“It was great the way her mind worked. No guilt, no doubts, no fear. None of my specialties. Just the shameless pursuit of immediate gratification. What a capitalist.”

Joel Goodson (played by Tom Cruise), Risky Business, 1983
As was described in the Introduction of this thesis, aggression and violence are globally regarded as potential disastrous phenomena for personal safety and well-being and therefore pose serious challenges to society. Even though there is growing scientific knowledge of the causal and perpetuating factors of aggression, including underlying neurobiological and neuropsychological processes, much is also still unclear about the exact nature of relations between aspects of neurocognition, neurochemistry and neuroanatomy on the one hand and specific characteristics of aggression on the other hand. This neuroscientific knowledge is especially valuable when it comes to its translation into interventions for aggression reduction, which is not always straightforward. In the present study the overall goal was to increase insight in specific dynamics between two aspects of social cognition - i.e. risky decision making and biased facial emotion perception - and aggression. In the Introduction section in the first chapter of this thesis, eight specific sub-aims were formulated, which were subsequently researched and discussed in the following six chapters, divided across two sections. In this final chapter, the main conclusions will be summarized and discussed pointwise on the basis of each of the eight initially formulated objectives, followed by an outline of the most important limitations and strengths of this thesis. To conclude this chapter, the clinical relevance of the present findings is contemplated on, accompanied by an elaboration of potential directions for future studies.

**Part I: Risky Decision Making**

**Objective 1: Systematically review the empirical literature on the potential relation between risky decision making and aggression to gain more understanding in the nature of this relation (Chapter 2).**

A systematic database search was conducted by two researchers in 2014, in compliance with the guidelines from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher et al., 2009). Four large databases were consulted using an extensive search strategy with various search terms to cover all research on risky decision making in relation to topics relevant for aggression. After a systematic screening and selection procedure 16 studies were included in qualitative synthesis (Zaza et al., 2000). The results of the quality assessment showed marginal quality for most studies, largely due to small study groups or lack of sufficient correction for biases. In the three studies with the best quality, a significant positive relationship was found between the tendency for increased risk taking during decision making tasks and measures of aggression (Best et al., 2002; Fishbein et al., 2009; Levi et al., 2010). In eight out of
eleven studies with a moderate study quality, significantly positive relationships were also found between increased risk taking and increased aggression. Most of these studies revealed that a history of violent (criminal) behaviour or current self-reported aggressive feelings were significantly correlated to more risk taking in decision making tasks (Fishbein et al., 2009; Fontaine & Nolin, 2012; Levi et al., 2010; Lewis et al., 2004; Prehn et al., 2013). In addition, one study showed that the induction of aggression led to increased risk taking on the Iowa Gambling Task (IGT) (Bobadilla et al., 2012), and in another study, risk-taking tendencies on the IGT significantly predicted seclusions for aggressive behaviour during detention (Bass & Nussbaum, 2010).

When looking at specific subtypes of aggression, only three studies discriminated between reactive and proactive aggression (Bobadilla et al., 2012; Broomhall, 2005; Fishbein et al., 2009). In all of these studies it was found that reactive aggression was related to more risk taking on the Cambridge Gambling Task (CGT) and IGT than proactive aggression. Nevertheless, there were substantial differences in the methodologies and definitions of reactive and proactive aggression between these studies.

Only a small number of studies investigated neuro-anatomical correlates, which provided some support for the hypothesis that orbitofrontal dysfunction might underlie both aggression and poor decision making. No studies were conducted in which both decision making and aggression were investigated in relation to hormones such as testosterone or cortisol. When looking into potential confounders in the studies on the relationship between risky decision making and aggression, results of the literature study showed that it seems unlikely that this relationship can be accounted for by factors such as age, intelligence, educational level, attention, (sub)cultural factors, gender, drug abuse, and use of medication, but the influence of psychopathic personality traits remained unclear.

Even though the results of the systematic literature review were contradicting on some points, all in all the findings most strongly pointed towards the existence of a relationship between risky decision making and reactive, impulsive aggression. This main conclusion marked the foundation for the chapters that followed, in which this relationship was further empirically explored in a sample of 125 adult Dutch male prisoners.
Objective 2: Assess if prisoners perform worse than normal controls on measures of risky decision making (hot) and classic (cool) executive functions (EF) (Chapter 3).

125 Participants were recruited in a large prison setting in the Netherlands (Penitentiary Institution Vught). They were accused of or convicted for committing criminal acts: 98 participants had at least one conviction for a violent crime in their lifetime, while 27 were convicted of non-violent crimes only. A healthy control group ($n = 32$) was formed by prison workers, who all were guaranteed not to have criminal records. Neuropsychological tests were assessed in both groups: hot EF were assessed with the IGT, while the Wisconsin Card Sorting Test (WCST) was completed as an indication of cool EF.

There was no significant difference between offenders and non-offenders in the distribution of their scores on the five blocks of the IGT, although some interesting trends were visible. When looking at the mean IGT curves across the five blocks of the two study groups, it was observed that the curve of the offender group was overall slightly lower than the curve of the non-offender group, which means that they made slightly more risky and less safe choices than the non-offender group during the whole task. Although this difference was not-significant when looking at the separate block-scores, the overall performance (reflected in the total NET-score) was significantly worse for offenders compared to non-offenders. The small effect size indicates that the difference was small. When looking at the shape of the learning curves, both groups showed parallel increasing curves during the first three blocks. The non-offender group curve showed a decrease and subsequent increase in the final two blocks, while the offender group curve remained flat during the last two blocks. Although the group comparisons show no clear statistically significant difference in the final two blocks specifically, it is still interesting that the flat shape of the offender curve does point to a lack of learning during the task. How the sudden changes in the curve of the non-offender group can be explained, is unclear. Since this group was relatively small ($n = 32$), individual scores have a higher impact on the group mean scores than in the offender group ($n = 125$) and therefore this sudden drop in the curve may merely resemble an artefact. Increasing the group size in the control group then might lead to a different statistical outcome, which can definitely be regarded as a shortcoming in this study.

In addition, a comparison between the groups on the WCST, showed that offenders generally performed worse than non-offenders on all measures, with effect sizes all in the medium range. When intelligence matched groups were created, the differences on the WCST and IGT disappeared, but this was also
(at least partly) due to a substantial decline in statistical power as result of the matching procedure.

These findings indicate that, on average, offenders have more problems in hot and cold EF compared to non-offenders, of which the problems in cold EF are the most pronounced. Considering that risky decision making (indicative of hot EF) is the central concept in the first part of this thesis, and that problems in cool EF in offenders are already widely recognized (Ogilvie et al., 2011), in this final chapter only the first will be further explored. The present findings on hot EF in offenders are supported by earlier studies. After all, results of the systematic literature review in chapter 2 revealed that all investigated forensic groups (offenders, forensic psychiatric patients, clinical psychopaths) showed worse performance on decision making tasks when compared to healthy controls. This suggests a relationship between antisocial behaviour/criminality and increased risk taking during decision making, regardless of specific aggression subtypes.

Looking more into the cognitive functions related to risk taking, it appears that multiple processes might be involved here. For example, risky decision making is dependent on concepts such as response inhibition, sensitivity for reward, perception of risks, risk preference (also known as risk attitude), sensation seeking, and how rewards are represented and remembered (Reyna & Huettel, 2014). Although decision making has been regarded as merely a matter of computing gains (rewards), losses and probabilities, a more integrated approach that also incorporates motivational and emotional elements seems to do more justice to the complex processes involved (Reyna & Huettel, 2014). Therefore, making risky decisions in daily life can be caused by multiple processes: this can be either due to a problem understanding or remembering response-reward contingencies or to understanding the risks but taking them anyway as a result of sensation seeking or disinhibition. Based on the outcomes of the traditional scoring method of the IGT it is not possible to differentiate between such aspects of decision making. In the present thesis these traditional outcome variables were applied. An alternative scoring method has been proposed to extract more information from IGT based on a mathematical model: the ‘Expectancy-Valence model’ (EVM) (Busemeyer & Stout, 2002; Yechiam et al., 2005; Yechiam et al., 2008; Yechiam, Veinott, Busemeyer, & Stout, 2007). The EVM derives three main components from the IGT: data on the relative impact of rewards and punishments, the rate that contingent payoffs are learned and the consistency between learning and responding during the task. According to this model subjects can either place more emphasis on immediate outcomes or give more weight to gains versus losses, whereas choice consistency
can also be measured. In an earlier study with offenders, the EVM was applied to unravel performance on the IGT (Yechiam et al., 2008). The findings from this study indicated that whereas all the criminal groups tended to select disadvantageously, the cognitive profiles exhibited by different offender groups were considerably different. Certain subpopulations - most significantly, drug and sex offenders - overweighed potential gains as compared with losses, similar to chronic cocaine abusers. In contrast, criminals who were convicted for assault or murder tended to make less consistent choices and to focus on immediate outcomes and, in these respects, were more similar to patients with orbitofrontal damage (Yechiam et al., 2008). However, the subgroups in this study were small (ranging from 4 to 22) and no studies are known to have replicated these findings thus far.

Apart from the various cognitive processes involved in risky decision making, performance on the IGT is also dependent on multiple other cognitive processes such as working memory, reversal learning and classic executive functions (Buelow & Suhr, 2009; Dunn, Dalgleish, & Lawrence, 2006), which further complicates clear interpretations of impaired performance on the task. For example, some studies have shown that performance on the IGT is also related to performance on the WCST, suggesting cool EF may also be of influence (Buelow & Suhr, 2009), although this relation between the two tests was not found in the present study. In addition, fMRI research has revealed that not only the orbitofrontal cortex (OFC) and ventromedial prefrontal cortex (vmPFC) are active during this task, but many other structures as well, and the simultaneous and countervailing interactions among different brain processes challenge interpretation (Lawrence, Jollant, O’Daly, Zelaya, & Phillips, 2009; Li, Lu, D’Argembeau, Ng, & Bechara, 2010; Reyna & Huettel, 2014). These include dorsolateral prefrontal areas (usually typically related to cool EF, but also related to cognitive control through inhibition or re-evaluation of emotional, rewarding or prepotent responses), posterior cingulate cortex (for representations of emotional states), ventral striatum including the nucleus accumbens (part of the reward circuit), anterior cingulate and the insula (Dunn et al., 2006; Lawrence et al., 2009; Li et al., 2010; Reyna & Huettel, 2014). The role of the latter structure is suggested to be more important than previously recognized (Paulus, Rogalsky, Simmons, Feinstein, & Stein, 2003), mainly for its role in experiencing aversive emotions during the perception of risk and potential losses, and thus relevant for actual risk avoidance (Reyna & Huettel, 2014).

Even though the IGT is globally the most applied task to assess risky decision making (Buelow & Blaine, 2015; da Mata et al., 2011), all of the foregoing can lead to the question if the cognitive constructs involved in IGT performance are not too
complex and multifaceted to apply this task to assess a single construct in empirical studies. There are other tasks designed to measure risky decision making that seem to be more straightforward. This includes, for example, the Balloon Analogue Risk Task (BART) (Lejuez et al., 2002), a computerized task in which participants pump up a balloon on a screen by pressing a key. For each pump, money is rewarded, but the larger the balloon grows, the greater the risk becomes that it pops, resulting in a loss of money. The amount of pumps is indicative of risk taking behaviour. Compared to the IGT this task is less complicated: it is not ambiguous and provides only one response option. Although both tasks are supposed to measure similar processes, they in fact appear to assess different aspects of decision making (Buelow & Blaine, 2015), probably due to a different learning process during the task (Bishara et al., 2009). In hindsight, adding the BART to the research protocol next to the IGT would have led to a stronger empirical foundation in the research of risky decision making in offenders. There are also options of including other, even more singular cognitive tasks (for examples in related domains, see Hoppenbrouwers et al. (2015), Ly et al. (2016) or Brazil et al. (2013)). However, such paradigms are often laboratory tasks, especially designed for fundamental research purposes, which makes it difficult to compare outcomes between studies. Furthermore, such tasks tend to be limited to such small pieces of complex cognitive processes, that the outcomes of such studies are even harder to transfer to daily life and clinical practice.

One final consideration on the aspect of risky decision making in the present study is that the variance in performance of offenders on the IGT was large. The group as a whole performed worse than non-offenders, but individual variability within the group was high. Apart from the fact that this had a large impact in the statistical outcomes, this also implies that there were many offenders with good performance on the IGT, while there was also a group with clear impaired performance. Due to the fact that little to no significant relations were found to other investigated variables, it remains unclear what characterizes the offenders who take extremely high risks on the task. Still, these offenders who actually performin an impaired range, might be especially at risk for future problems in daily life. Investigating this from another point of view, by comparing bad performers to good performers on a whole set of variables, might shed more light on the question if there are any other shared characteristics in this group. If so, bad performance on the IGT might be an important neurocognitive marker to be able to identify certain high risk endophenotypes within the offender population in the future (Insel & Cuthbert, 2015).
Objective 3: Investigate if the tendency for risky decision making in prisoners is more closely related to impulsive/reactive aggression than to instrumental/proactive aggression (Chapter 3).

"Aggression is of all times, species and cultures". Those words marked the start of this thesis. Being of all times implies that aggression can be functional in some circumstances, especially from an evolutionary perspective to guarantee survival (Wrangham, 2018). In mild forms certain aggressive traits can be regarded as dominance or assertiveness, which is often valued in Western cultures (Parham, Lewis, Fretwell, Irwin, & Schrimsher, 2015). However, when aggressive behaviour intentionally or unintentionally leads to the infliction of harm to others, it becomes dysfunctional and morally problematic, which is especially true for violent crimes.

Chapter 3 contains the main empirical research topic of part I of this thesis: the investigation of the relationship between risky decision making and aggression in prisoners. Aggression can be operationalised as a dichotomous concept (you can be either problematically aggressive or not aggressive) or a dimensional concept (you can be more or less aggressive with gradual changes, ranging from absolutely non-aggressive to extremely violent in the outer ends of the spectrum). In the present thesis, aggression was mainly regarded in the latter, dimensional manner, which can be seen in the way aggression was measured and statistically analysed. Instead of classifying participants as either problematic aggressors or non-aggressors, aspects of aggression were measured on gradual scales and correlations between those scores and other variables were calculated. Participating prisoners filled in three self-report questionnaires on aggression and hostility, and their behaviour was systematically observed for aggressive tendencies by staff members during four consecutive weeks. Although the outcomes of the systematic literature review pointed to the direction of a relationship between (reactive) aggression and risky decision making, this could not be corroborated by our empirical data. The regression analyses revealed that, apart from age and general intelligence, neither self-reported aggression (as one factor, based on all three self-report measures), staff-observed aggressive behaviour or convictions for violent crimes could be statistically predicted by performance on the IGT (or WCST). Against a priori expectations, this leads to the conclusion that aspects of aggression are not directly related to decision making difficulties or other executive deficiencies in the prisoner group, even though they on average performed worse than non-offenders.

Because there were too many subscale scores to include in the regression analyses based on those three instruments, an exploratory factor analysis was conducted with the aim to create new variables based on multiple instruments: favourably
one combined variable for reactive/impulsive aggression and one for proactive/ premeditated aggression. However, this factor analysis revealed that subscales of the self-report aggression questionnaires (IPAS-30, RPQ and AQ) were intercorrelated in such a manner that no separate factors could be distinguished that were consistent across the instruments showing specific subtypes of aggression. Therefore, it was not possible to validly differentiate between instrumental/proactive and impulsive/reactive aggression or to enter these aggression subtypes in the regression analysis. However, correlational analyses between all scale scores of the self-report questionnaires and outcomes on the neurocognitive instruments could be performed and these analyses indicated that those variables were not correlated in any way. Thus, even though the aggression subtypes were not included in the regression analyses, these additional findings confirmed that there was no relationship whatsoever between reactive or impulsive aggression and risky decision making in our offender sample, and this relationship was also absent between proactive or premeditated aggression and risky decision making. Combining these findings with those of the previous section leads to the conclusion that, in general, offenders show poorer performance on measures of risky decision making than non-offenders, but this difference can not be explained by a specific relationship between risky decision making and (subtypes of) aggression. Instead, it seems to be related to antisocial aspects in general. In line with this, other antisocial aspects have also been linked to risky decision making and disfunctioning of the vmPFC/OFC as well (Maes, Woyke, & Brazil, 2018; Ogilvie et al., 2011; Vassileva et al., 2007). These include: psychopathic personality (investigated in the present study as well, but also not related to risky decision making), general criminal behaviour (for which the same was true) and antisocial personality traits, such as disinhibition/impulsivity, sensation seeking, lack of emotional control, need for instant gratification, self-centeredness, immorality (all not investigated in the present study), etc (Luna, Padmanabhan, & Geier, 2014; Maes et al., 2018; Nystrom & Bengtsson, 2016).

Something completely different but equally important to point out in this section is that in the present study it was not investigated how decision making unravels during a state of heightened aggressive arousal. Apart from the question if these concepts are related in a general manner, it could be that offenders’ performance on the IGT is different during a calm or a negatively aroused state. In one earlier study it was attempted to assess decision making after induction of such a state (Bobadilla et al., 2012), by provoking either proactive or reactive aggression using the Taylor competitive reaction paradigm (Taylor, 1967). Only provoked reactive aggression was related to poorer performance on the IGT (Bobadilla et al., 2012).
This Taylor competitive reaction task only concerns non-serious aggression and has weak or no significant relation to real-life aggression (Ferguson & Rueda, 2009; Ritter & Eslea, 2005) and this study was performed in a student sample. Therefore it is hard to translate these findings to daily life decision making in offenders. As far as known, no studies have been performed in violent/offender samples in which angry/irritated states are induced before assessing decision making. And so, questions remain if risky decision making in this group is a general and persistent phenomenon, or if this can fluctuate dependent on mood and other (situational) factors. It must be said, however, that provoking aggression, even minimally, in such groups for scientific purposes comes with considerable ethical objections and safety issues, which makes conducting such studies challenging and questionable. With respect to situational factors, it is interesting to mention that in one study was found that recently released ex-prisoners showed more risk taking on a decision making task than offenders who were still imprisoned, which is explained by the fact that former prisoners have far more opportunity to engage in risky activities after release from prison than during their detention (Rolison et al., 2013). As far as known, this study has not been replicated yet.

**Objective 4: Determine if the relationship between risky decision making and aggression, when present, is mediated by neurochemical factors (testosterone and cortisol levels) (Chapter 3).**

In the previous section has already been established that there was no direct relationship whatsoever between risky decision making and aggression in our study sample. In spite of that, it was still relevant to investigate if both were related to hormonal processes, especially since this had not been studied altogether before. Endogenous testosterone and cortisol levels were determined by means of saliva samples, while prenatal testosterone exposure was based on the ratio between the length of the index and ring fingers (2D-4D digit ratio). Because only a small number of offenders agreed to provide saliva samples \((n = 38)\), it was decided to not include testosterone/cortisol measures in the regression model as a predictor. Separate analyses revealed that having high endogenous testosterone levels in combination with low cortisol levels was significantly correlated to worse performance on two IGT-measures (score on block 4 and total NET score) and to higher levels of self-reported aggression. No such relation was found for WCST variables, observed aggressive behaviour or number of lifetime convictions for violent crimes. No significant correlations were found for prenatal testosterone exposure and cognitive or aggressive measures.
The relationship between testosterone/cortisol ratios and self-reported aggression was based on the common factor that was derived from all self-report aggression questionnaires. Interestingly, when looking at the single scale scores of the self-report questionnaires separately, none of these significantly correlated to the mean testosterone/cortisol ratio, so it is hard to point out what this relationship exactly may characterise. In spite of this unclarity, these preliminary findings are still new and insightful with respect to biological markers that are involved in risk taking and aggression. This is fortified by the fact that these findings are in line with those of other studies (Apicella et al., 2014; Batrinos, 2012; Carré et al., 2013; Carré et al., 2017; Denson et al., 2013; Mehta & Beer, 2010; Mehta et al., 2015; Montoya et al., 2012; Op de Macks et al., 2016; Popma et al., 2007; Van Honk et al., 2004; Wagels et al., 2017), even though none of those investigated these factors altogether. Our finding that both aggression and risky decision making are related to the same hormonal (im)balances may point to a shared causal (hormonal) factor (probably through its effect on the functioning of the vmPFC/OFC), but no direct relationship (e.g. that a tendency for risky decision making directly causes an increase in aggressive behaviour). In that way, both are merely symptoms of the same underlying problem, but not directly related. This would explain why studies on the relationship between correlations between aggression and risky decision making to date revealed inconsistent findings.

Furthermore, a remarkable coincidental finding in this study was that testosterone levels were remarkably high in the offenders (i.e. 16.8 times larger) when compared to mean testosterone levels in a large group of males in a prior study (Giltay et al., 2012). These findings are not completely in line with those of earlier studies, which revealed that testosterone levels in prisoners were in the normal range (Chichinadze et al., 2010), although testosterone levels were significantly higher in offenders of violent crimes as opposed to committers of non-violent crimes, and higher amounts of aggression or misdemeanour in prison were significantly related to higher amounts of testosterone (Chichinadze et al., 2010; Dabbs, Carr, Frady, & Riad, 1995; Dabbs, Jr., Frady, Carr, & Besch, 1987). Thus, how these present findings can be explained, is still unclear and it must be noted that the sample size was small. These findings therefore first need further replication before drawing any firm conclusions.
Objective 5: Explore if there is a relationship between risky decision making and psychopathic personality traits in a prison population (Chapter 4).

Since psychopathic traits are closely linked to aggression and violence, it was investigated if there was a relationship between such traits and risky response styles in the same offender sample. Psychopathic traits were assessed by means of a self-report questionnaire, the Psychopathic Personality Inventory revised (PPI-R) (Lilienfeld & Widows, 2005). Unfortunately, no valid estimation of the two factors of psychopathy according to the model of Hare (1980) could be subtracted from the data. No convincing evidence was found for a relationship between measures on the IGT and WCST on the one hand and general psychopathy or any specific psychopathic trait on the other hand. If any correlation seemed present, it disappeared largely when correcting for other variables, of which age and history of drug abuse had the most pronounced contribution.

Objective 6: Assess the prevalence rates for (mild) TBI in Dutch prisoners and if TBI is related to measures of aggression and elevated tendencies for the making of risky decisions (Chapter 5).

In chapter 5, results are presented from the study on traumatic brain injury (TBI) in relation to risky decision making and aggression in the same offender sample. Lifetime histories of TBI were examined with a semi-structured interview. Results indicated that prevalence rates in the present Dutch sample were comparable to those in previous international studies: out of the total of 133 participants, 74 (55.6%) reported having sustained one or more incidents with potential TBI. Statistical comparisons between offenders with and without TBI indicated that the offenders with TBI reported significantly higher levels of (reactive) aggression on two out of three questionnaires than offenders without TBI, but no significant differences were found on staff observations of aggression or decision making. Although there was a trend showing higher conviction rates in offenders with TBI compared to the non-TBI group (i.e. TBI-offenders were, on average, convicted for four more crimes than non-TBI offenders), the distribution of those scores was not statistically significant between those groups.

One important issue to address here is the role that TBI can have as a common cause of risky decision making and (reactive) aggression. Even though traumatic brain injury is often linked to damage in OFC and vmPFC regions and risky decision making on the IGT (Bechara, Tranel, et al., 2000; Jonker et al., 2014), as well as to reactive aggression (Buckley et al., 2017; Wood & Worthington, 2017), in the present research support was found for the latter relationship only. Due to methodological
issues in the assessment of TBI, it is too early to draw definite conclusions on the nature of cognitive deficits that offenders with TBI may or may not suffer from. Furthermore, a diverse array of injuries all can be classified as TBI, with inflections to various parts of the brain, each with different implications for cognitive functioning. Therefore, it may be impossible to find shared cognitive problems in this group. Investigating cognition in TBI-patients with heterogeneous lesion locations as a single group may even level out important injury-cognition relations on a group level, and thus mask true and relevant individual impairments as a result of a specific lesion.

**Overall Conclusions, Part I**

Overseeing all of the findings in the first part of this thesis, it can be concluded that the hypothesis of a relationship between taking more risk during decision making and increased reactive aggression in prisoners should be rejected, even though this appeared to be plausible based on the systematic literature review. Instead of a direct relationship between the two, the present investigation provides support that they are indirectly related, as shared symptoms of a more general problem, for example because offenders altogether tend to make more risky decisions and this group is also known for higher aggression levels (thus, general antisocial traits are the common ground), or because both decision making and aggression are influenced by endogenous testosterone/cortisol levels (thus, sharing a common underlying neurochemical imbalance).

Finally, the finding that TBI is highly prevalent in prisoners and is related to higher levels of reactive aggression raises concern for the potential vulnerability of this group with respect to potential recidivism or interpersonal functioning in general.

**Part II: Facial Emotion Perception**

**Objective 7: Investigate if there is a hostile interpretation bias in facial emotion perception in violent offenders, and if this is related to aspects of impulsive/reactive aggression (Chapter 6).**

The second part of this thesis revolved around the relationship between (a hostile interpretation bias in) the perception of emotional facial expressions and (impulsive/reactive) aggression, including its potential as a target for treatment. In chapter 6 the outcomes of the study on this relationship are reported. In this study it was investigated if violent offenders \( n = 71 \) show a stronger tendency to interpret ambiguous facial expressions on a computer task as angry rather than
happy, compared to non-violent offenders \((n = 14)\) and to a control group of non-offenders \((n = 32)\). It was also investigated if hostile perception of facial expressions is related to specific characteristics of aggression, such as proactive and reactive aggression, which was assessed with self-report questionnaires and behavioural observations by staff members during four consecutive weeks. Comparison of the average scores on the emotion perception task between the violent offenders, non-violent offenders and normal controls showed a difference in the expected direction, with violent offenders rating faces as slightly more angry than the other two groups. However, there was no clear statistical evidence to suggest that this difference was meaningful. A power calculation indicated that the achieved sample sizes in the present study were too small to be able to detect plausible group differences on the emotion perception task. The fact that the study was underpowered and that there were unequal groups sizes makes it difficult to interpret the results of our statistical tests. Interestingly, when looking at the data in a more qualitative approach, a similar trend was visible in the curves that represent the percentage of angry responses on each frame of the emotion perception task: especially in the ambiguous faces, the curves showed a slight but consistent tendency for violent offenders to rate these faces as more angry than the non-violent offenders or normal controls. A regression analysis in the violent offender group showed that only age and a self-report measure of hostility predicted outcome on the emotion perception task. Other traits, such as psychopathic traits, intelligence, attention and a tendency to jump to conclusions were not associated with interpretation of anger in facial emotional expressions.

Even though no consistent evidence was found in the present study for a relationship between hostile perception biases and aggression, mounting other studies point to the opposite (Mellentin et al., 2015; Schönenberg & Jusyte, 2014; Smeijers et al., 2017; Wegrzyn et al., 2017), and this is further supported by numerous research performed on hostile attribution biases and reactive/impulsive aggression (Gagnon et al., 2017; Law & Falkenbach, 2018; Martinelli, Ackermann, Bernhard, Freitag, & Schwenck, 2018; Quan et al., 2019). Therefore, it can be concluded that there likely is a relationship between perceiving and interpreting others intentions as angry / hostile and feeling or behaving aggressively oneself, but that it was not possible to show this relationship in the emotion perception task and group sizes that were selected in the present study.

One potential explanation for this discrepancy between these null-findings and previous literature can be explained by the selected experimental paradigm to detect a specific hostile perception bias in the happy-angry spectrum. This task was
selected, because it was already modelled into an intervention procedure that had shown positive effects in aggression reduction in a previous study (Penton-Voak et al., 2013). The selection of this experimental paradigm also implicated certain shortcomings for the present study. For example, in this task was only assessed if there was a bias in the happy-angry spectrum, but not if the same was true in the distinction with other basic emotions. These emotions are not only conceptually and visually different, they are also each characterized by different neurological patterns. For example, experiencing or perceiving happiness is mostly related to activation in the right superior temporal gyrus, while this is the left inferior frontal gyrus for anger, and fear is related primarily to the left amygdala (Vytal & Hamann, 2010). Theoretically, it could be easier to differentiate anger from happiness than to differentiate fear from anger, because the latter are more similar. Therefore, a hostile perception bias in the fearful/angry dimension might be more pronounced than in the happy-angry spectrum. Moreover, when looking at daily life conflict situations, it can be argued that it would be more likely that the opponent in the conflict would experience fear rather than happiness in such situations, which makes the ability to adequately differentiate between anger and fear more relevant for daily life conflict situations than the happy-angry differentiation. In fact, other studies show a stronger hostile perception bias in forensic groups when differentiating with fearful rather than happy expressions (Schönenberg & Jusyte, 2014; Smeijers et al., 2017; Wegrzyn et al., 2017). Including other emotional dimensions in our investigation therefore, might have yielded different results in establishing baseline bias. On the other hand, with respect to the intervention part of the study, the angry-happy dimension did make sense. After all, learning to perceive happiness instead of anger in ambiguous or neutral faces would logically result in the most positive associated change in self-perceived anger, hostility and aggression as well as a similar change on a behavioural level.

Objective 8: Determine if a computerised training, designed to reduce hostile interpretation of emotion perception in ambiguous faces, is effective to restore a hostile interpretation bias in an adult prison population, and whether this leads to a subsequent decline of aggressive traits (Chapter 7).

In addition to the study on the relationship between a hostile interpretation bias and aggression in a prisoner sample, a double blind, randomised controlled intervention study was conducted, which is presented in chapter 7. In this study it was investigated if it was possible to change the perception of emotional expressions in offenders on the same emotion perception task, by providing feedback to the participant after each response (“Correct/Incorrect! That face was happy/angry”). This feedback
was provided in such a way that participants in the intervention condition would shift their threshold scores towards perceiving ambiguous faces as more happy rather than angry. In the placebo control condition participants followed the same training and also received feedback after each response, but this feedback would not result in a shift in threshold scores. At the start of the training the participants (who all were male prisoners) were randomly assigned to either the intervention group (n = 46), or the sham training control procedure (n = 44). Trainers did not know in which condition the participants were placed. Both groups followed the training during five days in one week, approximately 45 minutes a day. The results of the statistical analyses showed that the training procedure was highly effective in promoting the perception of happiness over anger in the training group as compared to the controls, independent of age or intelligence (p < 0.001). These training effects remained at six weeks post training (p < 0.001).

Moreover, it was investigated if this training would result in a subsequent decline of aggressive characteristics and an increase in prosocial behaviour up to six weeks post treatment. This was assessed with self-report questionnaires and staff observations, which started four weeks pre-treatment and lasted up to six weeks post-treatment. The linear regression analyses were statistically corrected for age and presence of (mild) intellectual disability. Contrary to expectations, there was no clear evidence of a change in measures of aggression and hostility, or of an increase in prosocial behaviour, although some trends were visible in this direction on weekly self-report measures.

It can be questioned why it was decided to conduct an intervention program even when no bias could be detected in the first place. The answer to this question lies partly in the practical fact that both studies were conducted simultaneously with different participant groups: the intervention study had already started before the results of the first ‘bias’ study were analysed. Moreover, even in the situation that no bias is present, it would still be interesting to find out if creating a ‘positive interpretation bias’ in the perception of facial expressions could lead to a change in emotional states and behaviour. The fact that the training was highly effective in changing perception of expressions towards such a positive bias in the experimental group certainly shows that this is a malleable process. Additionally, the finding that this effectivity is independent of intelligence levels, is also promising, especially since low intelligence is common in offender groups and offenders with mild intellectual disabilities may be less susceptive of classic, cognitive and verbal therapies. Whether these effects transfer to perception of other faces (in daily life) or remain limited to the experimental paradigm only is something that is unfortunately unknown and
can be questioned. Nonetheless, another study using the same emotion perception paradigm as the present study already found promising evidence that a training with this task results in a generalized effect to other faces (Dalili et al., 2017). Moreover, other comparable intervention studies revealed that such a training can result in an aggression reduction and changed activity in the orbitofrontal cortex (Cougle et al., 2017; Penton-Voak et al., 2013; Schönenberg et al., 2014; Smith, Dillon, & Cougle, 2018; Stoddard et al., 2016). It remains unclear how it can be explained that such training procedures were effective in previous studies but not in the present one.

Overall Conclusions, Part II
In line with the findings in the first section of this thesis on the relationship between decision making and aggression, the research on the relation between a hostile perception bias in facial expressions and aggression also yielded somewhat disappointing results. A priori hypotheses that this bias would be present in violent offenders and would be related to other measures of aggression, could not be clearly confirmed, although there was a trend in the comparison of violent offenders and non-violent offenders and healthy controls and performance on the emotion perception task was significantly associated to self-reported hostility. Furthermore, the training procedure, although highly effective in altering perception of emotional expressions, could not induce a decline in aggressive feelings or behaviour. However, the fact that it was possible to train offenders with a short and non-invasive program, to significantly alter their perception of facial expressions in a positive manner, and that these effects lasted up to six weeks after the training was completed, is promising for future studies.

Limitations and Strengths
In the previous chapters, multiple concerns have already been addressed with regard to limitations of the present research. Therefore, only the major considerations will be addressed here.

One major point of discussion regards the statistical power of this study. Even though a sample size of >100 is fairly large for studies with forensic ingroups, the capacity to statistically detect small true effects is still limited. This was further complicated by the large variance in the data in both parts of the present study. Those aspects increase the chance of making a Type II error (e.g. failing to detect a true effect and wrongfully accepting the null hypothesis). In addition, multiple analyses were performed in the same dataset (this was only true for the dataset from
the exploratory investigations), which on its turn elevates the risk for coincidental findings and thus making a Type I error (assuming an actually non-existing effect and wrongfully rejecting the null hypothesis). The fact that all significant relationships were in line with our a priori hypotheses (which was also true for a lot of the null-findings in this study) and that no seemingly random relations were present, makes the latter error less likely, though certainly not unthinkable. In addition, the size of the control group was considerably smaller than that of the offender sample \( (n = 125 \text{ and } 32 \text{ respectively}) \), which is also not ideal with respect to statistical power. Initially, testing a control group was not part of the study design. Later on in the testing stage it was decided to add a control group, and because of limited time and resources only a group of 32 controls could be tested, resulting in an underpowered study.

Furthermore, in studies such as the present, selection biases always need to be considered. It can be hard to find volunteers for research in offender populations, because antisocial individuals can be especially apprehensive when it comes to psychological research. The ones that do decide to participate might have less antisocial and more prosocial traits than the ones that decline, and thus this sample may be not fully representative of the general prison population. This makes generalization of findings to prisoners in general difficult. The fact that all data were collected in one instead of multiple prisons further complicates generalizability. On the other hand, data were collected from prisoners among different prison regimes (mostly from normal prison wards, but also from wards in the psychiatric centre that are especially geared towards the care for offenders with personality disorders, and wards for repeated offenders). This ensures that the samples at hand were drawn from a varied array of prison settings, in spite of the fact that these were all on the same grounds. Also, when looking at the criminal patterns of the offenders who participated in this study, it can be seen that most had repeated convictions, also for severe (violent) crimes. Thus, it can be argued that the present study sample includes a diverse pattern of offenders, matching general prison characteristics in regards of type and number of committed crimes.

Strengths in the present study can, first of all, be found in the fact that all aspects were measured in a clinical group that is known for (severe) problematic behaviour, but is not easily accessible for scientific research. Another strength can be found in the broad way in which aggression was measured. Not only was self-reported aggression measured in three ways, behavioural measures and criminal records were also included. Second, the incorporation of other factors, such as general intelligence, psychopathic personality traits, testosterone and cortisol values, and
indications for TBI provide a more thorough insight in the complex interplay between factors that may be associated to the relationship between decision making and aggression.

For the intervention study, specific strengths were the strong empirical design (double blind randomised controlled trial, which is not often possible in a prison setting) and the clear theoretical foundation for the intervention procedure.

**Implications for clinical practice and directions for future studies**

To conclude this thesis the reported findings are translated into implications for clinical practice and directions for future research.

In the first part of this thesis on risky decision making in Dutch prisoners and its relationship to aggression the investigation remained limited to fundamental exploratory research. No direct clinical outcome measures were included. However, the value of all of the above ultimately depends on its translatability to forensic clinical practice. Even though no specific relationship between decision making and aggression was found in the present study, it was evident that offenders had more problems with executive functions and decision making than non-offenders, which makes it imperative to investigate if and how these difficulties might be related to clinical factors. In this section it is discussed in what directions such a translation could be made and what is already known from previous studies on the matter.

Advancing from our present investigation, one could ask questions such as: 1) can performance on the IGT or WCST predict drop-out or treatment outcome of forensic psychiatric therapeutic programs, and if so 2) are such tools valuable for the detection of vulnerable groups in the early stages of treatment, 3) can alterations to existing programs be made in order to better suit the learning style of identified cognition-based risk groups and increase their benefit from those programs, or 4) is it possible to train offenders through specific cognitive function training techniques to take less risky decisions and/or to improve executive functions (favourably prior to participation in other cognitive behavioural treatment programs), and finally 5) would such a training be helpful to promote their resocialization after imprisonment and decrease recidivism rates (whether or not for violent crimes specifically)? Although it would be excessive to address all these questions in the present section, some relevant previous findings can be highlighted here. First of all, in a previous study it
has been found that prison inmates who made fewer safe choices on a risky decision making task improved less in terms of reactions to provocation after a standard correctional treatment program (Fishbein et al., 2009). Furthermore, in another study in Dutch prisoners drop-out from cognitive skills program was predicted by performance on a neuropsychological measure for attention (lower attention scores were related to higher dropout rates), even more so than self-reported motivation for treatment, but no relationship was found between performance on the modified version of the WCST and treatment engagement (Cornet, van der Laan, Nijman, Tollenaar, & de Kogel, 2015). No task for risky decision making was included in the latter study. Even though the foregoing points to the possibility that poor cognitive performance can lead to smaller gains during treatment, the opposite has been found as well. In a study by Mullin and Simpson (Mullin & Simpson, 2007) was found that prisoners who showed reduced attentional set-shifting and poor planning skills before treatment displayed the largest improvement on a behavioural level after treatment. Multiple studies showed that effective behavioural treatment in antisocial groups is accompanied by reduced activation in dorsolateral and ventromedial parts of the prefrontal cortex after treatment, suggesting that self-regulatory mechanisms developed towards increased efficiency in result of training (Cornet, de Kogel, Nijman, Raine, & van der Laan, 2015).

There have been studies conducted to train cold executive functions, such as working memory in order to improve self-control (Hofmann et al., 2012; Klingberg, 2010; Olesen et al., 2004), but the extent to which the improvements as result from training generalize and show positive transfer to daily life is strongly debated (Hofmann et al., 2012). With respect to risky decision making in particular it must be said that little is known about its effect on treatment outcome or if cognitive behavioural treatment can affect risky decision making. It can be argued that risk taking could be reduced by means of explicit learning styles, simply by increasing awareness about their decision making styles in individuals who tend to take risk, which in itself might lead to more conscious and deliberated decision making. This does, however, require a certain degree of cognitive control and self-insight, while it is said that this type of decision making often occurs on a more basic, intuitive level (Damasio, 1994). Therefore, it would be interesting to investigate if these decision making processes can also be influenced through implicit learning strategies on a basic level of information processing, such as cognitive and attentional bias modifications programs, including approach-avoidance modification training procedures. The latter is already successfully applied in therapeutic programs for addiction (Eberl et al., 2013; Manning et al., 2016; Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011). As far as known, no such intervention studies are performed to improve decision making in antisocial groups.
Other potential applications for clinical practice also arise in result of the present study. For example, could hormonal therapy aiming to optimise the testosterone/cortisol balance be effective for reducing aggression and improving healthy decision making? Drawing on evidence from animal studies, in which chronic treatment with testosterone propionate has resulted in a decrease in aggression (Martínez-Sanchis, Salvador, Moya-Albiol, González-Bono, & Simón, 1998), it can be suggested that this is a potential area for further investigation in aggressive humans. Hormonal therapies may also have potential for the improvement of social cognition, as a meta-analysis revealed that a single administration of intranasal oxytocin enhances the recognition of facial emotional expressions (Shahrestani, Kemp, & Guastella, 2013). Interesting to point out here is that behavioural interventions in itself can lead to a change in cortisol levels in antisocial groups, although this is mostly assessed in children (Cornet et al., 2015).

Another important consideration to make here involves the question whether neurocognitive measures of risky decision making can add anything substantial to existing taxation instruments to assess recidivism risks and, in result, improve ways to shape risk management for ex-prisoners who show a high 'risk taking profile.' As far as known, only two longitudinal studies have been performed to investigate the predictive abilities of the IGT for recidivism rates (Beszterczey et al., 2013; Miura & Fuchigami, 2017). In one of those studies poorer decision making scores on the IGT significantly predicted recidivism after 3 and 6 months follow up. Recidivists showed a distinct profile from non-recidivists and a normal control group during the final blocks of the IGT, characterized by a failure to learn from feedback and no shift towards selecting more advantageous cards (Beszterczey et al., 2013). In the other longitudinal study IGT scores could not predict future recidivism, however (Miura & Fuchigami, 2017). Interestingly, in a different study it has been found that in the period after release from prison, risk taking in ex-prisoners tends to increase (Rolison et al., 2013). Recidivism was not included as a variable in that study. These findings suggest that making risky choices may be related to increased risk for recidivism after release from prison and can have value for risk assessment in offender groups, although this needs further replication. This would certainly be an important matter to address in future studies.

In addition to all of the above, the subject of TBI in offenders needs explicit attention. The present study underlines that TBI is highly prevalent in forensic samples (often multiple injuries) and seems to be related to higher amounts of self-reported aggression. This makes this group extra vulnerable for future problems and adds to the importance of further investigation on TBI in forensic populations.
For example, it should be further investigated how the potential vulnerabilities and needs of offenders with TBI can be better identified in clinical practice. Not only is it important to adequately diagnose and treat TBI sufferers in offender groups, but this is also an important factor for prevention purposes. Although head injuries are often not preventable, it is important to create public awareness about potential long-term problems after (repeated) mild injuries, especially in relation to high impact contact sports for example (Asken et al., 2016; Haring et al., 2015; Kornguth, Rutledge, Perlaza, Bray, & Hardin, 2017).

To close this section, clinical implications drawn from the second part of this thesis on (influencing) hostile interpretation biases in emotion perception need to be addressed. The clinical advantages of the present training procedure are, first of all, that it is transdiagnostically applicable for all types of offenders, including those with low intellectual or verbal capacities, or who are not familiar with the language. Second, participants are not required to self-reflect or disclose much in therapy, making it less emotionally stressful than other therapies. Furthermore, it is a short program, does not require much skills from trainers and is easy to administer in groups. All of these features can make this training procedure especially attractive for clinical practice.

In chapter 7 of this thesis suggestions were made how a training procedure such as the present one can be made more effective, by incorporating faces of multiple persons, other emotional cues such as intonation of voice or bodily postures and, for example, by re-enacting complete social interactions through virtual reality or gaming procedures. Such virtual reality procedures enable incorporation of multiple social cognitive cues in the procedure, such as posture and intonation, and are interactive instead of static. As such, integrating different aspects of social information processing can make the occurrence of generalizability of trained skills to daily life much more plausible. Even more, in such procedures newly learned skills can be immediately applied by participants, which could further increase its effectivity. For scientific research, a downside of such a procedure would be that it is hard to identify the specific effective components.

In future research on such training procedures it should also be addressed whether such a training procedure adds anything to the already existing aggression treatment procedures. Are the single procedures equally effective, can they be regarded as alternative procedures, or should they be applied in combination with other (cognitive behavioural) techniques? In addition, even though training these emotion perception skills on an implicit cognitive level is important, it is also
interesting to find out if it can also be influenced by means of verbal procedures, merely by increasing awareness through psycho-education about biases in the perception of others’ emotions and intentions.

Another side to this study is its relevance for prison staff members. Is it possible to increase their awareness about potential consequences of their own facial expressions in relation to implicit messages they unintentionally may send to prisoners, and can staff member be trained to influence their own expressions in order to reduce conflict with prisoners? Could that lead to a decrease in incidents during incarceration and thereafter?

One last consideration to address here is how prisoners themselves experience such a training procedure. For this reason, evaluation forms were filled in by all participants in the intervention study, with questions about how they feel about a computerised training, if they are confident that such a training could be effective and if they would advise others to follow such a training. The results from that evaluation are not included in this thesis, because this research was more of a qualitative nature. Such questions are, however, especially relevant for clinical applications of such procedures. It is important that interventions have a certain appeal for participants in order to increase their motivation and commitment to therapy. When approaching the stage of actual clinical implementation of such procedures, these questions certainly need to be further addressed.

Final considerations

The start of this thesis marked the gravity of the main topic ‘aggression’, as it can have major negative consequences for individual wellbeing and society in general. Even though the specific focus on two pieces of the ‘aggression-puzzle’ can be important, the fact that aggression is the result of a complex interplay between various causal and perpetuating factors must not be overlooked and this helps to put the present findings in perspective. In this thesis is confirmed that the relationship between aspects of social cognition and aggression is not straightforward, which is equally true for many topics in psychological research. Nevertheless, the severity of the present subject requires both researchers and clinicians to continue their efforts towards understanding and reducing aggression, in which the current research made a small contribution.