trial to explore the effects of personalized lifestyle advices and tandem skydives on pleasure in anhedonic young adults. *BMC Psychiatry*, *16*, [182]. DOI: 10.1186/s12888-016-0880-z

**Vrijen, C.**, Hartman, C.A., Lodder, G.M.A., Verhagen, M., de Jonge, P., & Oldehinkel, A.J. (2016). Lower Sensitivity to Happy and Angry Facial Emotions in Young Adults with Psychiatric Problems. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01797

**Vrijen, C.**, Hartman, C.A., & Oldehinkel, A.J. (2016). Slow identification of facial happiness in early adolescence predicts onset of depression during 8 years of follow-up. *European Child & Adolescent Psychiatry*, *25*(11), 1255-1266. https://doi.org/10.1007/s00787-016-0846-1

### *2006*

**Vrijen, C.** (2006). Ryle and Collingwood: Their Correspondence and its Philosophical Context. *The British Journal for the History of Philosophy*, Vol. 14, no. 1.

McGuinness, B., & **Vrijen, C.** (2006). First thoughts: an unpublished letter from Gilbert Ryle to H.J. Paton. *The British Journal for the History of Philosophy*, Vol. 14, no. 4.

## Philosophy dissertation

**Vrijen, C.** (2007) The philosophical development of Gilbert Ryle: A study of his published and unpublished writings. (defended in June 2007)

Research Institute

SHARE