Chapter 8

General Discussion
The specific aims of this thesis were to investigate (1) the relation between positive and negative affect, and (2) the relation between three well-known risk factors for depression (i.e. subclinical levels of depression, high neuroticism and low extraversion) on the one hand, and affective reactivity to positive and negative events on the other hand. In this final chapter I integrate the attained knowledge related to these aims by discussing five topics: (1) the longitudinal relation between positive and negative affect, (2) within- and between-person differences in the cross-sectional relation between positive and negative affect, (3) the relation between depressive symptoms, neuroticism, and extraversion on the one hand, and affective reactivity to positive and negative events on the other hand, (4) the effects of major events on affect and depressive symptoms, and (5) the relevance of investigating positive aspects (i.e. positive affect and the effects of positive events on affect). After discussing these topics, I offer directions for future research, and general methodological food for thought. Finally, encountered difficulties with designing and analyzing Experience Sampling Methodology (ESM) studies are presented.

**Reflection on main findings**

**The longitudinal relation between positive affect and negative affect**

The major theoretical models of affect focus on cross-sectional relations between positive and negative affect (i.e. the relation between positive and negative affect at one point in time). Until now, the literature has lacked studies that investigate whether positive and negative affect are longitudinally related to each other (i.e. whether positive affect at one time point is related to negative affect at a subsequent time point and vice versa). To address this gap in the literature, chapter 2 of this thesis was devoted to investigate the longitudinal relation between positive and negative affect across measurement intervals of approximately two and a half years. More specifically, we investigated the longitudinal relation between anhedonia (i.e. lack of pleasure) and depressed affect in the longitudinal cohort study TRAILS.

In line with previous studies (Lewinsohn et al., 2003), we found a positive cross-sectional relation between depressed affect and anhedonia. More interestingly, depressed affect was also longitudinally related to anhedonia: depressed affect increased the risk of developing anhedonia several years later and vice versa. Hence, our results suggest that experiencing either anhedonia or depressed affect can initiate a negative spiral in which
a person ends up with a persistent cluster of both anhedonia and depressed affect, which may initiate other depressive symptoms in turn.

Our results support the idea that the independence of positive and negative affect is not so clear as previously thought: recent studies have provided no evidence for specific neurons and neurotransmitters for positive and negative affect, but have indicated that certain neurons and neurotransmitters are active in both positive and negative affective states (e.g. Barrett & Bliss-Moreau, 2009; Leknes & Tracey, 2008; Lindquist et al., 2015). However, as Barrett and Bliss-Moreau (2009) indicated (p. 186): “a description of psychological content will rarely ever shed light on the processes that caused it, in much the same way that the experience of the sun rising and setting is not evidence that the sun actually revolves around the earth”. Otherwise stated, the question whether positive and negative affect are independent dimensions of affective experiences will most likely not be answered by only investigating underlying processes of positive and negative affect. Studies that investigate the co-occurrence of specific affective states (e.g. sadness and motivational anhedonia), and at the same time take individual and contextual differences into account, will likely shed more light on the bipolarity debate.

Due to the observational nature of our study, it remains unknown whether anhedonia and depressed affect are causally related to each other, or whether mediating factors or common causes give rise to the (longitudinal) relation between anhedonia and depressed affect. One important third variable candidate might be low perceived control over a stressful situation, which has been associated with both depressed affect (i.e. increased perception of stress, feelings of hopelessness and despair) and anhedonia (Bogdan et al., 2012).

**Between- and within-person differences in the relation between positive affect and negative affect**

An impediment of the majority of affect models is not only their sole focus on cross-sectional rather than longitudinal relations between positive and negative affect, but also their neglect of between- and within-person differences in this relation. The majority of affect models assume that the described structure of affect applies to (1) every person, and (2) every moment. Chapter 3 of this thesis indicated that this assumption does not hold.

First, we found between-person differences in the extent to which positive and
negative affect were related to each other across the measurements using an ESM study (chapter 3). This supports previous studies showing that the relation between (1) valence and arousal, and (2) positive and negative affect differ between persons (Brose, Voelkle, Lövdén, Lindenberger, & Schmiedek, 2014; Coifman et al., 2007; Kuppens, 2008; Ong et al., 2006).

Second, we found that the relation between positive and negative affect varied from moment to moment as a function of arousal. High arousal was associated with a more inverse relation (i.e. a more unidimensional relation) between positive and negative affect. Our results are in line with the Dynamic Model of Affect (DMA, Reich et al., 2003). The DMA postulates that positive and negative affect are more inversely related to each other at moments of high stress. The idea behind the model is that stress narrows the attention and causes more quick-and-dirty decision making, in order to regulate negative emotions and to adjust rapidly to the uncertain stressful situation. At the same time, positive cues are neglected as a result of this regulation process. Several studies indeed found that at stressful moments the relation between positive and negative affect becomes more unidimensional (Davis et al., 2004; Pruchno & Meeks, 2004; Reich et al., 2003; Zautra et al., 2000). Since stress measures intertwine high arousal and unpleasantness (e.g. number of negative events or the occurrence of a specific stressor), it is unknown from these studies whether high arousal or unpleasantness plays a key role. The results of our study expands the DMA by showing that high arousal, regardless of unpleasantness, predicts a more inverse relation between positive and negative affect. Our results indicate that pleasant affective states high in arousal are also characterized by a stronger inverse relation between positive and negative affect, similar to stressful affective states.

In addition to the relation between positive and negative affect, individual differences in reactivity to positive and negative events are the topic of this thesis. One could contemplate on the meaning that arousal was associated with a more inverse relation between positive and negative affect for the effects of events on affect. The results suggest that positive and negative events only exert within-domain effects on affect (i.e. the effects of positive events on positive affect, and that of negative events on negative affect) when arousal is low, whereas positive and negative events exert both within-domain effects and cross-domain effects on affect (i.e. the effects of positive events on negative affect, and that of negative events on positive affect) when arousal is high.
The relation of depressive symptoms, neuroticism, and extraversion with reactivity to positive and negative events

Individuals high in depressive symptoms, neuroticism, and low in extraversion are characterized by high negative affect and low positive affect on average (chapter 5, 6 and 7). One possible explanation for these affective states is that these individuals are characterized by high reactivity to negative events and low reactivity to positive events. However, our findings in chapter 5, 6 and 7 contradict this explanation.

We found that individuals high in cognitive-affective depressive symptoms (e.g. sadness, loneliness, feelings of guilt and worthlessness, helplessness, self-blame and low self-esteem; chapter 5), high in neuroticism (chapter 6) and low in extraversion (chapter 7) did not show decreased reactivity to positive events in daily life. In contrast, we found tentative evidence that the effects of positive events on positive affect were slightly larger for individuals higher in cognitive-affective depressive symptoms in an early adolescent and young adult sample (see chapter 5 and 6).

In this thesis, individual differences with respect to negative events were only investigated in chapter 6. We found that the total neuroticism score was associated with a profile reflecting high reactivity to both positive and negative minor events in daily life. This is in line with the differential susceptibility theory proposing that certain behavioral, physiological or genetic factors, such as 5-HTTLPR or DRD4 polymorphisms or negative emotionality (Bakermans-Kranenburg et al, 2012; Li, 2013), are related to individual differences in 'generalized' sensitivity to events (i.e. highly sensitive to both negative and positive events). Our findings indicated that neuroticism is a factor that could predict whether an individual is highly sensitive to both negative and positive events (see chapter 6). Given that neuroticism is a heterogeneous concept with different facets (Ormel et al., 2004), we also examined whether different facets were differentially associated with reactivity to positive and negative events. From the different facets, only the depression facet was uniquely associated with a profile indicating high reactivity to both positive and negative events.

The finding that depressive symptoms and neuroticism were associated with slightly larger effects of positive events on positive affect is in line with previous studies showing that depressed patients have high affect reactivity to positive events (Bylsma et al., 2011; Peeters et al., 2003; Thompson et al., 2012). The results are not entirely comparable, however, because we only focused on the effects of positive events on positive affect.
and a combined measure of positive and negative affect, whereas previous studies have also investigated the effects of positive events on negative affect. This difference in main outcome may also explain why previous studies have shown a greater influence of depressive symptoms on the effects of positive events compared to our study, because the relation between depressive symptoms and reactivity to positive events is stronger with negative affect as outcome variable (Bylsma et al., 2011; Peeters et al., 2003; Thompson et al., 2012).

The counter-intuitive finding that individuals high in neuroticism and depressive symptoms showed high reactivity to positive events in daily life might be explained by the relatively scarcity of positive events in these individuals (chapter 5 and 6; Bylsma et al., 2011; Peeters et al., 2003). Therefore, the significance of positive events might be relatively high. According to Brock’s commodity theory ‘any commodity (i.e. messages, experiences, material objects, traits and skills) will be valued to the extent that it is scarce, unavailable, or difficult to attain’ (Brock & Brannon, 1992; Sehnert, Franks, Yap, & Higgins, 2014). Evidence for this theory is provided by communication and marketing studies. Intuitively, scarcity also appeared to be an important factor in the effects of events on affect. For example, people became increasingly happy if they received a good grade on an exam after a period of falling grades. Hence, it makes sense that positive events will exert greater influences on affect when these events do occur in individuals who experience low affect and a relative paucity of positive events compared to individuals in a more positive environment.

So far, I have described the influence of neuroticism and depressive symptoms on direct effects of positive events on positive affect. The pace in which positive affect returns to baseline levels after an increase (i.e. attractor strength) might, however, be more relevant. Possibly, individuals vulnerable to depression have a more reactive (positive) affective system to the environment with respect to direct influences, but at the same time have a deficiency in sustaining positive affect over time. In other words, they may benefit from positive events for a short time period, but not for a longer period because positive affect will return back to baseline levels quickly. Remarkably little attention has been paid to this topic so far, and therefore we investigated whether neuroticism and extraversion predicted prolonged effects of positive events on positive affect (see chapter 7). We found that neuroticism and extraversion were not related to prolonged effects of positive events on positive affect in two ESM studies. Our finding is in
contrast with studies showing that high neuroticism and low extraversion are negatively related to self-reported use of savoring and positively related to the use of dampening strategies (Bryant, 2003; Ng, 2012; Wood et al., 2003). Based on these studies, one would expect that positive affect would have a shorter duration in individuals high in neuroticism or low in extraversion, but this was not supported by our findings. Possibly, individuals have difficulties in accurately judging their use of these strategies in daily life. Alternatively, it could be that neuroticism and extraversion are only related to the duration of positive affect after the occurrence of specific events. For example, it has been found that neuroticism was only associated with dampening strategies after a self-relevant event (i.e. personal success) and not after a non-self-relevant event (i.e. survival of a friend) (Wood et al., 2003). In our study, the event measure was a mixture of different types of events.

The effects of major events on affect and depressive symptoms

The previous paragraphs have dealt with the effects of minor events on affect. We found that minor positive events only influenced affect in the short term, and that high reactivity to minor events was not associated with higher mean levels (i.e. baseline levels) of positive affect or lower mean levels of negative affect (chapter 6). This paragraph discusses whether major events are able to change baseline levels of affect and depressive symptoms over long periods of time.

Baseline affect is considered to be a trait characteristic of a person, and trait characteristics have long been considered as factors that cannot be influenced by the environment. There is, however, accumulating evidence that trait characteristics such as neuroticism are not so hard-wired as previously thought (Caspi, Roberts, & Shiner, 2005; Jeronimus, Riese, Sanderman, & Ormel, 2014; Roberts & DelVecchio, 2000; Specht, Egloff, & Schmukle, 2011). Baseline affect is also amenable to change: major stressful life events could suddenly alter levels of baseline affect with the onset of depression as one potential outcome (e.g. Clark, Diener, Georgellis, & Lucas, 2008; Headey, 2010). If major negative events could shift baseline affect for worse, one would expect that major positive events are able to shift baseline affect for better. It has been found that major positive events like marriage have positive effects on average happiness levels in healthy controls (e.g. Lucas, Clark, Georgellis, & Diener, 2003).

With respect to depressive symptoms, we found that an excess of major positive
events relative to major negative events over a two year period was associated with fewer depressive symptoms (chapter 4). However, experiencing a large excess of positive life events (more than three) did not have any additional beneficial effects anymore, rather the opposite. Thus, major positive events have the potential to reduce depressive symptoms over time, yet only when the total amount of events is limited. This finding could be explained by that every major event, positive or negative, evokes some arousal and requires effort to adjust to the new situation. Persistent exposure to high arousal could result in exhaustion of underlying systems and might consequently lead to a vicious circle of reduced energy and disturbed affect (e.g. Armon et al., 2008; Grossi, Perski, Evengard, Blomkvist, & Orth-Gomer, 2003; Ursin, 2000; Wyller, Eriksen, & Malterud, 2009). Hence, if a person experiences chronically heightened arousal levels, it is not favorable to induce even more arousal. It is therefore important to take into account the overall amount of events when examining the effects of positive events on mental health.

Taken together, it seems that major positive events are able to influence levels of affect and depressive symptoms for the better across long periods of time. The question remains how long the positive changes in baseline level of affect and depressive symptoms last over time. Could major positive events induce permanent changes in baseline levels of affect and depressive symptoms?

Several theories like the hedonic treadmill theory postulate that after the occurrence of a positive event, positive affect will return to baseline levels over time (Lucas et al., 2003). However, it has also been shown that there are large differences between individuals (Lucas et al., 2003). It is likely that a persistent change in affect will only occur when at the same time individuals are exposed to changes that increase positive affect on a daily basis, but to my knowledge this has not been directly tested yet. In our research department, a study has been designed that investigates whether a tandem skydive is able to boost the reward system and to set lifestyle changes into motion, which would increase the chance that the change in affect persists over long periods of time (Oldehinkel, 2012).
Does black overrule color?

One of the aims of this thesis is to bring the color back in the literature after the focus on black. In 1991, the positive psychology movement already called for more interest in positive aspects, but during the twenty-four years between 1991 and 2015, the main body of research and clinical practice still predominantly focused on negative events and negative affect. Have researchers deliberately neglected positive aspects, because they (or funding sources) do not see the relevance of bringing color back in the literature?

Baumeister and colleagues wrote an article entitled ‘Bad is stronger than good’ (Baumeister et al., 2001), in which they reviewed whether the effects of negative psychological experiences, such as daily and major life events, interpersonal interactions and learning processes, are larger than the effects of positive ones. They concluded that, with a few exceptions, negative triumphs over positive. We also found that the effects of major positive events on depressive symptoms are smaller than the effects of major negative events (chapter 4), and that the reactivity coefficients of minor negative events (i.e. the influence of negative events on affect) are larger than the coefficients of minor positive events in daily life (chapter 6).

The finding that effects of negative psychological processes are stronger than effects of positive psychological processes does not necessarily imply that focusing on positive aspects is not important. In our studies, neuroticism was equally strongly related to positive affect as to negative affect (chapter 6), and depressive symptoms was associated with low mean positive affect (chapter 5). Furthermore, we demonstrated that neuroticism is somewhat stronger related to the effects of positive events on affect than the effects of negative events on affect (chapter 6), and that depressive symptoms are associated with a lower proportion and intensity of positive events. Thus, the effects of positive events may be less strong, but not less relevant as both neuroticism and depressive symptoms are associated with positive events and positive affect.

Furthermore, as the general impression of people is that positive events are under more personal control than negative events (Wright, Zautra, & Braver, 1985), it may be easier to increase the experience of positive events than to decrease the experience of negative events. Hence, focusing on positive aspects may accomplish a mood brightening effect more easily. Until now, the main focus of cognitive treatments is to help depressed patients in dealing with negative environments and thoughts, and less to help patients find the sunshine in the dark cloud, that is generating positive
environments and thoughts, with some exceptions such as behavioral activation therapy (Dimidjian, Barrera, Martell, Muñoz, & Lewinsohn, 2011). Increasing positive events may set the stage for a spark of hope (“I am not doomed to darkness, because I see a ray of light”) and a positive spiral in depressed patients, whereas the reduction of negative events will probably not offer this hope (“It is not getting darker, but it will stay dark”). A study by Kramer and colleagues (2014) showed that providing patients feedback about their levels of positive affect after and during daily events and social interactions resulted in a decline in depressive symptoms. This suggests that increased attention to the effects of positive events and activities on positive affect could indeed set a positive spiral in motion in depressed patients.

Taken together, the answer to the question: ‘Does black overrule color?’ depends on the perspective. Black might overrule color concerning the size of the effects of events, but black and color are at least equal in terms of relevance.

**Future directions**

**Examining the generalizability of our findings to other populations**

It would be interesting to follow a high-risk sample before and after the transition to depression, in order to investigate whether certain reactivity patterns are present before the occurrence of depression or whether these patterns are a consequence of the disorder itself. This could provide clues on what to focus on in prevention programs. Our high-risk sample consisted of individuals with subclinical levels of depression, high neuroticism or low extraversion. However, the majority of these individuals will not develop depression, because many other factors are involved. To enhance the possibility that a sample includes individuals that will develop depression, it might be helpful to follow individuals with a ultra-high risk for developing a depressive disorder (e.g. those who have a combination of risk factors such as high neuroticism or a family member with depression and individuals who have experienced a major negative life event).

In addition, the focus of this thesis was on individuals aged 10 to 25 years, and half of the chapters included females only. Further investigation is required to examine the generalizability of our findings to other populations (e.g. male). It has been found that females are more reactive to different kinds of negative affective stimuli than males (Whittle, Yücel, Yap, & Allen, 2011). Furthermore, females tend to pay more attention
and better understand their emotions, which may result in adaptive emotion regulation strategies (Nolen-Hoeksema, 2012). However, the maladaptive emotion regulation strategy rumination is also more common in females (Tamres, Janicki, & Helgeson, 2002). Some females may become obsessed in analyzing their emotions which could result in excessive rumination, the belief that the emotion is a result of something in themselves of which they have lost control, and in turn depressive symptoms (Nolen-Hoeksema, 2012). It has indeed been found that rumination is a mediator of the higher prevalence of depression in women compared to men (Nolen-Hoeksema, 2012). Gender differences are most pronounced with respect to negative affect and negative affective stimuli (Whittle et al., 2011). There are relatively few studies investigating differences with respect to reactivity to positive stimuli and the regulation of positive affect, and the studies that did showed mixed findings (Whittle et al., 2011).

**Which specific positive events are especially beneficial? A person-fit approach**

I think that an important step in further research is to investigate which positive events are particularly effective in reducing depressive symptoms, increasing positive affect, and decreasing negative affect. There are indications that only interpersonal events and not achievement-related events could buffer the effects of negative events (Shahar & Priel, 2002), and that only controllable and internally caused positive events (e.g. playing a sport, game or cards with friends) are able to reduce negative affect (Strand, Reich, & Zautra, 2007). Unfortunately, these findings are based on group results, whereas my hypothesis is that individuals greatly differ in which specific positive events are most favorable.

The importance of personality characteristics with respect to the effects of negative events has already been demonstrated: dependent persons are vulnerable to negative interpersonal events, whereas self-critical persons are vulnerable to negative achievement-related events (Hammen, Marks, Mayol, & DeMayo, 1985). I expect that person characteristics could also predict which specific positive events generate a boost in positive affect or reduction in depressive symptoms in a certain person. To examine this specific hypothesis, future studies should measure positive events, depressive symptoms, and affect frequently over time within a high number of individuals. With a large sample, these studies are able to investigate whether person characteristics could predict specific reactivity patterns within persons.
Should we focus on the actual exposure or the interpretation of positive events?

In our studies, individuals high in depressive symptoms and neuroticism reported less (intense) positive events. It remains unknown whether individuals with depressive symptoms and high neuroticism are actually exposed to less (intense) positive events or whether they experience the positive events they encounter differently. Unfortunately, as our event measures were based on self-report, we were not able to provide more information on this issue. Self-report measures always have a certain degree of subjectivity. One might prefer to measure participants in Big Brother-like scenes in which a camera with microphone is worn by the participants. Then the researcher would have an overview of the actual events a participant encountered, which could be rated on ‘pleasantness’ by independent coders. In addition, participants have to indicate several times a day whether a positive event has occurred and how pleasant the event was. Such a study may provide clues on whether individuals high in depressive symptoms or neuroticism are actually exposed to less (intense) positive events as indicated by independent coders, or whether these individuals solely experience the events as less positive, which would be revealed by a discrepancy between the reports of the participants and the independent coders. Previous studies have already found that individuals with depressive symptoms and high neuroticism have the tendency to degrade the pleasantness of an event (Gentzler et al., 2014; Ng, 2012; Raes et al., 2012). For example, they may think that a colleague asked them to go out for dinner because the colleague was not able to find another person to have dinner with, or they may think that an achievement is worthless, because everyone could do it. The new study would provide clues on whether an approach that stimulates individuals to actively elicit positive events by their behavior (i.e. behavioral activation therapy) (Cuijpers, van Straten, & Warmerdam, 2007) is able to increase the experience of positive events in addition to an approach that teaches individuals to take off their black glasses and wear their pink glasses in the interpretation of positive events (cognitive approach) (Menne-Lothmann et al., 2014). The feasibility of this study is questionable since a lot of effort from the participants and researchers is required, and the study compromises the privacy of a participant.

Before increasing the exposure (i.e. actual exposure and appraisal) to positive events, it is necessary to find out whether increasing the exposure to positive events may actually help individuals to decrease mean negative affect and increase mean positive affect. As mentioned before, positive events are likely meaningful for individuals high
in neuroticism and depressive symptoms, because of their scarcity. Future studies could find out whether an increase in exposure to positive events will decrease the value of positive events, and which strategies could help to retain their value in individuals vulnerable for depression or depressed patients.

Arousal and the influence of positive events
We found that positive and negative affect are more inversely related to each other during high levels of arousal. This finding is based on an observational study. Future studies should experimentally manipulate the degree of arousal during an event, to elucidate whether increasing arousal during the experience of positive events really results in a decrease in negative affect in concert with an increase in positive affect. If this is the case, prevention and intervention strategies for depression might be improved by incorporating strategies that increase arousal during the experience of minor positive events, and hence exert a ‘double positive effect’ on well-being (i.e. a decrease in negative affect in concert with an increase in positive affect), especially when this effect could be sustained over time.

The role of positive automatic thoughts
This thesis ignores an important third factor in the relation between events and affect, that is implicit and conscious thoughts. Negative thoughts play an important role in theories about depression and neuroticism. For example, Beck’s cognitive model of depression postulates that dysfunctional beliefs (i.e. negative biased schemas, assumptions and attitudes towards oneself) are central to depression (Beck, 1967). In response to negative events, these dysfunctional beliefs are activated and could initiate negative automatic thoughts, which in their turn could increase negative affect (Beck, 1967). Negative affect could also evoke negative automatic thoughts, resulting in a vicious circle (Teasdale, 1988). Remarkably, also in the field of thoughts, most attention has been paid to black (i.e. negative) rather than color (i.e. positive).

The presence of positive automatic thoughts during the occurrence of a positive event (e.g. the thought ‘I am fun to be with’ when a friend ask you to join him on a weekend trip) might predict whether a positive event will influence affect. In addition, the presence of positive automatic thoughts during a negative event, might predict that positive affect will not decline. This maintenance of positive affect during negative events is beneficial,
since it has also been shown to be an important predictor of resilience (Coifman et al.,
2007; Tugade & Fredrickson, 2004). The relevance to investigate positive automatic
thoughts in the context of vulnerability for depression has also been shown by studies
finding that the general tendency to have positive automatic thoughts (i.e. trait positive
automatic thoughts) has been negatively related to neuroticism, and positively to
e extraversion (Lightsey et al., 2013). Furthermore, these thoughts mediated the relation
between both neuroticism and extraversion, and trait positive affect (Lightsey et al.,
2013). Trait positive automatic thoughts have also been inversely related to depressive
symptoms (Ingram, Atkinson, Slater, Saccuzzo, & Garfin, 1990). Until now, studies
have focused on trait positive automatic thoughts rather than on momentary positive
automatic thoughts. An avenue for future research is to investigate how to measure
thoughts and affect separately from each other. Next steps would be to examine whether
the presence of momentary positive automatic thoughts could predict whether positive
or negative events influence positive and negative affect in concert or solely positive or
negative affect, and whether techniques that increase positive automatic thoughts will
help individuals with depression to boost their affect.

**Increasing perceived control**

Another interesting challenge for future research is to develop strategies that are able to
enhance perceived control in individuals high in depressive symptoms or neuroticism.
This is relevant for several reasons. First, positive events are judged as being (more)
positive when people believe that they caused them themselves (Strand et al., 2007).
Second, perceived control not only increases the pleasantness of events, it also predicts
positive affect beyond the pleasantness of the positive events themselves (Langston,
1994). Third, individuals high in depressive symptoms and neuroticism have low
perceived control over events including positive ones (Edmondson & MacLeod, 2014).
Taken together, increasing the perception of control over positive events, might be an
important strategy to increase both the experience of positive events and positive affect
in individuals high in depressive symptoms and neuroticism.
General methodological food for thought

The need to discriminate between within-person effects and between-person effects

As mentioned before, we found that at the within-person level, higher levels of momentary arousal were associated with a stronger inverse relation between positive and negative affect. We, however, did not find a significant association between mean arousal and the relation between positive and negative events at the between-person level (chapter 3). Otherwise stated, individuals with higher mean arousal did not show a stronger inverse relation between positive and negative affect. These findings illustrate that within-person effects do not necessarily imply between-person effects.

Researchers sometimes tend to misinterpret their statistical analyses and infer data obtained from the group level (i.e. between-person level) to the individual (within-person level), also known as the ecological fallacy (Clancy, Berger, & Magliozi, 2003). As of yet, the focus has been too much on between-person effects. This by no means implies that every study should now examine within-person effects instead of between-person effects, but rather that every researcher should be aware of this ecological fallacy and should decide which approach helps to further the theory more or has the most clinical importance.

An example that nicely demonstrates the differences between within-person and between-person effects is given by Snippe (2014, p. 123): the finding that individuals who eat cake more often are on average not happier than other individuals does not imply that at moments when a person eats a cake, he or she will not feel happier. This example highlights the importance of differentiating within-person and between-person effects. One could conclude from this example that it is important to investigate within-person effects instead of between-person effects. However, one might discuss what the clinical relevance is of knowing how happy a person feels when he or she is eating a cake. From the example it would seem that eating cake only resulted in a very short increase in happiness and not in an increase in the long-term or average level of happiness given the absence of between-person effects. One might suggest that in this case the focus should be on investigating factors that increase average levels of happiness, because this is actually more of significance for well-being or psychopathology than short-lived peaks in happiness. An example in which the interest lies more in within-person effects of cake on momentary happiness is a study in which a dietician examines whether eating a cake results in a short-lived peak in happiness in the client. Taken together, every researcher
should ask: is it more relevant for this specific topic to focus on within-person effects or between-person effects?

**Correlation is not causation, even in time series analysis**

In this thesis, we relied on observational studies, which are not able to clarify causal relationships. Researchers should be aware of the trap of reporting correlations as causal relations with words as ‘result in’, ‘lead to’, ‘effects of.. on’. It is important to prevent these causal phrases when causal relations were not studied. Several conditions must be fulfilled for a causal relationship between X and Y: (1) association: there is a significant co-variation between X and Y, (2) temporal precedence: X should occur before Y, (3) isolation: there are no other plausible explanations for the co-variation between X and Y (e.g. third variables) (Kline, 2011). Longitudinal observational studies, including diary studies, are able to fulfill the first two criteria, but not the third. Several researchers incorrectly claim that they investigated causal effects in daily life studies with techniques such as Granger causality tests (Granger, 1969). Causality cannot be tested in non-experimental daily life studies, because effects are not tested in isolation (Kline, 2011). Precisely this lack of isolation (i.e. the possibility that other variables explain the co-variation between the two variables) is a characteristic of diary studies since the measurements take place in spontaneous contexts without control of the researcher. An experimental design, in which the potential cause is manipulated, is required to demonstrate a causal relation.

Unfortunately, it is not easy to manipulate potential causes. In experimental studies, affect is often induced by imagining situations, reading self-evaluation statements, watching images or films, listening to stories or music, experiencing success or failure, social interactions and receiving gifts. The results of a meta-analysis on these different induction procedures showed that the average effect sizes of the induction procedures are medium to large, but that there are large differences in the effectiveness of each procedure (Westermann, Spies, Stahl, & Hesse, 1996). A problem is that the effect of positive events on positive affect is difficult to investigate, because positive mood induction procedures usually fail to enhance positive affect to a sufficient extent, in particular positive affect high in arousal. Furthermore, it is questionable whether the results based on these mood induction procedures can be generalized to daily life, because of the artificial induction and the controlled environment. Hence, my ideal study would
be a combination of an experimental design and daily life study to investigate whether a ‘cause’-effect relation revealed in daily life is also found in a controlled lab situation in which the cause is experimentally manipulated, or an experimental manipulation of a certain factor in daily life. For example, one could induce negative affect in a controlled lab situation, and investigate whether there are differences in reward responsiveness between a condition in which an arousal-reducing strategy (e.g. relaxation exercises) is applied and a control condition in which this strategy is absent. Furthermore, by investigating arousal, negative affect, stress, positive affect and the (successful) use of the arousal-reducing strategy in daily life, one could examine whether findings in the lab situation can also be translated to daily life.

The importance of investigating smaller parts of a heterogeneous construct

Both neuroticism and depression are heterogeneous constructs with different underlying facets or symptoms (Chen et al., 2000; Kendler et al., 1996; Ormel et al., 2013). In this thesis, we divided these constructs into smaller parts in the hope to ease the interpretation of the results. When it comes to depression, we investigated the core symptoms depressed mood and anhedonia specifically in chapter 2, and made a distinction between neurovegetative-somatic symptoms (i.e. appetite or weight change, sleep problems, fatigue) and cognitive-affective symptoms (i.e. depressed mood, anhedonia, feeling worthless, guilt, suicidal ideation) in chapter 4. These latter two dimensions have been found to be differentially associated with, amongst, others co-morbid problems, clinical characteristics of depression, and personality traits (Lux & Kendler, 2010). In our study, we also found that these two dimensions were differentially associated with the effects of positive and negative life changes. With regard to neuroticism, we found that specific facets of neuroticism were differentially related to affect reactivity to positive and negative events. We found that in particular the depression facet was associated with a “for better and for worse” profile: high reactivity to both positive and negative events in daily life. Taken together, our studies demonstrated the relevance of investigating smaller parts of heterogeneous constructs. However, one should consider that investigating symptoms or facets also has the disadvantage that the whole concept might fade from sight as a result, since the whole is more than the sum of its parts. For example, the occurrence of a few symptoms on their own does not result in loss of functioning in depression, but it is the specific interplay between different symptoms
that constitutes the depressive syndrome.

The reliability and validity of the selection of our items

In all of our studies, we only used a few specific items to measure the desired constructs. The composition of the set of items was not always based on recommended steps for test construction (Oosterveld & Vorst, 1996), such as concept analysis of the construct under investigation, the collection of all available items and the application of psychometric analyses on all items to select the most valid and reliable items.

In the chapters in which we used the TRAILS study (chapters 2 and 4), we were not able to influence the choice of questionnaires, because the data had already been collected. In chapter 4, we used a 10-item scale (the Turning Point Questionnaire) to measure positive life changes (5 items) and negative life changes (5 items). The questions were general, for example: ‘There has been a change in your family for the better’. These questionnaires broadly measured our desired constructs, but in future research, I would like to more extensively test whether all change domains are captured in the questionnaire and whether questionnaires asking about more specific life changes are better measures.

The study in chapter 2 is limited by using single items to measure anhedonia and depressed mood. The view that single-item measures cannot provide a reliable measure of relatively complex psychological constructs because they are prone to measurement errors has long dominated the literature. The aim of a multiple-item scale, which should be characterized by a random selection of items in which each item explains some unique variance of the measured construct, is to average these measurement errors out, and hence increase the reliability and the validity of the measure (Diamantopoulos, Sarstedt, Fuchs, Wilczynski, & Kaiser, 2012). In practice, the use of multiple-item scales in research has become an end in itself, at the cost of reduced validity. By including semantically redundant items, the internal reliability of multiple-item scales is often artificially high, because Cronbach’s alpha, the most frequently used measure of internal reliability, is higher by definition when more items are included in the scale. A single-item measure is able to outperform multiple-item measures if the measured construct is concrete and translatable to one good item. Thereby, it is important that the construct is accurately described and explained to the respondents to prevent that the single-item measure is too vague. If the construct is very complex, a multiple-item scale is the preferential choice (Diamantopoulos et al., 2012).
It is the question whether the single items used in chapter 2 were able to adequately capture the constructs depressed mood and anhedonia, in particular the anhedonia item (i.e. “I enjoy very little”). A multiple-item scale would have been able to measure lack of pleasure on different domains and to distinguish between motivational anhedonia (i.e. the lack of motivation to obtain a pleasant outcome) and consummatory anhedonia (i.e. low hedonic response to rewards) (Treadway & Zald, 2011). However, it has to be kept in mind that we were interested in the relatively narrow constructs of depressed mood and anhedonia in this chapter. These narrow constructs are better captured by single items than the overarching constructs of positive and negative affect that should be measured with a multiple-item scale in my opinion.

In our ESM studies, we used six items to measure negative affect, and four to six items to measure positive affect (chapters 3, 5, 6 and 7). One could wonder whether this number of affect items is enough to capture the broad constructs of negative and positive affect. We have selected our specific items based on (1) previous ESM studies, (2) the wish to measure both low arousal and high arousal affective states, and (3) the careful consideration of which affect items would show enough variation in daily life. I think that our measure included the most important affective states, because approximately 75% of our participants indicated that all their affective states could be captured with our selection of affective items (van Wees, 2013). However, these findings should be interpreted with caution, since participants filled out this question retrospectively after the ESM period. For the majority of the affective states that were indicated as absent (i.e. grief, grumpiness, anxiety, discontent, anger and amorousness, van Wees, 2013), there were states that closely resembled the missing states.

Surprisingly, the use of single items or scales with only a couple of items appears to receive less criticism when they are used in ESM studies, with the argument that the number of measurements is large. This is a valid argument when the main outcome is an aggregated measure of the single item, such as a mean score, and hence the largest part of measurement error will be averaged out. However, most ESM studies did not use aggregated scores as main outcome, but used the information at each moment in their calculation models, and hence measurement error at each moment could influence the results. In chapter 3, we measured arousal with a single item, because our opinion was that psychological arousal is a narrow concept that could be measured by a single item. To reduce measurement error, we gave participants detailed instructions on the meaning
of the concept of arousal and the use of the scale.

As ESM studies are quite burdensome for participants due to the frequent and total number of measurements, the inclusion of too many items in ESM measurements could also result in less reliable responses as a result of a too high burden for participants. Hence, ESM studies rather rely on specific selected items than on standardized questionnaires. The consequence is that there are large differences between ESM studies in how different constructs are measured. This is problematic, because this makes it difficult to compare the results of different studies. It is likely that the relation between positive and negative affect and the impact of events on affect is influenced by which affect items one selects.

**Self-report measures**

It has to be mentioned that this thesis is solely based on self-report measures. Self-report measures traditionally receive a lot of criticism, because they are plagued by different biases (Schwarz, 1999). For example, depressed affect itself could result in over-report of negative events, or recall biases occur in which responses are based on the most intense or most recent experiences (i.e. peak-and-end effects). Traditional self-reports in which traits are measured are viewed as the most biased self-reports, whereas momentary self-reports based on ESM measurements are thought to be the least biased method (Conner & Barrett, 2012). However, ESM measurements also rely on recall to some extent and can therefore be biased. For example, the number of positive events that were reported retrospectively over part of the day can be biased by high momentary negative affect, especially in a signal-content ESM design with less frequent measurements. As I mentioned earlier in the discussion, one might choose to film participants after which independent raters could score all events that occurred. However, affective states remain difficult to measure objectively.

**Methodological challenges in ESM research**

ESM studies have several highly appealing advantages, for example: (1) high ecological validity, (2) the ability to examine temporal order of effects and to reliably examine within-person processes, (3) no or less limitations by retrospective memory biases (Shiffman et al., 2008). At the same time, ESM designs have a number of inherent obstacles that need to be addressed. My knowledge has expanded after designing,
conducting and analyzing our first ESM study, and I would like to provide the most important experiences, food for thought, and insights.

**The design of the study**

A crucial question at the start of an ESM study is: what is the optimal frequency of the ESM measurements? The answer depends on several factors, but one of the most important factors is the amplitude and speed of fluctuations of the variables of interest. If the fluctuations occur rapidly, then the ESM measurements should also follow each other rapidly to appear in the data. In other words, it is important to base the frequency of measurements in ESM studies on the temporal dynamics of the variable of interest (Ebner-Priemer & Sawitzki, 2007). In addition, the frequency should be adapted to the possibilities of the participants: to fill out questionnaires each fifteen minutes of the day is not feasible, especially not for a couple of days. This would intervene too much with daily life and result in a large problem for statistical analysis: a high number of missing values.

In addition, a researcher has to choose between an event-based design (i.e. design in which measurements are based on the occurrence of certain events) or a signal-based design (i.e. design in which measurements occur at fixed or random time points). This choice is again dependent on the variance of the variables of interest. For example, if you are interested in negative social interactions and you expect that these interactions will not occur very often in a day, an event-based design might be more appropriate. If a signal-based design is chosen, then the choice is whether time-points should be fixed or random. The advantage of random time-points is that the measurements occur randomly. Hence, the chance is higher that participants are measured in different environments and situations, and that participants will adapt their behavior and activities to a lesser extent. An advantage of fixed-time points is that time-series analyses can be performed. These allow for the possibility of reliably studying patterns of change over time and the temporal order of effects within-persons.

Another question in designing an ESM study is related to power: how many measurements and persons are enough? It is difficult to provide one clear answer, because several factors, such as the size of the hypothesized effect and the complexity of the model have to be taken into account. Furthermore, the type of analysis plays a role. Recommendations for the sample size in multi-level analysis depend on several factors,
including the investigation of cross-level interactions. An example of a cross-level interaction is the effect of the between-person factor neuroticism on the within-person effects of positive events on positive affect. Articles provide different guidelines, but in general they indicate that the number of participants is of greater importance than the number of observations per person in multi-level analysis (Hox, 2010). One guideline, when cross-level interactions are the main outcome, is the 50/20 rule: a minimum of 50 participants with 20 observations (Hox, 2010). In contrast, in pure time-series analysis in which the interest is solely in processes within-persons, the number of observations per person is important since the reliability of the results is higher with larger numbers of observations. Time-series analysis requires at least 30 observations per person, but this number also depends amongst others on the number of lags (Lütkepohl, 2006).

**Idiographic or nomothetic approach?**

In this thesis, we used both idiographic and nomothetic analyses. In idiographic analyses the study of an individual is the key focus, whereas nomothetic analyses are focused on the study of classes of individuals.

Idiographic analyses, including time series analysis, are the first choice when the primary interest lies in unraveling patterns within individuals. Idiographic analyses are tailored to the specific person of investigation. That is, the set of predictors, number of lags and the covariance structure are explicitly selected for that specific person. A nomothetic approach is often applied with respect to the choice of measures. In other words, a standardized set of items, which is considered to be reliable based on nomothetic research, is used to measure certain constructs in idiographic study designs. I noticed large differences in the reliability of scales between individuals (e.g. Cronbach’s alpha ranged from 0.29 to 0.85 for positive affect), which led me to think that more attention has to be paid to select person-specific measurements.

The disadvantage of idiographic analyses is that the results cannot be generalized to other individuals. Ideally, one would like to cluster individuals based on their specific patterns over time, and investigate whether individuals within a cluster share resilience or risk for developing psychopathology. This requires a combined idiographic-nomothetic approach, in which a high number of individuals undergo intensive longitudinal measurements. A combined idiographic-nomothetic approach is not new; several studies have already investigated processes within-persons over time, and have
used between-person variables such as personality to predict individual differences in within-persons processes. Most studies (like chapters 3, 5 and 7 in this thesis) used multi-level analyses to this end. Multi-level analyses are essentially more nomothetic than idiographic, because the analyses are not person-tailored (i.e. the same covariance structure, the same number of lags and so on are applied to all individuals), and therefore not an ideal means to investigate processes within an individual.

The idiographic approach is rapidly gaining popularity and sometimes its application almost seems to have become an end in itself, rather than a considered choice for a specific research question. As the data collection is inherently burdensome for the participants, the analyses relatively complex, and the costs relatively high, every researcher should question whether the gains of this approach outweigh the costs.

The reliability of n = 1 results

The main aim of science is to make predictions or inferences about a population, based on analyses of a random sample of that population. A p-value is nothing else than the probability of getting the results you did (or more extreme results) given that the null hypothesis is true in the population. In other words, it is the probability that the results of the analysis of the random sample could have occurred by chance when there is no relationship at all between the variables in the population. One important assumption in order to perform ‘inferential’ statistics is that the sample is a good representation of the population to which it is being generalized. In between-person analyses the random sample should consist of representative individuals from the population, whereas in within-person analyses the sample should consist of representative time-points of the period of interest (e.g. the whole life of a participant if one is interested in trait characteristics, or the period after a major event if one is interested in a high risk period for the development of depression).

It is questionable whether researchers consider this ‘representativeness assumption’ before performing analyses within persons. If the goal is to make inferences about the characteristics of a person, it is of great importance that the time period which is subjected to statistical analysis, is representative for that person. In almost all cases making inferences is the (underlying) goal, but it is often doubtful whether the ‘representativeness assumption’ is met, in particular in studies with very short measurement periods of a couple of days. Although the well-known advantage of ESM is that it ‘captures the
film rather than a snapshot of daily reality', it still only captures a relatively short film of the life of a participant. Hence, my opinion is that analyses are only reliable if the measurement period is large enough, when there are enough observed data (not too many missings), and when the expectation is that the measurement period is a good reflection of the period to which generalizations are made. Providing patients feedback about their results should therefore only be done if these criteria are met. I think it is crucial for the field of ESM studies that future studies examine how reliable the patterns are across time. This could be done by collecting ESM data in two different periods, and to examine whether the same patterns are found in the two different periods.

Another key issue with regard to the reliability is that the intensive repeated measurements could have influenced the way in which the participants filled out the questions. A couple of studies have investigated this issue, but did not find evidence that behavior or experience is affected by the repeated measurements (see for an overview Shiffman et al., 2008). However, the study of Conner and Reid (2012) indicated that intensive mobile self-reporting of happiness (i.e. 6 times per day for a period of two weeks) is detrimental for individuals high in depressive symptoms and neuroticism, because they showed decreased momentary happiness over time with more frequent measurements. In our own evaluation study of the UPPER study (van Wees, 2013), 82% of the participants indicated that their emotions were influenced by the measurements; of these participants the majority (74%) indicated that this was both in a positive and negative direction. Taken together, it appeared that frequent measurements could influence behavior and affect, and hence the reliability of the results.

**Room for change – the effects of floor effects in the data**

We ran into the problem of floor effects in negative affect by analyzing the effects of positive events on negative affect. Floor effects not only impede the examination of positive events on negative affect, but also hinder the investigation of other dynamics, such as negative affect inertia (i.e. the autocorrelation of negative affect). The occurrence of floor effects in our data is not a characteristic of our specific sample, but is typical for a population of healthy individuals (e.g. Bylsma et al., 2011; Jacobs et al., 2011; Thompson et al., 2012). Remarkably, the majority of researchers appear to ignore this fact, and continue to analyze differences between individuals with low baseline negative affect (i.e. 'healthy' individuals) and individuals with high baseline negative affect (i.e.
individuals high in depressive symptoms or neuroticism) on negative affect dynamics. This is problematic, because claims that these groups differ in negative affect dynamics may be false if these effects are driven by floor effects and associated ‘room for change’ in negative affect. The same problems occur in patient samples, for example with ceiling effects in negative affect in depressed patients. In my opinion, the problem of floor and ceiling effects is underestimated and should receive more attention in the literature.

**Concluding remarks**

The title of this thesis (i.e. every dark cloud has a colored lining) suggests that darkness (i.e. negativity or unpleasantness) and color (i.e. positivity or pleasantness) are closely related to each other. We found that depressed affect increased the risk of developing anhedonia (i.e. lack of pleasure) several years later and vice versa. Our findings further showed that the momentary relation between negative and positive affect depends on the momentary degree of arousal. The ‘dark cloud’ in the title is a metaphor for an individual characterized by darkness (an individual high in depressive symptoms or neuroticism). Using the appealing ESM technology, we found that individuals high in depressive symptoms or neuroticism experienced low levels of positive affect and less (intense) positive events, but that they did show slightly higher reactivity to positive events than individuals with low depressive symptoms and neuroticism. Future research should focus on how the exposure to positive events (i.e. actual exposure and appraisal) could be enhanced in individuals high in depressive symptoms and neuroticism without reducing the meaning of positive events, and which specific positive events are able to change positive and negative affect in concert. This can be achieved by investigating the role of arousal, positive automatic thoughts and perceived control in the effects of events on affect using a combined idiographic-nomothetic and experimental daily-life approach.