Neurophysiological Correlates of Aggression Related Biased Cognitive Processing in Healthy Adults

Ibrahim Qasem Hakami

Neurophysiological Correlates of Aggression Related Biased Cognitive Processing in Healthy Adults

Ibrahim Qasem Hakami

Cover design by Fatimah Majrashi, F_majrashii@hotmail.com

Layout by Ridderprint BV, www.ridderprint.nl

Printed by Ridderprint BV, www.ridderprint.nl

ISBN: 978-94-6416-020-8

The research presented in this dissertation was financially supported by Imam Muhammad ibn Saud Islamic University

Neurophysiological Correlates of Aggression Related Biased Cognitive Processing in Healthy Adults

Neurofysiologische correlaten van aan agressie gerelateerde vertekende cognitieve verwerking bij gezonde volwassenen

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. R.C.M.E. Engels

en volgens het besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op woensdag 23 september 2020 om 9:30

door

Ibrahim Qasem Hakami

geboren te Jizan, Saoedi-Arabië

Erasmus University Rotterdam

Promotiecommissie:

Promotor:	Prof. dr. I.H.A. Franken
Overige leden:	Prof. Dr. J.W. van Strien Prof. Dr. M.J. Wieser Dr. I.A. Brazil
Copromotoren:	Dr. F.M. van der Veen Dr. J. Huijding

Contents

CHAPTER 1

General Introduction

Introduction

Sometimes, aggression as a violent act can be justified and can even be considered a healthy reaction to harmful events. Moreover, individuals often use aggression as a tool with which to defend themselves. However, when aggression becomes a default reaction, especially in everyday life, it can be maladaptive and may lead to a variety of problems. In fact, acts of aggression affect millions of people throughout the world and are associated with significant personal distress and health issues for both the victims and aggressors. For instance, aggression, along with anger, and hostility, represents one of the most widely studied psychosocial risk factors related to coronary heart disease and premature mortality (Gallo & Matthews, 2003; Kop, 1999; Krantz & McCeney, 2002; Rozanski, Blumenthal, & Kaplan, 1999; Smith & Ruiz, 2002). Aggression is currently not an official disorder included in the International Classification of Diseases and Related Health Problems (10th edition; ICD 10) or the Diagnostic and Statistical Manual of Mental Disorders (DSM-5). However, aggression can be a symptom of a number of mental health disorders, such as intermittent explosive disorder (IED; American Psychiatric Association [APA], 2013), depression (Fava, 1998; Mammen, Kolko, & Pilkonis, 2002), anxiety disorders, antisocial personality disorder (Eronen, Angermeyer, & Schulze, 1998; Moran, 1999), post-traumatic stress disorder (PTSD; Beckham, Moore, & Reynolds, 2000; Ohayon & Shapiro, 2000), borderline personality disorder (Sanislow, Grilo, & McGlashan, 2000; Skodol et al., 2002), and alcohol dependence disorder (Giancola et al., 2009; see Kohn & Asnis, 2003; Swann, 2003).

Why it is important to study aggression?

Aggression has typically been viewed as a crucial component of human behavior due to its functionality as an adaptive device or emergency mechanism; arguably, mankind as a social group could never have survived without aggression (Ellis, 1976). Additionally, Ellis (1976) considered this form of aggression that serves humanity as an adaptive device or emergency mechanism to be positive due to its capacity to promote protection, happiness, social acceptance, preservation, and intimate relationships. Negative aggression, on the other hand, has some destructive and damaging consequences to humans (Barker et al., 2008; Tremblay & Nagin, 2005). For this reason, humans need to learn to control their aggression in such a way that meets social norms. For example, the use of aggressive behavior on a daily basis has decreased in modern societies. Nevertheless, aggressive acts still account for an estimated 1.43 million deaths each year worldwide (Siever, 2008).

As previously mentioned, aggression can have negative consequences for both the aggressor and his or her surroundings. For instance, aggression is a key symptom or characteristic of (developing) anti-social behavior and certain types of psychopathology, and early aggression in terms of presenting in youth may be a good predictor of anti-social behavior and certain types of psychopathology (Blair, 2010; Cleary & Nixon, 2012; Hubbard, McAuliffe, Morrow, & Romano, 2010). Moreover, aggressive behavior can lead to social isolation (Richman & Leary, 2009; Twenge, Baumeister, Tice, & Stucke, 2001). Specifically, in workplace environments, aggression can lead to unwanted conditions, such as uncooperative behavior (Niemann, Wisse, Rus, Van Yperen, & Sassenberg, 2014), decrease in productivity, termination of employment, and even homicide (Pearson & Porath, 2005; Schat & Kelloway, 2000; Schat & Kelloway, 2003). In school settings, aggressive behavior can lead to maladjustment and have a negative impact on the child itself, their classmates and teachers (McConaughy & Skiba, 1993; Wilson & Lipsey, 2006).

Domestic violence, sexual assault, homicide, and school shootings are all clear examples of aggressive acts that lead to serious consequences for the victims (Leary, Kowalski, Smith, & Phillips, 2003; Leary, Twenge, & Quinlivan, 2006). Moreover, aggressive behavior also poses a significant economic burden on society. For instance, a study by Phaedra Corso and colleagues (2007) demonstrated that in 2000, the total cost of interpersonal aggression as a violent act in the United States was $37 billion. This total included medical costs as well as costs incurred due to loss of productivity.

As previously mentioned, humans have the capacity for aggression and the potential for damage; therefore, it is important that researchers focus on studying aggression in order to understand the reason behind this behavior's to persistence within a global context as well as its etiology. It is also important to note that researchers recognize aggression is normally distributed as a continuum within healthy samples (Anderson & Huesmann, 2007; Bowins, 2016), and aggressive behaviors with negative consequences are certainly not exclusively displayed by individuals from clinical samples. Therefore it is essential to study this behavior using non-clinical samples. In fact, many authors have argued that studying the normal range of a given behavior like aggression is necessary to better understand its extreme cases (Anderson, 2012).

Aggression definition

To better understand aggression, it must be clearly defined. In general, human aggression can be defined as any act that harms another individual who is motivated to avoid this harm (Baron & Richardson, 2004). These aggressive acts can be direct (e.g., insults and threats), or indirect (e.g., rumors and gossip). Another distinction that is often made is one between proactive and reactive subtypes of aggression (Roth & Struber, 2009). Reactive aggression can be seen as an emotionally charged response to provocations or frustration (Dodge & Coie, 1987; Kockler, Stanford, Meloy, Nelson, & Sanford, 2006: Stanford et al. 2003). Proactive aggression, on the other hand, is characterized as a conscious and planned act used for personal gain (Blair, Peschardt, Budhani, Mitchell, & Pine, 2006; Blair, 2001; Dodge & Coie, 1987). Although there is a relative agreement that aggression refers to observable behavior, the terms “anger", “aggression", “hostility", “impulsivity", and “behavior” have been used relatively interchangeably by some clinicians and researchers. Yet, it is clear that these concepts are distinct from each other (Suris et al., 2004). It is also important to note that, in the literature, often a distinction is made between state and trait aggression. Trait aggression is considered to be a part of an individual’s personality and is therefore a long-term characteristic that shows through their behavior. State aggression, on the other hand, is a passing condition that individuals experience for a short period of time. Importantly, after this state has passed, individuals will return to their normal condition.

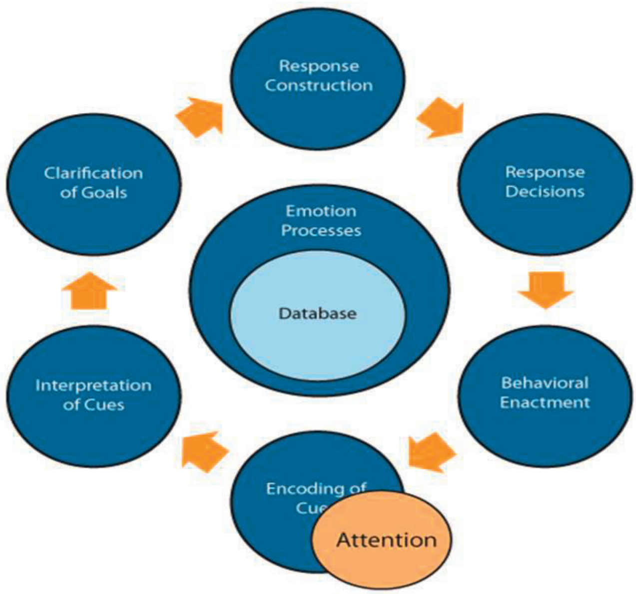
When it comes to the persistence of aggression throughout life, during childhood, aggressive behavior is seen as a part of the normal developmental process (Greydanus, Pratt, Greydanus, & Hoffman, 1992) that tends to disappear when children grow up (e.g., Hawley & Vaughn, 2003; Tremblay et al., 2004). Although this may be true for most children, in some, aggressive behavior persists and reaches problematic proportions as they grow older (e.g., Sawyer et al., 2001; Verhulst, Root, Gullickson, & Ramser, 1996). One consistent finding within aggression research is that aggression is a stable behavior that begins early in life (Huesmann, Eron, & Dubow, 2002; Huesmann & Moise, 1998; Juon, Doherty, & Ensminger, 2006; Loeber & Dishion, 1983; Olweus, 1979; Sharp, 2002; Tremblay, 2000; Zumkley & Frączek, 1992). However, when aggression continues to persist in adolescence, this behavior becomes typically associated with gang activities, cooperative stealing, truancy, and other manifestations of participation in a delinquent subculture (Lopez & Emmer, 2002).

Adulthood aggression, on the other hand, can escalate to include extreme acts such as assault, robbery, rape, and homicide.

Given all the potentially negative outcomes of aggressive behaviors mentioned above, a better understanding of the social, cognitive, and emotional processes involved in the persistence of aggression is required, through which researchers can identify targets for new effective treatment or improve existing treatments (Coccaro, Fanning, Keedy, & Lee, 2016).

The social information-processing (SIP) model and aggressive behavior

Aggression is complex and, for this reason, various theories have been proposed to explain it. Research developments in past years have led to a deeper understanding of the factors involved in aggressive behavior. One theory that has inspired much research is the social cognitive theory. In general, the social cognitive theory suggests that the manner in which one cognitively processes situational input is a strong determinant of one's reaction to a situation. A leading social-cognitive model is the SIP model proposed by Crick and Dodge (1994); it serves as the basis for many studies on social cognition and aggressive behavior. According to this model, individuals enter an unambiguous social situation with collective social knowledge and a record of their previous social experiences. Within a social situation, these individuals receive a massive number of social cues as input, and their behavioral response is a reaction to how they process these cues (Figure 1). That is, the behavioral response depends upon (1) which cues are encoded and which are not, (2) what attributions are made based on the encoded cues, (3) what goals are selected regarding a particular situation, (4) what response options are thought to be available, (5) which response is selected, and (6) behavioral enactment (Erdley, Rivera, Shepherd, & Holleb, 2010). Note, however, that although Crick and Dodge proposed six sequential processing steps, they did not view the nature of the social information processing steps as strictly linear. Instead, they argued that each step in this model may affect another step through a chain of feedback loops. For example, an individual goes with others to a social gathering, they first encode and interpret social cues. In the first two steps, these individuals are guided by their collective social knowledge, which is based on their previous experiences. An individual's knowledge of a given social situation plays an important role in the attributions that individuals make, for example, the interpretation of a colleague's intent. For instance, an individual with a history of being repeatedly mistreated by his colleagues is more likely to attribute an act, for example, a colleague breaking his pencil, to hostile intentions instead of attributing his colleague's act to an unintentional mistake (Erdley et al., 2010). In the third step of the SIP model, an individual decides on possible goals for a given social situation. For instance, the same individual can either retaliate or decide to keep the peace with his colleague. The goal that an individual decide upon as having the highest priority, is likely to produce related behavioral strategies. In the fourth step of this model, an individual starts to engage in response to options, searching his long term memory for a possible behavioral solution. For instance, if the individual selected to retaliate, the response options may include breaking the colleague's pencil in return or punching them in the face. In the fifth step, an individual chooses a specific behavioral response, while the sixth step involves enacting this chosen response (Erdley et al., 2010).



Model of social information processing (SIP) adapted from Crick & Dodge, 1994.

According to the SIP model, aggression can originate in biases formed during Step 2 (interpretation) of processing social situations. For example, imagine a scenario in which a colleague does not smile back at you, as you smile at him or her while you are passing his or her office. Incorrectly interpreting the colleague's intentions (not smiling) as being mad at you could lead to a different response than interpreting that the colleague was busy with his or her work and that he or she never saw you passing. This example shows how interpreting others' intentions as hostile could lead to justifying aggressive responses (Dodge & Coie, 1987). For this reason, a study by De Castro and colleagues (2002) suggested that the way aggressive individuals interpret social situations (Step 2) could play an important role in the persistence of aggression. In line with this, the interpretation of social situations has been the primary topic of empirical aggression studies focusing on the SIP model.

Hostile-Attribution Bias and aggression

Children who usually interpret others' intentions as hostile show what researchers call the Hostile-Attribution Bias (HAB) (Dodge, 2006; Nasby, Hayden, & DePaulo, 1980). It is known that children who show HAB respond more quickly with aggressive behavior (Dill, Anderson, Anderson, & Deuser, 1997). In fact, many studies have confirmed that aggressive behavior in children is related to biased attribution (Brugman et al., 2014; Crick & Dodge, 1996; De Castro et al., 2002; Dodge & Coie, 1987). It is not only children who show HAB, who respond more quickly with aggressive behavior, but adults and adolescents showing HAB are also more likely to interpret an ambiguous action as hostile due to their hostile attribution of intent (Dodge et al., 2015). In fact, Martinelli, Ackermann, Bernhard, Freitag, and Schwenck (2018) found in a systematic review of children and adolescents that having a strong HAB is strongly related to more reactive aggressive behavior. Moreover, in a more recent meta-analytical review using over 25 studies, it has been found higher levels of HAB were associated with higher levels of aggressive behavior in adults (Tuente, Bogaerts, & Veling, 2019). For this reason, previous researchers developed what is called the Cognitive Bias Modification procedures (CBM), in order to reduce and modify aggressive behavior by targeting biased information processing during treatment (Wilkowski & Robinson, 2010).

What is CBM and how does it reduce aggressive behavior?

CBM refers to a procedure that is designed with the intention to modify unwanted cognitive biases using systematic practice in an alternative processing style (Koster, Fox, & MacLeod, 2009). More specifically, CBM is a computer-based treatment that aims to modify automatic, cognitive biases associated with clinical disorders. Specific cognitive biases are manipulated by exposing participants to certain tasks in which a good performance requires the desired processing style. Rather than explicitly working with the patient on thought content and meaning, such tasks implicitly train patients to alter their information processing over many trials and often multiple sessions in order to perform better at the given task (Koster el., 2009).

More recent studies have adapted the cognitive bias modification of interpretation (CBM-I) to target hostile attributions and associated aggressive behavior. For instance, a study by Hawkins and Cougle (2013), who randomly assigned 135 undergraduate students to either positive training, negative training, or a control condition, gave insight into how CBM-I can modify HAB. The positive training in this study led to a decrease in HAB, while the negative training led to an increase in HAB. Along with this, participants in the positive training also showed fewer angry responses in reaction to an insult than participants in the other conditions. Another study by Vassilopoulos, Brouzos, and Andreou (2015) in which they trained a sample of 10–12-year-old children using a three-session HAB training program, found that positive training led to a decrease in HAB, while negative training led to an increase in HAB. Furthermore, AlMoghrabi, Huijding, & Franken (2018) randomly assigned 40 healthy adult male participants to a single session of positive training to increase HAB or a single session of negative training to decrease HAB. Their results revealed that positive training led to an increase in HAB while negative training seemed to have no effect on HAB. These study results support that HAB can be modified using CBM-I procedures, and that positive HAB training may be a promising treatment option for reducing aggression.

Recent meta-analysis and reviews that focused on mental health outcomes suggested that CBM procedures typically are (moderately) effective in reducing cognitive bias (Cristea, Mogoașe, David, & Cuijpers, 2015; Hallion & Ruscio, 2011; Krebs et al., 2017). Although these studies only focused on anxiety and depression outcomes, they suggest that CBM is promising but that there is room for improvement and refinement. It is possible that these results were (moderately) effective in reducing cognitive biases because researchers are not using CBM training in the optimal fashion. Instead of attenuating negative cognitive biases, CBM procedures might only train participants to avoid negative stimuli (Cisler & Koster, 2010; Koster, Baert, Bockstaele, & De Raedt, 2010). This raises the issue of identifying the factors that mediate changes in information processing during CBM. One important factor seems to be the way individuals make use of provided feedback during CBM. Since learning from feedback is of crucial importance to CBM, it might be useful to monitor the way in which it is used to learn in CBM.

Social feedback importance and the neural bases of aggression

Throughout one's life, feedback is an essential part of education, training, and personal development. For example, individuals use social feedback in their daily life, including the emotional facial expressions of others (e.g., happy or angry) to generate and evaluate multiple solutions to a social problem. In fact, happy-appearing facial expressions elicit accepting behavior, while an angry expression elicits feelings of rejection (Seidel, Habel, Kirschner, Gur, & Derntl, 2010; for a review, see Blair, 2003). Interestingly, facial expressions as a form of social feedback can function as signals to others and elicit a specific behavioral response that adapts to the norms and values of the individual's culture (Frith, 2009; Keltner & Haidt, 1999). For example, an angry facial expression may elicit a signal for a given individual that he or she is no longer welcome in the group (socially rejected). The socially rejected individual may use this social feedback (angry facial expression) to better adjust his or her behavior to the group's social norms and thereby increase his or her chance of being included in the group on future occasions. However, for some individuals and in some situations, receiving negative social feedback can result in aggression toward the individuals who previously rejected them (Chester et al., 2014; Chester & DeWall, 2015; DeWall & Bushman, 2011; Leary, Twenge, & Quinlivan, 2006; Riva, Romero Lauro, DeWall, Chester, & Bushman, 2015; Twenge et al., 2001). It must be mentioned that during normal development, individuals tend to learn that not every conflict is a provocation that requires defense and thus acquire more experience with nonaggressive solutions (Hubbard et al., 2010). Reactively aggressive individuals, however, usually do not learn from feedback and continue to behave aggressively (Matthys, Vanderschuren, Schutter, & Lochman, 2012).

Recent studies investigating the neural basis of processing negative social feedback as a form of social rejection have used computerized social rejection paradigms to examine individuals' cognitive biases (Gunther Moor, Crone, & van der Molen, 2010; Kujawa, Arfer, Klein, & Proudfit, 2014; Somerville, Heatherton, & Kelley, 2006; Sun & Yu, 2014; Van der Molen et al., 2014; Van der Molen, Dekkers, Westenberg, Van der Veen, & van der Molen, 2017). For instance, Kujawa et al. (2014) designed a social feedback task (the Island Getaway Task; IGT) in which participants competed against other virtual players to become the last remaining player on the island. To do this, participants voted their co-players on and off the island while receiving both positive and negative feedback from them.

Although these studies investigated the neural basis of negative social feedback, the underlying neural mechanisms of aggression following negative social feedback as a form of social rejection are still largely undiscovered. In fact, little is known about how do differences in aggression and anger affect how people learn from feedback. So far, neuroimaging studies point to the Anterior Cingulate Cortex (ACC) as the brain region associated with social rejection (Cacioppo et al., 2013). More specifically, researchers suggest a potential role of the dorsolateral prefrontal cortex (DLPFC) as an important brain region responding to aggression regulation (Achterberg, van Duijvenvoorde, Bakermans-Kranenburg, & Crone, 2016).

In the last two decades, functional magnetic resonance imaging (fMRI) studies have uncovered a hypo-functionality of the ACC and DLPFC is related to aggression in humans (Ortega-Escobar & Alcázar-Córcoles, 2016; Pawliczek et al., 2013; Van der Gronde, Kempes, van El, Rinne, & Pieters, 2014). These two brain regions are essential for the regulation of the amygdala and hypothalamus, which have been associated with aggressive behavior (Nelson & Trainor, 2007).

Deficits within these two regions may lead to reduced control and thus aberrant activity of the amygdala and hypothalamus, which can lead to more or less aggressive behavior (Ortega-Escobar & Alcázar-Córcoles, 2016; Sterzer, Stadler, Krebs, Kleinschmidt, & Poustka, 2005; Van der Gronde et al., 2014). More specifically, Sterzer and colleagues (2005) have attempted to understand the reduced ACC activation observed during participants' viewing of negative pictures as an interference of emotional states with cognitive processing, which results in a failure to cognitively control and regulate emotional behavior. Although both proactive and reactive aggression have been associated with the hypo-responsiveness of the ACC and DLPFC (Anderson & Kiehl, 2014; Dambacher et al. 2015; Patrick, 2008; Perach-Barzilay et al. 2012), they differ in their dependency on the activation of subcortical structures. This is due to the fact that proactive aggression is premeditated in nature and is thought to be regulated by higher-order prefrontal cortical systems (i.e., ACC and DLPFC), resulting in less dependency on the amygdala and hypothalamus. Reactive aggression, on the other hand, is more impulsive in nature and is thought to be dependent on the hypothalamic and limbic systems, resulting in a decrease in prefrontal cortical regulation (Nelson & Trainor, 2007). In addition to this, there are some studies suggesting the anterior midcingulate cortex (AMC), a sub-region of the ACC, to be one of the neural nodes underlying the experience of anger and hostility (Nakagawa et al., 2017), which typically seen as an example of reactive aggression, because it is believed to be the reason behind expecting negative experiences in future social events. Accordingly, people's perception of negative social feedback from others can elicit aggression as a response in which these people defend themselves from expected further harm.

Finally, not only is the ACC involved in negative social feedback and social rejection, but it has also been suggested to be the neural node underlying aggression. For example, aggressive individuals typically show decreased activation in the ACC, and the ACC has been associated with self-monitoring and behavioral regulation (Davidson, Putnam, & Larson, 2000). Although fMRI studies have uncovered a hypo-functionality of the ACC in relation to aggression in humans, such studies are still expensive to conduct (Crosson et al., 2010). Furthermore, most MRI scanners are large, fixed machines situated in hospitals, making them inconvenient for some research purposes. Therefore, researchers often prefer to use inexpensive and portable but nonetheless high-quality accurate brain activity measuring devices that can help shed light on the early stages of information processing (Kujawa et al., 2014) and reflect ACC activation (Holroyd & Coles, 2002).

Event-Related Potentials (ERPs) to measure feedback in the early stages of information processing

One way to study feedback, other than using an fMRI, is to use electroencephalogram (EEG) event-related potentials (ERPs) because they can shed light on the time course of the early stages of information processing (Kujawa et al., 2014). ERPs have provided scientists with a better method for studying cognitive processes in general (Coles, 1989). ERPs are electrical potentials produced in the brain in response to specific events or stimuli (Fabiani et al., 2007). For example, when individuals come across a specific event, a series of neural units are triggered to process this event. Researchers call these electrical potentials ERPs and refer to their segments as components (e.g., N400, ERN, and FRN). An N400, for instance, is an ERP component particularly sensitive to modulations of the meanings of presented stimuli at a semantic and associative level (Kutas & Federmeier, 2000). Researchers have defined the N400 as negative deflections that manifest 400 ms after stimulus presentation (White, Crites, Taylor, & Corral, 2009). Along with this, N400s have been found to be modulated by violations at the level of semantics or meaning. In particular, N400s are supposed to reflect the ease with which a stimulus is integrated into a given context (Kutas & Federmeier, 2000). This is also based on the cognitive effort required by individuals to access information stored in their long-term memory (Kutas & Federmeier, 2000). For example, items incongruent in one context compared to congruent items can elicit N400s with larger amplitudes. In addition, the N400s are elicited not only by semantically incongruent sentence endings (Kutas & Hillyard, 1980, 1984) but also by the second word or target in semantically incongruent word pairs (Bentin, McCarthy, & Wood, 1985). There are some findings that point also to the importance of the emotional context itself for the N400. For instance, Herbert, Junghofer, and Kissler, (2008) found that the emotional content of words significantly reduced the N400 amplitude, and this was more the case for pleasant rather than unpleasant adjectives. Therefore, the N400 can be used to understand the neural underpinnings of the ways in which ambiguous words are interpreted and, thus the neural basis of HAB. However, before this can be done, it needs to be ascertained whether the N400 modulation is affected by the emotional content of a word and not its physical appearance. For example, most ERP research examining emotional processing utilizes stimuli that physically differ from each other (Delaney-Busch & Kuperberg, 2013; Kutas & Hillyard, 1980; Holt et al., 2009). Therefore, investigating the role of an emotional versus non-emotional context under more stringent conditions is a necessary first step toward determining whether the N400 can be a useful measure of emotions in the context of aggression.

There is another component of the ERP found to be associated with performance monitoring that can shed light on the early stages of information processing (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, & Meyer, 1993). Studies call this component Error-Related Negativity (ERN; Falkenstein et al., 1991; Gehring et al., 1993). According to the error detection hypothesis, the ERN is a response-locked ERP in the form of a negative deflection that peaks between 0 and 100 ms after an error commission (Gehring, Liu, Orr, & Carp, 2012). The ERN is found to be reduced in types of psychopathology that are closely related to reactive aggression (Hall, Bernat, & Patrick, 2007; Santesso, Segalowitz & Schmidt, 2005; Stieben et al., 2007), and its source is thought to be located in the ACC (Carter et al., 1998; Dehaene, Posner, & Tucker, 1994). The ERN is assumed to signal other brain regions to inhibit or correct errors in progress, and to avoid future errors via enhanced control strategies (Gehring, et al., 1993). Thus, ERN provides an index of internal feedback processing.

Similar to the ERN, there is another component also known to reflect ACC activation (Holroyd & Coles, 2002), called the Feedback-Related Negativity (FRN). FRN is a feedback-locked ERP component in the form of a negative-going deflection located at front-central electrode sites and peaks between 250 and 350 ms after the onset of feedback stimulus (Gehring, et al., 2012; Miltner, Braun, & Coles, 1997). The FRN amplitude has been found to be sensitive to negative feedback (Hajcak, Moser, Holroyd, & Simons, 2006;

Holroyd & Coles, 2002; Kujawa et al., 2014; Yeung & Sanfey, 2004). Finally, both the ERN and FRN components have been associated with feedback processing in the ACC (Holroyd & Coles, 2002; Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Miltner, et al., 1997; Ullsperger & Von Cramon, 2003), which is an area thought to be essential to response reinforcement associations (Rushworth, Behrens, Rudebeck, & Walton, 2007) and performance monitoring (e.g., Holroyd & Coles, 2002). Therefore, the ERN and FRN components are very interesting candidates for examining the neural underpinnings of (social) feedback learning in aggressive individuals in the context of CBM. This is specifically interesting because previous evidence has shown that aggressive individuals appear to have difficulties with learning from previous experiences (e.g., Matthys et al., 2012).

A search of the literature revealed that most studies that have examined HAB and aggression have only been performed in criminal and psychopathic individuals, institutionalized in prisons or mental security facilities, whereas studies on aggressive individuals within the general population are less common, especially those using EEG techniques. For example, there is considerable evidence that late negative ERP amplitudes observed in psychopaths during linguistic tasks are possibly related to larger N400s in this sample (Kiehl, Laurens, Bates, & Liddle, 2006; Niznikiewicz, et al., 1997; Williamson, Harpur, & Hare, 1991). However, to the best of my knowledge, there are only two studies that examined the relation between N400 and aggression during linguistic tasks in healthy male adult samples (Gagnon et al., 2016; Gagnon, et al., 2017). In fact, the aggression studied in most of the previous studies (target; behavior aggression) was only symptomatic of some other disorder. This raises the question of whether previous results of aggression studies can be generalized to aggressive individuals in the general population. Given that aggression is normally distributed as a continuum within healthy samples (Anderson & Huesmann, 2007; Bowins, 2016), it is essential to study this behavior using non-clinical samples.

As for ERN and FRN, the morphology and topography of these two components are highly similar. Moreover, these two components reflect the activity of feedback processing during learning. For instance, FRN reflects the internalization of external feedback, and ERN reflects an internal feedback loop (Holroyd & Coles, 2002). In line with this, there is evidence from reinforcement learning tasks that the magnitude of ERN increases as FRN decreases (Holroyd & Coles, 2002; Nieuwenhuis, Yeung, Van Den Wildenberg, & Ridderinkhof, 2003). For example, Holroyd and Coles (2002) tested changes in ERN and FRN amplitudes during probabilistic learning: participants had to learn through trial and error, with a companion of feedback (i.e., rewarded vs penalized), which button to press in a two-choice decision task. The results showed an increase in ERN amplitude with learning and a decrease in FRN amplitude with learning. It has been suggested that the increase in ERN with learning reflected the development of an internal representation of the correct response (Holroyd & Coles, 2002). That is, learning the correct response can increase the mismatch signal after the incorrect response is presented. On the other hand, the decrease in the FRN with learning might have been due to the redundant information value of the feedback stimulus. Therefore, we can use ERN and FRN as indicators of learning and tools with which to monitor the process of change in information processing during the course of CBM training. A larger shift from FRN to ERN in this context would indicate a high degree of internalizing and possibly more successful learning following training.

Aims and Hypothesis

The major aim of this dissertation was to use ERPs to examine aggression in healthy male adults to shed light on the early stages of information processing to identify the factors (i.e., feedback) that mediate changes in information processing among this sample.

More specific aims include the following:

1- To examine whether the neuronal reflection of feedback processing (FRN) is associated with self-report measures of aggression, and to provide an assessment of the validity of the IGT paradigm to determine the neural correlates of HAB in aggression.

2- To examine whether the N400 is a useful measure of emotions in the context of aggression by examining (under stringent conditions) whether the negative emotional content of a word can elicit a larger

N400 effect as compared to the neutral emotional content of a word, and to provide an assessment of the validity of the homonyms task.

3- To examine whether higher scores of self-report aggression measures are associated with a greater Aggression Related Interpretation Bias (ARIB) in the Aggressive Interpretation Task (AIT) while using behavioral and EEG methods. The AIT is a lexical decision task in which word pairs are primed, and participants are asked to decide whether the second word (target) is a Dutch word or a non-word. The first word (cue) is an ambiguous or unambiguous word. Target words are either associated words in a violent context, associated words in a neutral context, un-associated words, and non-words.

4- To examine whether self-report aggression and anger are related to FRN and ERN during an attribution training for facial expressions; to determine whether FRN and/or ERN are related to changes in HAB from pre-to post-training; and to provide an assessment of the validity of the Face task (FT; facial expressions training) paradigm to determine the neural correlates of HAB in aggression.

The importance of the current dissertation

The current dissertation is important in replicating previous findings that showed associations between behavioral and psychophysiological reflections of HAB as well as self-report measures of aggression on the other hand in healthy male adult samples. Furthermore, the current dissertation can further inform the development of prevention efforts and future intervention studies, stressing the important role of feedback processing in the development and persistence of the HAB.

Overview of the current dissertation studies

Study 1: The primary objective of this study was to examine whether the neuronal reflection of feedback processing (FRN) was associated with self-report measures of aggression. This study was based on the hypothesis that aggressive behavior persists when individuals have a learning deficit (e.g., Matthys et al., 2012), which can be measured via the FRN. Therefore, it used a social feedback paradigm (the island getaway; IGT) in order to measure this ERP process. The Island Getaway task (IGT) is a computer aided design task inspired by the television game show "Survivor" and consists of both a social feedback and a peer rejection element (Reijntjes, Stegge, Terwogt, Kamphuis, & Telch, 2006a; Reijntjes, Stegge, & Terwogt, 2006b; Kujawa et al., 2014). The current study hypothesized that FRN is mainly related to the extent to which an individual will learn and adapt based on social feedback. We expected that aggressive participants would be less sensitive to negative social feedback and therefore that participants who score high on reactive aggression would have a smaller FRN amplitude following the rejection feedback.

Study 2: The primary objective of this study was to examine whether the N400 is a useful measure of emotions in the context of aggression. Therefore, under stringent conditions we examined whether the negative emotional content of a word can elicit a larger N400 effect when compared to the neutral emotional content of a word. Specifically, we were interested in studying whether homonym words with either no emotional meaning or a negative emotional meaning differed with respect to the N400 response. A secondary objective for this study was to examine whether the affect (positive or negative) and anxiety level (trait or state) of the participant had an effect on the magnitude of the N400 effect, since previous studies had found an influence of mood on the N400 effect. This study employed the homonyms task, in which the emotional words used were identical to the neutral words in form and articulation and only differed in semantic content, which we made dependent on the context in which the word was presented. For the brain measures, we expected a significant enhancement of the N400 effect to be found for the homonym within an emotional context in comparison to the homonym embedded in a neutral context. More specifically, we expected the N400 to be elicited exclusively by the emotional content of the word. Finally, we expected to find an association between affect and anxiety as well as between affect and the N400.

Study 3: The primary objective of this study was to examine a healthy undergraduate population with implicit measures of bias to ascertain whether higher scores of self-report aggression measures were associated with greater

ARIB. ARIB was examined using the so-called AIT. Participants in this study performed a lexical decision task and viewed different types of target stimuli (associated words in a violent context, associated words in a neutral context, un-associated words, and non-words). We expected higher scores on the self-report aggression measures to be associated with a faster reaction time (RT) to violent context targets but not to neutral context targets. For the brain measures, we expected higher scores on the self-report aggression measures to be associated with smaller N400 amplitudes only in violent context targets and not in neutral context targets.

Study 4: The primary objective of this study was to examine whether self-report aggression and anger are related to FRN and ERN during an attribution training for facial expressions and to determine whether FRN and/or ERN are related to changes in HAB from pre-to post-training; and to provide an assessment of the validity of the FT paradigm to determine the neural correlates of HAB in aggression. HAB was examined by using the FT experimental paradigm developed by Penton-Voak et al. (2013). In this chapter, we expected higher scores of the self-report aggression measures to be associated with smaller ERN and FRN responses. In addition, we expected larger FRN and ERN responses to be associated with a bigger change in HAB throughout the training. We also expected participants with higher reactive aggression scores to perceive anger facial expressions more often than those with low scores during the facial expressions training. Finally, we expected these participants to alter their perceptions of ambiguous emotional expressions during the training and perceive more positive facial expressions.

CHAPTER 2

Can We Get Along? The relationship between Feedback Related Negativity and Reactive Aggression in Healthy Individuals During Social Feedback Processing

THIS CHAPTER HAS BEEN SUBMITTED FOR PUBLICATION AS:

HAKAMI, I., F.M. VAN DER VEEN, FRANKEN., I.H.A. & HUIJDING, J.

CAN WE GET ALONG? THE RELATIONSHIP BETWEEN FEEDBACK RELATED NEGATIVITY AND REACTIVE AGGRESSION IN HEALTHY INDIVIDUALS DURING SOCIAL FEEDBACK PROCESSING.

Abstract

Recent research suggests that aggressive individuals fail to learn from social feedback. This study tested whether a brain measure of feedback processing (Feedback-Related Negativity: FRN) during a social feedback task (the Island Getaway Task: IGT) was related to self-report measures of reactive aggression, trait anger, and aggression (RTA). Male university students (18-29 years; N = 188) played the IGT with virtual peers, which involved receiving rejection and acceptance feedback from others. Results did not show a difference between positive and negative feedback or the association between FRN and other RTA. Behavior (voting) was not associated with reactive aggression or FRN. To conclude, our results suggest that the IGT as a tool to test behavioral and brain responses to social feedback among the adult healthy male population, depend on the physical appearance of the feedback stimuli.

Introduction

Reactive aggressive behavior is part of normal and adaptive development, it is common in pre-school aged children and tends to subside over time (e.g., Hawley & Vaughn, 2003; Tremblay et al., 2004). However, in approximately 12 to 14% of the children, this behavior persists and reaches problematic proportions (e.g., Sawyer et al., 2001; Verhulst, Root, Gullickson, & Ramser, 1996). During our psychological development, social feedback on our behavior teaches us to pay attention to cues that disambiguate the situation (e.g., an emotional facial expression, a blush) and to interpret these situations in a non-hostile way. Over time, most individuals acquire this ability, learning from their mistakes as they go along. However, some individuals do not acquire this ability fully, with aggressive behavior as a result. This aggressive behavior poses a significant social and economic burden on society. For instance, the estimated cost of interpersonal violence in the U.S. alone in 2004 was more than $300 billion per year, including medical costs and costs due to loss of productivity (Waters et al., 2004). These adverse consequences of aggression underline the importance of understanding what causes the aggressive behavior to persist.

There is some evidence that individuals showing reactive aggression may be characterized by a failure to learn from previous experiences (e.g., Matthys, Vanderschuren, Schutter, & Lochman, 2012). It has been found, that aggressive individuals who look at affective pictures show a decreased activation in both the anterior cingulate cortex (ACC) and the orbitofrontal cortex (OFC) in response to strong negative affective pictures (Stadler et al., 2007). Both these brain structures are involved in self-monitoring (Davidson, Putnam, & Larson, 2000). More importantly, several studies highlighted the fundamental role of the ACC in conflict processing and error monitoring (Holroyd & Yeung, 2012; Holroyd & McClure, 2015; Shenhav, Botvinick, & Cohen, 2013).

There are -at least- two important electrophysiological indices of self-monitoring. The first is the Error-Related Negativity (ERN; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, & Meyer, 1993), a negative-going event related potential (ERP) component time-locked to (incorrect) responses at the fronto-central midline, which peaks between 0 and 100 ms after an error has been committed (e.g., Gehring, Liu, Orr, Carp, 2012), and which has been associated with the functioning of the ACC. It has been found that the ERN is reduced in individuals with behavioral problems associated with reactive aggression. For instance, ERN amplitude is reduced in under socialized children (Santesso, Segalowitz & Schmidt, 2005), and in adults (Hall, Bernat, & Patrick, 2007), and children with externalizing problems (Stieben et al., 2007).

A second electrophysiological response that is more associated with self-monitoring, is the so-called feedback-related negativity (FRN). The FRN is a fronto-central negative ERP component that occurs 200–300 ms after negative feedback is presented (Holroyd & Coles, 2002; Talmi, Atkinson, & El-Deredy, 2013).On the other hand, another positive neural component occurs between 200 and 300 ms after positive feedback is presented. This component is referred to as the reward positivity (RewP) in which is superimposed on the FRN and thus contributes to the net effect that is measured (Baker & Holroyd, 2011; Holroyd, Pakzad-Vaezi, & Krigolson, 2008). In fact, empirical evidence seems to show that especially reward feedback elicits a positive deflection. This positive deflection seems to be absent in response to the loss feedback trials, however it is added to the "baseline" response that is negative in polarity (Holroyd et al., 2008; Proudfit, 2015). Reinforcement Learning Theory (RLT) posits that the ERN reflects error-monitoring after the learning phase (i.e., the phase in which a person has learned the correct and incorrect responses). In this phase, there is no or little uncertainty about the correct action. In contrast, the FRN reflects the internalization of external feedback, which takes place before the learning phase. In this phase, there is still uncertainty about the correct response and responses are guided by the external feedback. In fact, it has been hypothesized that, during learning there is a shift from FRN (learning phase) to ERN (post-learning phase; Eppinger, Kray, Mock, & Mecklinger, 2008; Holroyd & Coles, 2002; Pietschmann, Simon, Endrass, & Kathmann, 2008; Pietschmann, Endrass, Czerwon, & Kathmann, 2011). In line with this notion, the FRN is thought to be a reinforcement learning signal that is used to modify behavior with negative outcomes and reinforce the behavior with positive outcomes (Holroyd & Coles, 2002). Since reinforcement learning is crucial for displaying appropriate social behavior, we hypothesize that particularly the FRN is a highly relevant brain measure for understanding the specific learning processes that are involved in aggressive behavior.

One way to investigate feedback processing in an ecologically valid way is by using emotional facial expressions as feedback. Emotional facial expressions represent social support versus disapproval and are strong stimuli in social evaluation. Similar to non-social feedback (i.e., financial rewards), also social feedback, such as positive or negative facial expressions, can elicit an FRN. The non-social FRN is known to reflect a binary (good/bad) evaluation of the outcome (Hajcak, Moser, Holroyd, & Simons, 2007), and in analogue to this, the social feedback should also be explicit and clear and contain explicit positive or negative expressions. Holroyd and Coles (2002) stated that the FRN represents a reward prediction error, that is, a signed value corresponding with the difference between the amount of reward obtained and the prior expected value of the reward. In the same vein, implicit social norms could warrant acceptance instead of social rejection or fair instead of unfair offers in bargaining situations. If acceptance is expected but rejection received, one would anticipate that the FRN should be more negative to reflect this negative prediction error.

Several electroencephalogram (EEG) studies have been conducted and found some evidence for the idea that FRN is relevant for understanding social feedback processing. For example, the FRN has been found to be more sensitive to the expectation of social feedback than the valence of that feedback (e.g., acceptance vs rejection). In particular, the FRN is larger when social feedback is unexpected than when the feedback is expected, both for positive and negative feedback (Dekkers, van der Molen, Moor, van der Veen, & van der Molen, 2015; Van der Molen et al., 2014; Van der Molen, Dekkers, Westenberg, Van der Veen, & van der Molen, 2017; Van der Veen, van der Molen, Franken, & van der Molen, 2016). Despite the fact that the FRN is a promising marker for studying social feedback learning, to date, there is little research on the FRN in the context of aggression. There is some evidence that the FRN is related to aggression in non-social tasks, although the direction of effects is inconsistent. Some studies found that the FRN amplitude during simple monetary gambling tasks was inversely related to self-report aggression (Ma et al., 2016; Yi et al., 2012), suggesting that reported aggression associated with decreased feedback sensitivity. In contrast, violent juvenile offenders showed a larger FRN amplitude than a control group, suggesting an increased sensitivity for feedback (Vila-Ballo et al., 2015). We only found one study addressing the relationship between the FRN and aggression during a social task. In this study, Pincham and colleagues (2015) found that

FRN amplitude was reduced during a modified Taylor Aggression Paradigm and was negatively related to self-report hostility.

As reviewed above, there is some evidence that the FRN is sensitive to social feedback and expectations, but there is hardly any data regarding the FRN of aggressive individuals in the context of social feedback learning. Therefore, the primary objective of this study is to examine the relation between FRN and aggression in the context of social feedback and to test the idea that reactive aggressive individuals are characterized by a deviant social feedback processing. More specifically, we examined whether the FRN responses are smaller for people who score high on the reactive aggression measure during a social feedback (using the Island Getaway Task). The Island Getaway task (IGT) is a computerized task inspired by the game show Survivor and consists of both a social feedback and a peer rejection element (Kujawa, Arfer, Klein, & Proudfit, 2014). In this task, participants play a game with virtual peers, in which participants vote to reject and accept co-players, but also receive a combination of rejection and acceptance feedback from other co-players themselves. In a pilot study (Kujawa et al., 2014), including 20 children and adolescents, it was demonstrated that participants alter their voting behavior in response to peer feedback, and that rejection feedback was associated with larger FRN than acceptance feedback. In addition, participants who showed larger FRN responses to social feedback were less likely to reject co-players. As such, this task seems to be sensitive to the role of the FRN in adaptively responding to social challenges. We hypothesized that FRN is mainly related to the extent to which an individual will learn and adapt based on social feedback. We expected that aggressive participants would be less sensitive to negative social feedback and therefore that participants who score high on reactive aggression would have a smaller FRN amplitude following the rejection feedback.

Material and Methods

Participants

A total number of 200 participants from the Erasmus University Rotterdam (EUR) student population were tested, of which 188 participants (age range 18-29; mean age 20.85 years; *SD* = 2.19) had a complete data set. Participants were recruited from or within the proximity of EUR and received payment of 25 euro for their voluntary participation. All participants had no history of psychiatric illness, history of neurological disease, and were free from the use of a psychoactive medication, which can affect their cognitive performance. This study was conducted according to the local and ethical guidelines of the Institute of Psychology at Erasmus University Rotterdam.

Self-report measures

The Reactive-Proactive Questionnaire (RPQ) is a self-report instrument that is designed to measure and to distinguish between reactive and proactive aggression (Raine et al., 2006). The RPQ consists of 23 items: 11 items measuring reactive aggression, and 12 items measuring proactive aggression. These items are rated on three-point scale: 0 = (never), 1 = (sometimes), or 2 = (often). According to Raine et al., 2006 findings, RPQ is a reliable and valid self-report instrument that can be used to assess aggression subscales: reactive aggression, proactive aggression, and total aggression. In this study, we examined healthy college-aged participants by using the Dutch version of RPQ. The RPQ also has been tested in the Dutch language, and it has been proven a reliable and valid instrument (Cima, Raine, Meesters, & Popma, 2013). Cronbach's alpha in the current sample was 0.96 for reactive aggression subscale, indicating good internal consistency.

For further exploratory analysis, the State-Trait Anger Expression Inventory-2 (STAXI-2) was used. The STAXI-2 is a questionnaire meant to assess the intensity and frequency of anger through 57 items on a 4-point Likert scale. These items are spread through six different dimensions: state anger, trait anger, anger expression in, anger expression out, anger control in, and anger control out. Cronbach's alpha in the current investigation for the entire list is 0.75, which indicates acceptable internal consistency.

For further exploratory analysis, the Trait Aggression (AVL) was used. The AVL is the Dutch version of Buss Perry's Aggression Questionnaire. This questionnaire consists of 29 items scored on a 5-point Likert scale, which are further divided into 4 dimensions: physical aggression, verbal aggression, anger, and hostility. AVL Cronbach's alpha in the current sample for the entire list 0.80, indicating good internal consistency.

Stimuli and experimental paradigm

The Island Getaway task (IGT) is a computerized task inspired by the television game show "Survivor" and consists of both a social feedback and a peer rejection element (Reijntjes, Stegge, Terwogt, Kamphuis, & Telch, 2006a; Reijntjes, Stegge, & Terwogt, 2006b; Kujawa et al., 2014). In this task, participants were led to believe that they are competing with twelve virtual players (6 Males, and 6 Females) to be the last player on a Hawaiian Island. Ages of the co-players ranged from 18 to 24, and locations were distributed throughout the Netherlands. Co-player photographs were taken from stock photographs available online from the Radboud Faces Database (RFD). All photographs that were taken from RFD have the same neutral expression, and all of them have the same natural background. All the (RFD) photographs that we used in the IGT task were validated with respect to the shown facial expression, the intensity of expression, clarity of expression, genuineness of expression, attractiveness, and valence. (Langner et al., 2010). Before the start of the actual task, participants were asked to provide basic demographic information (name, age, school, hometown, and Nationality). Then, participants were instructed on how to play the game. During the IGT task, participants competed with each other over eight rounds. In addition to that, participants voted to accept and reject other co-players. The feedback stage consisted of a fixation mark "+" that presented for 1000 ms, followed by the co-player's profile picture that stated his or her name for 1000 ms, and another fixation mark "+" for 1000 ms and then the feedback stimuli. As soon as the participant finishes receiving feedback from all the co-players, he was presented with the co-player that was voted of the island and removed from the list. At the start of every new round, using the keyboard, the participants responded to some personal questions regarding life experiences, interests, or beliefs. These personal questions were used to make the task more realistic and engaging for participants, as they believe that their acceptance or rejection from other co-players, was based on their answers. In the last round, participants were told that they survived the game along with the last there surviving co-players. Finally, it is important to mention that previous studies (Cao, Gu, Bi, Zhu, & Wu, 2015; Kujawa et al., 2014) have used a green checkmark to indicate that the co-player voted to keep the participant and a red X to indicate that the co-player voted for the participant to leave the game. However, in order to make the feedback signals more neutral and to avoid physical differences between the feedback signals, the current study employed a yellow rectangular shape to indicate acceptance feedback, whilst a blue rectangular shape denoted that the player was voted off the island. This feature was counterbalanced for every other participant, where yellow and blue switched from positive to negative feedback and vice versa. See Figures 1 and 2 for a schematic illustration of the IGT task.

Procedure

At the start of the visit, participants were instructed to fill out three measures (RPQ, AVL, and, STAXI-2). Then, participants were seated in front of a computer screen in a comfortable chair in a light and sound-attenuated EEG room. The EEG equipment was attached, and a brief instruction was given about the task before the measurements started. Then, to reinforce the social nature of the task, participants were reminded that they are going to play a multiplayer online game with 12 other players. Subsequently, participants started the experiment by reading the task instructions that appeared on the screen and they were given a practice trial followed by the real experiment.

Electroencephalogram acquisition and analyses

EEG were recorded continuously at a sampling frequency of 512 Hz with a 32 - channel Active Two system (BioSemi, Amsterdam, the Netherlands), using Ag-AgCl electrodes mounted in an elastic cap. Six electrodes were added. Two electrodes were placed on both mastoids that served as reference electrodes. Two electrodes were placed on the outer canthi of each eye to measure the horizontal electro-oculogram (HEOG), and the final two electrodes were placed in the supraorbital region and in the infraorbital region of the left eye to measure the vertical electro-oculogram (VEOG), which was used to correct for eye movements. The passive electrode of Driven Right Leg (DRL) and the active electrode of Common Mode Sense (CMS) were added to use as a feedback loop as amplifier reference. Offline analysis of the EEG time series was performed using Brain Vision Analyser 2 (Brain Product GmbH, Munich, Germany). FRN data were filtered with standard filter settings of 0.1-30 Hz. Segments containing artifacts were rejected using a semi-automatic artifact rejection procedure with a maximum allowed voltage 75 μV. Epochs containing artifacts other than eye blinks (e.g., muscular activity, clipping, and movement artifacts) were removed from the data. Other ocular artifacts were corrected by using the (Gratton, Coles, & Donchin, 1983) procedure. Bad channels were interpolated with neighbouring channels. The continuous EEG was segmented into an epoch that starts at -200 ms before the onset of the feedback and last until 800 ms after feedback. EEG was averaged separately for rejection and acceptance trials. Baseline was defined as the 200 ms preceding stimulus onset. FRN amplitude was computed as an area window between 225-300 ms after feedback, at Fz, FCz, and Cz.

Statistical analyses

Statistical analyses were performed using SPSS 25. To measure the voting behavior, we computed the percentage 'vote off' responses, and we used Pearson correlations to examine whether there is an association between voting behavior and reactive aggression scores. To test whether FRN negative social feedback was associated with a larger FRN amplitude, which is maximal at fronto-central positions, a 2-feedback type (positive vs negative) × 3 electrode positions (Fz, FCz, and Cz), repeated-measures ANOVA was employed. We also performed another GLM analyses with FRN amplitude as the dependent variable, and feedback level and electrode positions as within-subject variables, but now with Reactive aggression, Trait Anger, and Trait Aggression added as covariates sequentially. When the assumption of sphericity was violated, Greenhouse-Geisser corrected degrees of freedom were used to compute the corrected *p*-value but uncorrected degrees of freedom are reported. Effects size is reported as partial eta squared (η_p^2) and effects were considered significant at $p < .05$.

Results

Behavioral Results

On average, participants voted co-players out of the game 35.9% of the time (SD = 14.95). The voting behavior did not correlate significantly with reactive aggression scores, $r(188) = -.018$, $p = .803$, which means that reactive aggression was not related to voting behavior . See table 1 for the correlations.

Feedback Related Negativity Analysis

Table 2 shows means and standard deviations for FRN amplitudes at different electrode positions and for different categories. First, a repeated-measures GLM analysis was performed to assess the relationship between feedback and the FRN amplitude (dependent variable) where electrodes (Fz, FCz, Cz) and feedback (positive feedback, negative feedback) are the within-subject factors. A main effect was found for electrodes, $F(2, 374) = 203.11$, $p < .001$, $\eta_p^2 = .521$, but not for feedback $F(1, 187)$

= .739, p = .391, η_p^2 = .004. The pairwise comparisons show that Fz (M = 2.39, SD = 3.53) is significantly smaller than Cz (M = 5.20, SD = 4.02), $p < .001$, moreover Fz is significantly smaller than FCz (M = 3.97, SD = 3.97), $p < .001$. In addition, Cz is significantly larger than FCz, $p < .001$. However, the FRN amplitude was not different in the negative feedback condition compared to the positive feedback condition. See figure 3 for a visual representation of the pairwise comparisons. Finally, no significant two-way interaction was found, $F(2, 374) = 1.25$, p = .286, η_p^2 = .007.

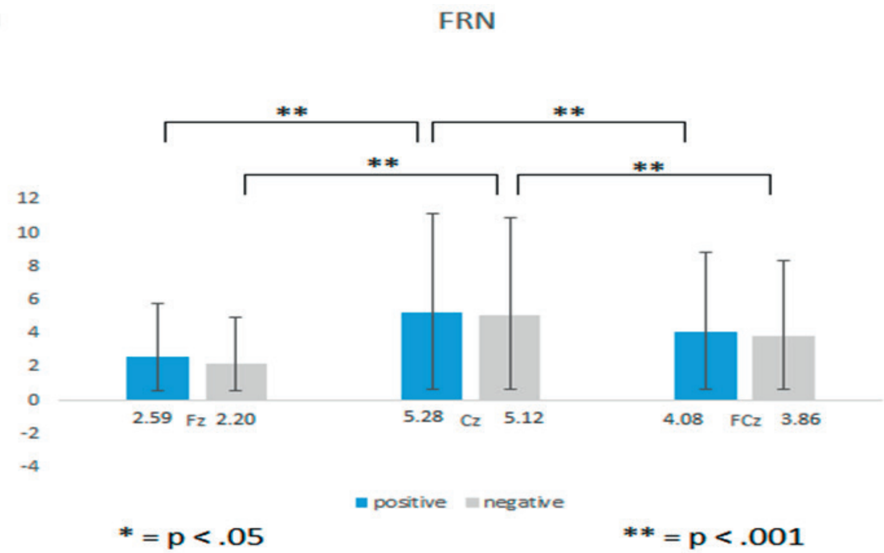


Figure 3. *The mean FRN amplitude measured at Fz, Cz, and FCz.*

Table *2* –

Means, standard deviations, and 95% CI of FRN amplitude at different electrodes and feedback types (Fz positive, Fz negative, Cz positive, Cz negative, FCz positive, and, FCz negative)

Electrodes	Mean	Std. Deviation
Fz positive	2.59	4.15
	[1.99, 3.19]	
Fz negative	2.20	4.09
	[1.61, 2.79]	
Cz positive	5.28	4.43
	[4.64, 5.91]	
Cz negative	5.12	4.64
	[4.46, 5.79]	
FCz positive	4.30	4.72
	[3.46, 4.71]	
FCz negative	3.77	4.42
	[3.21, 4.50]	

Reactive Aggression (RA) Analysis

In order to examine the effect of reactive aggression on FRN amplitude, a repeated measures GLM was carried out with feedback type and electrodeposition as within-subject factors, and self-report RA measure as the covariate. No main effect was found for reactive aggression (RA) as a covariate, $F(1, 186) = 3.38$, $p = .067$, $\eta_p^2 = .018$. In addition, there was no significant two-way interaction between electrodes and RA, $F(2, 372) = 2.04$, $p = .131$, $\eta_p^2 = .011$, nor for feedback and RA, $F(1, 186) = 1.58$, $p = .210$, $\eta_p^2 = .008$. Finally, there was no significant three-way interaction, $F(2, 372) = 1.33$, $p = .265$, $\eta_p^2 = .007$.

Trait Anger (TA) Analysis

In order to examine the effect of trait anger on FRN amplitude, a repeated measure GLM was carried out with feedback type and electrodeposition as within-subject factors, and TA as the covariate. No main effect was found for trait anger (TA) as a covariate, $F(1, 186) = 1.60$, $p = .207$, $\eta_p^2 = .009$. Just like the previous covariate analyses, there was no significant two-way interactions between electrodes and TA, $F(2, 372) = 1.23$, $p = .293$, $\eta_p^2 = .007$, nor between feedback and TA, $F(1, 186) = 1.68$, $p = .196$, $\eta_p^2 = .009$. As to the three-way interaction, no significance was found, $F(2, 372) = .984$, $p = .375$, $\eta_p^2 = .005$.

Trait Aggression (AVL) Analysis

In order to examine the effect of AVL on FRN amplitude, a repeated measure GLM was carried out with feedback type and electrodeposition as within-subject factors and AVL as the covariate. No main effect was found for AVL as a covariate, $F(1, 186) = .332$, $p = .565$, $\eta_p^2 = .002$. No significant two-way interactions were found between electrode and AVL, $F(2, 372) = 1.68$, $p = .186$, $\eta_p^2 = .009$, nor between feedback and AVL, $F(1, 186) = .004$, $p = .950$, $\eta_p^2 = .000$. Finally, there was no significant three-way interaction, $F(2, 372) = .090$, $p = .914$, $\eta_p^2 = .000$.

Table 1-

Correlations between the FRN amplitudes at different electrodes and participants' negative voting behavior

	Fz positive	Fz negative	FCz positive	FCz negative	Cz positive	Cz negative
Reactive Aggression	.181	-.116	.157	.077	.134	.017
Negative voting behavior	-.009	.082	-.051	-.127	-.034	.064

Note: * = $p < .05$; ** = $p < .01$

Discussion

The aim of the current study was to investigate whether the FRN amplitude in a social feedback task (Island Getaway Task) was related to self-report measures of reactive aggression, and related concepts such as trait anger. In contrast to our expectations, results showed that there is no difference between positive and negative feedback. Moreover, there were no associations between FRN and other self-report anger and aggression measures. Finally, participants' voting behavior was not associated with FRN amplitudes or the self-report measures of aggression and anger.

Our finding that FRN amplitudes were not different for negative as compared to positive feedback is consistent with previous studies that did not report an FRN differentiation effect in healthy controls who are similar to the current study sample characteristics (Bolling et al., 2011; Dekkers et al., 2015; Leitner, Hehman, Jones, & Forbes, 2014). For instance, a study by Leitner et al. (2014) who used similar social feedback to the current study paradigm in which undergraduate psychology students were led to believe that other individuals were viewing their photo and deciding whether to accept or reject them. Results for this study could not find any feedback effect positive (acceptance) or negative (rejection) on the FRN. Still, the absence of the feedback effect in the current study seems inconsistent with Kujawa et al. (2014), who designed the IGT paradigm, and found that FRN amplitudes are larger for negative (rejection) compared to positive (acceptance) feedback.

Although this may be true, it is important to realize that there is a different study, that also employed the same IGT paradigm and found a contradictory result in which FRN amplitudes were larger for positive (acceptance) feedback compared to negative (rejection) feedback (Cao et al., 2015). One important difference that has to be mentioned that Cao et al. (2015) study differed in crucial aspects in comparison to the original IGT paradigm. That is, their version of the IGT paradigm did not present the participants nor other co-players with additional personal information or questions that were gradually exchanged during the experiment and participants were competing against only a single pseudo-participant. This apparently shows the importance of ERP's paradigm designing procedure when studying social feedback in the context of aggression. On the other hand, the current study followed the original IGT paradigm designing procedure by Kujawa et al. (2014), but differed from Kujawa et al. 2014 study in terms of the sample selection and feedback stimuli presentation. While the current study only examined healthy adult students, who are between (18-29 years old; 100% male *N*= 188), Kujawa et al. 2014 only examined children and adolescents who are between (10-15 years old; 57.9 % male *N*= 19). However, the current study used symbolic feedback in which a coloured rectangle was shown to indicate acceptance/rejection feedback and to avoid physical differences between the feedback signals. Moreover, the colour that signalled negative or positive feedback was counterbalanced across participants. Kujawa et al. (2014), on the other hand, used a more complex feedback signal that is, a green checkmark to indicate acceptance and a red X to indicate rejection. In addition to that, recent studies that employed the IGT task in early adolescence and emerging adulthood samples used complex feedback in which a green thumbs-up appears to indicate acceptance feedback, and red thumbs down appears to indicate rejection feedback (Ethridge et al., 2017; Kujawa, Kessel, Carroll, Arfer, & Klein, 2017; Pegg et al., 2019). All these studies consistently showed a more positive ERP to positive feedback in the FRN/RewP time window. Combining our results, in which no difference can be found with symbolic counterbalanced feedback, with these results it could be argued that the different physical appearance strongly affects the amplitude of the FRN or RewP.

Further evidence for different ERPs following physically different social and non-social feedback were found in a recent study by (Pfabigan, Gittenberger, & Lamm, 2019), in which either social complex, social non-complex, non-social complex, or non-social non- complex were used to provide performance feedback. Results of this study showed that FRN amplitudes were sensitive to all factors, with a larger FRN after social compared to non-social stimuli and larger FRN amplitudes after complex positive compared to non-complex positive stimuli. Therefore, it should be concluded that results with the IGT strongly depend on the physical appearance of the feedback stimulus and that the use of neutral symbolic feedback seems to lead to suboptimal results.

Our findings that there is no association between FRN and other self-report anger and aggression measures, can be explained by the fact that most of the previous studies who found the effect of the FRN on aggression used participants who scored higher on aggression measures as compared to the current study or used a clinical sample (Krämer, Büttner, Roth, & Münte, 2008; Krämer, Munte, Richter, & Kopyciok, 2009; Wiswede et al., 2011). Most importantly, all these studies provoked their samples while testing them. (Bertsch, Bohnke, Kruk, & Naumann, 2009; Krämer, Jansma, Tempelmann, & Münte, 2007; Lotze, Veit, Anders, & Birbaumer, 2007). Therefore, it could be argued that if the current study would have used a clinical population or a population with higher aggression scores, and would have used a provocation task inside the lab, the chance of finding the effect of the FRN on aggression might have been higher. However, it is important to realize that the current study presented results should be treated with caution. Specifically, that the IGT paradigm might not be capable to measure FRN association with aggression among adult healthy male populations.

Finally, the current study could not find an association between aggression and self-report measures of aggression and anger on the one hand, and the likeliness to reject co-players on the (IGT) on the other hand, possibly, the low reactive aggression scores among our participants can possibly explain this finding. In fact, the average score for participants within the current sample in reactive aggression is 8.07 (*SD* = 3.25), which is comparable to the standard group of non-offender's adult study by (Cima et al., 2013). Possibly, the relation between voting behavior and aggression might only become visible only when participants cross a certain threshold, that is, only participants who score high on aggression might show changes in feedback processing that are detectable in self-report measures of aggression and anger. Future research including more extreme groups or clinical populations has to show whether this is the case.

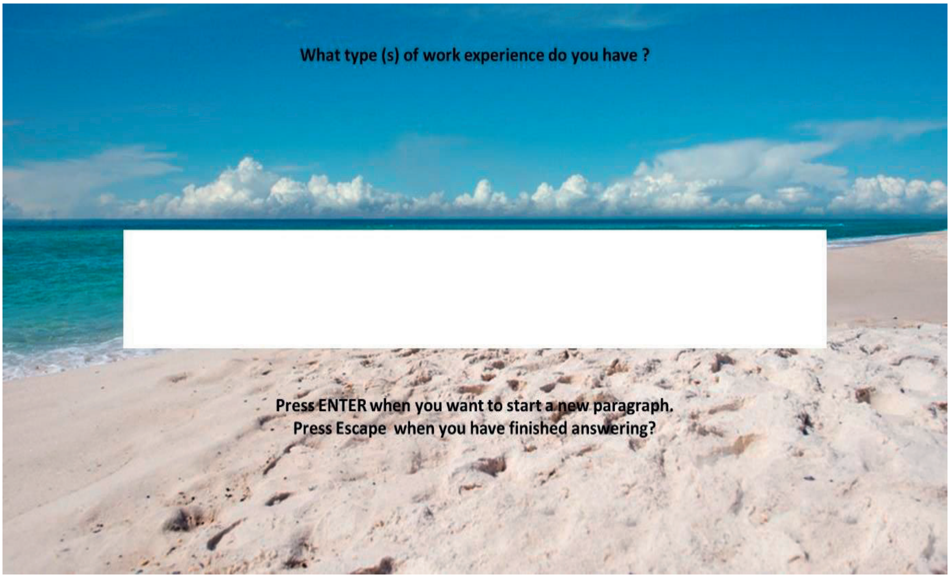
Limitations and implications for future research

The current study has a number of strengths and limitations. One obvious strength is the use of a relatively large sample size. However, there are also limitations. One important limitation is, that the current only used a non-clinical sample to measure aggression, which can explain the low scores on the self-report measures. The use of clinical samples could result in a better distribution of the aggression measures. In addition, the use of self-report measures could have led to socially desirable answers, also resulting in low aggression scores.

Conclusion

To conclude, our results showed that there is no difference between positive and negative feedback. Moreover, there is no association between FRN and other self-report anger and aggression measures. Finally, participants' voting behavior was not associated with FRN amplitudes or the self-report measures of aggression and anger. Finally, our results suggest that the brain responses in the IGT used as a tool to test behavioral and brain responses to social feedback in an adult healthy male population, depend on the physical appearance of the feedback stimuli and at the use of neutral symbolic feedback.

One of the co-players profiles



An example of personal questions regarding life experiences, interests, or beliefs that the participants need to answer while playing the game

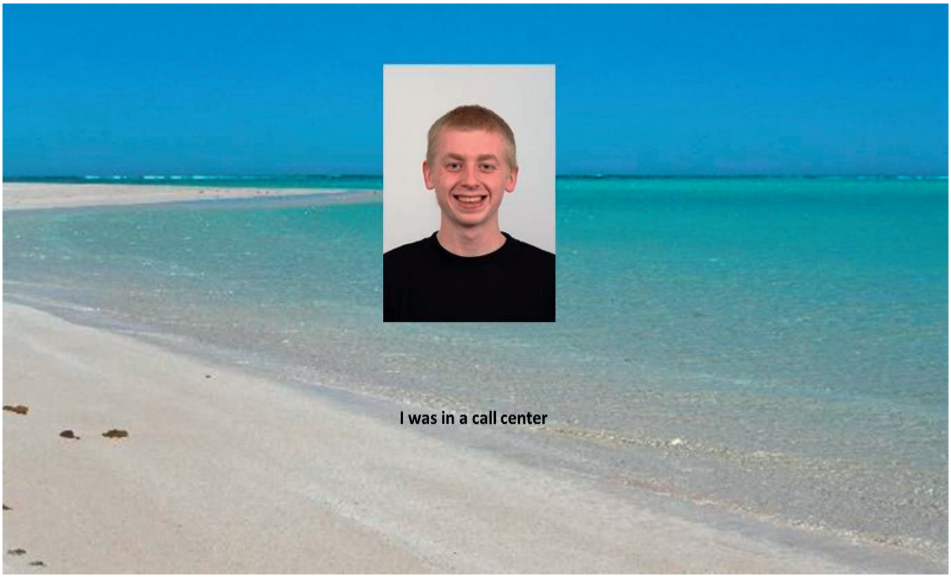


Figure 1. An example of an answer from one of the co-players in response to the personal questions

Figure 2 The IGT feedback stage

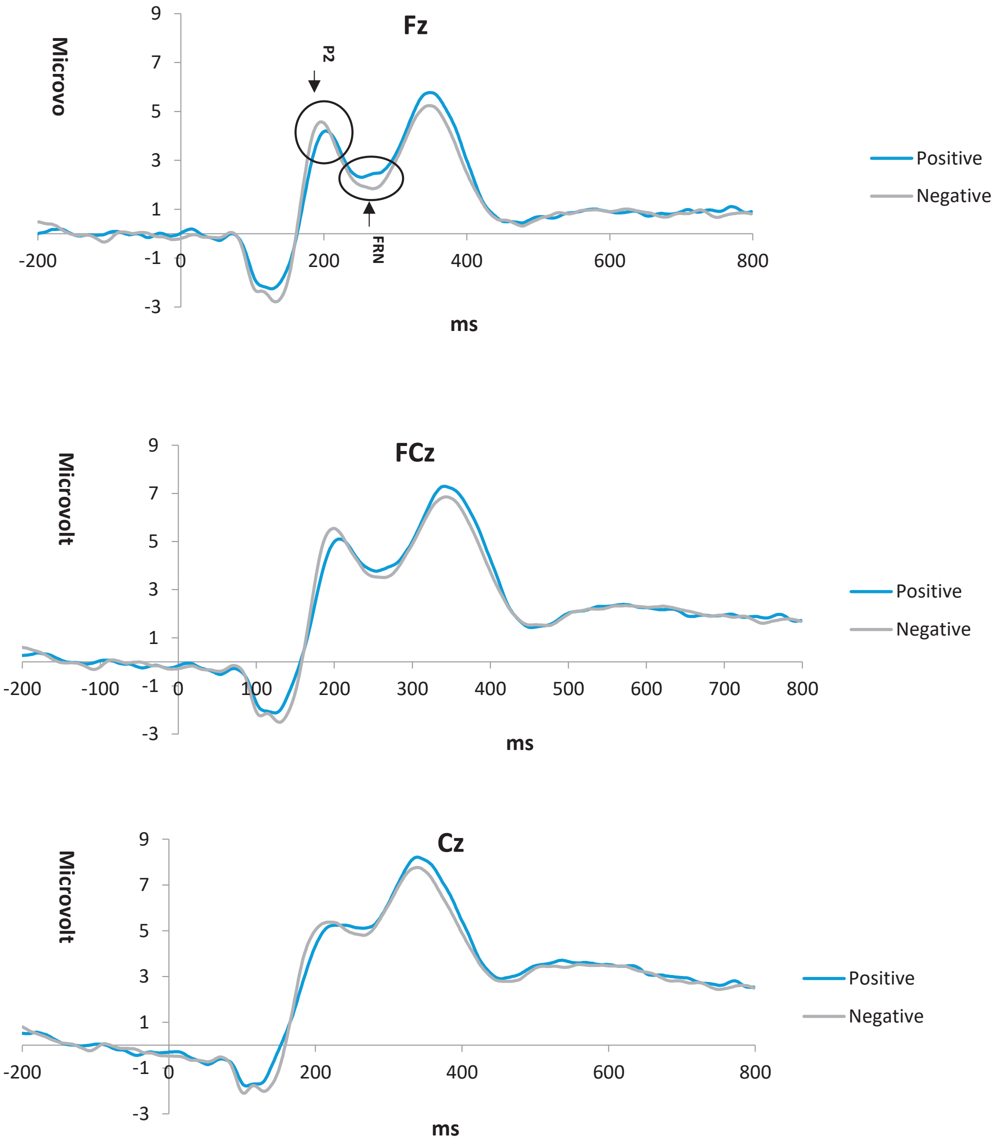


Figure 4. Grand-average ERPs using a (0.1–30 Hz bandpass filters settings) on electrode positions Fz, FCz, and Cz in response to negative and positive feedback

CHAPTER 3

Emotional modulation of the N400: Manipulating The Emotional Meaning of Homonyms

THIS CHAPTER HAS BEEN SUBMITTED FOR PUBLICATION AS:

HAKAMI, I., F.M. VAN DER VEEN., YU YUEN, CHAK., WIESER, M. J., VAN STRIEN, J. W., & FRANKEN, I.H.A. EMOTIONAL MODULATION OF THE N400: MANIPULATING THE EMOTIONAL MEANING OF HOMONYMS.

Abstract

There is evidence that the N400 ERP component elicited by semantic aspects of words is modulated by emotion. In addition, it is known that the context of words can affect the N400 amplitude of ambiguous words. Under stringent conditions (using homonyms) we investigated the relationship between semantic content (emotional vs non-emotional) and the N400. The meaning of the ambiguous word (neutral vs. emotional) was determined by the semantic context they appeared in. The participants viewed the Homonyms task ambiguous words in two different (neutral vs. negative) emotional contexts while their brain activity was measured with electroencephalogram (EEG). The main finding confirms the hypothesis that homonyms embedded in an emotionally negative context, which causes the homonym to gain an emotionally negative meaning itself, elicit a more negative N400 compared to homonyms presented in a neutral context. This suggests that the N400 is modulated by the emotional context.

Introduction

There is broad agreement that emotional stimuli elicit robust electrophysiological responses (Lang, Bradley, & Cuthbert, 1998; Schupp et al., 2007). It seems that non-verbal emotional stimuli (e.g., pictures) generally have stronger effects on the Event related potentials (ERPs) than words (Citron, 2012). However, language is a prominent and important way of communicating, giving rise to psycholinguistics studies addressing emotional processing. Electrophysiological studies addressing the processing of language have shown that words with emotional content elicit clearly defined brain responses in terms of both early and late ERPs (Citron, 2012; Kissler, Assadollahi, & Herbert, 2006).

One of the most important waves reflecting semantic processing of words is the N400. The N400 response emerges 300 ms after word onset with a peak near 400 ms, and ending around 600 ms (Kutas & Federmeier, 2011; Lau, Phillips, & Poeppel, 2008). As for its distribution, much remains unclear about the exact anatomical localization of the N400 effect. For example, the N400 response when elicited by single words has a more anterior distribution with a maximum over frontal or central sites (Bentin, McCarthy, &Wood, 1985; Bentin, 1987; McCarthy & Nobre, 1993). However, when elicited by semantic incongruities in sentences, the N400 is largest over the centro-parietal regions (Johnson & Hamm, 2000; Kutas & Hillyard, 1982; Kutas, Hillyard, & Gazzaniga, 1988). This is because the N400 does not only reflect activity in a single source but activity in a broad neural network including the left superior temporal gyrus, left temporal and prefrontal areas (Kutas & Federmeier, 2011). Therefore, the N400 has been frequently related to the semantic integration of a word within a larger context (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980, 1984). In fact, Kutas & Hillyard (1980) were among the first to find that an inappropriate word at the end of a sentence is able to modulate the N400. A larger N400 is observed if the last word is semantically inappropriate. Aside from the increased N400 activity in the situation of semantic incongruity, Kutas & Hillyard (1980) also found that grammatical errors did not provide the same strong effect on the N400 as inappropriate word endings, further showing that the N400 is particularly associated with the semantic content rather than with the grammar.

The enhanced N400, that has been associated with a semantic anomaly or violation of semantic expectations has been confirmed and further explored by later studies (e.g., Curran, Tucker, Kutas, & Posner, 1993; Moreno & Vazquez, 2011. Two studies (Berkum, Hagoort, & Brown, 1999; Holt, Lynn, & Kuperberg, 2009) presented participants with non-constraining, neutral contexts sentences (for example: Nancy's son ended up just like his father) followed by second sentences that have either a neutral, positive or negative critical word (for example: He was already a husband/millionaire/criminal by age 25). They found a small, localized, posterior N400 effect for emotional words (both negative and positive words). Delaney-Busch & Kuperberg (2013) further studied this matter, but instead of a neutral first sentence, they used first sentences that were pleasant, unpleasant or neutral. The second sentence contained a pleasant, unpleasant or neutral critical word, which was either congruently or incongruently matched with the first sentence. They found similar results as Holt et al. (2009), but their findings also suggest that an emotionally charged context (the first sentence is pleasant or unpleasant) causes small N400 ERPs, regardless of the content of the critical word. This finding points to the importance of the emotional context itself on the N400.

Several other recent studies have employed single words, rather than sentences as a context for ambiguous words (e.g., Atchley & Kwasny, 2003; Chwilla & Kolk, 2003; Klepousniotou, Pike, Steinhauer, & Gracco, 2012; Kotchoubey & El-Khoury, 2014; Lee & Federmeier, 2006; Meyer & Federmeier, 2007; Pylkkanen, Llinas, & Murphy, 2006; Titone & Salisbury, 2004). These studies highlighted the importance of the context in studying the electrophysiological (N400) processing of ambiguous words. For example, Titone & Salisbury (2004), investigated a semantic task that manipulated local and global contextual cues. In this study, word triplets were presented to participants, in which the second word was a homograph. While the first word created a natural, dominant meaning, or subordinate meaning biased global context. The third word created a dominant or subordinate biased local context that was either congruent or incongruent with the first word global context (for example, the relation between finances, banks, and money). Results from this study showed moderate modulation of N400 to ambiguity in the different hemispheres in which, the right hemisphere N400 was larger than the left when the context was appropriate for subordinate homograph associates, whereas the left hemisphere N400 was larger for dominant homograph associates in inappropriate subordinate context than the right. A larger activation of the dominant as compared to the subordinate word meaning was also seen in N400 data of Klepousniotou et al., (2012).

It has to be noted that not all of the previous studies that tested the importance of the emotional context, on the N400 points into the same direction. For instance, Herbert, Junghofer, and Kissler, (2008) found that the emotional content of words significantly reduced the N400 component, and this was more the case for pleasant than unpleasant adjectives. Nevertheless, it can be concluded that the N400 seems to be modulated by both semantic and emotional content.

Despite the evidence that context is relevant for the electrophysiological (N400) processing of ambiguous words, the role of an emotional vs non-emotional context has not been investigated yet. Moreover, ERP research that has been studying emotional processing, tended to use stimuli that always physically differ and therefore it cannot be completely ruled out that the modulation is affected by the physical appearance of a word than the emotional content itself. Therefore, the current study examined the N400 to test whether the negative emotional content of a word can elicit an N400 effect. This experiment will be done under stringent conditions, in which the negative emotional words we use are homonyms. Homonyms are identical to the neutral words in form and articulation, but only differ in semantic content, which we made dependent on the context in which the word is presented. Excluding other factors is an important advantage of using homonyms, since this means that any differences found can only come from the semantic content of the word. By putting the critical homonyms last in a series of three successive words, we are more likely to interpret the meaning of the word according to the first two words. In other words, the critical words are identical to each other, but the first two words create the context, setting the stage for the meaning of this critical word. For instance, the Dutch word equivalents of “leg” can be presented as the first word to create the context, followed by the word “head” and then the word “arm” can be presented as the third word in this series which is the neutral content condition. As for the emotional (negative) content condition, the Dutch word equivalents of “crazy” can be presented as the first word to create the context, followed by the word “dirty”. Finally, the word “poor” can be presented as the third word in this series. We hypothesize that a significant enhancement of the N400 effect will be found for the homonym within the negative emotional context in comparison with the homonym that is embedded in a neutral context. This N400 modulation will be elicited exclusively by the negative emotional content of the word which is determined by the emotional context. The homonyms used in the present study either have a neutral meaning or an emotional negative meaning. A secondary, more exploratory question was whether affect (positive or negative) and/or anxiety level (trait or state) of the participant had an effect on the magnitude of the N400 effect since Chwilla, Virgillito, & Vissers, (2011) found an influence of mood on the N400 effect. More specifically, we hypothesized that PANAS and anxiety self-report measures would affect the on-line semantic processing in language comprehension, as tapped by the N400.

Material and Methods

Participants

Twenty-two healthy undergraduate psychology students of the Erasmus University Rotterdam, the Netherlands participated in this study. They could sign up for the experiment on-line and would be contacted either by phone or e-mail. Participants received course credits for the completion of the experiment. Data of one participant was removed because of poor EEG signal quality (less than 50% of the segments were artifact-free). This resulted into a final group of twenty-one participants (eighteen female participants) and an average age of 19.81 (*SD* = 2.38). This study was conducted according to the local and ethical guidelines of the Institute of Psychology at Erasmus University Rotterdam.

Self-report measures

Positive Affect Negative Affect Schedule (The PANAS). The Positive Affect (PA) Negative Affect (NA) Schedule is a questionnaire that is designed to measure the affect. The PANAS consists of 20 five-point Likert –scale items. These items are rated on a five-point scale: 1= very few, 2 = a little) 3 = moderate, 4 = many, and 5= very many. The reliability of the PANAS appears to be high from previous studies (Ostir, Smith, Smith, & Ottenbacher, 2005). The PA scale has a reliability of .79 and the NA scale has a reliability of .85 (Engelen, Victoir, Van Diest, & Van den Bergh, 2006). For this study, the reliability was also determined and Cronbach's α .80 on PA and .91 on NA.

State-Trait Anxiety Inventory (STAI), The State-Trait Anxiety Inventory is a questionnaire that is designed to measure the degree of anxiety in adults. The STAI consists of 40 four-point scale items in which two aspects of anxiety are distinguished. The state anxiety measures the temporary / current anxiety level, and the trait anxiety measures long-term / general anxiety level. These items are rated on a four-point scale: 1= almost never, 2= sometimes, 3 = often, and 4 = almost always. The reliability of the

STAI appears to be high from previous studies (Defares, Ploeg, & Spielberger, 1980; Spielberger, Gorsuch, & Lushene, 1970). For the A-Trait scale Cronbach's alphas ranged from .84 to .95 and the test–retest reliability correlations varied from .75 (118 days interval) to .92 (1_ h interval) (Defares et al., 1980). For this study, the reliability was also determined and Cronbach's α .92.

Stimuli and experimental paradigm

The Homonyms task was presented by E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). Black-colored words in the upper case were shown on the computer screen with light grey background in clusters of three, with a total of fifty clusters (twenty-five for the negative emotional content condition and twenty-five for the neutral content condition). There were two conditions in the experiment, the neutral content condition (e.g., the Dutch equivalents of leg - head – *arm*) and the emotional (negative) content condition (e.g., the Dutch equivalents of needy - meagre - *poor*), which all participants went through. See Table 1 for examples of used word clusters and see Appendix 1 for the complete list of used clusters (in Dutch). Both two conditions of the Homonyms task were presented at the same block. However, participants were counterbalanced on the two conditions (emotional negative context vs neutral context first) to avoid order effects. The sequence and duration of one trial were as follows. The participants were asked to focus their attention on the fixation cross in the middle of the screen (500 ms). After this, a blank screen appeared (500 ms), followed by the first word of the cluster (500 ms). A second blank screen appeared (1000 ms), followed by the second word of the cluster (500 ms). A third blank screen appeared (1000 ms), followed by the third and critical word of the cluster (500 ms). Finally, the trial ended with a blank screen (3000 ms), and the second trial followed. See Figure 1 for visual representation for the task. No behavioral response of the participants was required. The whole experiment took approximately ten minutes.

Procedure

Upon arrival, the participants were asked to fill in their age, gender and the Dutch, short version questionnaires of the Positive Affect Negative Affect Scale (PANAS) to measure the general affect and the State-Trait Anxiety Inventory (STAI) to measure the general anxiety. After this, the participants were informed about the procedure and brought to the room specifically for the EEG recording. The participants were asked to set on a comfortable chair in front of the computer screen. Then, participants were asked to read the instructions on the screen, after which an example of the stimuli was shown. The experimenter then asked whether there were any questions and confirm the participants' understanding of the experiment. After this, the experimenter instructed the participants to pay attention to the words on the computer monitor and to try to refrain from moving too much or clench the teeth together tightly.

Electroencephalogram acquisition and analyses

The BioSemi ActiveTwo System was used to record the EEG signals and electrooculography (EOG) signals with a sampling rate of 500Hz and online filter-settings of 0-135Hz. Thirty-two active Ag/AgCl electrodes attached to the participant's cap was used to measure the activity on the scalp. Six additional reference electrodes were placed on both mastoids, on both eyes' outer canthi (for the horizontal electro-oculogram) and on the infraorbital and supraorbital regions of the left eye (for the vertical electro-oculogram). Data were off-line referenced to the averaged mastoids. After this data were filtered using a low cut-off of 0.05 Hz and a high cut-off of 35 Hz and Butterworth Zero Phase. VEOG and HEOG ocular activity was corrected using the Gratton & Coles (1983) methodology. All activities exceeding +/-100 µV were removed. Before calculating the averages, a 100 ms pre-stimulus baseline correction was applied. The ERPs were measured on the homonyms (i.e., last word in a series of three words, see Figure 1). The N400 was defined as the mean activity in the 300-500 ms range on the electrodes Fz, CZ, and Pz. The current study selection of the three electrodes (Fz, Cz, and Pz) is in line with previous studies that found the N400 to be maximal over midline electrodes, specifically the centro-parietal sites (Kutas & Hillyard, 1980; Kutas & Hillyard, 1982; Kutas et al., 1988).

Statistical analyses

Statistical analyses were performed using SPSS 25. ANOVA repeated measures was performed on the data. As discussed before in the introduction, much remains unclear about the exact anatomical localization of the N400 effect. With no conclusive knowledge about any specific brain areas that are particularly active when a N400 modulation occurs, the focus in this study was on the midline electrodes Fz, Cz and Pz. This resulted into a 3 (electrodes) x 2 (word content) ANOVA. Finally, we used Pearson correlation to test whether there is a significant correlation between the different amplitudes of the electrodes (negative emotional words minus neutral words) and our participants' scores on the self-report questionnaires. When the assumption of sphericity was violated, Greenhouse-Geisser corrected degrees of freedom were used to compute the corrected *p*-value but uncorrected degrees of freedom are reported. Effects size is reported as partial eta squared (η_p^2) and effects were considered significant at $p < .05$.

Results

N400-effect

The mean N400 amplitude for the negative emotional word content was consistently more negative than the mean for the neutral word content for all three electrodes (see Table 2), consistent with the global distribution of the N400. See Figure 2 for the grand average of the ERPs in the two conditions. Most importantly, the main effect of the word content was significant $F(1,20) = 5.80$, $p = .026$. In other words, the critical words in the negative emotional content condition differed significantly from the critical words in the neutral content condition. The homonyms in the negative emotional context had significantly more negative ERPs than the same words in the neutral context. The interaction effect between electrodes and word content was not significant $F(2,40) = 1.28$, $p = .28$.

Correlations between general affect/anxiety and event-related potential measures

To see whether any significant correlations would be found, we focused on the correlations between the difference scores of the N400 (negative emotional words minus neutral words) and the scores on the self-report questionnaires. No significant correlations were found between the ERP measures of the three electrodes with positive affect, negative affect, anxiety state and anxiety trait.

Discussion

The current study investigated the effect of semantic (negative emotional vs non-emotional) on the N400. By using homonyms, the current study was able to eliminate some major confounding factors such as the physical appearance of a word to investigate emotional modulation of the N400. Since the critical words are identical to each other, they only differed in their semantic (neutral vs. emotional) content. We hypothesized that a significant enhancement of the N400 effect would be found for the homonym in the negative emotional content condition, in which the critical word is embedded in the negative emotional context. Indeed, we found a significant main effect of emotion, suggesting that the negative emotional content enhances the N400 effect even under very stringent conditions. Furthermore, the question whether affect and/or anxiety influenced the emotional processing was explored. No significant correlations between the PANAS and anxiety on the one hand and the N400 on the other hand were found.

The finding that the N400 response to homonyms in the two conditions differed significantly points to the influence of the semantic content of the critical word on the N400. In the present experiment, this semantic content depended on the context in which the critical word was embedded, creating either a negative or neutral meaning of the homonym. In contrast to previous studies that used sentences, conditions were created by noncritical words that had a negative meaning. We believe this study provides more insight into the N400 effect, showing the importance of the context on the emotional content of a word. Previous studies of N400 contextual effects, found that the N400 amplitude was also increased for visual objects that were incongruent with the semantic expectancies provided by the context scene (Ganis & Kutas, 2003). For instance, Hannula, Federmeier, & Cohen, (2006), also showed in their study that tested faces in the same context have a smaller N400 negativity compared to the context in the study-phase. Further research can build on the present findings, which pointed to the importance of context on the N400 effect. Unlike the earlier ERPs, late ERP modulations seem to be more influenced by the context in which the word is embedded.

There are also some alternative explanations. It could be the case that the emotional meaning of the word is less or more frequently used. In that case, the enhanced N400 represents a frequency effect. Alternatively, it could be that the meaning in the emotional or neutral case is sometimes somewhat harder to grasp. It is difficult to rule this out as there is no research on this issue. If this the case, the enhanced N400 could also indicate that it is harder to access the meaning of the word. Further, we don't know whether the effect is specific for semantic processing. It could be that any kind of stimulus after the emotional words lead to different processing as compared to the neutral condition (i.e., emotional priming; Gendron, Lindquist, Barsalou, & Barrett, 2012).

Not addressed in the present study is the modulation of the N400 by positive/pleasant contexts on the meaning of the word. The expectation would be to find similar results as those for negative contexts, which would be in line with the current body of research on the N400 in which no clear differences have been found between pleasant and unpleasant word meanings. Nevertheless, as the matter remains inconclusive, positive (pleasant) and negative (unpleasant) contexts having an effect on the N400 remains a tentative conclusion that needs to be further analyzed. We would like to suggest to also use homonyms when studying this, but also realize that using homonyms might not be possible in all languages.

Contrary to our expectation, the current study could not find an association between the PANAS on the one hand and the N400 on the other hand. This result is not in line with Chwilla et al. (2011), who used Mood Induction Procedure (MIP) to induce different emotional states (i.e., happy vs. sad) in participants by presenting them with fragments from movie clips in the context of a happy movie or a sad movie. Participants in Chwilla et al.'s study watched these clips before reading sentences while their ERPs were recorded. The results from this study showed reduced N400 amplitudes for the sad mood condition compared to the happy mood. Moreover, Chwilla et al.'s study supported the idea that emotions may impact processing styles. However, it must be noted that the current study differs from Chwilla et al.'s in terms of the use of MIP. While Chwilla et al. used movie clips to manipulate and induce mood to prolong the intended mood during the whole experiment, the current study only measured the participants' positive/negative affect without manipulating or inducing the participants' mood. It is possible that the association between mood and N400s only becomes visible when a participant's mood is manipulated before an experiment.

The finding that no correlations were found between anxiety and N400 amplitude is not in line with Yu et al. (2018), who tested the emotional memory bias in both high and low trait and high and low state anxiety individuals in a classic recognition paradigm using event-related potentials (ERPs). Their results showed that higher N400 amplitudes elicited by negative words in a high level of state anxiety as compared to low state levels. However, it has to be mentioned that the study of Yu et al. (2018) differed from the current study in terms of its sample characteristics. For example, the average score of participants who were high in trait anxiety in the Yu et al. study was about 50 while, the average score of the low scoring group was about 34.

In addition to that, Yu et al. (2018) employed an experimental operation between the learning and testing phases in order to induce/calm state anxiety levels. Following that, they separately calculated participants' scores pre-test and post-test. It is important to realize that studies have shown that this experimental operation works effectively in inducing state anxiety (negative mood states; Quigley et al., 2012; Waechter & Stolz, 2015). With this in mind, the induced pre-test mean scores for Yu et al. study were about 39 for trait anxiety while the induced trait anxiety post-test mean scores were about 41. Also, the calm state anxiety group mean was about 36 pre-and post-test. The current study, however, only had one group of participants and did not induce nor measured anxiety post the experiment. In fact, the average trait anxiety scores for participants in the current sample were about 38, and for the state anxiety the scores were about 31 which both were much lower than Yu et al. s' anxiety scores. It is possible that the absence of high anxiety scores along with not inducing anxiety can be the reason why the current study could not find the association between N400 and the subjective measures of anxiety.

Limitations and implications for future research

A limitation of this study is the relatively small sample size (21 healthy undergraduates). This number has potential implications for generalizability. It should be noted however, that, our sample size was sufficient to show the effect of emotion, suggesting that the emotional content of a word elicits an N400 effect even under very stringent conditions. Since it has been shown that the N400 can be used as a useful measure of emotions, future studies could benefit from the use of the Homonyms task in emotional processing research in general. Finally, it is possible that including participants scoring high on anxiety measures might reveal the association between affect and anxiety on the one hand and the N400 on the other hand. Therefore, future studies should include participants with more extreme anxiety scores in order to study the influence of mood on the N400.

Conclusion

Despite the limitations, the current study was able to find a significant main effect of emotion, suggesting that the emotional content of a word elicits an effect on N400 amplitude, even under very stringent conditions. One the other hand, no significant correlations between PANAS, anxiety and N400 were found. This study result supports the hypothesis that homonyms embedded in an emotionally negative context causes the homonym to gain an emotionally negative meaning itself, and elicit a more negative N400 compared to homonyms presented in a neutral context. This result suggests that the N400 is modulated by the emotional context.

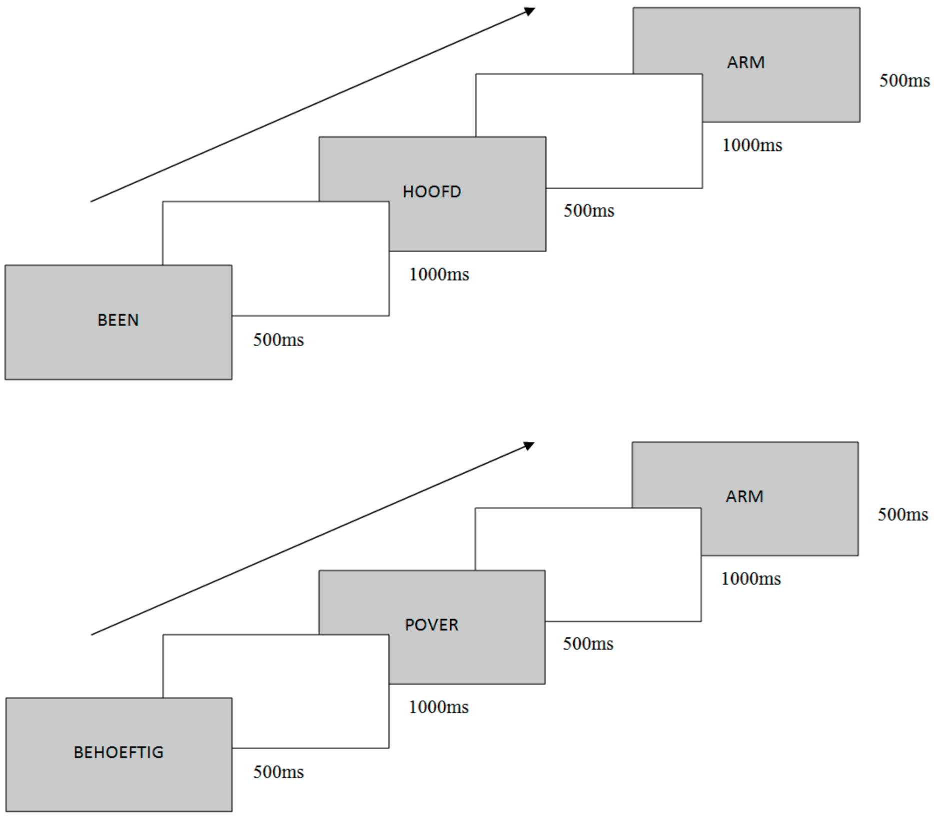


Figure 1. Example of a neutral (upper figure) and emotional (lower figure) word cluster. After the screen with the fixation cross and the blank screen, the noncritical neutral words would be shown with blank screens before the target stimulus (the homonym) would be shown

3

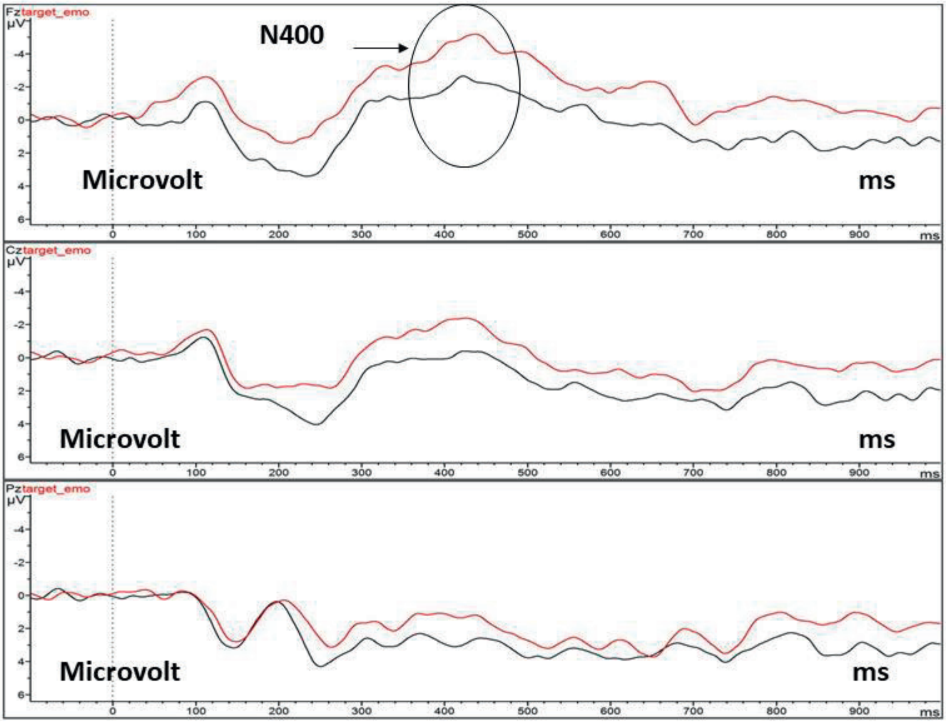


Figure 2. Grand-average ERPs on different electrodes. The Red color is the emotional content condition and the black color is the neutral content condition

Table 1

Examples of the used word clusters

Content	Word 1	Word 2	Critical word (homonym)
Emotional	Behoeftig (*Needy*)	Pover (*Meagre*)	Arm (*Poor*)
Neutral	Been (*Leg*)	Hoofd (*Head*)	Arm (*Arm*)
Emotional	Onaangenaam (*Unpleasant*)	Pijn (*Pain*)	Zeer (*Hurt*)
Neutral	Erg (*Very*)	Heel (*Highly*)	Zeer (*Extremely*)
Emotional	Doden (*Dead people*)	Lichamen (*Bodies*)	Lijken (*Corpses*)
Neutral	Imiteren (*Imitate*)	Spiegelen (*Mirror*)	Lijken (*Resemble*)
Emotional	Mep (*Smack*)	Trap (*Kick*)	Sla (*Hit*)
Neutral	Broccoli (*Broccoli*)	Prei (*Leek*)	Sla (*Lettuce*)
Emotional	Meningsverschil (*Discord*)	Ruzie (*Argument*)	Mot (*Row*)
Neutral	Vlieg (*Fly*)	Mug (*Mosquito*)	Mot (*Moth*)

Table 2

Means and standard deviations electrodes for the N400

Electrode		Mean	Standard Deviation
Fz	Neutral	-1.71	3.34
	Emotional	-3.94	3.81
Cz	Neutral	.17	3.39
	Emotional	-1.52	3.69
Pz	Neutral	2.85	3.64
	Emotional	1.62	3.52

Appendix 1

Used word clusters of both conditions (emotional and neutral word content)

Content	Word 1	Word 2	Critical word
Emotional	Klap	Stoot	*Beuk*
Neutral	Eik	Es	*Beuk*
Emotional	Duw	Trap	*Stomp*
Neutral	Rond	Bot	*Stomp*
Emotional	Vervelend	Eng	*Naar*
Neutral	Toe	Gaan	*Naar*
Emotional	Aangeschoten	Dronken	*Lam*
Neutral	Geit	Schaap	*Lam*
Emotional	Doden	Lichamen	*Lijken*
Neutral	Imiteren	Spiegelen	*Lijken*
Emotional	Vernietigen	Doden	*Afmaken*
Neutral	Klaar	Eindigen	*Afmaken*
Emotional	Schop	Stoot	*Trap*
Neutral	Trede	Ladder	*Trap*
Emotional	Knal	Pistool	*Schot*
Neutral	Hek	Raster	*Schot*
Emotional	Naar	Bang	*Eng*
Neutral	Smal	Nauw	*Eng*
Emotional	Behoeftig	Pover	*Arm*
Neutral	Been	Hoofd	*Arm*
Emotional	Onaangenaam	Pijn	*Zeer*
Neutral	Erg	Heel	*Zeer*
Emotional	Verlinken	Verraden	*Klikken*
Neutral	Tikken	Klakken	*Klikken*
Emotional	Wijf	Trut	*Muts*
Neutral	Hoed	Pet	*Muts*
Emotional	Ontneem	Jat	*Steel*
Neutral	Halm	Stengel	*Steel*

Emotional	Lomp	Gemeen	*Bot*
Neutral	Zenuw	Weefsel	*Bot*
Emotional	Mep	Trap	*Sla*
Neutral	Broccoli	Prei	*Sla*
Emotional	Meningsversch il	Ruzie	*Mot*
Neutral	Vlieg	Mug	*Mot*
Emotional	Kapot	Scherf	*Stuk*
Neutral	Deel	Part	*Stuk*
Emotional	Gemeen	Stiekem	*Vals*
Neutral	Onzuiver	Onecht	*Vals*
Emotional	Etteren	Pussen	*Zweren*
Neutral	Verklaren	Beloven	*Zweren*
Emotional	Billen	Kont	*Achterste*
Neutral	Voorste	Middelste	*Achterste*
Emotional	Verraden	Vernederd	*Genaaid*
Neutral	Geborduurd	Gehaakt	*Genaaid*
Emotional	Idioot	Hufter	*Eikel*
Neutral	Kastanje	Noot	*Eikel*
Emotional	Domoor	Stommelin g	*Rund*
Neutral	Kalf	Koe	*Rund*
Emotional	Domper	Vernederin g	*Afgang*
Neutral	Vertrek	Weggaan	*Afgang*

CHAPTER 4

No Effect of Self-Report Aggression Measures on The N400 in A Lexical Decision Task with Associated Words in a Violent Context

THIS CHAPTER HAS BEEN SUBMITTED FOR PUBLICATION AS:

HAKAMI, I., F.M. VAN DER VEEN, FRANKEN., I.H.A. & HUIJDING, J. NO EFFECT OF SELF-REPORT AGGRESSION MEASURES ON THE N400 IN A LEXICAL DECISION TASK WITH ASSOCIATED WORDS IN A VIOLENT CONTEXT.

Abstract

Recent studies have shown that reactive aggressive behavior in children is related to a biased interpretation (i.e., in a hostile or aggressive way) of ambiguous stimuli. However, most of the previous studies only focused on the behavioral measures in children and on aggression in general. Therefore, the current study focused on the relationship between aggression related interpretational bias (ARIB) and aggression in an adult population using both behavioral and electro-cortical measures. In this study, we examined (N=93, M=21.86 years) whether higher scores on self-report aggression measures were associated with greater ARIB in a sample of healthy male undergraduate students. Participants performed a lexical decision task in which they had to decide whether the second word (target) in a word pair was a word or a non-word. Associated, un-associated, in ambiguous (related in violent or neutral contexts)-or unambiguous contexts, were presented. Associated words led to faster reaction time (RT) and non-words were associated with both slower RTs and a larger N400 amplitude. However, Target type words associated in a violent context did not significantly influence RT and N400 measures. Moreover, there was no significant association between RT, N400, and self-report aggression measures. Finally, our results showed that the aggressive interpretation task is only able to find the associations between ARIB in terms of RT and N400 and self-report aggression measures in a forensic population but might not be sensitive enough to measure ARIB in healthy populations.

Introduction

Aggressive behavior is the first leading cause of mortality in the United States, resulting in approximately 50,000 deaths and 2.2 million injuries per year (Corso, Mercy, Simon, Finkelstein, & Miller, 2007). Aggressive behavior also affects communities and societies, leading to losses in business sectors, and to personal health costs. Aggression can also increase the burden on the healthcare and justice systems (Corso et al., 2007). One important distinction that is often made in the literature is between proactive and reactive aggression (Roth & Strüber, 2009). While proactive aggression can be characterized as planned and goal-directed, reactive aggression can be characterized as aggression that is committed in anger as retaliation for a perceived negative experience (Murray-Close & Ostrov, 2009). According to the social information-processing model (SIP), aggressive behavior can emerge as a result of people's attention to hostile cues in ambiguous situations, which might lead to interpreting other people's intentions in an aggressive manner (Crick & Dodge, 1994). These hostile attentional and interpretation biases might lead to aggressive behavior in social situations. As reported by the SIP model, ARIB is specifically associated with aggression. For instance, a recent meta-analytical review for children and adolescents found that aggression, in general, is positively correlated with the ARIB, however, the relation between ARIB and aggression depends on the level of emotional engagement. That is, the relationship is stronger for more reliable ARIB measures, but is not stronger for reactive aggression as compared to aggression in general (Verhoef, Alsem, Verhulp, & De Castro, 2019). A number of studies have also shown that specifically reactive aggression in children is related to ARIB (step 2 in the SIP model; Arsenio, Adams, & Gold, 2009; Crick & Dodge, 1996; De Castro, Merk, Koops, Veerman, & Bosch, 2005; Dodge & Coie, 1987; Dodge et al., 2015). On the other hand, there are much less studies that focused on studying adults ARIB and aggression subtypes (reactive and proactive; Bailey & Ostrov, 2008; Lobbestael, Cima, & Arntz, 2013; Miller & Lynam, 2006; Murray-Close, Ostrov, Nelson, Crick, & Coccaro, 2010), wherein, other adult studies only focused on examining ARIB and aggression in general (Coccaro, Noblett, & McCloskey, 2009; Gagnon, McDuff, Daelman, & Fournier, 2015; Helfritz-Sinville & Stanford, 2014; Matthews & Norris, 2002).

One common way to study ARIB and aggression is the use of vignettes as stimuli. For instance, a recent meta-analytical review of over twenty-five studies found that nineteen studies used some form of hypothetical-situation vignettes (Tuente, Bogaerts, & Veling, 2019). Although, using vignettes as stimuli to examine ARIB within the children population has been an effective way in accounting for aggression (e.g., see Arsenio, Adams, & Gold, 2009; Crozier et al., 2008; Lösel, Bliesener,& Bender, 2007; Quiggle, Garber, Panak, & Dodge, 1992; Shahinfar, Kupersmidt, & Matza, 2001; Zelli, Dodge, Lochman, & Laird, 1999), adult studies on the other hand only reported small to medium associations between ARIB and aggression (Tuente et al., 2019). In addition, not all studies that used vignettes as stimuli (Helfritz-Sinville & Stanford, 2014; Miller & Lynam, 2006) were successful in finding the associations between ARIB and aggression. However, we found one new interesting study that used Event-Related Potential (ERP) in order to examine whether aggressive and nonaggressive participants (between 18 and 55 years) differed in their brain responses (Gagnon et al., 2016). Participants in this study were presented with social scenarios depicting a hostile versus non-hostile context (first context sentences) followed by an ambiguous behavior (second context sentence). In the second context sentence, an individual carried out a social provocation act that is directed at the participant whose intention is ambiguous. Following that, a third sentence was presented (target sentence) ending in a target word that let the participant know whether the intention behind the behavior was hostile or non-hostile. Each scenario of this study contained a hostile or non-hostile context that can match or mismatch with the target sentence. All the match and mismatch scenarios had the same ambiguous behavior (second context sentence), and the same (target sentence), but they differed with respect to the hostility of the context (context sentence). Gagnon et al. 2016 hypothesized that aggressive individuals would be more likely to interpret others' ambiguous behavior as hostile, and they would develop stronger expectations about hostile intent from others in a hostile context than nonaggressive individuals. Along with that, the N400 signal following a hostile expectancy violation would be stronger in aggressive than non-aggressive individuals. However, in scenarios where hostile intention follows an ambiguous provocative behavior in a non-hostile context, nonaggressive individuals would show a stronger N400 as compared to aggressive individuals, because of their tendency to infer a non-hostile intent from others. Results of this study partly confirmed these hypotheses by demonstrating that in a hostile social context, the presentation of a critical word that violated hostile expectations regarding the intent behind the ambiguous behavior elicited significantly larger N400 responses in aggressive as compared to nonaggressive individuals.

Although Gagnon et al. 2016 used ERP's, specifically the N400 component to study ARIB association with aggression, the use of vignettes as stimuli might only reveal participants' explicit biases (Dovidio, Kawakami, & Gaertner, 2002), specifically within children's studies (Lansu, Cillessen, & Bukowski, 2013; Lansu, Cillessen, & Karremans, 2012). Due to this fact, researchers are still questioning whether the use of vignettes as stimuli is ideal to measures children's cognitive biases (see, e.g., Cillessen & Bellmore, 2010). It is possible that, the association between ARIB and aggression should be done with an implicit measure that can manifest ARIB spontaneously without the participant controlling such biases (Greenwald & Banaji, 1995). We found one former study that used such an implicit measure of ARIB within aggression research (Cima, Vancleef, Lobbestael, Meesters, & Korebrits, 2014). Cima et al. 2014 used a lexical decision paradigm in order to examine ARIB among 88 male aggressive adults to see whether there is a relationship between ARIB and different types of aggression. ARIB was examined using the Aggressive Interpretation Task (AIT). In this task, word pairs were primed, and participants were asked to decide whether the second word (target) was a Dutch word or a non-word. The first word (cue) was an ambiguous or unambiguous word. Target words were either associated words in a violent context, associated words in a neutral context, un-associated words, and non-words. The underlying assumption for the AIT task is that people with higher (reactive) aggression scores would show an ARIB and therefore they will be faster to decide if a target is a word or a nonword if the target is related in a violent context to the ambiguous cue, compared to the other cue-target word combinations. The participants in this study were either at risk for aggressive behavior (at-risk group), currently incarcerated (delinquent group) or healthy controls. The results showed no relationship between reactive aggression and the ARIB among the healthy control group. However, a relationship was found in the at-risk group in which more ARIB was associated with higher reactive aggression. Moreover, in the delinquent group, they found that ARIB was significantly related to proactive aggression. However, Cima et al. 2014, only examined the behavioral measures of male aggressive adults without focusing on the electro-cortical reflections of the ARIB while using the (AIT).

With this in mind, it would be interesting to see whether the previous findings of Gagnon et al. 2016 are specific to their paradigm and whether similar biases can be found while using Cima et al. 2014 lexical decision paradigm the (AIT) within a sample of healthy male undergraduate students. Therefore, the primary goal of the current study was to examine whether higher scores of self-report aggression measures were associated with greater ARIB in a sample of healthy male undergraduate students. Similar to Gagnon et al. 2016, we focused on both the behavioral and electro-cortical reflections of the ARIB, in order to understand the neuronal processes underlying this bias. ARIB was examined by using the (AIT), which was designed to measure a possible ARIB towards aggression (Cima et al., 2014). We expected higher scores on the self-report aggression measures to be associated with a faster RT to violent associated targets, but not to neutral associated targets. This was based on the expectation that participants with higher self-report aggression scores will interpret the ambiguous primes more as the ‘violent’ meaning of the word and therefore they are more primed for a target that is associated in a violent context.

In addition to behavioral measures, we also examined the N400 component of the EEG. The N400 ERP component is a negative-going peak occurring about 400 ms after the presentation of the word (Kutas & Hillyard, 1980). The N400 amplitude has been shown to be sensitive not only to the degree of semantic unexpectedness but can also be reduced by stimulus repetition, the frequency of a stimulus word in its language, and the expectancy of a word in a sentence context (Kutas & Federmier, 2000). Semantic priming studies have shown that the N400 reflects the amount of activation of semantic memory triggered by the presentation of the prime. We expected higher scores on the self-report aggression measures to be associated with smaller N400 amplitudes only in violent context targets and not in neutral context targets. This was based on the expectation that participants with higher self-report aggression scores will interpret the ambiguous primes more as the 'violent' meaning of the word and therefore they are more primed for a target that is associated in a violent context.

Material and Methods

Participants

A total number of 100 male students of which 93 participants (age range 19-30; mean age 21.86 years; SD = 2.19; 93) who completed the RPQ were included in this study. One of the participants had to be removed from the analysis due to an insufficient number of trials per condition. All participants for this study were recruited from or within the proximity of Erasmus University Rotterdam and university websites: psyweb.nl and euro system. They provided digital or written consent and received either €25 or 2 credit points for their participation in the experiment. All participants completed a general mental health questionnaire. No participants had a history of psychiatric illness, history of neurological disease, and none of the participants did use psychoactive medication, which could have affected their cognitive performance. The research was conducted according to the guidelines of the Ethics Committee of the Netherlands Institute for Psychologists (NIP).

Self-report measures

The Reactive-Proactive Questionnaire (RPQ) is an instrument that is designed to measure and differentiate between reactive and proactive aggression (Raine et al., 2006). The RPQ consists of 23 behavioral items: 11 items measuring reactive aggression, 12 items measuring proactive aggression. These behavioral items are rated on a three-point scale: 0 = (never), 1 = (sometimes), or 2 = (often). According to the findings of Raine et al. 2006, the RPQ is a reliable and valid self-report instrument that can be used to assess different types of aggression: reactive aggression, proactive aggression, and total aggression. In this study, we used the Dutch version of RPQ, which has been proven to be a reliable and valid instrument (Cima, Raine, Meesters, & Popma, 2013). Cronbach's alpha in the current sample for the entire list was 0.81, indicating satisfactory internal consistency. Cronbach's alpha in the current investigation for reactive aggression is 0.77, which indicates acceptable internal consistency. Cronbach's alpha in the current investigation for proactive aggression is 0.64, which indicates questionable internal consistency.

For further exploratory analysis, we also used the State-Trait Anger Expression Inventory-2 (STAXI-2). The STAXI-2 is a self report measure meant to assess the intensity and frequency of anger through 57 items on a 4-point Likert scale. The STAXI-2 items are spread through six different dimensions: state anger, trait anger, and anger expression in, anger expression out, anger control in, anger control out (Hovens, Rodenburg, & Lievaart, 2015). Cronbach's alpha in the current investigation for the entire list is 0.76, which indicates acceptable internal consistency. The current study only used trait anger from STAXI-2, Cronbach's alpha in the current investigation for trait anger sub-scale is 0.79 which indicates acceptable internal consistency.

Stimuli and experimental paradigm

The ARIB was measured by the AIT. During the AIT task, the first word of a pair was presented, and participants were asked to decide as quickly as possible whether the second word (Target) in a word pair was an existing Dutch word or a non-word. First, participants were presented with a cue (ambiguous or unambiguous in their meaning), 750 milliseconds later, a target was presented. In the original task by Cima et al. 2014, there were four types of target stimuli (associated words in a violent context, associated words in a neutral context, un-associated words, and non-words). See S4 Appendix for schematic illustration. However, the current study distinguished six types of word pairs: 1- Ambiguous prime violence, and target violence associated. 2- Ambiguous prime violence, and target neutral associated. 3- Ambiguous prime violence, and target neutral un-associated. 4- Unambiguous prime neutral, and target neutral associated. 5- Unambiguous prime neutral, and target neutral un-associated. 6- Unambiguous prime neutral, and target non-word. In total, the AIT consisted of 288 trials divided into three lists. Each list contained 96-word pairs in which 48-word pairs are real words and the other 48 are None-words (fillers). All the cues that the AIT used in the current experiment were tested previously by Cima et al. 2014, and they included only words with an aggressive rating above 80%.

Before the start of the electroencephalogram (EEG) measurement, participants performed twelve practice trials. In the actual task, participants were shown a fixation point for 500 ms, following that, cues were presented for 500 ms. Following that, a blank page was presented for 750 ms, and then target words were presented for (5000 ms). As for the inter-trial-interval (ITI), a black empty screen was presented in all trials for 500 ms. Finally, participants' answerers were delivered through the keyboard, in which participants had to push (P) if the letter series in front of them was a real word and (Q) if the letter series in front of them was not a real word.

Procedure

At the start of the visit, participants were instructed to complete two questionnaires (RPQ, and STAXI-2). Then, participants were seated in front of a computer screen in a comfortable chair in a light and sound-attenuated EEG room. The EEG equipment was attached, and a brief instruction was given about the task before the measurements started. Then, participants started the experiment by reading the task instructions that appeared on the screen and they were given practice trials followed by the AIT.

Electroencephalogram acquisition and analyses

EEG signals were recorded continuously at a sampling frequency of 512 Hz with a 32 - channel Active Two system (BioSemi Company, Amsterdam, the Netherlands), using Ag-AgCl electrodes mounted on an elastic cap. Six electrodes were added. Two electrodes were placed on both mastoids that served as reference electrodes. Two electrodes were placed on the outer canthi of each eye to measure the horizontal electrooculogram (EOG), and the final two electrodes were placed in the supraorbital region and in the infraorbital region of the left eye to measure the vertical EOG. The EOG signal was used to correct for eye movements and blinks. The passive electrode of Driven Right Leg (DRL) and the active electrode of Common Mode Sense (CMS) were added to use as a feedback loop as amplifier reference. Offline analysis of the EEG time series was performed using Brain Vision Analyser 2 (Brain Product GmbH, Munich, Germany). Data were referenced to the average of the recording from the left and right mastoids electrodes and band-pass filtered with cut-offs of 0.1 and 30 Hz.

Segments containing artifacts were rejected using a semi-automatic artifact rejection procedure with a maximum allowed voltage 75 µV. Epochs containing artifacts other than eye blinks (e.g., muscular activity, clipping, and movement artifacts) were removed from the data. Ocular artifacts were corrected by using a well-established procedure developed by (Gratton, Coles, & Donchin, 1983). Bad channels were interpolated with neighboring channels. The continuous EEG was segmented into an epoch that started at -200 ms, before the onset of the target stimulus and lasting until 1200 ms. Since N400 has been found to be maximal over midline electrodes, specifically the centro-parietal sites (Beres, 2017; Johnson & Hamm, 2000; Kutas & Hillyard, 1980; Kutas & Hillyard, 1982; Kutas, Hillyard, & Gazzaniga, 1988; Swaab, Ledoux, Camblin, & Boudewyn, 2012), the current study selected these three midline electrodes (Fz, Cz, and Pz). Therefore, the N400 was computed on the three mentioned midline electrodes and based on visual inspection was scored as the average amplitude in a window between 300-450 ms after the onset of the target word, the time at which the N400 was maximal.

Finally, the average number of trials that were kept during automatic artifact rejection per condition were 20.04 for Ambiguous prime violence, and target violence associated, and 20.27 for Ambiguous prime violence, and target neutral associated. As for the third condition, Ambiguous prime violence, and target neutral un-associated, the average was 20.24, and for Unambiguous prime neutral, and target neutral associated the average was 20.01. Finally, the average for Unambiguous prime neutral, and target neutral un-associated was 40.19, and for Unambiguous prime neutral, and target non-word the average was 117.77. According to recent literature, the total number of trials (24 targets ×93 participants = 2232 trials) is sufficient for the measurement of most N400 effects (Duncan et al., 2009).

Statistical analyses

Statistical analyses were performed using SPSS 23 for the Reaction Times (RT), using a 1x3 GLM repeated measures analysis was conducted with target type as the within-subject factor (Associated in a Violent context, Associated in a Neutral context and Un-associated). RT's faster than 150 ms, and slower than 1500 ms were excluded from RT analyses because they are not realistic given the relative simplicity of the task and the task instruction (respond as fast as possible). Moreover, all trials with incorrect answers (*M*= 14, 38) have been removed from the RT analysis which is an acceptable number of incorrect answers if we kept in mind that the (AIT) instructions (respond as fast as possible). Finally, the average number of trials that were excluded (due to incorrect answers and short vs long responses) from the analysis were (22, 38). We also performed additional GLM analyses for all self-report aggression measures by adding them sequentially as covariates. Moreover, we performed a Pearson correlation between the various self-report aggression measures. For a further examination of the effect of priming and word vs. non-word decisions, RT was examined by a design with target type as a within-subject factor (Associated, Un-associated, and Non-word).

For the N400, all responses were factored into the analyses including responses faster than 450 ms. In order to examine whether the violent vs. neutral context influences the N400 amplitude and whether the aggression self-report measures in any way influence the N400. The current study performed statistical analyses for the mean N400, in which a 3x3 GLM repeated measures analysis with electrodes (Pz, Fz, and Cz) and target type (Associated in a Violent context, Associated in a Neutral context and Un-associated) as within-subject factors. Then, we performed additional GLM analyses for all self-report aggression measures by adding them sequentially as covariates. To show the classical effect of word-pair association and word-non-word differences on N400 amplitudes. A GLM design with target type (Associated, Un-associated, and Non-word) and electrodes as within-subject factors examined the N400 amplitudes. Finally, to make sure that we really measure the N400 effects, we performed a follow up analysis in different electrode positions to quantify the distribution of our effects. The reported degree of freedom of the F-ratio was corrected by using the Greenhouse-Geisser method when the sphericity assumption was violated.

Results

Reaction Times (ambiguous primes)

A GLM for repeated measures analysis was conducted with target type as the within-subject factor (Associated in Violent context, Associated in Neutral context and Un-associated) showed a significant main effect of target type, $F(2, 184) = 6.86$, $p = .001$, $\eta_p^2 = .069$. See Table 1, for descriptive statistics.

Table 1

Descriptive Statistics for reaction times during the different target types in an ambiguous prime

Target Type	Mean [95% CI]	SEM	SD
Associated in Violent context	672.56 [696.77, 648.35]	12.35	119.05
Associated in Neutral context	673.14 [696.44, 649.84]	11.89	114.68
Un-associated	689.75 [712.82, 666.68]	11.77	113.47

The pairwise comparisons revealed a significant difference between associated in violent context and Un-associated, $p = .005$, between Associated in Neutral context and Un-associated, $p = .006$, but not between Associated in Violent context and Associated in Neutral context, $p = .912$. See (Fig 1), for a visual presentation.

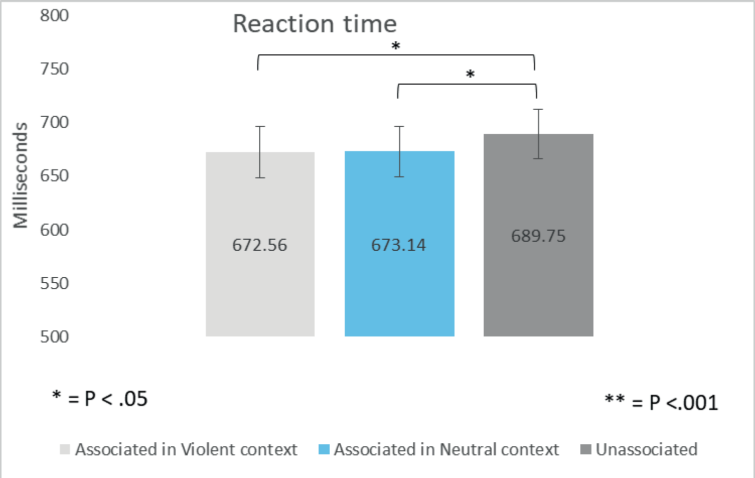


Fig 1 – The mean reaction times including 95% confidence interval error bars during the three different target types in an ambiguous prime.

With the addition of reactive aggression as a covariate, no significant main effect of this covariate was found, $F(1, 91) = .43$, $p = .516$, $\eta_p^2 = .005$, and there was no significant interaction between target type and reactive aggression, $F(2, 182) = .80$, $p = .451$, $\eta_p^2 = .009$. With the addition of proactive aggression as a covariate no significant main effect of proactive aggression was found, $F(1, 91) = .33$, $p = .569$, $\eta_p^2 = .004$, nor was there a significant interaction between target type and proactive aggression, $F(2, 182) = 1.29$, p = .268, $\eta_p^2 = .014$. With the addition of trait anger as a covariate, we found no significant main effect $F(1, 91) = .01$, $p = .946$, $\eta_p^2 = .000$, nor a significant two-way interaction between target type and trait anger, $F(2, 182) = 1.66$, $p = .193$, $\eta_p^2 = .018$. See Table 2 for descriptive statistics of the aggression measures. As for the correlations between the various self-report measures, we have found that self-report aggression measures were significantly correlated in our study. For example, reactive aggression and proactive aggression were significantly correlated, $r = 0.54$, $p < 0.01$. Moreover, reactive aggression was also significantly correlated with Trait anger $r = 0.65$, $p < 0.01$. Proactive aggression was also significantly correlated with Trait anger $r = 0.28$, $p < 0.01$.

Table 2

Descriptive statistics of the reactive aggression, proactive aggression on three different studies

	The current study Mean	SD	Gagnon et al., 2016 Mean	SD	Cima et al., 2014 Mean	SD	Cima et al., 2013 Mean	SD
Reactive Aggression	8.49 [9.18, 7.8]	3.41	13.10	4.34	8.60	3.95	7.98	4.10
Proactive Aggression	2.91 [3.38, 2.44]	2.35	3.81	2.67	2.30	2.05	2.37	3.04
Trait Anger	16.54 [17.32,15.76]	3.85	-	-	-	-	-	-

Reaction Times (un-ambiguous primes)

To examine the effect of priming and word/non-word decisions, the RT was examined with a design with target type as a within-subject factor (Associated, Un-associated, and Non-word). See Table 3 for the descriptive statistics. A significant main effect of target type was found, $F(2, 184) = 204.70$, $p < .001$, $\eta_p^2 = .690$. Pairwise comparisons show that there is a significant difference between Un-associated and Non-word, $p < .001$, between Un-associated and Associated, $p < .001$, and between Associated and Non-word, $p < .001$. See (Fig 2) for a visual presentation of the pairwise comparisons.

Table 3

Descriptive statistics of reaction times during the different target types in un-ambiguous prime

Target Type	Mean [95% CI]	SEM	SD
Un-associated	658.94 [679.29, 638.60]	10.38	100.10
Associated	639.18 [660.19, 618.17]	10.72	103.40
Non-word	735.10 [756.86, 713.34]	11.10	107.03

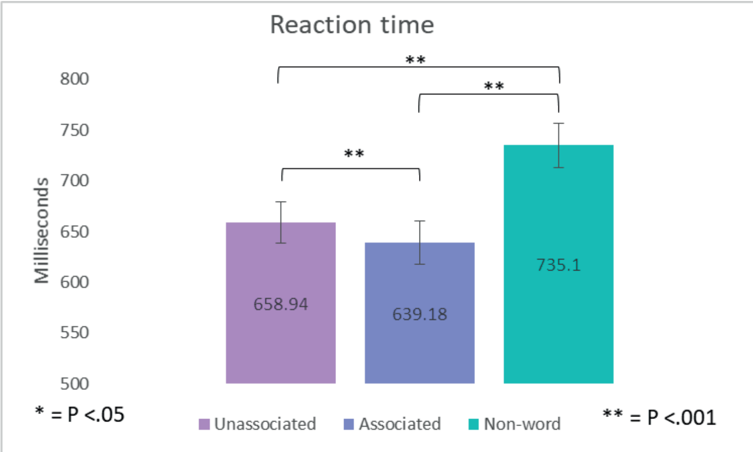


Fig 2 - The mean reaction times including 95% confidence interval error bars during the different target types in un-ambiguous prime.

N400 electrodes (ambiguous primes)

To study the effect of target type on the N400, a 3x3 GLM for repeated measures analysis with electrodes (Fz, Cz, and Pz) and target type (Associated in Violent context, Associated in Neutral context and Un-associated) as within-subject factors was conducted. See Table 4 for the descriptive statistics of the electrodes.

Table 4

Mean amplitude 300-450ms after target word onset for ambiguous primes in different electrodes, Fz, Cz, and Pz

Electrodes	Mean [95% CI]	SEM	SD
Pz	0.53 [1.18, -0.12]	0.33	3.36
Fz	-1.99 [-1.34, -2.64]	0.33	2.97
Cz	-1.99 [-1.30, -2.68]	0.35	3.33

There was a significant main effect of electrode position, $F(2, 184) = 54.73$, $p < .001$, $\eta_p^2 = .373$, but not for target type, $F(2,184) = .55$, $p = .576$, $\eta_p^2 = .006$. The two-way interaction between electrodes and target type was not significant, $F(4, 368) = .41$, $p = .803$, $\eta_p^2 = .004$. Follow-up analyses showed that there was a significant difference between Pz and Fz, $p < .001$, Pz and Cz, $p < .001$, but no significant difference for Fz and Cz, $p = .982$. See (Fig 3), for a visual representation for the differences between the N400 electrodes.

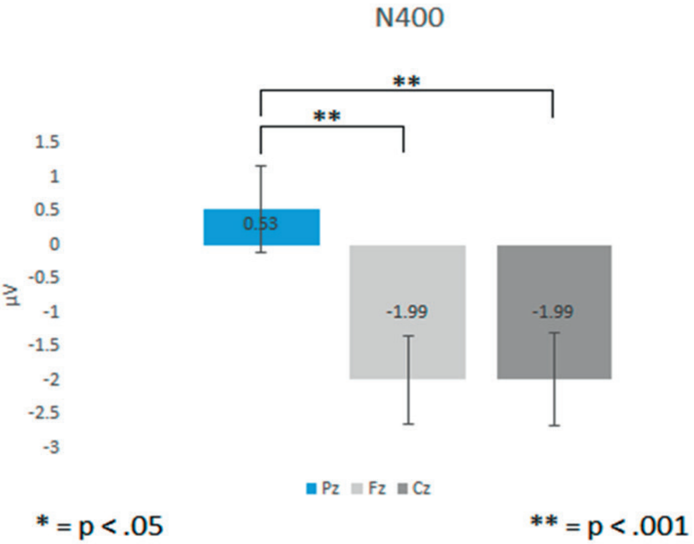


Fig 3 – N400 measures at electrode position Pz, Fz, and Cz.

Reactive aggression was included as a covariate in this analysis, but no significant main effect of reactive aggression was found, $F(1, 91) = .12$, $p = .733$, $\eta_p^2 = .001$. Similarly, the analysis revealed a non-significant two-way interaction between electrodes and reactive aggression, $F(2, 182) = .01$, $p = .995$, $\eta_p^2 = .000$, between target type and reactive aggression, $F(2, 182) = 1.96$, $p = .144$, $\eta_p^2 = .021$, as well as a non-significant three-way interaction between electrodes, target type, and reactive aggression, $F(4, 364) = .26$, $p = .907$, $\eta_p^2 = .003$. Next, proactive aggression was added as a covariate. Similar to reactive aggression, there was no significant main effect of proactive aggression, $F(1, 91) = .02$, $p = .877$, $\eta_p^2 = .000$. In addition, there were no significant two-way interactions between electrodes and proactive aggression, $F(2, 182) = .08$, $p = .927$, $\eta_p^2 = .001$, between target type and proactive aggression, $F(2, 182) = 2.36$, $p = .098$, $\eta_p^2 = .025$, or between electrodes, target type, and proactive aggression, $F(4, 364) = .41$, $p = .802$, $\eta_p^2 = .004$. The addition of trait anger as a covariate did not result in a significant main effect of trait anger, $F(1, 91) = .03$, $p = .860$, $\eta_p^2 = .000$. The two-way interactions between electrodes and trait anger, $F(2, 182) = .02$, $p = .978$, $\eta_p^2 = .000$, and between target type and trait anger, $F(2, 182) = .34$, $p = .709$, $\eta_p^2 = .004$, turned out to be non-significant. Similarly, there was a non-significant three-way interaction between electrodes, target type, and trait anger, $F(4, 364) = .04$, $p = .997$, $\eta_p^2 = .000$.

N400 electrodes (un-ambiguous primes)

To further examine the effect of priming and word/non-word decisions, the N400 amplitude was examined with a design with target type (Associated, Un-associated, and Non-word) and electrodes as within-subject factors. See Table 5 for the descriptive statistics. There was a significant main effect of electrodes, $F(2, 184) = 66.34$, $p < .001$, $\eta_p^2 = .419$. Pairwise-comparisons show significant differences between Pz (*M* = 1.43; *SEM* = .40) and Fz (*M* = -1.61; *SEM* = .32), $p < .001$, between Pz and Cz (*M* = -1.23; SEM = .37), $p < .001$, but not between Fz and Cz, $p = .364$. In addition, a significant main effect of target type was found, $F(2, 184) = 33.47$, $p < .001$, $\eta_p^2 = .267$. Pairwise comparisons showed a significant difference between Un-associated (*M* = -.10; *SEM* = .38) and Non-word (*M* = -1.78; *SEM* = .34), $p < .001$, between Non-word and Associated (*M* = .48; *SEM* = .38), $p < .001$, but not between Un-associated and Associated, $p = .142$. Lastly, there was a significant two-way interaction between electrodes and target type, $F(4, 368) = 9.99$, $p < .001$, $\eta_p^2 = .098$.

4

Table 5

Descriptive statistics of the electrodes during the different target types in un-ambiguous prime

Electrodes	Target Type	Mean [95% CI]	SEM	SD
	Un-associated	1.82 [2.70, 0.94]	.45	4.32
Pz	Associated	1.99 [2.87, 1.11]	.45	4.34
	Non-word	.48 [1.24, -.28]	.39	3.78
	Un-associated	-1.32 [-.60, -2.05]	.37	3.57
Fz	Associated	-.43 [.33, -1.19]	.39	3.72
	Non-word	-3.08 [-2.39, -3.77]	.35	3.35
	Un-associated	-.81 [.03, -1.65]	.43	4.13
Cz	Associated	-.12 [.74, -.98]	.44	4.23
	Non-word	-2.75 [-1.99, -3.51]	.39	3.77

Follow-up analysis revealed that, on Pz, there is a significant difference between Un-associated targets and Non-word, $p < .001$, between Associated and Non-words, $p < .001$, but not between Un-associated and Associated, $p = 1.000$. The pairwise comparisons within Fz show that there is a significant difference between Un-associated and Associated, $p = .015$, between Associated and Non-words, $p < .001$, and between Un-associated and Non-word, $p < .001$. The pairwise comparisons within Cz show that there is a significant difference between Un-associated and Non-word, $p < .001$, between Associated and Non-words, $p < .001$, but not between Un-associated and Associated, $p = .144$. To summarize, for all analyzed electrodes we could detect a significantly larger N400 for non-words as compared to words, but only within Fz could a significantly higher N400 for Un-associated as compared to Associated words could be found. See (Fig 4), for a visual representation.

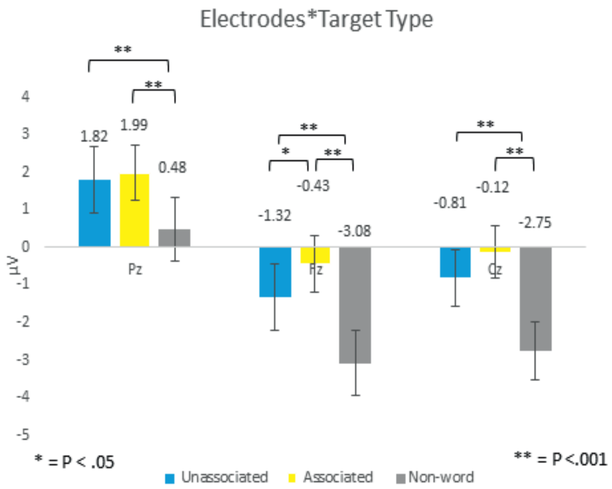


Fig 4 - The mean reaction times including 95% confidence interval error bars during the three different target types in un-ambiguous prime within the three different electrode positions

Discussion

The aim of the current study was to examine the relationship between aggression and ARIB in adults. Although, we found the expected shorter RT for associated words, there was no significant evidence of an effect of target type words associated in a violent context on the RTs or any relation with self-report measures of aggression. Concerning the electrophysiological responses, for all electrodes we were able to detect a significantly larger N400 for non-words as compared to words. However, only on the Fz electrode we found a significantly higher N400 for un-associated as compared to associated words. Finally, the N400 was not related to the self-report measures of aggression.

Reaction times

In general, our finding of faster RT for associated words as compared to un-associated words and non-words shows that the participants performed the task as expected. However, the current study could not find the predicted association between reactive aggression and RT. This result is in line with Cima et al. 2014 that no relationship was found between reactive aggression and ARIB among individuals who scored similarly to the current study participants in reactive aggression measures. For example, in the healthy control group, no relationship was found in Cima et al. 2014, between reactive aggression and ARIB. However, they did find a relationship in the at-risk group. They found that more ARIB was associated with higher reactive aggression. Moreover, in the delinquent group, they also found that ARIB was significantly related to proactive aggression in which more ARIB was associated with higher proactive aggression. See table 2 for means comparison. It may be the case that the AIT is only able to measure ARIB in a forensic population, but might not be sensitive enough to measure ARIB in healthy populations.

As for our results that self-report aggression measures were significantly correlated to each other. Although these concepts (reactive, proactive aggression, and trait anger) are somewhat related to each other, they still conceptually distinct from each other (Parrott & Giancola, 2007). In fact, self-report anger and behavioral aggression often correlate only weakly in both laboratory, and real-life situations (Kassinove, Roth, Owens, & Fuller, 2002; Giumetti & Markey, 2007; Nesbit, Conger, & Conger, 2007). The relationship between anger and aggression is likely to be influenced by several variables that need to be explored further in empirical research (Hortensius, Schutter, & Harmon-Jones, 2011).

N400-effect

In line with our expectation, we were able to detect a significantly larger N400 for non-words as compared to words. However, we only found a larger N400 for un-associated words on Fz as compared to associated words. This result is more or less in line with previous findings, since it is known that the N400 scalp distribution seems to depend on the used task. For example, when elicited by single words (similar to the current study task), the N400 has a more anterior distribution with a maximum over frontal or central sites (Bentin, McCarthy, &Wood, 1985; Bentin, 1987; McCarthy & Nobre, 1993). In contrast, when elicited by semantic incongruities in sentences, the N400 seems to be largest over the centro-parietal regions (Johnson & Hamm, 2000; Kutas & Hillyard, 1982; Kutas, Hillyard, & Gazzaniga, 1988). In addition to that, a larger N400 signal for non-words could be explained by participants trying to understand or pronounce the non-words, a process known to generate a larger N400 signal as a response (Deacon, Dynowska, Ritter, & Grose-Fifer, 2004). As mentioned in the introduction, the N400 amplitude has been proven to be sensitive not only to the degree of semantic unexpectedness but can also be reduced by stimulus repetition, the frequency of a stimulus word in its language, and the expectancy of a word in a sentence context (Kutas & Federmier, 2000).

Reactive aggression, proactive aggression, and trait anger failed to interact with differences in N400 waveforms. Our results are not in line with the study of Gagnon et al. 2016, who found that the presentation of a critical word that violated hostile expectations elicited significantly larger N400 response in aggressive than nonaggressive individuals. However, it must be noted that the current study differs from Gagnon et al. 2016, in terms of its sample characteristics. While the sample in the Gagnon et al. 2016 study was recruited from three different populations, that is, university students, the general population, and patients consulting in a personality disorder clinic, the current study sample was recruited only from non-clinical group university students. Moreover, participants in the sample in Gagnon et al. 2016 were much more aggressive individuals that scored higher than the current study sample on reactive and proactive aggression as measured by the French version of the RPQ. Participants' scores in the current sample are closer to the standard group of none offenders' adults from Cima et al. 2013. See Table 2. The absence of high aggressive scores in the current sample might be the reason why there is no association between N400 and the self report measures of aggression. Recently it has been argued that the role of the reactive and proactive aggression on the ARIB is more prominent in populations in which higher reactive-proactive aggression scores are found (Tuente et al., 2019).

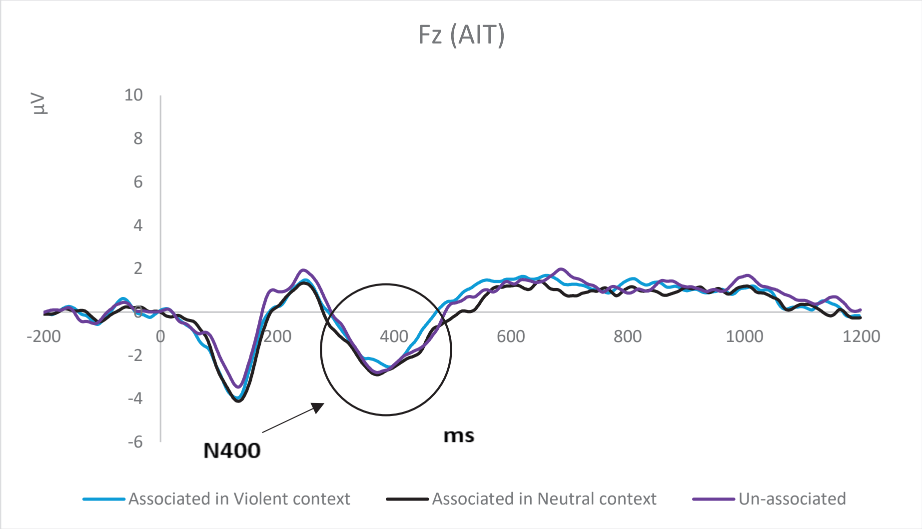
Limitations and implications for future research

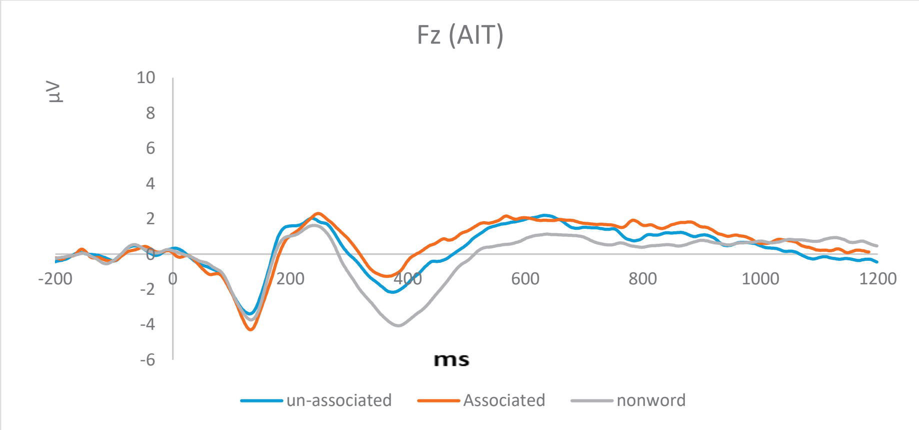
Despite the fact that this study paradigm was successful in its findings of longer RT and larger N400 for un-associated words and non-words as compared to associated words, and that these findings are in line with Cima et al. 2014 in the healthy control group, it should be noted that this study has at least one limitation. The distribution of aggression scores in the current sample is relatively narrow and high scores are more or less absent in the current sample. Future studies should include more extreme aggression scores in order to study the relation between ARIB and aggression in adults. Finally, aggression outside clinical and/or forensic context is a big enough problem for society as a whole, which needs a better understanding of the underlying processes in "healthy" people.

Conclusion

To conclude, the ARIB was not associated to the self-report aggression measures in the current study population. However, we have managed to find shorter reaction times for associated words as compared to un-associated words and non-words, but the nature of the relationship in terms of violence had no influence on the speed. Finally, our results showed that the AIT is only able to measure the associations between ARIB in terms of RT and N400 and self-report aggression measures in a forensic population but might not be sensitive enough to measure ARIB in healthy populations.

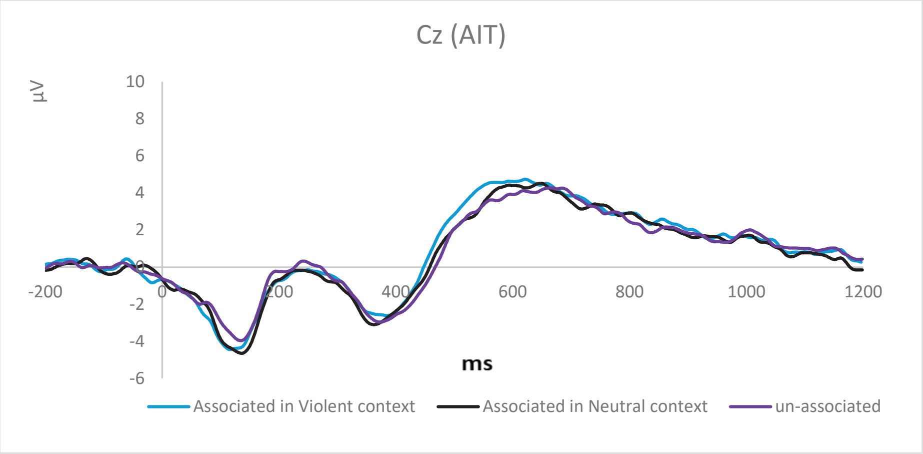
Appendix 1

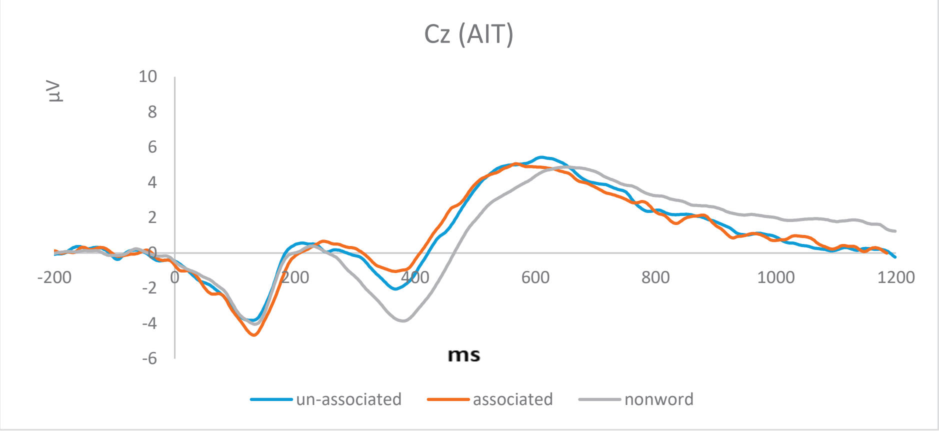




Grand-average ERPs on electrode position Fz.

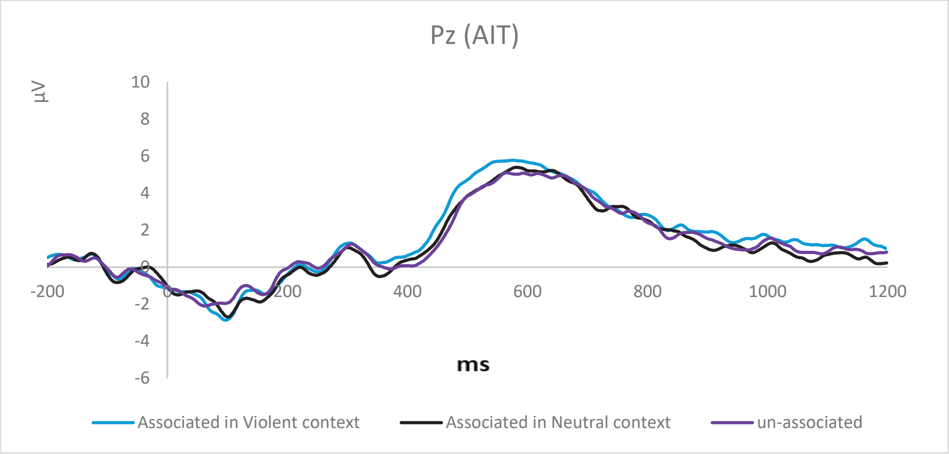
Appendix 2

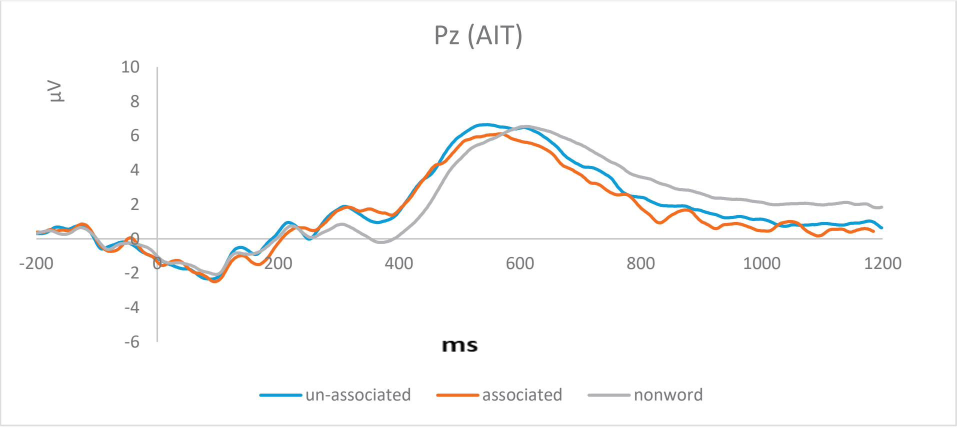




Grand-average ERPs on electrode position Cz.

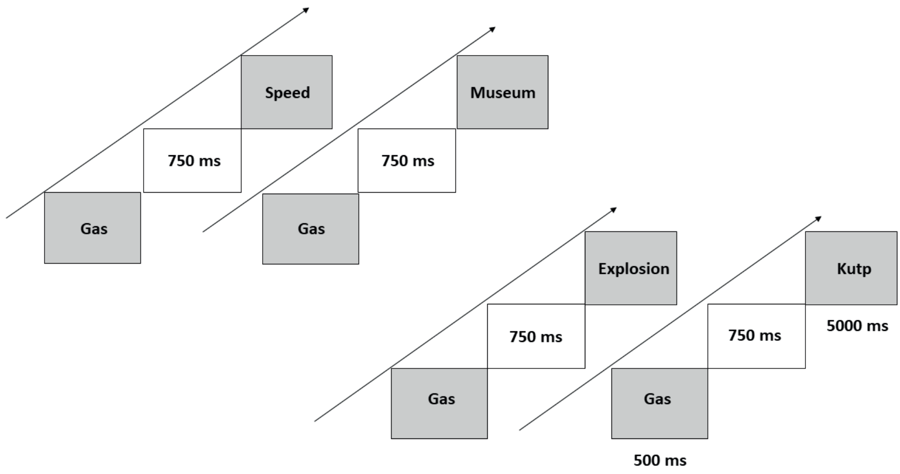
Appendix 3





Grand-average ERPs on electrode position Pz.

Appendix 4



The upper figure is an example of a target stimuli that is Associated in a Neutral context. The next example (upper figure) is for a target stimuli that is Un-associated. The third example is for a target stimuli that is Associated in a violent context. The final example is for a Non-word which is the experiment control condition target stimuli.

CHAPTER 5

Cognitive Behavioral Modification to modulating Negative and Positive Attributions in Individuals scoring High on Reactive Aggression Measures

THIS CHAPTER HAS BEEN SUBMITTED FOR PUBLICATION AS: HAKAMI, I., F.M. VAN DER VEEN, FRANKEN., I.H.A. & HUIJDING, J. COGNITIVE BEHAVIORAL MODIFICATION TO MODULATING NEGATIVE AND POSITIVE ATTRIBUTIONS IN INDIVIDUALS SCORING HIGH ON REACTIVE AGGRESSION MEASURES.

Abstract

Learning from errors and feedback is important in Cognitive Bias Modification procedures (CBM). We studied whether neurophysiological markers of feedback (FRN) and error (ERN) processing during an attributions training for facial expressions, is related to self-report aggression and anger before training, and to changes in hostile attribution bias (HAB) from pre-to post-training. During a Face Task (FT) prototypical happy and angry images were used to alter participants' attributions of ambiguous emotional expressions. The results showed that positive training succeeded in increasing positive attributions. Self-report trait aggression also showed a modest relation to lower HAB before the training. However, the correlations between HAB before the training and reactive aggression and trait anger were not significant. More importantly, the FRN and ERN during the attributions training for facial expressions were not related to self-report aggression and anger, indicating that individuals scoring higher on anger or aggression do not show impaired error and feedback processing during this CBM-I procedure. The FRN and ERN were not related also to changes in (HAB) from pre-to post-training. Finally, the FRN amplitudes were larger in response to negative feedback as compared to positive feedback, and ERN amplitudes were larger in response to incorrect responses as compared to the correct responses. This suggests that the FRN and ERN are not sensitive to the effectiveness of CBM-I procedure in a sample from a healthy population who score low on self-report aggression. The currently used approach might be better suitable for clinical populations or in a combination with a provocation manipulation.

Introduction

Aggressive behavior often has serious emotional and social consequences among victims and people witnessing aggression. On a societal level, this behavior has severe economic consequences. For instance, in 2004 the World Health Organization (WHO) estimated the cost of interpersonal violence in the U.S. at more than $300 billion per year, including medical costs and costs due to loss of productivity (Waters et al., 2004). Even though we know much about the correlates of aggressive behavior, relatively little is known regarding the factors that cause such behavior to persist. Therefore, understanding the etiology of aggressive behavior has been an important focus in past research (Bailey & Ostrov, 2008). A relevant distinction that is often made is between proactive and reactive subtypes of aggression (Roth & Strueber, 2009), in which reactive aggression is categorized as a behavior committed in anger as retaliation for a perceived negative experience. On the other hand, proactive aggression is characterized by planned and goal-directed behavior (Murray-Close & Ostrov, 2009). While there is overlap between reactive and proactive aggression, the Social Information Processing (SIP) model provides a useful explanation of specific cognitive processes that are assumed to underlie each type of aggression (Crick & Dodge, 1994; Dodge, 1986). This model proposes that behavioral responses to social situations depend on a sequence of information-processing steps, including (1) encoding of external and internal cues, and (2) attributions of those cues (Crick & Dodge, 1994). During these two steps, individual differences occur as they selectively attend to cues, which are then subject to attribution.

In general, reactive aggressive individuals tend to interpret the others' intentions as hostile, which is often referred to as Hostile Attribution Bias (HAB; Nasby, Hayden, & DePaulo, 1980). For this reason, HAB plays a central role across different information-processing theories of aggression (Guerra & Huesmann, 2004). In fact, reactive aggression as a behavior is found to be associated with this HAB, in adults. That is, higher levels of this bias are found to be associated with higher levels of aggressive behavior (see e.g., Martinelli, Ackermann, Bernhard, Freitag, & Schwenck, 2018; Tuente, Bogaerts, & Veling, 2019 for a recent review). For instance, one study showed a group of male offenders in a correctional facility and a group of matched controls a set of pictures of faces showing different proportions of angry, happy and fearful expressions (Schönenberg & Jusyte, 2014). Participants were asked to judge for each of these ambiguous stimuli whether the face depicted anger, happiness or fear. Results showed that aggressive individuals attributed ambiguous facial cues as more hostile (i.e., angry) than the control group. Moreover, aggressive individuals showed a tendency to overrate the perceived intensity of anger.

The findings mentioned in the previous paragraph inspired the development of a number of interventions aimed at reducing reactive aggressive behavior by manipulating social information processing. One possible method to target information processing directly is a Cognitive Bias Modification training to target the attribution bias (CBM-I). CBM-I aims at modifying attribution bias by exposing the participants several times to ambiguous situations and training them to interpret these situations either in a negative or positive way, by providing feedback on their responses (Vassilopoulos, Brouzos, & Andreou, 2015). For instance, in one of the first CBM-I studies in the context of aggression 135 undergraduate students were randomly assigned to a positive training, negative training or a control condition. Results showed that the positive training led to a decrease in HAB while the negative training led to an increase in HAB. Interestingly, participants in the positive training also reported fewer angry responses in reaction to an insult than participants from the other conditions. This result supports the idea that HAB can be modified using CBM-I procedures and that this may affect anger and aggression (Hawkins & Cougle, 2013).

Recently, other studies have also supported that HAB can be modified using CBM-I procedures (e.g., AlMoghrabi, Huijding, & Franken, 2018; Penton-Voak et al., 2013). For example, in one study 40 healthy adult male participants were randomly assigned to either a single session of positive training to increase positive attributions or negative training to increase HAB. The results revealed that positive training led to an increase in positive attributions and the change in these attributions was associated with lower anger and verbal aggression scores after training. The negative training seemed to have no effect on HAB (AlMoghrabi et al., 2018). In another study, 40 youths at risk of criminal offending were trained to perceive happiness over anger in ambiguous facial expressions. Immediately after training, and two weeks later, participants who received the training showed reduced self-report and observer-report aggressive behavior, compared to participants who received a control training. This suggests that training the attribution of low-intensity (ambiguous) expressions can affect the attribution of emotions outside the lab, and have a positive effect on mood and behavior (Penton-Voak et al., 2013). The potential importance of attributions of facial expressions in the context of aggression is further supported by results showing that aggressive individuals attribute neutral and ambiguous facial expressions more negatively than nonaggressive individuals do (Schönenberg & Jusyte, 2014).

Since aggression studies revealed that changes in information processing may mediate changes in behavior (Dodge, 2006). This raises the question of what factors influence the effectiveness of CBM-I training. One factor seems to be the way people make use of feedback during CBM-I. Since learning from errors and feedback is probably crucially important in CBM-I, it seems logical to examine how error and feedback processing is related to the effectiveness of CBM-I. Concerning this issue, it is interesting to note that aggressive individuals appear to have difficulties with learning from previous experiences (e.g., Matthys, Vanderschuren, Schutter, & Lochman, 2012). In addition, aggressive individuals typically show decreased activation in the anterior cingulate cortex (ACC; Stadler et al., 2007), which is the brain structure involved in conflict and error monitoring (Holroyd & Yeung, 2012; Holroyd & McClure, 2015; Shenhav, Botvinick, & Cohen, 2013). This suggests that CBM-I procedures could be less effective in high versus low aggressive individuals, and may have implications for the design of CBM-I studies in the context of aggression.

An effective way to study error and feedback processing is by recording Event-Related Potentials (ERP's) while participants perform a learning task (Chavarriaga, Sobolewski, & Millán, 2014; Ferrez & Millan, 2008). When people make or perceive an error during such a task, an error-related potential (ERN) can be detected in the electroencephalogram (EEG) due to the person recognizing that error (Falkenstein, Hoormann, Christ, & Hohnsbein, 2000). The ERN is a negative ERP potential wave, time-locked to (incorrect) responses at the frontocentral scalp, which usually appears between 0 and 100 ms after an error has been committed (Gehring, Liu, Orr, & Carp, 2012). On the other hand, another fronto-central negative potential has also been observed after a correct response, referred to as correct-related negativity (CRN; Mathalon et al., 2002). However, due to this fact, the functional significance of ERN remains controversial among researchers, but evidence suggests that ERN indexes general behavioral monitoring within the brain (Van Veen & Carter, 2002; Bates, Liddle, Kiehl, & Ngan, 2004). In fact, the ERN is found to be reduced in individuals with behavioral problems that are strongly associated with reactive aggression (Hall, Bernat, & Patrick, 2007).

Closely related to the ERN, is the feedback-related negativity (FRN). The FRN is a fronto-central negative deflection that occurs 200–300 ms after negative feedback is presented (Holroyd & Coles, 2002; Talmi, Atkinson, & El-Deredy, 2013). The FRN reflects the internalization of external feedback (i.e., before learning), and the ERN reflects internal (error) monitoring. Taken together, these findings suggest that it might be interesting to study feedback processing on the behavioral and neuronal level in a healthy population varying in their self-report reactive aggression during the completion of a CBM-I training.

The primary objective of this study was to examine whether self-report aggression and anger are related to the FRN and ERN during an attribution training for facial expressions, and whether the FRN and ERN are related to changes in HAB from pre- to post training. Based on the earlier finding that aggression may be related to weaker error and feedback processing, we expected that higher self-report aggression would be related to smaller ERN and FRN responses. In addition, we expected that stronger FRN and ERN responses would be related to a bigger change in HAB over the training. As a prerequisite for examining these questions, we expected participants who score high on reactive aggression (RA) to perceive anger facial expressions more than people who score low on the reactive aggression during facial expressions training. We also expected that participants would alter their perception of ambiguous emotional expressions during the training and perceive more positive facial expressions.

To our knowledge this is the first study to examine error and feedback processing during a CBM-I. Therefore, the second objective of this study was to test the sensitivity of FRN and ERN to errors and feedback during a CBM-I task. More specifically, we examined whether FRN and ERN responses differed depending on the accuracy of responses and the type of feedback. Based on previous studies (e.g., Miltner, Braun, & Coles, 1997; Nieuwenhuis, Holroyd, Mol, & Coles, 2004; Yeung & Sanfey, 2004) we expected to find larger FRN amplitudes in response to negative feedback as compared to positive feedback, and larger ERN amplitudes in response to incorrect responses as compared to correct responses (Falkenstein et al., 1991; Gehring et al., 1993).

Material and methods

Participants

A total number of 95 participants from the Erasmus University Rotterdam (EUR) student population were tested, of which 94 participants (age range 18-28; mean age 20.7 years; SD = 2.0) were included. One participant had to be removed from the analysis due to a very noisy signal. All participants for this study were recruited from or within the proximity of Erasmus University Rotterdam and university websites: psyweb.nl and euro system. They provided digital or written consent and received either €25 or 2 credit points for their participation in the experiment. All participants completed a general mental health questionnaire. No participants had a history of psychiatric illness, history of neurological disease, and none of the participants did use psychoactive medication, which could have affected their cognitive performance. The research was conducted according to the guidelines of the Ethics Committee of the Netherlands Institute for Psychologists (NIP).

Self-report measures

The Reactive Proactive Aggression Questionnaire

Originally developed by Raine et al. (2006), the Reactive Proactive Aggression Questionnaire (RPQ) is a self-report instrument designed to distinguish between reactive aggression and proactive aggression in children, adolescents, and adults. The RPQ consists of 23 items, of which there are 11 of them assess reactive aggression, and 12 items assess proactive aggression. Each item is rated using a 3-point Likert scale (0 = never, 1 = sometimes, 2 = often) and results in a total score that ranges from 0 to 46. The totaled scores provide a measure of reactive aggression or proactive aggression, as well as total aggression. Higher scores indicate higher levels of aggression. In this study, we used the Dutch version of the Reactive-Proactive aggression questionnaire (RPQ), which has been proven to be a reliable and valid instrument (Cima, Raine, Meesters, & Popma, 2013). Cronbach's alpha for the entire list were 0.73, indicating an acceptable internal consistency. The current study only used reactive aggression (RA) list, Cronbach's alpha in the reactive aggression list was 0.69, which indicates a questionable internal consistency.

The State-Trait Anger Expression Inventory-2

The Dutch State-Trait Anger Expression Inventory-2 (STAXI-2; Hovens, Lievaart, & Rodenburg, 2014) is a self-report instrument specifically designed to measure the experience of anger, the tendency to express anger, and the tendency to control anger in adolescents (16-19), adults, and psychiatric patients. The STAXI-2 is scored on a 4-point Likert scale (1 = almost never, 2 = sometimes, 3 = often, 4 = almost always) and consists of 57 items and 6 subscales: State Anger, Trait Anger, Anger Expression-Out, Anger Expression-In, Anger Control-Out, Anger Control-In. For each subscale, the scores are summed such that higher scores indicate more aggression. Cronbach's alpha for the entire list were 0.73, indicating an acceptable internal consistency. The current study only used Trait Anger (TA) list, Cronbach's alpha in the Trait Anger list was 0.79, which indicates an acceptable internal consistency.

The Trait Aggression Questionnaire (Agressie Vragenlijst; AVL)

Trait Aggression Questionnaire (Meesters, Muris, Bosma, Schouten, & Beuving, 1996) is the Dutch version or Buss Perry's Aggression Questionnaire. This questionnaire consists of 29 items scored on a 5-point Likert scale, ranging from: 1 = completely disagree to 5 = fully agree. AVL is divided into 4 dimensions: physical aggression, verbal aggression, anger, and hostility. Morren & Meesters (2002), found support for validity in Dutch adolescent offenders (all men) through relevant correlations with measurement instruments for aggressive behavior symptoms, anger as disposition, social desirability, and age. Cronbach's alpha in the current sample was 0.79, indicating acceptable internal consistency.

Stimuli and task

The Face Task (FT) was constructed by Penton-voak and colleagues (2013). In this task, 40 composite images were generated using 20 faces for a male individual showing a happy facial expression and, another 20 faces for the same individual showing an angry facial expression. Similar to the Penton-voak et al. (2013) study, we used these composite images as endpoints to generate a linear morph sequence that consists of images that change step-by-step from unambiguously happy to unambiguously angry, with neutral images in the middle. See Appendix 1. This morph sequence had 15 images with an equal spacing that we used as experimental stimuli. The procedure for the baseline and test phases of this task were largely similar. During all phases each image from the morph sequence was presented three times in a random order, resulting in 45 trials per phase. Each trial started with the presentation of a fixation cross (1500-2500 ms. randomly jittered) in the middle of the screen, followed by an image of the morph sequence. After 150 ms, the image was replaced by a mask of visual noise (150 ms) after which a prompt appeared asking the participant to indicate whether the presented image showed a happy or an angry face. As grading responses to such morph continua tend to shift monotonically from one response to another over the sequence (Young et al., 1997), the balance point, at which just as much happiness or anger is perceived in the presented faces can easily be calculated for each participant from the number of "happy" responses as a percentage of the total number of trials. The individual balance points that calculated from the baseline phase were used to provide participants feedback on their response during the test phase (during the baseline phase participants received no feedback). However, to train participants towards a more positive attribution of the faces, the two images nearest the balance point that the participant would have classified as "angry" at baseline were now considered "happy" for purposes of feedback. In order to make the feedback signals more neutral and to avoid physical differences between the feedback signals, we employed a yellow rectangular shape to indicate “correct” feedback, and a blue rectangular shape to indicate “incorrect” feedback. This feature was counterbalanced for every other participant, where yellow and blue switched from “correct” to “incorrect” feedback and vice versa. FT task contains six training blocks in which every face from the 15- morph sequence is presented twice (in random order) within each block, for a total of 180 training trials. At the end of the test phase, the participant's balancing point was calculated again by using the same method as in the baseline phase, in order to establish whether the training modified the participant's perception of ambiguous facial emotional expressions. Finally, Penton-Voak et al. (2013) showed that FT was able to alter participant perception of ambiguous emotional expressions. In addition to this, the FT testing phase demonstrated that feedback had shifted participants’ balance points in the direction of training to choose happiness over anger facial expressions (Penton-Voak et al., 2013).

Procedure

Before taking part in the experiment, the recruited participants gave their consent by signing up on either psyweb.nl or the euro system website. When on site, the participants were informed about the sequence of events that would take place in addition to what they entailed. First, the participants went through an intake procedure that included the self-report questionnaires RPQ, STAXI-2, and AVL. After the intake, the participants were taken to the EEG room, then the participants were seated in front of a computer screen in a comfortable chair in a light and sound-attenuated room. All the EEG equipment was attached including the electrode with which the heart rate was measured. Then, a brief instruction was given about the task before the measurements started. Lastly, participants started the experiment by reading the task instructions that appeared on the screen and they were given practice trials followed by FT and Island Getaway (IGT). The latter task was only used for an exploratory purpose and discussion of this task is beyond the scope of this manuscript.

Electroencephalogram acquisition and analyses

Continuous electroencephalography (EEG) was recorded at a sampling frequency of 512 Hz using a 32-channel Ag/AgC1 electrode cap and the Biosemi Active-Two with Actiview acquisition software (BioSemi, Amsterdam, the Netherlands). In addition, six electrodes were added for different purposes. Two electrodes were placed on both mastoids to serve as reference electrodes. Another two were placed on the outer canthi of each eye to measure the horizontal electro-oculogram (HEOG). In order to measure the vertical electro-oculogram (VEOG) and to correct for eye movements, the final two electrodes were placed on the supraorbital region and in the infraorbital region of the left eye. For the creation of the feedback loop as amplifier reference, the active electrode of Common Mode Sense (CMS) and the passive electrode of Driven Right Leg (DRL) were used. The analysis of the data acquired occurred offline using the Brain Vision Analyzer 2 (Brain Product GmbH, Munich, Germany). Data were filtered offline with a bandpass filter of 0.1 – 30 Hz. Segments containing artifacts were rejected using a semi, automatic artifact rejection procedure with a maximum allowed voltage 75 µV. Epochs containing artifacts other than eye blinks (e.g., muscular activity, clipping, and movement artifacts were removed from the data. Ocular artifacts were corrected by using the (Gratton, Coles, & Donchin, 1983) procedure. Bad channels were interpolated with neighboring channels. For the FRN components, the continuous EEG was segmented into an epoch that starts at -200 ms before the onset of the feedback and lasting until 800 ms after feedback. EEG was averaged separately for positive and negative trials. Following that, baseline correction was applied using the 100 ms pre-stimulus interval. FRN amplitude was computed as an area between 230-310 ms after feedback, at Fz, FCz, and Cz. For the ERN components, the ERPs (ERN, and CRN) were computed within an epoch starting -200 ms prior to the response and lasting until 800 ms after the response. The baseline correction was applied to begin at -198.00 ms and to end at -100.00 ms. ERN amplitude was computed as an area between - 20-80 ms following the error response. We also define the CRN as the area between -20 and 80 ms following the correct response.

Results

Preliminary analyses

As a first step we tested whether the Face Task was successful in changing attributions of facial expressions, using a repeated-measures ANOVA with the percentage of images classified as "happy "as the dependent variable, and assessment (pre-training, post-training) as an independent variable. In line with our expectation, this analysis showed that participants mean percentage of faces classified as “happy” significantly increased from pre- training (M = 6.74, SD = 1.65) to post training (M = 8.25, SD = 1.75), $F(1.93) = 141.89$, $p < .001$, $\eta_p^2 = .604$. See table 1 for the descriptive statistics.

Table 1

The means, 95% confidence intervals, and the standard deviations of the two conditions

	Mean [95% Confidence Interval]	Standard Deviation
Pre training	6.74 [6.41, 7.07]	1.65
Post training	8.25 [7.90, 8.60]	1.75

Next, in order to test the hypothesis that a higher score on RA, TA, and AVL would correlate with a higher tendency to perceive ambiguous emotional expressions as angry (HAB) before the training we performed a simple correlation. In line with our expectations, higher self-report aggression on the AVL showed a modest but significant relation to lower HAB pre training, $r(94) = -.22$, $p < 0.05$. However, the correlations between HAB pre training and RA and TA were not significant for both $r < 0.01$, all p's < -.08.

Finally, to test whether self-report anger and aggression influenced the change of HAB over training, a repeated-measures GLM was performed with HAB as the dependent variable, training condition (pre-training, post-training) as the within-subject factor and RA,TA, and AVL as a covariate. The analyses where RA, TA, and AVL were entered sequentially as covariates did not yield main or interaction effects (all p's > .082). In other words, the effects of training were not influenced by RA, TA, and AVL.

Feedback Related Negativity (FRN)

In order to test whether reactive aggression and other self-report questionnaires influenced FRN amplitude, we performed another GLM analyses with FRN amplitude as the dependent variable, and feedback level and electrode positions as within-subject variables, with all self-report questionnaires added as covariates sequentially (RA, TA, and AVL). No main effects were found, and, none of the covariates yielded significant two-way (all p's > .921), or significant three-way interaction effects (all p's > .781). In other words, neither RA, TA, nor AVL influenced the FRN amplitude.

To address the research question, whether feedback processing as indexed by the FRN is related to the effectiveness of the training, we first calculated HAB change scores by subtracting the HAB scores pre-training from the HAB scores post-training. Then, we did a repeated-measures GLM analysis with FRN amplitude as the dependent variable, and feedback level and electrode positions as within-subject variables, wherein the HAB change scores was inserted as a covariate. No main effect was found of the HAB change scores, $F(1, 92) = .501$, $p = .481$, $\eta_p^2 = .005$. In addition, there was no significant two-way interaction between electrodes and HAB change scores, $F(2, 184) = .747$, $p = .475$, $\eta_p^2 = .008$, nor between feedback and HAB change scores, $F(1, 92) = .600$, $p = .414$, $\eta_p^2 = .006$. There was also no significant three-way interaction, $F(2, 184) = 1.98$, $p = .140$, $\eta_p^2 = .021$.

To assess whether the FRN is sensitive to feedback during the FT, a repeated-measures ANOVA was performed with FRN amplitude as the dependent variable, and Electrodes (Fz, FCz, Cz) and Feedback (positive feedback, negative feedback) as the within-subject factors. A main effect was found for electrodes, $F(2, 186) = 13.99$, $p <$

.001, η_p^2 = .131. The pairwise comparisons show that Fz (*M* = -.652, *SD* = 3.67) is significantly different from Cz (*M* = -1.50, *SD* = 4.03), *p* = .001, whilst Fz is not significantly different from FCz (*M* = -.797, *SD* = 4.04), *p* = .872. On the other hand, Cz is significantly different from FCz, *p* < .001. A main effect was also found of feedback type, *F*(1, 93) = 12.85, *p* = .001, η_p^2 = .121, where FRN amplitude for positive feedback (*M* = .032, *SD* = 3.27) is smaller or less negative than negative feedback (*M* = -2.00, *SD* = 5.77). See table 2 for the descriptive statistics.

Table 2

The means, 95% confidence intervals, and standard deviations per electrode and feedback type

	Mean [95% Confidence Interval]	Standard Deviation
Fz positive	.81 [.147, 1.47]	3.22
Fz negative	-2.11 [-3.29, -.94]	5.71
FCz positive	.32 [-.39, 1.04]	3.49
FCz negative	-1.91 [-3.16, -.66]	6.07
Cz positive	-1.03 [-1.76, -.30]	3.55
Cz negative	-1.97 [-3.21, -.73]	6.01
Positive Feedback	.032 [-.639, .702]	3.27
Negative Feedback	-2,00 [-3.19, -.815]	5.77

A significant two-way interaction was found, *F*(2, 186) = 30. 62, *p* < .001, η_p^2 = .248. Pairwise comparisons show that within the Fz electrode that FRN amplitude following positive feedback (*M* = .810, *SD* = 3.21) was significantly smaller or less negative than negative feedback (*M* = -2.11, *SD* = 2.71), *p* < .001. However, on Cz the FRN amplitude for positive feedback (*M* = -1.03, *SD* = 3.55) was not significantly smaller than for negative feedback (*M* = -1.97, *SD* = 6.01), *p* < .114. Finally, FCz positive feedback (*M* = .323, *SD* = 3.50) was significantly smaller than FCz negative feedback (*M* = -1.91, *SD* = 6.06), *p* < .001. See figure 1 for a visual representation of the differences.

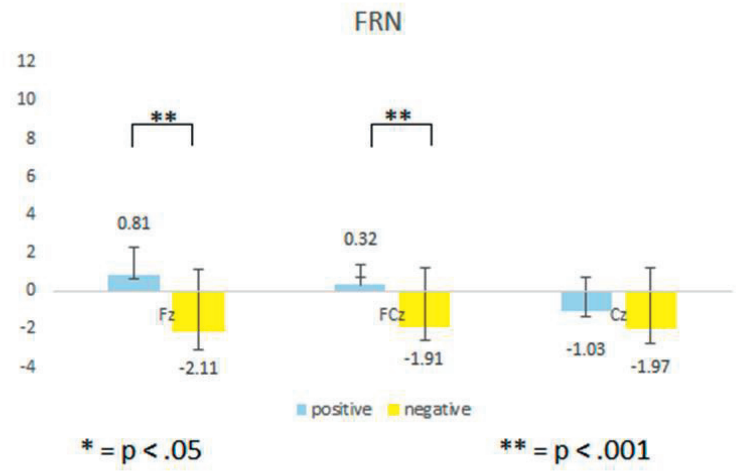


Figure 1 - FRN amplitudes on electrode positions Fz, Cz, and FCz in response to positive and negative feedback.

Error Related Negativity (ERN)

In order to test whether reactive aggression and other self-report questionnaires influenced ERN amplitude, we performed another GLM analyses with ERN amplitude as the dependent variable, and feedback level and electrode positions as within-subject variables, with all self-report questionnaires added as covariates sequentially (RA, TA, and AVL). No main effects were found, and none of the covariates yielded significant two-way interaction effects (all *p*'s > .206), or significant three-way interaction effects (all *p*'s > .967). In other words, neither RA, TA, nor AVL influenced the ERN amplitude.

To address the research question, whether feedback processing as indexed by the ERN is related to the effectiveness of the training, we first calculated HAB change scores by subtracting the HAB scores pre-training from the HAB scores post-training. Then, we did a repeated-measures GLM analysis with ERN amplitude as the dependent variable, and feedback level and electrode positions as within-subject variables, wherein the HAB change scores was inserted as a covariate. No main effect was found of the HAB change scores, $F(1, 92) = 1.24$, $p = .276$, $\eta_p^2 = .013$. In addition, there was no significant two-way interaction between electrodes and HAB change scores, $F(2, 184) = .224$, $p = .799$, $\eta_p^2 = .002$, nor between feedback and HAB change scores, *F*(1,

92) = 3.55, p = .063, η_p^2 = .037. There was also no significant three-way interaction, $F(2, 184)$ =.402, p = .670, η_p^2 = .004.

To assess whether the ERN is sensitive to error during the FT, a repeated-measures ANOVA was performed with ERN amplitude as the dependent variable, and Electrodes (Fz, FCz, Cz) and Error (correct feedback 'CRN', incorrect feedback 'ERN') as the within-subject factors. No main effect was found for electrodes, $F(2, 186)$ = 1.78, p = .170, η_p^2 = .019. A main effect was found of feedback type, $F(1, 93)$ = 31.16, p < .001, η_p^2 = .251, where incorrect response (M = -3.33, SD = 3.69) elicited a significantly larger ERN than correct response (M = -1.32, SD = 1.96), p < .001. See table 3 for the descriptive statistics.

Table 3

The means, 95% confidence intervals, and standard deviations per electrode and feedback type

	Mean [95% Confidence Interval]	Standard Deviation
Fz CRN	-1.17 [-1.56, -.778]	1.88
Fz ERN	-3.25 [-4.05, -2.45]	3.91
FCz CRN	-1.35 [-1.78, -.922]	2.10
FCz ERN	-3.48 [-4.27, -2.69]	3.86
Cz CRN	-1.44 [-1.90, -.995]	2.21
Cz ERN	-3.25 [-3.99, -2.51]	3.61
CRN Feedback	-1.32 [-1.72, -.922]	1.96
ERN Feedback	-3.33 [-4.08, -2.57]	3.69

No significant two-way interaction was found between electrodes and feedback, $F(2, 186) = 1.97$, $p = .141$, $\eta_p^2 = .021$. See figure 2 for a visual of the differences.

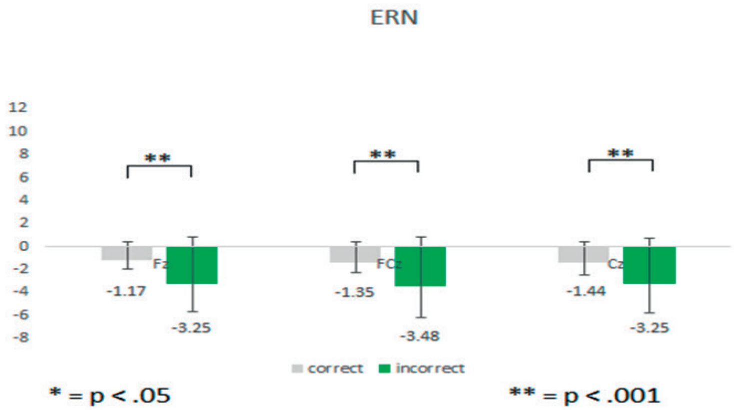


Figure 2 - ERN amplitudes on electrode positions Fz, Cz, and FCz in response to correct and incorrect feedback.

Discussion

The current study examined whether self-report aggression and anger are related to the FRN and ERN during an attribution training for facial expressions, and whether the FRN and ERN are related to changes in HAB from pre-to post-training. As a prerequisite for examining these questions, we first checked whether the training indeed made participants perceive more ambiguous faces as positive, and whether our measures of aggression and HAB were meaningfully related. As expected, we found that training succeeded in increasing positive attributions of ambiguous emotional expressions. In addition to that, the AVL showed a modest but significant relation to lower HAB before the training in which lower self-report scores were associated with classifying more faces as happy. The correlations between HAB before the training and RA and TA were not significant. However, it should be noted that the measure of RA showed low reliability. On the other hand, TA measure represents a slightly different concept than aggression. Although, aggression as a concept is related to anger, they still conceptually distinct for each other (Parrott & Giancola, 2007). In support of the usability of FRN and ERN in the context of CBM-I the FRN amplitudes were larger in response to negative feedback as compared to positive feedback, and ERN amplitudes were larger in response to incorrect responses as compared to the correct responses. Most importantly, the FRN and ERN during the attributions training for facial expressions were not related to self-report aggression and anger, and the FRN and ERN were not related to changes in HAB from pre-to post-training.

The finding that the FT positive training was successful in changing participants' attributions of facial expressions to a more positive one is in line with the previous study of Penton-Voak et al. (2013) that showed that HAB of facial expressions can be trained. Replicating this finding is important because it indicates that the FT procedure that we used provided a meaningful context for exploring error and feedback processing during CBM-I. In further support of the validity of the currently used operationalization of HAB, and in line with previous findings, lower self-report trait aggression scores were associated with classifying more faces as happy (e.g., Tuente et al., 2019). However, self-report trait anger and reactive aggression scores were not significantly related to HAB. This result is in contrast with predictions from cognitive models of anger (Wilkowski & Robinson, 2010) and previous research (e.g., Dill, Anderson, & Deuser, 1997; Epps & Kendall, 1995; Hall & Davidson, 1996). One explanation may be that most of the previous studies compared clinically aggressive individuals to non-aggressive individuals. Given that the current sample consisted of non-clinically aggressive undergraduate students low scores and a restriction of the range may at least partly explain the current lack of relation between our measures reactive aggression and trait anger on the one hand, and our measure of HAB and other outcomes on the other hand. The finding that reactive aggression was not significantly related to HAB scores may also not be so surprising given that this self-report measure is of questionable reliability. Finally, it is possible also that the FT task might only be sensitive enough to measure meaningful individual differences in HAB between (groups of) participants who differ more extremely on aggression measures that is not the case within healthy populations.

Interestingly, there is no association between FRN amplitudes and self-report measures of aggression in our non-clinical group. This result is not in line with previous studies that showed that aggressive individuals appear to have difficulties with learning from previous experiences (Matthys et al., 2012). In fact, the current results indicate that individuals who scored higher on aggression do not show impaired error and feedback processing during this CBM-I procedure. This result is partially in line with

Krämer, Büttner, Roth, and Münte, (2008) who examined participants who performed a modified Taylor Aggression Paradigm against two block-wise alternating fictitious opponents who played either fairly (low provocation) or unfairly (high provocation). In this study, they found that FRN was larger for participants with high trait aggression, as compared to low trait aggressive participants, but only in the provocation block. In the low provocation block no effects of trait aggression were found. In fact, most of the previous studies that found the effect of aggression on FRN used participants who scored higher on aggression measures as compared to the current study or used a clinical sample (Krämer et al., 2008; Krämer, Munte, Richter, & Kopyciok, 2009; Wiswede et al., 2011). Moreover, all of these studies provoked participants in laboratories (Bertsch, Bohnke, Kruk, & Naumann, 2009; Krämer, Jansma, Tempelmann, & Münte, 2007; Lotze, Veit, Anders, & Birbaumer, 2007). So, while the current findings are encouraging with respect to the efficiency of CBM procedures in the context of aggression, the current study results need to be replicated in a clinical population or a population with higher aggression scores, as well as in the context of a provocation manipulation, in order to make sure that the effectiveness of a CBM-I training in the context of aggression is truly unaffected by the increasing levels of aggression. Finally, the finding that FRN amplitudes were larger for negative compared to positive feedback is consistent with many previous studies (e.g., Miltner et al., 1997; Nieuwenhuis et al., 2004), suggesting that FRN may serve as a trait or state variable indicator associated with whether participants are more receptive to positive or negative reinforcement. In addition, these study results support the hypothesis that FRN reflects common neural processes.

The current study also did not find a significant association between ERN amplitude and self-report measures of aggression, which again seems to support the usefulness of CBM-I training procedures in the context of aggression. However, this result must also be taken with caution since most of the studies that examined the association between ERN and aggression are still not fully conclusive either. In fact, a few of these studies showed diminished error monitoring in psychopaths or individuals with antisocial personality disorder, and others found no differences (Kiehl, Smith, Hare, & Liddle, 2000; Munro et al., 2007; Brazil et al., 2011; Von Borries et al., 2010). Some of these studies, for instance, point to altered error monitoring in an affective context (Munro et al., 2007) or in learning tasks (Von Borries et al., 2010). For example, Munro et al. (2007) results showed reduced ERN amplitudes in violent offenders as compared to a control group, following errors in a flanker task with emotional faces but not with neutral letter stimuli. This apparently shows the emotional context is important when studying the association between ERN and aggression. In addition to that, Von Borries et al. 2010 report reduced ERN along with impaired learning and diminished accuracy in a probabilistic learning task within the psychopathology group individuals as compared to the control group. It might be that aggression related HAB assessed with the FT task is not very sensitive to relatively small individual differences in self-report measures of aggression. In fact, the current sample means scores in reactive aggression were (*M*= 7.72 *SD* =3.06) which is below average according to the non-offender adult norm scores in (Cima et al., 2013) indicates the relatively little variation in the current sample. Therefore, replication studies of the FT paradigm should compare more extreme groups, or include a bigger range of aggression scores in order to find the significant association between ERN amplitude and self-report measures of aggression. An alternative view that could explain this finding is the possibility that HAB may be influenced by a lot more than aggression, and therefore researchers need extreme differences in self-report aggression before the association becomes reliably visible over the "noise" that is caused by all the other variables influencing HAB. On the other hand, the current study found larger ERN amplitudes in response to incorrect as compared to correct responses are consistent with many previous findings who established the sensitivity of the ERN to errors (e.g., Falkenstein et al., 1991; Gehring et al., 1993).

Finally, the current study did not find the expected association between FRN and ERN and changes in HAB from pre-to post-training. These results can be explained by the fact that within the current study, all participant's HAB scores were very close to each other. In fact, these scores were not very negative to begin with. Therefore, not too much variation in changes over training can be seen due to the restriction of the range. An alternative view that could explain this finding is the design of the FT procedure. That is, the current study training task was built in a fashion that the participants would only be likely to make errors in response to ambiguous faces. This procedure resulted in participants making low numbers of errors. In fact, participants in the current study had a mean of 20.76 incorrect trials out of (180 trials). It is important to note that previous studies exist that find stable, grand-average ERPs with 10 or fewer trials for the ERN (see Pontifex et al., 2010; Steele et al., 2016). Although potentially true, it presupposes that these studies' results were only based on calculating the ERN stability. For instance, within these studies, ERN was calculated as a high correlation between ERPs averaged over a small number of trials, and ERPs averaged over a large number of trials. Nevertheless, the ability of standard analyses to find between-condition or between-group differences can be lost with small numbers of trials (Gehring et al., 2012). Other studies have indeed found that the ability to detect a between-group difference on the ERN is a function of the number of trials included in the ERP average and the error rate across groups (Duncan et al., 2009; Fischer et al., 2017).

Limitations and implications for future research

Besides the strengths of a relatively large sample size and finding the expected differences between positive and negative feedback on FRN and between correct and incorrect responses on ERN, this study has some limitations. First, as mentioned before, the chosen sample of healthy undergraduate students might be suboptimal to study the effects of aggression on the FRN and ERN signal. Future studies may benefit from making use of paradigms such as the Taylor Aggression Paradigm (Taylor, 1967) instead of self-report measures. For instance, a paradigm that is designed to frustrate and to promote reactively aggressive behavior in the lab could give a clearer indication of how FRN and ERN are associated with measures of aggression.

Conclusion

To conclude, as expected, we found that training succeeded in increasing positive attributions of ambiguous emotional expressions. In addition, the AVL showed a modest relation to lower HAB before the training. However, the correlations between HAB before the training and RA and TA were not significant. More importantly, the FRN and ERN during the attributions training for facial expressions (the FT) were not related to self-report aggression and anger measures, indicating that individuals scoring higher on anger or aggression do not show impaired error and feedback processing during this CBM-I procedure. Along with that, the FRN and ERN were also not related to changes in HAB from pre-to post-training. Finally, in line with previous findings, the FRN amplitudes were larger in response to negative feedback as compared to positive feedback, and ERN amplitudes were larger in response to incorrect responses as compared to the correct responses. Our results suggest that the FRN and ERN amplitudes were not sensitive to the effectiveness of this CBM-I procedure in a sample of individuals from the healthy population scoring low on self-report aggression. Finally, the currently used approach might be better suitable for clinical populations and probably with provocation manipulation.

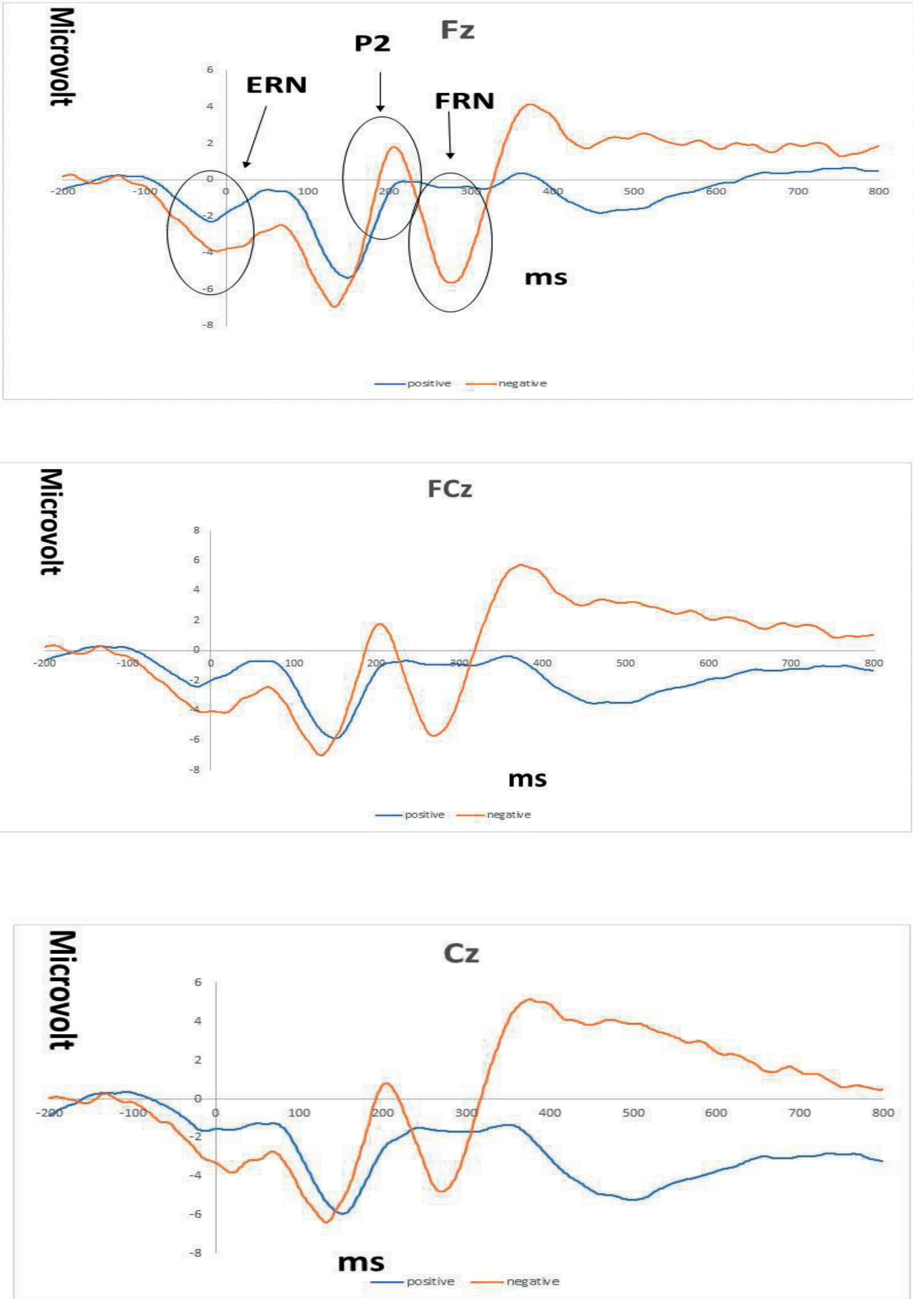
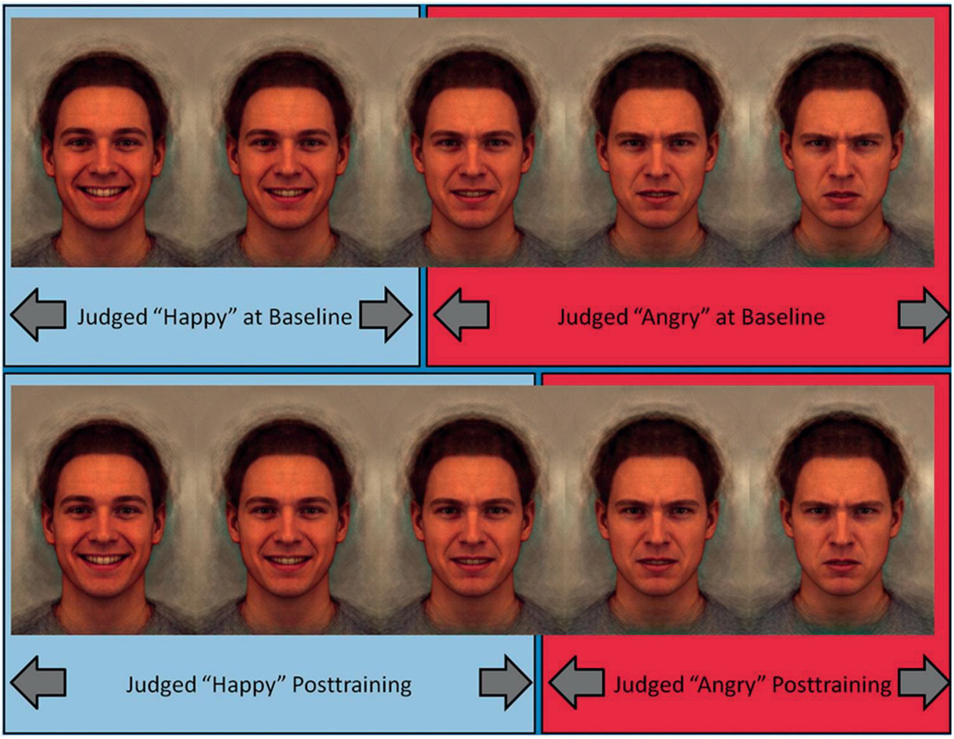


Figure. 3. Grand-average ERPs on electrode positions Fz, FCz, and Cz in response to negative and positive feedback.



Appendix 1 an illustration adapted from Penton-voak et al. (2013), shows how the morph sequence of images changes step-by-step from unambiguously happy to unambiguously angry, with neutral images in the middle.

CHAPTER 6

Summary and Discussion

Introduction

The major aim of this dissertation was to use ERPs to examine aggression in healthy male adults to shed light on the early stages of information processing to identify the factors (i.e., feedback) that mediate changes in information processing among this sample. The first study examined whether the neuronal reflection of Feedback Related Negativity (FRN) processing was associated with self-report measures of aggression. In addition to this, the first study aimed to provide an assessment of the validity of the island getaway task (IGT) paradigm to determine the neural correlates of Hostile Attribution Bias (HAB) in aggression. The second study examined whether the N400 can be a useful measure of emotions in the context of aggression by examining whether the negative emotional content of a word can elicit a larger N400 effect compared to the neutral emotional content of a word. In addition to this, the second study aimed to provide an assessment of the validity of the homonyms task. After investigating the role of an emotional versus non-emotional context, the current dissertation concluded that the N400 can be a useful measure of emotions in the context of aggression. Therefore, the third study examined whether higher scores of self-report aggression measures were associated with greater Aggression-Related Interpretation Bias (ARIB). The fourth and final study examined whether self-report aggression and anger were related to the FRN and Error Related Negativity (ERN) during an attribution training for facial expressions; to determine whether the FRN and/or ERN are related to changes in HAB from pre- to post-training; and to provide an assessment of the validity of the Face task (FT) paradigm to determine the neural correlates of HAB in aggression.

The current dissertation aimed to replicate previous studies and test the associations between HAB and aggression. HAB is defined as a tendency to attribute the (ambiguous) behavior of others to harmful aggressive intent (Milich and Dodge, 1984). By now, it is well established that aggressive individuals show HAB (Arsenio, Adams, & Gold, 2009; Brugman et al., 2014; De Castro, Veerman, Koops, Bosch, & Monshouwer, 2002; De Castro, Merk, Koops, Veerman, & Bosch, 2005; Dodge & Pettit, 2003; Dodge et al., 2015; Hubbard, McAuliffe, Morrow, & Romano, 2010). The fact that HAB is strongly linked to aggressive behavior, (Tuente, Bogaerts, & Veling, 2019; Verhoef, Alsem, Verhulp, & De Castro, 2019) has led to the development of interventions to prevent and reduce aggressive behavior by targeting social information processing. Among these interventions is Cognitive Bias Modification (CBM) training. CBM is a computer-based treatment that trains participants to modify their automatic, cognitive biases. Studies focusing on mental health in general have revealed that CBM procedures are moderately effective in reducing cognitive bias (Cristea, Mogoașe, David, & Cuijpers, 2015; Hallion & Ruscio, 2011; Krebs et al., 2017). Although encouraging, this does suggest that there is still room for improvement and refinement. One reason for which current CBM interventions are moderately effective could that researchers are not using the most optimal CBM procedure. For instance, instead of reducing negative cognitive biases in participants, CBM procedures may simply train them to avoid negative stimuli (Cisler & Koster, 2010; Koster, Baert, Bockstaele, & De Raedt, 2010). This raises the question of which factors mediate changes in information processing. One such factor which can be examined by ERPs, is the way in which participants learn from feedback during CBM training. Therefore, functional neuroimaging methods, such as electroencephalogram (EEG), may be useful in determining the underlying neural processes associated with HAB. To the best of my knowledge, only a small number of studies have used EEG to investigate neuro-cognitive processes involved with HAB (e.g., Gagnon et al., 2016; Gagnon et al., 2017; Godleski, Ostrov, Houston, & Schlienz, 2010; Moser, Hajcak, Huppert, Foa, & Simons, 2008).

Summary of the current dissertation and general discussion

Chapter 2 Can We Get Along? The Relationship between Feedback-Related Negativity and Reactive Aggression in Healthy Individuals During Social Feedback Processing

Previous studies have shown that reactive aggressive individuals are characterized by deviant social feedback processing (e.g., Matthys, Vanderschuren, Schutter, & Lochman, 2012) which may hinder their ability to learn from social feedback and adjust their behavior in a more pro-social manner. Previous ERP studies have also shown that ERN and FRN are directly linked to feedback processing in the anterior cingulate cortex (ACC; Holroyd & Coles, 2002; Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Miltner, et al., 1997; Ullsperger & Von Cramon, 2003). The ACC brain area also has been linked to associative learning (Rushworth, Behrens, Rudebeck, & Walton, 2007) and performance monitoring (e.g., Holroyd & Coles, 2002), making ERN and FRN components perfect candidates for examining learning deficits in aggressive individuals. This chapter examined whether the neuronal reflection of social feedback processing (FRN) was associated with self-report measures of aggression. Specifically, we examined the relationship between FRN and aggression in the context of social feedback by examining whether higher scores of self-report aggression measures were associated with smaller FRN during a social feedback task (IGT) in a sample of healthy male students. The Island Getaway task (IGT) is a computerized task inspired by the game show survivor and consists of both a social feedback and a peer rejection element (Kujawa et al., 2014).

The results of this study showed that FRN amplitudes did not differ between negative feedback and positive feedback. Moreover, there was no association between the FRN amplitudes and self-report measures of aggression. Finally, participants' voting behavior was not associated with reactive aggression or FRN. Surprisingly, in this study, there was an absence of differences in FRN amplitudes for negative and positive feedback, whereas several other studies have found such an effect. For example, previous studies that used social feedback tasks showed that FRN becomes more negative in response to rejection than acceptance (e.g., Kujawa, Arfer, Klein, & Proudfit, 2014; Sun & Yu, 2014). However, the results of other studies testing the validity of the IGT (Babinski, Kujawa, Kessel, Arfer, & Klein, 2019; Cao et al., 2015; Ethridge et al., 2017; Kujawa et al., 2014; Kujawa, Kessel, Carroll, Arfer, & Klein, 2017; Pegg et al., 2019) are not fully conclusive. For example, larger FRN amplitudes for negative as compared to positive feedback were found by Kujawa et al. (2014) and others, whereas Cao et al. (2015) found larger negative FRN amplitudes for positive as compared to negative social feedback. On the other hand, we had different findings in which FRN amplitudes were not different for negative feedback as compared to positive feedback. These results can be explained by the use of symbolic feedback in the IGT task that the current study employed. For instance, within the IGT version that we employed, a colored rectangle was shown to indicate acceptance/rejection feedback, and to avoid physical differences between the feedback signals. However, recent studies that have used the IGT have employed more complex feedback, in which a green thumbs-up appears to indicate acceptance feedback and a red thumbs down appears to indicate rejection feedback (Babinski et al., 2019; Cao et al., 2015; Ethridge et al., 2017; Kujawa et al., 2014; Kujawa et al., 2017; Pegg et al., 2019). All these studies consistently showed a more positive ERP to positive feedback within the FRN time window. Therefore, combining our results, in which no difference between positive and negative feedback could be found with symbolic feedback, we concluded that the different physical appearance of feedback might affect the amplitude of FRN. This is in line with the finding that FRN seems to be larger after social compared to non-social stimuli as well as larger after complex positive compared to non-complex positive stimuli (Pfabigan, Gittenberger, & Lamm, 2019). Therefore, it can be concluded that FRN results from the IGT are dependent on the physical appearance of the feedback stimulus and that the use of neutral symbolic feedback seems to lead to similar FRN amplitudes for positive and negative feedback. This has implications for the interpretation of the FRN findings of the previous studies.

The result that there was no association between FRN amplitudes and self-report measures of reactive aggression trait anger, as well as trait aggression, will be discussed below along with all of the other four experimental questions about the association between the EEG measures (i.e., N400, ERN, and FRN) and the self-report measures of aggression. Finally, the current study also found that participants' voting behavior was not associated with reactive aggression or FRN amplitudes. However, it should be noted that the participants in the study described in this chapter scored average in terms of their self-report measures of aggression and anger compared to non-offender adults in the study by (Cima et al., 2013). Therefore, the current chapter concluded that the relationship between voting behavior and aggression possibly only becomes visible among participants who score high on aggression.

Finally, the current version of the IGT is not sensitive enough to find associations in a healthy population, and the conclusion that can be drawn from this study is that the brain responses in the IGT, used as a tool to test behavioral and brain responses to social feedback in an adult healthy male population, depend on the physical appearance of the feedback stimuli.

Chapter 3 Emotional modulation of the N400: Manipulating the emotional meaning of homonyms

Previous ERP studies have shown that the N400 elicited by semantic aspects of words could possibly be modulated by the emotional context of the words (Atchley & Kwasny, 2003; Chwilla & Kolk, 2003; Delaney-Busch & Kuperberg, 2013; Herbert, Junghofer, & Kissler, 2008; Holt, Lynn, & Kuperberg, 2009; Klepousniotou, Pike, Steinhauer, & Gracco, 2012; Kotchoubey & El-Khoury, 2014; Lee & Federmeier, 2006; Meyer & Federmeier, 2007; Pylkkanen, Llinas, & Murphy, 2006; Titone & Salisbury, 2004). However, it should be mentioned that not all previous studies testing the importance of the emotional context on the N400 point into the same direction, where some of these studies point to larger brain N400 amplitude of the dominant as compared to the subordinate word meaning (Titone & Salisbury, 2004; Klepousniotou et al., 2012), whereas others point to smaller N400 amplitude for pleasant than unpleasant words (Herbert et al., 2008). Nevertheless, these studies concluded the N400 amplitude seems to be modulated by both semantic and emotional content. Although there is evidence that context is relevant to the processing of ambiguous words as reflected in the N400, the role of an emotional versus non-emotional context has not yet been investigated. In addition to this, previous ERP studies examining emotional processing have tended to use stimuli that physically differ from each other. For example, the critical words used for the emotional and neutral conditions in some of these studies differed quite clearly from each other (Sugar/dog; Kutas & Hillyard, 1980; Snake/diamond/bouton Delaney-Busch & Kuperbug, 2013). Therefore, the results of previous studies cannot rule out with certainty whether their ERP results were affected by the physical appearance of a word instead of only by the emotional content itself. Examining the role of an emotional versus non-emotional context (under stringent conditions; where all the stimuli physical appearances are identical to each other) can reveal whether the N400 can be used as a useful measure in other contexts (i.e., an aggression context). Therefore, this study examined whether the negative emotional content of a word can elicit a larger N400 effect. Specifically, it examined whether homonym words with either no emotional meaning or a negative emotional meaning differed with respect to the N400 response. The homonyms used in this study were identical to the neutral words in form and articulation and only differed in semantic content, which we made dependent on the context in which the word was presented. This study also examined whether affect (positive or negative) and/or the anxiety level (trait or state) of the participant had an effect on the amplitude of the N400.

In this study, we found a significant effect of emotion, suggesting that the negative emotional content of a word elicits a larger N400 effect, as compared to the neutral emotional content of words. This result replicates previous findings by showing that the negative emotional content of a word can elicit a larger N400 effect (Delaney-Busch & Kuperberg, 2013; Holt et al., 2009). However, this study did not find correlations between affect and anxiety as well as the N400 amplitudes. It is known that most previous studies that have managed to find the N400 association with anxiety used samples who were clearly at the clinical level of anxiety (Chwilla, Virgillito, & Vissers, 2011: Yu et al. 2018). These studies also used anxiety induction procedures to increase their participants' anxiety levels. Inducing anxiety can be done through employing an experimental procedure between learning and testing phases or by using a Mood Induction Procedure (MIP), in which for instance movie clips are used to manipulate participants' moods (Westermann, Spies, Stahl, Gün, & Hesse, 1996). Concluding from this, the results of the current study are not surprising, because we did use any MIP in order to increase our sample's anxiety levels. Moreover, our sample was limited to adult students without a history of anxiety and who had been recruited at the university. Thus, future studies should include populations with more extreme anxiety scores, and to use MIP in order to increase their chances of finding the N400 association with anxiety. The main conclusion from this study is that N400 amplitude is affected by the emotional context of the word itself and not by the differences in their physical appearance, a factor earlier studies could not rule out. In addition, this study also showed the validity of the homonyms task and its potential to be used in emotional processing research in general. Finally, it must be mentioned that the current dissertation employed a different task in the subsequent study, in which we utilized more stimuli or words that are clearly associated in aggression context in order to examine the N400 association with self-report aggression measures.

Chapter 4 No Effect of Self-report Aggression Measures on The N400 in A lexical Decision Task with Associated Words in a Violent Context.

A lot of studies that focused on the relationship between aggression-related interpretational bias (ARIB) and aggression subtypes (reactive and proactive) studied children (Arsenio, Adams, & Gold, 2009; Crick & Dodge, 1996; De Castro, Merk, Koops, Veerman, & Bosch, 2005; Dodge & Coie, 1987; Dodge et al., 2015; Verhoef, Alsem, Verhulp, & De Castro, 2019). In addition, most adult studies have only focused on examining ARIB and aggression in general without making the distinction between the reactive and proactive subtypes (Coccaro, Noblett, & McCloskey, 2009; Gagnon, McDuff, Daelman, & Fournier, 2015; Helfritz-Sinville & Stanford, 2014; Matthews & Norris, 2002). One common way to study the association between ARIB and aggression is to use vignettes as stimuli. Interestingly, while studies using this method have quite consistently reported associations between ARIB and aggression in children, adult studies, have only reported a small to medium associations (Tuente et al., 2019), or found no significant relation at all (Helfritz-Sinville & Stanford, 2014; Miller & Lynam, 2006). The use of vignettes as stimuli might only reveal participants' explicit biases (Dovidio, Kawakami, & Gaertner, 2002), specifically within children's studies (Lansu, Cillessen, & Bukowski, 2013; Lansu, Cillessen,

& Karremans, 2012). Due to this fact, researchers still questioning whether the use of vignettes as stimuli is ideal to measures children's cognitive biases (see, e.g., Cillessen & Bellmore, 2010). Moreover, adult studies that used ERP's, specifically the N400 component, only employed some sort of vignette as stimuli (Gagnon et al., 2016; Gagnon, et al., 2017), in order to examine ARIB association to aggression in healthy samples. It is possible that the association between ARIB and aggression might be better studied with implicit measures that can manifest ARIB spontaneously without the participant controlling such biases (Greenwald & Banaji, 1995). We found only one study that used such an implicit measure of ARIB in aggression research (Cima, Vancleef, Lobbestael, Meesters, & Korebrits, 2014), which examined the behavioral measures of male aggressive adults by using the Aggressive Interpretation Task (AIT), without focusing on the electro-cortical reflections of the ARIB. Therefore, the current study examined healthy undergraduate populations with implicit measures of bias using the AIT. The AIT is a lexical decision task in which the first word of a pair is presented and participants are asked to decide as quickly as possible (750 ms) whether the second word (target) in a word pair is an existing Dutch word or a non-word. This study divided the AIT into six types of word pairs: (1) ambiguous prime violence and target violence associated; for example, the prime (Gas) can be presented as an ambiguous violent prime that can be followed with the target (Explosion), which is also associated in the violence context; (2) ambiguous prime violence and target neutral associated; for example, the same prime (Gas) can be presented as an ambiguous violent prime that can be followed with the target (speed), which is a target associated in the neutral context; (3) ambiguous prime violence and target neutral un-associated; (4) unambiguous prime neutral and target neutral associated; (5) unambiguous prime neutral and target neutral un-associated;(6) unambiguous prime neutral and target non-word. Specifically, we examined whether higher scores of self-report aggression measures were associated with greater ARIB in a sample of healthy male undergraduate students. We found that associated words led to faster reaction times (RT) compared to non-words. Non-words were also associated with a larger N400 amplitude compared to words. Target type words associated in a violent context did not influence reaction time and N400 measures. Moreover, there was no association between reaction times, N400, and self-report measures of aggression. These results can be seen as a replication and extension of the findings of Cima and colleagues, who also could not find a relationship between self-report aggression measures and ARIB among individuals with similarly low aggression scores (Cima, Vancleef, Lobbestael, Meesters, & Korebrits, 2014). On the other hand, we did find a significantly larger N400 for non-words compared to words for all electrodes. However, we only found a significantly higher N400 for un-associated compared to associated words on the Fz electrode. This result is also a replication of previous findings that revealed the N400 elicited by a single word seems to have a greater anterior distribution with a maximum over frontal or central sites (Bentin, McCarthy, & Wood, 1985; Bentin, 1987; McCarthy & Nobre, 1993). Importantly, the N400 in this study was not associated with self-report measures of aggression and anger. This result will be discussed in more detail below.

The main conclusion that can be drawn from this study is that the AIT is only able to find the associations between ARIB in terms of RT, N400, and self-report aggression measures in a forensic population and might not be sensitive enough to reliably detect more subtle associations that may be present in healthy populations. Therefore, further studies are needed to be carried out among highly aggressive (e.g., forensic) participants in order to understand the relationship between aggression and ARIB in more detail.

Chapter 5 Cognitive Behavioral Modification to Modulating Negative and Positive Attributions in Individuals scoring High on Reactive Aggression Measures

Previous studies examining the effect of CBM on mental health in general revealed that CBM procedures are only moderately effective in reducing cognitive bias (Cristea et al. 2015; Hallion & Ruscio, 2011; Krebs et al. 2017) suggesting there is room for improvement. One possible reason for the moderate effects could be that these studies did not yet use the most optimal CBM training procedure. For instance, instead of reducing negative cognitive biases in participants, CBM procedures may simply train them to avoid negative stimuli (Cisler & Koster, 2010; Koster, Baert, Bockstaele, & De Raedt, 2010). However, there are different factors that can influence the effectiveness of CBM training. For instance, learning from errors and feedback is an important part of CBM training. Examining how error and feedback processing is related to the effectiveness of CBM training, might provide important clues about how to improve CBM procedures. Along with this, it is important to note that reactively aggressive individuals appear to be characterized by a failure to learn from previous experience (e.g., Matthys et al., 2012), which possibly suggests that CBM training could be less effective in high versus low aggressive individuals and may also have implications for the design of CBM studies in the context of aggression. Previous studies have also revealed that FRN reflects the internalization of external feedback, while ERN reflects internal monitoring (Eppinger, Kray, Mock, & Mecklinger, 2008; Holroyd & Coles, 2002; Pietschmann, Simon, Endrass & Kathmann, 2008; Pietschmann, Endrass, Czerwon, & Kathmann, 2011). This could help examine feedback processing on the behavioral and neuronal levels during the completion of CBM training. Therefore, this study examined whether self-report measures of aggression and anger were related to FRN and ERN during an attribution training for emotional facial expressions, as well as whether FRN and ERN were related to changes in HAB from pre-to post-training.

The results of this study show that the training was successful in increasing the positive attributions of ambiguous emotional facial expressions. This result is in line with those of other studies (AlMoghrabi, Huijding, & Franken, 2018; Hawkins & Cougle, 2013; Penton-Voak et al., 2013; Vassilopoulos, Brouzos, & Andreou, 2015), which also showed that HAB of facial expressions can be trained. This result also suggests that the FT procedure we used provides a meaningful context for exploring error and feedback processing during CBM-I training. Furthermore, our result that lower self-report trait aggression scores were associated with classifying more faces as happy was also in support of the validity of the used operationalization of HAB and in line with previous findings (e.g., Tuente et al., 2019). However, the correlations between HAB before the training and reactive aggression and trait anger were not significant which is in contrast with the cognitive models of anger (Wilkowski & Robinson, 2010) and previous research (e.g., Dill, Anderson, Anderson, & Deuser, 1997; Epps & Kendall, 1995; Hall & Davidson, 1996), which has proposed that participants with high trait anger have a cognitive processing bias that makes them more likely to attribute ambiguous situations as having hostile intents. However, these studies only compared clinically aggressive individuals to non-aggressive individuals. In fact, it was only in studies using criminal and psychopathic individuals as a sample where it was found that HAB manifests itself in the perception of emotionally ambiguous faces. For example, in a recent study Smeijers and colleagues (2017) that tested 142 participants who had been recruited from clinics of forensic psychiatry, it was found that HAB appeared to be a characteristic of the pathological aggression of these samples.

Concerning electrophysiological responses, this study found the expected larger FRN in response to negative feedback when compared to positive feedback (Miltner, Braun, & Coles, 1997; Nieuwenhuis, Holroyd, Mol & Coles, 2004; Yeung & Sanfey, 2004). In addition, ERN amplitudes were larger for incorrect trials than for correct trials (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993; Gehring, Liu, Orr, & Carp, 2012). However, both the FRN and ERN amplitudes were not associated with self-report measures of aggression, indicating that individuals scoring higher on anger or aggression did not show impaired error and feedback processing during this CBM-I procedure. This result must be considered with caution, since most of the studies that have examined the association between ERN and FRN as well as aggression are also not fully conclusive. For example, most of the previous studies that found the effect of aggression on FRN provoked their participants in laboratories (Bertsch, Bohnke, Kruk, & Naumann, 2009; Krämer, Jansma, Tempelmann, & Münte, 2007; Lotze, Veit, Anders, & Birbaumer, 2007). Provocation situations in laboratories can occur in the form of verbal aggression or physical aggression (Anderson & Bushman, 2002). ERN studies, on the other hand, revealed a diminished error monitoring in psychopaths or individuals with antisocial personality disorder, and other studies found no differences (Kiehl, Smith, Hare, & Liddle, 2000; Munro et al., 2007; Brazil et al., 2011; Von Borries et al., 2010). In addition, the sample that was used in this chapter only consisted of non-clinically aggressive undergraduate students who scored relatively low on self-report measures of aggression and anger. Therefore, further studies need to replicate the current study findings among a clinical population or a population with higher aggression scores in order to clearly make sure that the effectiveness of a CBM-I training in the context of aggression is truly unaffected by the increasing levels of aggression. This result will also be discussed along with the results from Chapter 4 below.

Finally, this study aimed to examine whether self-report aggression and anger were related to FRN and ERN during an attribution training for facial expressions. Furthermore, it aimed to determine whether the FRN and/or ERN were related to changes in HAB from pre-to post-training and finally to provide an assessment of the validity of the FT paradigm to determine the neural correlates of HAB in aggression. The main conclusion that could be drawn from this study was that the FRN and ERN amplitudes were not sensitive enough to the effectiveness of this CBM-I procedure for the sample of individuals from the healthy population scoring low on self-report aggression. Additionally, the approach the current study used might benefit from using clinical populations and probably a provocation manipulation. As for the validity of the FT paradigm, this study concluded that it might only be sensitive enough to measure meaningful individual differences in HAB between (groups of) participants who differ more extremely on aggression measures, which is not the case within healthy populations.

Overall, one of the most major conclusions of the current dissertation was that we were not very successful in finding an association between EEG, performance measures (i.e., N400, ERN, and FRN), and self-report measures of aggression in diverse populations of healthy subjects. Although it is always difficult to explain null findings, below I will discuss some general points that could explain those found throughout the studies.

The use of healthy subjects to study aggressive behavior

As discussed in the introduction, many authors have argued that studying the normal range of a given behavior such as aggression is necessary to better understand its extreme cases (Anderson, 2012). For example, within the psychological literature of psychopathology, much of the existing findings are based on research within healthy samples. In fact, psychopathology studies of depression have shown that healthy and clinical samples are equivalent to each other in terms of their findings (Vredenburg, Flett, & Krames, 1993). Also, the analogue studies for obsessive-compulsive disorders (OCD) seem to give the same results. For instance, from reviewing the available research on OCD, a recent study concluded that analogue samples are highly relevant in understanding OCD symptoms (Abramowitz, et al., 2014). These findings are not new per se in fact, two decades ago, Gibbs (1996) also found that studying healthy samples is relevant to understand psychopathology, in this case OCD symptoms.

Like most, if not all psychopathological symptoms and behavior in general, aggressive behavior is normally distributed as a continuum within healthy samples (Anderson & Huesmann, 2007; Bowins, 2016). Accordingly, researchers agree that the study of aggression using non-clinical samples is highly relevant because it gives insight into the reasons behind the aggressive behavior's persistence, and etiology. However, the results of the current dissertation did not show any associations between ERP components and self-report measures of aggression. Given the preliminary evidence for an association between HAB and aggression as well as the notion that aggressive behavior is distributed as a continuum within healthy samples, our results were somewhat unexpected. However, most studies reporting the association between HAB and aggression only examined participants with higher clinical levels of aggression (e.g., Bowen, Roberts, & Kocian, 2016; Coccaro, Noblett, & McCloskey, 2009; Edwards & Bond, 2012; Gagnon, McDuff, Daelman, & Fournier, 2015; Gagnon et al., 2016; Gagnon et al., 2017; Lobbestael, Cima, & Arntz, 2013; Matthews & Norris, 2002; Neumann, Malec, & Hammond, 2015; Pert, Jahoda, & Squire, 1999; Tremblay & Belchevski, 2004). In combination with previous studies on the associations between HAB, and aggression, the findings in the current dissertation suggest that aggressive behavior as a form of psychopathology might be an exception regarding the extent to which similar processes underly the construct of interest in healthy and clinical samples, at least with respect to HAB. That is, only within highly aggressive populations in which aggressive behavior is above the clinically significant levels can HAB can be found. Therefore, we would like to suggest that future studies include populations in which (participants) aggressive behavior is above clinically significant levels. It is possible that the low aggression scores the participants scored in this dissertation were the main reason behind our not finding association between the EEG measures (i.e., N400, ERN, and FRN) and self-report measures of aggression.

To provoke or not to provoke

Since it remains unclear whether the reaction to provocation is universal across different subtypes of aggression (i.e., reactive vs proactive) and whether healthy individuals differ from each other in terms of their reactivity to such signals (Weidler et al., 2019), the current dissertation did not use aggression provocation techniques. We were interested in identifying the factors (i.e., feedback) that mediate changes in information processing among healthy male adults. However, it must be noted that one possible explanation for the fact that we did not find the expected associations between the used ERP and performance measures as well as the self-report measurers of aggression could be that HAB for aggressive individuals possibly can only be examined inside labs when specific aggression provocation techniques are used. In fact, within ERP lab studies, the association between HAB and aggression has been found consistently in those using aggression provocation techniques (Bertsch, Bohnke, Kruk, & Naumann, 2009; Krämer, Jansma, Tempelmann, & Münte, 2007; Krämer, Büttner, Roth, & Münte, 2008; Krämer, Munte, Richter, & Kopyciok, 2009; Lotze, Veit, Anders, & Birbaumer, 2007; Wiswede et al., 2011). It has been known that provocation situations activate a heightened internal state in which the immediate judgment of the participant is to interpret the behavior of the target as hostile. After provocation, participants are more likely to demonstrate aggressive behavior. However, it has to be mentioned that this effect can only be achieved, if the provocation is sufficiently strong (Bettencourt & Miller, 1996). Based on the current findings, it is recommended to include provocation situations in future studies, especially when not using very aggressive individuals. For example, provocation situations in experiments can include mild insults, slights, and other forms of verbal or physical aggression and interference with one's attempts to attain an important goal (Anderson & Bushman, 2002). For instance, in one fMRI study, an anagram task was presented and the participants were instructed to loudly declare their answers even if they did not know the correct answer. After the participants started answering the task, the experimenter politely asked the participants (two times) to speak louder. On the third time, the experimenter started to insult the participants by saying in an annoyed voice, "Look, this is the third time I've had to say this! Can't you follow directions?" The results, relative to a baseline, showed increased activation in many brain regions that had been active during the exposure of these participants to angry faces and during these participants' recall of anger-inducing experiences (Cohen et al., 2001). In addition, laboratory studies have demonstrated that believing that another person's actions are intended to be provocative can result in aggression (Epstein & Taylor, 1967; Greenwell & Dengerink, 1973; Ohbuchi & Kambara, 1985). In fact, aggressive individuals have been found to react to provocation with greater aggression compared to nonaggressive individuals within controlled laboratory situations (Berman, McCloskey, Fnning, Schumacher, & Coccaro, 2009). These results combined with our own could also lead to the hypothesis that the association between HAB and aggression can only be found using provocation paradigms. However, there are some obvious ethical considerations that prevent researchers from using aggressive provocation techniques. For example, suppose that a participant receives a severe insult from the experimenter intended to influence his/her brain activity during the experiment. Provoking techniques that include severe insults, punching, and kicking are not acceptable from the experimenter. Therefore, aggression researchers have developed a repertoire of tasks that purportedly measure aggression. These tasks are safe for participants and researchers, in addition to the fact that they are ethically and legally acceptable within the confines of a laboratory setting (McCarthy & Elson, 2018).

Strengths and Limitations of the current dissertation

The methodological shortcomings specific to each of the four studies are discussed in their corresponding discussion chapters. This section will provide an opportunity to explain the circumstances surrounding some of these issues and reflect on key decisions made during the course of the dissertation. One major strength of the current dissertation was the replication of previous ERP studies in every experiment except in the IGT task experiment. This strength was due to the design of the current dissertation's tasks. For example, within these tasks, we included elements that made them much more realistic and engaging for participants

The Dutch emotional words mentioned in Chapter 3 were identical to the neutral words in form and articulation, since we used homonyms. These words only differed in semantic content, which we made dependent on the context in which the word was presented. The use of homonyms in this study was an important advantage over other studies which have tended to use stimuli physically different from each other. The current use of homonyms meant that any differences found could only come from the semantic content of the word. Future studies could benefit from the use of this task (after including relatively larger numbers of stimuli per condition) in emotional processing research in general, as well as in HAB research more specifically. However, researchers should realize that using homonyms might not be possible in all languages.

It is worth acknowledging that the AIT task in the current dissertation was tested with EEG for the first time. Unfortunately, the AIT study was not sufficiently sensitive to measure ARIB in healthy individuals who scored low in the self-report measures of aggression. Therefore, further research needs to be carried out among highly aggressive participants (e.g., forensic) or in the context of provocation, in order to examine whether this task is valid in measuring the relationship between aggression and ARIB. Finally, future studies could also benefit from the use of facial expression stimuli instead of using words. That is because faces are recognized based on features and the structure of these features whereas recognition of words is dependent on understanding the letters that make up these words, and the structure of these letters (Ventura, 2014). For example, across cultures, people recognize common facial features, such as eyes, lips, and cheeks, but, in order to facilitate emotion recognition, they must process the full face shape (Ventura, 2014). Indeed, faces provide an important social cue and are therefore fundamental in human interaction (Argyle, 1994) as they are more universally recognized than words.

Future directions

First, it is important to mention that it is very difficult to mimic real-life aggression within laboratory settings since this behavior is a result of a complex interaction between trait-like vulnerabilities related to self-control and emotion regulation (Buckholtz, 2015). The biggest challenge for researchers is to recreate the dynamic nature of physical aggression within the static nature of laboratory-based assessments (Poldrack et al., 2017). Therefore, in the context of provocation, it would be interesting to employ some of the paradigms the current dissertation used, such as IGT, the homonyms, AIT, and FT, not only for clinical populations but for mentally healthy populations displaying unusually elevated aggression, such as sports fans, as well (Cikara, Botninick, & Fiske, 2011; van der Meij et al., 2015).

Future studies could also benefit from the use of paradigms that are designed to frustrate or promote both reactively and proactively aggressive behavior. Since laboratory aggression tasks have been proven to have constantly good external validity regarding manifestations, laboratory studies could provide researchers with a clearer indication of the underlying neural mechanisms of HAB (Anderson & Bushman, 1997; Giancola & Parrott, 2008). Second, in combination with laboratory studies, future studies could also benefit exploring HAB by observing aggressive behavior within its natural environment, which gives a more accurate measure of aggression than self-report measures. Finally, it would also be useful to investigate the links between HAB and aggression measures using samples of adults with a greater number of aggression problems, such as sub-samples of adults at higher risk for developing these problems (e.g., lower socioeconomic status, higher neighborhood violence; Guerra, Huesmann, & Spindler, 2003). Future research in this area could expand our understanding of how this bias might change within these sub-samples over the course of time.

To extend upon the IGT task from Chapter 2, future studies should possibly employ more complex feedback, such as using a green thumbs-up to indicate acceptance feedback, and a red thumbs down to indicate rejection feedback (Ethridge, et al., 2017; Kujawa et al., 2017; Pegg et al., 2019). It is important to note that the differences between FRN amplitudes can be attributed to physical appearance and not the meaning of the feedback. This has important implications for the interpretation of the previous work. In addition, future studies could benefit from changing the virtual opponent's facial expression in every feedback round, in which a new positive or negative facial expression presented when a participant has been voted off or voted on. These changes might be very meaningful within paradigms such as the IGT, which involve strategic play and in which such facial expression actions might elicit either anger or aggression depending on how the participants interpret these cues. Finally, although the current dissertation could not find associations between behavioral and psychophysiological reflections of HAB and self-report measures of aggression in healthy male adult samples, it still replicated previous findings showing that the negative emotional content of a word can elicit a larger N400 effect. Moreover, the current dissertation is a replication of previous findings that showed that the N400 elicited by a single word appears to have a more anterior distribution with a maximum over frontal or central sites. In addition, we also managed to replicate previous findings showing that CBM training is successful in modifying its targeted bias in a positive way. Along with this, this dissertation was also successful in the replication of previous findings showing larger FRN amplitudes in response to negative feedback as compared to positive feedback as well as, larger ERN amplitudes for incorrect trials as compared to correct trials. Overall, this dissertation addresses a gap in the HAB literature and provides important information on the development of prevention efforts and future intervention studies, stressing the important role of feedback processing in the development and persistence of HAB.

REFERENCES

Abramowitz, J. S., Fabricant, L. E., Taylor, S., Deacon, B. J., McKay, D., & Storch, E. A. (2014). The relevance of analogue studies for understanding obsessions and compulsions. *Clinical Psychology Review*, *34*(3), 206-217.

Achterberg, M., van Duijvenvoorde, A. C., Bakermans-Kranenburg, M. J., & Crone, E. A. (2016). Control your anger! The neural basis of aggression regulation in response to negative social feedback. *Social cognitive and affective neuroscience*, *11*(5), 712-720.

AlMoghrabi, N., Huijding, J., & Franken, I. H. (2018). The effects of a novel hostile interpretation bias modification paradigm on hostile interpretations, mood, and aggressive behavior. *Journal of behavior therapy and experimental psychiatry*, *58*, 36-42.

Anderson, C. A., & Bushman, B. J. (1997). External validity of "trivial" experiments: The case of laboratory aggression. *Review of General Psychology*, *1*(1), 19-41.

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual review of psychology*, *53*.

Anderson, C. A., & Huesmann, L. R. (2007). Human aggression: A social-cognitive view. *The Sage handbook of social psychology*, 259-287.

Anderson, D. J. (2012). Optogenetics, sex, and violence in the brain: implications for psychiatry. *Biological psychiatry*, *71*(12), 1081-1089.

Anderson, N. E., & Kiehl, K. A. (2014). Psychopathy and aggression: when paralimbic dysfunction leads to violence. In *Neuroscience of aggression* (pp. 369-393). Springer, Berlin, Heidelberg.

Argyle, M. (1994). Face, gaze and other non-verbal communication. *The psychology of interpersonal behavior , 5th ed., Penguin Psychology, London, UK*, 23-55.

Arsenio, W. F., Adams, E., & Gold, J. (2009). Social information processing, moral reasoning, and emotion attributions: Relations with adolescents' reactive and proactive aggression. *Child development*, *80*(6), 1739-1755.

Atchley, R. A., & Kwasny, K. M. (2003). Using event-related potentials to examine hemispheric differences in semantic processing. *Brain and Cognition*, *53*(2), 133-138.

Bailey, C. A., & Ostrov, J. M. (2008). Differentiating forms and functions of aggression in emerging adults: Associations with hostile attribution biases and normative beliefs. *Journal of Youth and Adolescence*, *37*(6), 713-722.

Barker, E. D., Boivin, M., Brendgen, M., Fontaine, N., Arseneault, L., Vitaro, F., ... & Tremblay, R. E. (2008). Predictive validity and early predictors of peer-victimization trajectories in preschool. *Archives of General Psychiatry*, *65*(10), 1185-1192.

Baker, T. E., & Holroyd, C. B. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological psychology*, *87*(1), 25-34.

Baron, R. A., & Richardson, D. R. (2004). *Human aggression*. Springer Science & Business Media.

Bates, A. T., Liddle, P. F., Kiehl, K. A., & Ngan, E. T. (2004). State dependent changes in error monitoring in schizophrenia. *Journal of Psychiatric Research*, *38*(3), 347-356.

Beckham, J. C., Moore, S. D., & Reynolds, V. (2000). Interpersonal hostility and violence in Vietnam combat veterans with chronic posttraumatic stress disorder: A review of theoretical models and empirical evidence. *Aggression and violent behavior*, *5*(5), 451-466.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and clinical Neurophysiology*, *60*(4), 343-355.

Bentin, S. (1987). Event-related potentials, semantic processes, and expectancy factors in word recognition. *Brain and language*, *31*(2), 308-327.

Beres, A. M. (2017). Time is of the essence: A review of electroencephalography (EEG) and event-related brain potentials (ERPs) in language research. *Applied psychophysiology and biofeedback*, *42*(4), 247-255.

Berkum, J. J. V., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of cognitive neuroscience*, *11*(6), 657-671.

Berman, M. E., McCloskey, M. S., Fanning, J. R., Schumacher, J. A., & Coccaro, E. F. (2009). Serotonin augmentation reduces response to attack in aggressive individuals. *Psychological Science*, *20*(6), 714-720.

Bertsch, K., Böhnke, R., Kruk, M. R., & Naumann, E. (2009). Influence of aggression on information processing in the emotional Stroop task-an event-related potential study. *Frontiers in behavioral neuroscience*, 3, 28.

Bettencourt, B., & Miller, N. (1996). Gender differences in aggression as a function of provocation: a meta-analysis. *Psychological bulletin*, *119*(3), 422.

Blair, R. J. R. (2001). Neurocognitive models of aggression, the antisocial personality disorders, and psychopathy. *Journal of Neurology, Neurosurgery & Psychiatry*, *71*(6), 727-731.

Blair, R. J. R. (2003). Facial expressions, their communicatory functions and neuro–cognitive substrates. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 561-572.

Blair, R. J. R., Peschardt, K. S., Budhani, S., Mitchell, D. G. V., & Pine, D. S. (2006). The development of psychopathy. *Journal of child psychology and psychiatry*, *47*(3-4), 262-276.

R

Blair, R.J.R. (2010). Psychopathy, frustration, and reactive aggression: The role of ventromedial prefrontal cortex. *British Journal of Psychology, 101*, 383-399.

Bolling, D. Z., Pitskel, N. B., Deen, B., Crowley, M. J., McPartland, J. C., Kaiser, M. D., ... & Pelphrey, K. A. (2011). Enhanced neural responses to rule violation in children with autism: a comparison to social exclusion. *Developmental Cognitive Neuroscience*, *1*(3), 280-294.

Bowen, K. N., Roberts, J. J., & Kocian, E. J. (2016). Decision making of inmates: Testing social information processing concepts using vignettes. *Applied Psychology in Criminal Justice*, *12*(1), 1-17.

Bowins, B. (2016). *Mental Illness Defined: Continuums, Regulation, and Defense*. Routledge.

Brazil, I. A., Mars, R. B., Bulten, B. H., Buitelaar, J. K., Verkes, R. J., and De Bruijn, E. R. (2011). A neurophysiological dissociation between monitoring one's own and others' actions in psychopathy. *Biol. Psychiatry* 69, 693–699.

Brugman, S., Lobbestael, J., Arntz, A., Cima, M., Schuhmann, T., Dambacher, F., & Sack, A.T. (2014). Identifying cognitive predictors of reactive and proactive aggression. *Aggressive Behavior, 9999*, 1-14.

Buckholtz, J. W. (2015). Social norms, self-control, and the value of antisocial behavior. *Current Opinion in Behavioral Sciences*, *3*, 122-129.

Cacioppo, S., Frum, C., Asp, E., Weiss, R. M., Lewis, J. W., & Cacioppo, J. T. (2013). A quantitative meta-analysis of functional imaging studies of social rejection. *Scientific reports*, *3*(1), 1-3.

Cao, J., Gu, R., Bi, X., Zhu, X., & Wu, H. (2015). Unexpected acceptance? Patients with social anxiety disorder manifest their social expectancy in ERPs during social feedback processing. *Frontiers in psychology*, *6*, 1745.

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD (1998) Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science 280:747–749.*

Chavarriaga, R., Sobolewski, A., & Millán, J. D. R. (2014). Errare machinale est: the use of error-related potentials in brain-machine interfaces. *Frontiers in neuroscience*, *8*, 208.

Chester, D. S., Eisenberger, N. I., Pond Jr, R. S., Richman, S. B., Bushman, B. J., & DeWall, C. N. (2014). The interactive effect of social pain and executive functioning on aggression: an fMRI experiment. *Social cognitive and affective neuroscience*, *9*(5), 699-704.

Chester, D. S., & DeWall, C. N. (2015). The pleasure of revenge: retaliatory aggression arises from a neural imbalance toward reward. *Social cognitive and affective neuroscience*, *11*(7), 1173-1182.

Chwilla, D. J., & Kolk, H. H. (2003). Event-related potential and reaction time evidence for inhibition between alternative meanings of ambiguous words. *Brain and Language*, *86*(2), 167-192.

Chwilla, D. J., Virgillito, D., & Vissers, C. T. W. (2011). The relationship of language and emotion: N400 support for an embodied view of language comprehension. *Journal of cognitive neuroscience*, *23*(9), 2400-2414.

Cikara, M., Botvinick, M. M., & Fiske, S. T. (2011). Us versus them: Social identity shapes neural responses to intergroup competition and harm. *Psychological science*, *22*(3), 306-313.

Cillessen, A. H., & Bellmore, A. D. (2014). Social skills and social competence in interactions with peers.

Cima, M., Raine, A., Meesters, C., & Popma, A. (2013). Validation of the Dutch Reactive Proactive Questionnaire (RPQ): Differential correlates of reactive and proactive aggression from childhood to adulthood. *Aggressive Behavior*, *39*(2), 99-113.

Cima, M., Vancleef, L. M. G., Lobbestael, J., Meesters, C., & Korebrits, A. (2014). Don't you dare look at me, or else: Negative and aggressive interpretation bias, callous unemotional traits and type of aggression. *Journal of Child and Adolescent Behavior*, *2*(2), 128.

Cisler, J. M., & Koster, E. H. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical psychology review*, *30*(2), 203-216.

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language, 122*(3), 211-226.

Cleary, A., & Nixon, E. (2012). Early adult outcomes for Irish children with behavioral difficulties. *International Journal of Social Psychiatry*, *58*(6), 643-651.

Coccaro, E. F., Noblett, K. L., & McCloskey, M. S. (2009). Attributional and emotional responses to socially ambiguous cues: Validation of a new assessment of social/emotional information processing in healthy adults and impulsive aggressive patients. *Journal of Psychiatric Research*, *43*(10), 915-925.

Coccaro, E. F., Fanning, J. R., Keedy, S. K., & Lee, R. J. (2016). Social cognition in intermittent explosive disorder and aggression. *Journal of psychiatric research*, *83*, 140-150.

Cohen, R. A., Paul, R., Zawacki, T. M., Moser, D. J., Sweet, L., & Wilkinson, H. (2001). Emotional and personality changes following cingulotomy. *Emotion*, *1*(1), 38.

Coles, M. G. (1989). Modern mind-brain reading: psychophysiology, physiology, and cognition. *Psychophysiology*, *26*(3), 251-269.

Corso, P. S., Mercy, J. A., Simon, T. R., Finkelstein, E. A., & Miller, T. R. (2007). Medical costs and productivity losses due to interpersonal and self-directed violence in the United States. *American journal of preventive medicine*, *32*(6), 474-482.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological bulletin*, 115(1), 74.

Crick, N.R., & Dodge, K.A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development, 67*(3), 993-1002.

Cristea, I. A., Mogoașe, C., David, D., & Cuijpers, P. (2015). Practitioner review: Cognitive bias modification for mental health problems in children and adolescents: A meta-analysis. *Journal of Child Psychology and Psychiatry*, *56*(7), 723-734.

Crozier, J. C., Dodge, K. A., Fontaine, R. G., Lansford, J. E., Bates, J. E., Pettit, G. S., & Levenson, R. W. (2008). Social information processing and cardiac predictors of adolescent antisocial behavior. *Journal of Abnormal Psychology*, *117*(2), 253.

Crosson, B., Ford, A., McGregor, K. M., Meinzer, M., Cheshkov, S., Li, X., ... & Briggs, R. W. (2010). Functional imaging and related techniques: an introduction for rehabilitation researchers. *Journal of rehabilitation research and development*, *47*(2), vii.

Curran, T., Tucker, D. M., Kutas, M., & Posner, M. I. (1993). Topography of the N400: brain electrical activity reflecting semantic expectancy. *Electroencephalography and Clinical Neurophysiology, 88*(3), 188-209.

Dambacher, F., Schuhmann, T., Lobbestael, J., Arntz, A., Bruman, S., & Sack, A.T. (2015). Reducing proactive aggression through non-invasive brain stimulation. *Social Cognitive and Affective Neuroscience, 10*, 1303-1309.

Davidson, R. J., Putnam, K. M., & Larson, C. L. (2000). Dysfunction in the neural circuitry of emotion regulation--a possible prelude to violence. *science*, *289*(5479), 591-594.

Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology*, *41*(1), 60-74.

Defares, P. B., Van der Ploeg, H. M., & Spielberger, C. D. (1980). Handleiding bij de Zelf-beoordelings Vragenlijst ZBV. Een nederlandstalige bewerking van de Spielberger State-Trait Anxiety Inventory.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*(5), 303-305.

DeWall, C. N., & Bushman, B. J. (2011). Social acceptance and rejection: The sweet and the bitter. *Current Directions in Psychological Science*, *20*(4), 256-260.

Dekkers, L. M., van der Molen, M. J., Gunther Moor, B., van der Veen, F. M., & van der Molen, M. W. (2015). Cardiac and electro-cortical concomitants of social feedback processing in women. *Social cognitive and affective neuroscience*, *10*(11), 1506-1514.

Delaney-Busch, N., & Kuperberg, G. (2013). Friendly drug-dealers and terrifying puppies: affective primacy can attenuate the N400 effect in emotional discourse contexts. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(3), 473-490.

Dill, K. E., Anderson, C. A., Anderson, K. B., & Deuser, W. E. (1997). Effects of aggressive personality on social expectations and social perceptions. *Journal of Research in Personality*, *31*(2), 272-292.

Dodge, K. A. (1986). A social information-processing model of social competence in children. In M. Perlmutter (Ed.), *Minnesota symposium on child psychology* (Vol. 18, pp. 77–125). Hillsdale, NJ: Erlbaum.

Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of personality and social psychology*, *53*(6), 1146.

Dodge, K.A., & Pettit, G.S. (2003). A biopsychosocial model of the development of chronic conduct problems in adolescence. *Developmental Psychology, 39*, 349-371.

Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Development and psychopathology*, *18*(3), 791-814.

Dodge, K. A., Malone, P. S., Lansford, J. E., Sorbring, E., Skinner, A. T., Tapanya, S., ... & Bacchini, D. (2015). Hostile attributional bias and aggressive behavior in global context. *Proceedings of the National Academy of Sciences*, *112*(30), 9310-9315.

Dovidio, J.F., Kawakami, K., Gaertner, S.L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82, 62–68.

DSM-5 American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders. *Arlington: American Psychiatric Publishing*.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., ... & Van Petten, C. (2009). Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883-1908.

Edwards, R., & Bond, A. J. (2012). Narcissism, self-concept clarity and aggressive cognitive bias amongst mentally disordered offenders. *Journal of Forensic Psychiatry & Psychology*, *23*(5-6), 620-634.

Ellis, A. (1976). Healthy and unhealthy aggression. *Humanitas*.

Engelen, U., De Peuter, S., Victoir, A., Van Diest, I., & Van den Bergh, O. (2006). Verdere validering van de Positive and Negative Affect Schedule (PANAS) en vergelijking van twee Nederlandstalige versies. *Gedrag en gezondheid*, *34*(2), 61-70.

Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia*, *46*(2), 521-539.

Epps, J., & Kendall, P. C. (1995). Hostile attributional bias in adults. *Cognitive Therapy and Research*, *19*(2), 159-178.

Epstein, S., & Taylor, S. P. (1967). Instigation to aggression as a function of degree of defeat and perceived aggressive intent of the opponent 1. *Journal of personality*, *35*(2), 265-289.

Erdley, C. A., Rivera, M. S., Shepherd, E. J., & Holleb, L. J. (2010). Social-cognitive models and skills. In *Practitioner's guide to empirically based measures of social skills* (pp. 21-35). Springer, New York, NY.

Eronen, M., Angermeyer, M. C., & Schulze, B. (1998). The psychiatric epidemiology of violent behavior. *Social Psychiatry and Psychiatric Epidemiology*, *33*(1), S13-S23.

Ethridge, P., Kujawa, A., Dirks, M. A., Arfer, K. B., Kessel, E. M., Klein, D. N., & Weinberg, A. (2017). Neural responses to social and monetary reward in early adolescence and emerging adulthood. *Psychophysiology, 54*(12), 1786-1799.

Fabiani, M., Gratton, G., & Federmeier, K. D. (2007). Event-related brain potentials: Methods, theory, and applications.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and clinical neurophysiology*, *78*(6), 447-455.

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biological psychology*, *51*(2-3), 87-107.

Fava, M. (1998). Depression with anger attacks. *The Journal of clinical psychiatry*.

Ferrez, P. W., & Millán, J. D. R. (2008). Error-related EEG potentials generated during simulated brain–computer interaction. *IEEE transactions on biomedical engineering*, *55*(3), 923-929.

Fischer, A. G., Klein, T. A., & Ullsperger, M. (2017). Comparing the error-related negativity across groups: The impact of error-and trial-number differences. *Psychophysiology*, *54*(7), 998-1009.

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453-3458.

Gagnon, J., McDuff, P., Daelman, S., & Fournier, S. (2015). Is hostile attributional bias associated with negative urgency and impulsive behaviors? A social-cognitive conceptualization of impulsivity. *Personality and Individual Differences*, *72*, 18-23.

Gagnon, J., Aubin, M., Emond, F. C., Derguy, S., Bessette, M., & Jolicoeur, P. (2016). Neural mechanisms underlying attribution of hostile intention in nonaggressive individuals: An ERP study. *International Journal of Psychophysiology, 110*, 153-162.

Gagnon, J., Aubin, M., Emond, F. C., Derguy, S., Brochu, A. F., Bessette, M., & Jolicoeur, P. (2017). An ERP study on hostile attribution bias in aggressive and nonaggressive individuals. *Aggressive behavior*, *43*(3), 217-229.

Gallo, L. C., & Matthews, K. A. (2003). Understanding the association between socioeconomic status and physical health: do negative emotions play a role?. *Psychological bulletin*, *129*(1), 10.

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research*, *16*(2), 123-144.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological science*, *4*(6), 385-390.

Gehring, W. J., Liu, Y., Orr, J. M., and Carp, J. (2012). "The error-related negativity (ERN/Ne)," in *Oxford Handbook of Event-Related Potential Components*, eds S. J. Luck and E. Kappenman (New York, NY: Oxford University Press), 231–291.

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion*, *12*(2), 314.

Giancola, P. R., & Parrott, D. J. (2008). Further evidence for the validity of the Taylor aggression paradigm. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, *34*(2), 214-229.

Giancola, P. R., Levinson, C. A., Corman, M. D., Godlaski, A. J., Morris, D. H., Phillips, J. P., & Holt, J. C. (2009). Men and women, alcohol and aggression. *Experimental and clinical psychopharmacology*, *17*(3), 154.

Gibbs, N. A. (1996). Nonclinical populations in research on obsessive-compulsive disorder: A critical review. *Clinical Psychology Review*, *16*(8), 729-773.

Giumetti, G. W., & Markey, P. M. (2007). Violent video games and anger as predictors of aggression. *Journal of Research in Personality*, *41*(6), 1234-1243.

Godleski, S. A., Ostrov, J. M., Houston, R. J., & Schlienz, N. J. (2010). Hostile attribution biases for relationally provocative situations and event-related potentials. *International Journal of Psychophysiology, 76*(1), 25-33.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and clinical neurophysiology*, *55*(4), 468-484.

Greydanus, D. E., Pratt, H. D., Greydanus, S. E., & Hoffman, A. D. (1992). Corporal punishment in schools: A position paper of the Society for Adolescent Medicine. *Journal of adolescent health*.

Greenwell, J., & Dengerink, H. A. (1973). The role of perceived versus actual attack in human physical aggression. *Journal of Personality and Social Psychology*, *26*(1), 66.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, *102*(1), 4.

Guerra, N. G., Rowell Huesmann, L., & Spindler, A. (2003). Community violence exposure, social cognition, and aggression among urban elementary school children. *Child development*, *74*(5), 1561-1576.

Guerra, N. G., & Huesmann, L. R. (2004). A cognitive-ecological model of aggression. *REVUE INTERNATIONALE DE PSYCHOLOGIE SOCIALE.*, *17*, 177-204.

Gunther Moor, B., Crone, E. A., & van der Molen, M. W. (2010). The heartbrake of social rejection: heart rate deceleration in response to unexpected peer rejection. *Psychological science*, *21*(9), 1326-1333.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905-912.

Hall, P., & Davidson, K. (1996). The misperception of aggression in behaviorally hostile men. *Cognitive Therapy and Research*, *20*(4), 377-389.

Hall, J. R., Bernat, E. M., & Patrick, C. J. (2007). Externalizing psychopathology and the error-related negativity. *Psychological science*, *18*(4), 326-333.

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological bulletin*, *137*(6), 940.

Hannula, D. E., Federmeier, K. D., & Cohen, N. J. (2006). Event-related potential signatures of relational memory. *Journal of cognitive neuroscience*, *18*(11), 1863-1876.

Hawkins, K. A., & Cougle, J. R. (2013). Effects of interpretation training on hostile attribution bias and reactivity to interpersonal insult. *Behavior Therapy*, *44*(3), 479-488.

Hawley, P. H., & Vaughn, B. E. (2003). Aggression and adaptive functioning: The bright side to bad behavior . *Merrill-Palmer Quarterly*, *49*(3), 239-242.

Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology*, *45*(3), 487-498.

Helfritz-Sinville, L. E., & Stanford, M. S. (2014). Hostile attribution bias in impulsive and premeditated aggression. *Personality and individual differences*, *56*, 45-50.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review*, *109*(4), 679.

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*(5), 688-697.

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behavior s by anterior cingulate cortex. *Trends in cognitive sciences*, *16*(2), 122-128.

Holroyd, C. B., & McClure, S. M. (2015). Hierarchical control over effortful behavior by rodent medial frontal cortex: A computational model. *Psychological review*, 122(1), 54.

Holt, D. J., Lynn, S. K., & Kuperberg, G. R. (2009). Neurophysiological correlates of comprehending emotional meaning in context. *Journal of cognitive neuroscience*, *21*(11), 2245-2262.

Hortensius, R., Schutter, D. J., & Harmon-Jones, E. (2011). When anger leads to aggression: Induction of relative left frontal cortical activity with transcranial direct current stimulation increases the anger–aggression relationship. *Social cognitive and affective neuroscience*, *7*(3), 342-347.

Hovens, J. E., Lievaart, M., & Rodenburg, J. J. (2014). STAXI-2: Vragenlijst over boosheid. *Amsterdam: Hogrefe*.

Hubbard, J.A., McAuliffe, M.D., Morrow, M.T., & Romano, L.J. (2010). Reactive and proactive aggression in childhood and adolescence: precursors, outcomes, processes, experiences, and measurement. *Journal of Personality, 78*(1), 95-118.

Huesmann, L. R., & Moise, J. (1998). The stability and continuity of aggression from early childhood to young adulthood. *Youth violence: Prevention, intervention, and social policy*, 73-95.

Huesmann, L. R., Eron, L. D., & Dubow, E. F. (2002). Childhood predictors of adult criminality: are all risk factors reflected in childhood aggressiveness? *Criminal Behavior and Mental Health*, *12*(3), 185-208.

Juon, H. S., Doherty, E. E., & Ensminger, M. E. (2006). Childhood behavior and adult criminality: Cluster analysis in a prospective study of African Americans. *Journal of quantitative criminology*, *22*(3), 193-214.

Johnson, B. W., & Hamm, J. P. (2000). High-density mapping in an N400 paradigm: evidence for bilateral temporal lobe generators. *Clinical Neurophysiology*, *111*(3), 532-545.

Kassinove, H., Roth, D., Owens, S. G., & Fuller, J. R. (2002). Effects of trait anger and anger expression style on competitive attack responses in a wartime prisoner's dilemma game. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, *28*(2), 117-125.

Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition & Emotion*, *13*(5), 505-521.

Kiehl, K. A., Smith, A. M., Hare, R. D., & Liddle, P. F. (2000). An event-related potential investigation of response inhibition in schizophrenia and psychopathy. *Biological psychiatry*, 48(3), 210-221.

Kiehl, K. A., Laurens, K. R., Bates, A. T., & Liddle, P. F. (2006). Psychopathy and semantic processing: An examination of the N400. *Personality and Individual Differences*, *40*(2), 293-304.

Kissler, J., Assadollahi, R., & Herbert, C. (2006). Emotional and semantic networks in visual word processing: insights from ERP studies. *Progress in brain research*, *156*, 147-183.

Klepousniotou, E., Pike, G. B., Steinhauer, K., & Gracco, V. (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain and language*, *123*(1), 11-21.

Kockler, T. R., Stanford, M. S., Meloy, J. R., Nelson, C. E., & Sanford, K. (2006). Characterizing aggressive behavior in a forensic population. *American journal of orthopsychiatry*, *76*(1), 80-85.

Kohn, S. R., & Asnis, G. M. (2003). Aggression in psychiatric disorders. In *Neurobiology of Aggression* (pp. 135-149). Humana Press, Totowa, NJ.

Kop, W. J. (1999). Chronic and acute psychological risk factors for clinical manifestations of coronary artery disease. *Psychosomatic Medicine*, *61*(4), 476-487.

Kotchoubey, B., & El-Khoury, S. (2014). Event-related potentials indicate context effect in reading ambiguous words. *Brain and cognition*, *92*, 48-60.

Koster, E. H., Fox, E., & MacLeod, C. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *Journal of abnormal psychology*, *118*(1), 1.

Koster, E. H., Baert, S., Bockstaele, M., & De Raedt, R. (2010). Attentional retraining procedures: Manipulating early or late components of attentional bias?. *Emotion*, *10*(2), 230.

Krantz, D. S., & McCeney, M. K. (2002). Effects of psychological and social factors on organic disease: a critical assessment of research on coronary heart disease. *Annual review of psychology*, *53*(1), 341-369.

Krämer, U. M., Jansma, H., Tempelmann, C., & Münte, T. F. (2007). Tit-for-tat: the neural basis of reactive aggression. *Neuroimage*, *38*(1), 203-211.

Krämer, U. M., Büttner, S., Roth, G., & Münte, T. F. (2008). Trait aggressiveness modulates neurophysiological correlates of laboratory-induced reactive aggression in humans. *Journal of Cognitive Neuroscience*, *20*(8), 1464-1477.

Krämer, U. M., Münte, T. F., Richter, S., & Kopyciok, R. P. (2009). Oscillatory brain activity related to control mechanisms during laboratory-induced reactive aggression. *Frontiers in behavioral neuroscience*, *3*, 46.

Krebs, G., Pile, V., Grant, S., Degli Esposti, M., Montgomery, P., & Lau, J. Y. (2018). Research Review: Cognitive bias modification of interpretations in youth and its effect on anxiety: a meta-analysis. *Journal of Child Psychology and Psychiatry*, *59*(8), 831-844.

Kujawa, A., Arfer, K. B., Klein, D. N., & Proudfit, G. H. (2014). Electrocortical reactivity to social feedback in youth: a pilot study of the Island Getaway task. *Developmental cognitive neuroscience*, *10*, 140-147.

Kujawa, A., Kessel, E. M., Carroll, A., Arfer, K. B., & Klein, D. N. (2017). Social processing in early adolescence: Associations between neurophysiological, self-report, and behavioral measures. *Biological psychology, 128*, 55-62.

Kutas, M., & Hillyard, S. A. (1980). Reading between the lines: Event-related brain potentials during natural sentence processing. *Brain and Language*, *11*(2), 354-373.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Kutas, M., & Hillyard, S. A. (1982). The lateral distribution of event-related potentials during sentence processing. *Neuropsychologia*, *20*(5), 579-590.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161-163.

Kutas, M., Hillyard, S. A., & Gazzaniga, M. S. (1988). Processing of semantic anomaly by right and left hemispheres of commissurotomy patients: Evidence from event-related brain potentials. *Brain*, *111*(3), 553-576.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, *4*(12), 463-470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry, 44*(12), 1248-1263.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and emotion*, *24*(8), 1377-1388.

Lansu, T. A., Cillessen, A. H., & Bukowski, W. M. (2013). Implicit and explicit peer evaluation: Associations with early adolescents' prosociality, aggression, and bullying. *Journal of Research on Adolescence*, *23*(4), 762-771.

Lansu, T. A., Cillessen, A. H., & Karremans, J. C. (2012). Implicit associations with popularity in early adolescence: An approach–avoidance analysis. *Developmental Psychology*, *48*(1), 65.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews: Neuroscience, 9*(12), 920-933.

Leary, M.R., Kowalski, R.M., Smith, L., & Phillips, S. (2003). Teasing, rejection, and violence: Case studies of the school shootings. *Aggressive Behavior* 29(3), 202-214.

Leary, M. R., Twenge, J. M., & Quinlivan, E. (2006). Interpersonal rejection as a determinant of anger and aggression. *Personality and social psychology review*, *10*(2), 111-132.

Lee, C. L., & Federmeier, K. D. (2006). To mind the mind: An event-related potential study of word class and semantic ambiguity. *Brain Research*, *1081*(1), 191-202.

Leitner, J. B., Hehman, E., Jones, J. M., & Forbes, C. E. (2014). Self-enhancement influences medial frontal cortex alpha power to social rejection feedback. *Journal of cognitive neuroscience*, *26*(10), 2330-2341.

Lobbestael, J., Cima, M., & Arntz, A. (2013). The relationship between adult reactive and proactive aggression, hostile interpretation bias, and antisocial personality disorder. *Journal of personality disorders*, *27*(1), 53-66.

Loeber, R., & Dishion, T. (1983). Early predictors of male delinquency: a review. *Psychological bulletin*, *94*(1), 68.

Lopez, V. A., & Emmer, E. T. (2002). Influences of beliefs and values on male adolescents' decision to commit violent offenses. *Psychology of Men & Masculinity*, *3*(1), 28.

Lopez, N. L., Vazquez, D. M., & Olson, S. L. (2004). An integrative approach to the neurophysiological substrates of social withdrawal and aggression. *Development and Psychopathology*, *16*(1), 69-93.

Lösel, F., Bliesener, T., & Bender, D. (2007). Social information processing, experiences of aggression in social contexts, and aggressive behavior in adolescents. *Criminal Justice and Behavior*, *34*(3), 330-347.

Lotze, M., Veit, R., Anders, S., & Birbaumer, N. (2007). Evidence for a different role of the ventral and dorsal medial prefrontal cortex for social reactive aggression: an interactive fMRI study. *Neuroimage*, *34*(1), 470-478.

Ma, R., Jia, H., Yi, F., Ming, Q., Wang, X., Gao, Y., ... & Yao, S. (2016). Electrophysiological responses of feedback processing are modulated by MAOA genotype in healthy male adolescents. *Neuroscience letters*, *610*, 144-149.

Mammen, O. K., Kolko, D. J., & Pilkonis, P. A. (2002). Negative affect and parental aggression in child physical abuse. *Child Abuse & Neglect*, *26*(4), 407-424.

Martinelli, A., Ackermann, K., Bernhard, A., Freitag, C. M., & Schwenck, C. (2018). Hostile attribution bias and aggression in children and adolescents: A systematic literature review on the influence of aggression subtype and gender. *Aggression and violent behavior*, *39*, 25-32.

Mathalon, D. H., Fedor, M., Faustman, W. O., Gray, M., Askari, N., & Ford, J. M. (2002). Response-monitoring dysfunction in schizophrenia: an event-related brain potential study. *Journal of abnormal psychology*, *111*(1), 22.

Matthews, B. A., & Norris, F. H. (2002). When Is Believing "Seeing"? Hostile Attribution Bias as a Function of Self-Reported Aggression 1. *Journal of Applied Social Psychology*, *32*(1), 1-31.

Matthys, W., Vanderschuren, L. J., Schutter, D. J., & Lochman, J. E. (2012). Impaired neurocognitive functions affect social learning processes in oppositional defiant disorder and conduct disorder: Implications for interventions. *Clinical child and family psychology review*, *15*(3), 234-246.

McCarthy, R. J., & Elson, M. (2018). A conceptual review of lab-based aggression paradigms. *Collabra: Psychology*, *4*(1).

McCarthy, G., & Nobre, A. C. (1993). Modulation of semantic processing by spatial selective attention. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *88*(3), 210-219.

McConaughy, S. H., & Skiba, R. J. (1993). Comorbidity of externalizing and internalizing problems. *School Psychology Review*, *22*(3), 421-436.

Meesters, C., Muris, P., Bosma, H., Schouten, E., & Beuving, S. (1996). Psychometric evaluation of the Dutch version of the Aggression Questionnaire. *Behavior research and therapy*, *34*(10), 839-843.

Meyer, A. M., & Federmeier, K. D. (2007). The effects of context, meaning frequency, and associative strength on semantic selection: Distinct contributions from each cerebral hemisphere. *Brain Research*, *1183*, 91-108.

Miller, J. D., & Lynam, D. R. (2006). Reactive and proactive aggression: Similarities and differences. *Personality and Individual Differences*, *41*(8), 1469-1480.

Milich, R., & Dodge, K. A. (1984). Social information processing in child psychiatric populations. *Journal of abnormal child psychology*, *12*(3), 471-489.

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *Journal of cognitive neuroscience*, *9*(6), 788-798.

Moran, P. (1999). The epidemiology of antisocial personality disorder. *Social psychiatry and psychiatric epidemiology*, *34*(5), 231-242.

Moreno, E. M., & Vazquez, C. (2011). Will the glass be half full or half empty? Brain potentials and emotional expectations. *Biological Psychology, 88*(1), 131-140.

Morren, M., & Meesters, C. (2002). Validation of the Dutch version of the Aggression Questionnaire in adolescent male offenders. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, *28*(2), 87-96.

Moser, J. S., Hajcak, G., Huppert, J. D., Foa, E. B., & Simons, R. F. (2008). Interpretation bias in social anxiety as detected by event-related brain potentials. *Emotion, 8*(5), 693-700.

Munro, G. E., Dywan, J., Harris, G. T., McKee, S., Unsal, A., and Segalowitz, S. J. (2007). ERN varies with degree of psychopathy in an emotion discrimination task. *Biol. Psychol.* 76, 31–42.

Murray-Close, D., & Ostrov, J. M. (2009). A longitudinal study of forms and functions of aggressive behavior in early childhood. *Child Development*, *80*(3), 828-842.

Murray-Close, D., Ostrov, J. M., Nelson, D. A., Crick, N. R., & Coccaro, E. F. (2010). Proactive, reactive, and romantic relational aggression in adulthood: Measurement, predictive validity, gender differences, and association with intermittent explosive disorder. *Journal of Psychiatric Research*, *44*(6), 393-404.

Nakagawa, S., Takeuchi, H., Yasuyuki, T., Nouchi, R., Sekiguchi, A., Kotozaki, Y., Miyauchi, C.M., Iizuka, K., Yokoyama, R., Shinada, T., & Yamamoto, Y. (2017). The anterior midcingulate cortex as a neural node underlying hostility in young adults. *Brain Structure and Function, 222*, 61-70.

Nasby, W., Hayden, B., & DePaulo, B. M. (1980). Attributional bias among aggressive boys to interpret unambiguous social stimuli as displays of hostility. *Journal of abnormal psychology*, *89*(3), 459.

Nelson, R.J., & Trainor, B.C. (2007). Neural mechanisms of aggression. *Nature Reviews Neuroscience, 8*(7), 536-546.

Nesbit, S. M., Conger, J. C., & Conger, A. J. (2007). A quantitative review of the relationship between anger and aggressive driving. *Aggression and Violent Behavior*, *12*(2), 156-176.

Neumann, D., Malec, J. F., & Hammond, F. M. (2015). The association of negative attributions with irritation and anger after brain injury. *Rehabilitation psychology*, *60*(2), 155.

Niemann, J., Wisse, B., Rus, D., Van Yperen, N.W., & Sassenberg, K. (2014). Anger and attitudinal reactions to negative feedback: the effects of emotional instability and power. *Motivation and Emotion, 38*, 687-699.

Nieuwenhuis, S., Yeung, N., Van Den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cognitive, affective, & behavioral neuroscience*, *3*(1), 17-26.

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience & Biobehavioral Reviews*, *28*(4), 441-448.

Niznikiewicz, M. A., O'donnell, B. F., Nestor, P. G., Smith, L., Law, S., Karapelou, M., ... & McCarley, R. W. (1997). ERP assessment of visual and auditory language processing in schizophrenia. *Journal of Abnormal Psychology*, *106*(1), 85.

Ohayon, M. M., & Shapiro, C. M. (2000). Posttraumatic Stress Disorder in the General Population. *Comprehensive psychiatry*, *41*(6), 469-478.

Ohbuchi, K., & Kambara, T. (1985). Attacker's intent and awareness of outcome, impression management, and retaliation. *Journal of Experimental Social Psychology*, *21*(4), 321-330.

Olweus, D. (1979). Stability of aggressive reaction patterns in males: a review. *Psychological bulletin*, *86*(4), 852.

Orobio de Castro, B., Veerman, J.W., Koops, W., Bosch, J.D., & Monschouwer, H.J. (2002). Hostile attribution of intent and aggressive behavior: a meta-analysis. *Child Development, 73*(3), 916-934.

Orobio de Castro, B., Merk, W., Koops, W., Veerman, J. W., & Bosch, J. D. (2005). Emotions in social information processing and their relations with reactive and proactive aggression in referred aggressive boys. *Journal of Clinical Child and Adolescent Psychology*, *34*(1), 105-116.

Ortega-Escobar, J., & Alcázar-Córcoles, M.A. (2016). Neurobiology of aggression and violence. *Anuario de Psicología Jurídica*, 26, 60- 69.

Ostir, G. V., Smith, P. M., Smith, D., & Ottenbacher, K. J. (2005). Reliability of the Positive and Negative Affect Schedule (PANAS) in medical rehabilitation. *Clinical Rehabilitation*, *19*(7), 767-769.

Parrott, D. J., & Giancola, P. R. (2007). Addressing "The criterion problem" in the assessment of aggressive behavior: Development of a new taxonomic system. *Aggression and Violent Behavior*, *12*(3), 280-299.

Patrick, C.J. (2008). Psychophysiological correlates of aggression and violence: an integrative view. *Philosophical Transactions of Royal Society B, 363*, 2543-2555.

Pawliczek, C.M., Derntl, B., Kellermann, T., Gur, R.C., Schneider, F., & Habel, U. (2013). Anger under control: neural correlates of frustration as a function of trait aggression. *PloS ONE, 8*(10), 1-10.

Pearson, C.M., & Porath, C.L. (2005). On the nature, consequences and remedies of workplace incivility: no time for "nice"? Think again. *Academy of Management Executive, 19*(1), 7-18.

Pegg, S., Ethridge, P., Shields, G. S., Slavich, G. M., Weinberg, A., & Kujawa, A. (2019). Blunted Social Reward Responsiveness Moderates the Effect of Lifetime Social Stress Exposure on Depressive Symptoms. *Frontiers in behavioral neuroscience*, *13*, 178.

Penton-Voak, I. S., Thomas, J., Gage, S. H., McMurran, M., McDonald, S., & Munafò, M. R. (2013). Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. *Psychological Science*, *24*(5), 688-697.

Perach-Barzilay, N., Tauber, A., Klein, E., Chistyakov, A., Ne'eman, R., & Shamay-Tsoory, S.G. (2012). Asymmetry in the dorsolateral prefrontal cortex and aggressive behavior: a continuous theta-burst magnetic stimulation study. *Social Neuroscience 8*(2), 178-188.

Pert, C., Jahoda, A., & Squire, J. (1999). Attribution of intent and role-taking: Cognitive factors as mediators of aggression with people who have mental retardation. *American Journal on Mental Retardation*, *104*(5), 399-409.

Pfabigan, D. M., Gittenberger, M., & Lamm, C. (2019). Social dimension and complexity differentially influence brain responses during feedback processing. *Social neuroscience*, *14*(1), 26-40.

Pietschmann, M., Simon, K., Endrass, T., & Kathmann, N. (2008). Changes of performance monitoring with learning in older and younger adults. *Psychophysiology*, *45*(4), 559-568.

Pietschmann, M., Endrass, T., Czerwon, B., & Kathmann, N. (2011). Aging, probabilistic learning and performance monitoring. *Biological psychology*, *86*(1), 74-82.

Pincham, H. L., Wu, C., Killikelly, C., Vuillier, L., & Fearon, R. P. (2015). Social provocation modulates decision making and feedback processing: Examining the trajectory of development in adolescent participants. *Developmental cognitive neuroscience*, *15*, 58-66.

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, *52*(4), 449-459.

Poldrack, R. A., Monahan, J., Imrey, P. B., Reyna, V., Raichle, M. E., Faigman, D., & Buckholtz, J. W. (2017). Predicting violent behavior: what can neuroscience add? *Trends in cognitive sciences*, 22(2), 111-123.

Pontifex, M. B., Scudder, M. R., Brown, M. L., O'Leary, K. C., Wu, C. T., Themanson, J. R., & Hillman, C. H. (2010). On the number of trials necessary for stabilization of error-related brain activity across the life span. *Psychophysiology*, *47*(4), 767-773.

Pylkkänen, L., Llinás, R., & Murphy, G. L. (2006). The representation of polysemy: MEG evidence. *Journal of cognitive neuroscience*, *18*(1), 97-109.

Quiggle, N. L., Garber, J., Panak, W. F., & Dodge, K. A. (1992). Social information processing in aggressive and depressed children. *Child development*, *63*(6), 1305-1320.

Quigley, Leanne, Andrea L. Nelson, Jonathan Carriere, Daniel Smilek, and Christine Purdon. "The effects of trait and state anxiety on attention to emotional images: An eye-tracking study." *Cognition & emotion* 26, no. 8 (2012): 1390-1411.

Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., ... & Liu, J. (2006). The reactive–proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, *32*(2), 159-171.

Reijntjes, A., Stegge, H., Terwogt, M. M., Kamphuis, J. H., & Telch, M. J. (2006a). Children’s coping with in vivo peer rejection: An experimental investigation. *Journal of Abnormal Child Psychology*, *34*(6), 873-885.

Reijntjes, A., Stegge, H., & Meerum Terwogt, M. (2006b). Children's coping with peer rejection: The role of depressive symptoms, social competence, and gender. *Infant and Child Development: An International Journal of Research and Practice*, *15*(1), 89-107.

Richman, L.S., & Leary, M.R. (2009). Reactions to discrimination, stigmatization, ostracism, and other forms of interpersonal rejection: a multimotive model. *Psychological Review, 116*, 365-383.

Riva, P., Romero Lauro, L. J., DeWall, C. N., Chester, D. S., & Bushman, B. J. (2015). Reducing aggressive responses to social exclusion using transcranial direct current stimulation. *Social cognitive and affective neuroscience*, *10*(3), 352-356.

Roth, G., & Strueber, D. (2009). Neurobiological aspects of reactive and proactive violence in antisocial personality disorder and" psychopathy". *Praxis der Kinderpsychologie und Kinderpsychiatrie*, *58*(8), 587-609.

Rozanski, A., Blumenthal, J. A., & Kaplan, J. (1999). Impact of psychological factors on the pathogenesis of cardiovascular disease and implications for therapy. *Circulation*, *99*(16), 2192-2217.

Sanislow, C. A., Grilo, C. M., & McGlashan, T. H. (2000). Factor analysis of the DSM-III-R borderline personality disorder criteria in psychiatric inpatients. *American Journal of Psychiatry*, *157*(10), 1629-1633.

Santesso, D. L., Segalowitz, S. J., & Schmidt, L. A. (2005). ERP correlates of error monitoring in 10-year olds are related to socialization. *Biological Psychology*, *70*(2), 79-87.

Sawyer, M. G., Arney, F. M., Baghurst, P. A., Clark, J. J., Graetz, B. W., Kosky, R. J., ... & Rey, J. M. (2001). The mental health of young people in Australia: key findings from the child and adolescent component of the national survey of mental health and well-being. *Australian & New Zealand Journal of Psychiatry*, *35*(6), 806-814.

Schat, A.C.H., & Kelloway, E.K. (2000). The effects of perceived control on the outcomes of workplace aggression and violence. *Journal of Occupational Health Psychology, 4*, 386-402.

Schat, A.C.H., & Kelloway, E.K. (2003). Reducing the adverse consequences of workplace aggression and violence: the buffering effects of organizational support. *Journal of Occupational Health Psychology, 8*(2), 110-122.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime User's Guide. *Psychology Software Tools*: Pittsburgh, PA.

Schönenberg, M., & Jusyte, A. (2014). Investigation of the hostile attribution bias toward ambiguous facial cues in antisocial violent offenders. *European archives of psychiatry and clinical neuroscience*, *264*(1), 61-69.

Schupp, H. T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2007). Selective visual attention to emotion. *Journal of neuroscience*, *27*(5), 1082-1089.

Seidel, E. M., Habel, U., Kirschner, M., Gur, R. C., & Derntl, B. (2010). The impact of facial emotional expressions on behavioral tendencies in women and men. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(2), 500.

Sharp, C. (2002). Sex Differences in Antisocial Behavior : Conduct Disorder, Delinquency and Violence in the Dunedin Longitudinal Study. By T. Moffitt, A. Caspi, M. Rutter and P. Silva.(Pp. 278;£ 14.95/$21.95.) Cambridge University Press: Cambridge. 2001. *Psychological Medicine*, *32*(8), 1475-1478.

Shahinfar, A., Kupersmidt, J. B., & Matza, L. S. (2001). The relation between exposure to violence and social information processing among incarcerated adolescents. *Journal of Abnormal Psychology*, *110*(1), 136.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217-240.

Siever, L. J. (2008). Neurobiology of aggression and violence. *American Journal of Psychiatry*, *165*(4), 429-442.

Skodol, A. E., Gunderson, J. G., Pfohl, B., Widiger, T. A., Livesley, W. J., & Siever, L. J. (2002). The borderline diagnosis I: psychopathology, comorbidity, and personaltity structure. *Biological psychiatry*, *51*(12), 936-950.

Smeijers, D., Rinck, M., Bulten, E., van den Heuvel, T., & Verkes, R. J. (2017). Generalized hostile interpretation bias regarding facial expressions: characteristic of pathological aggressive behavior. *Aggressive behavior*, *43*(4), 386-397.

Smith, T. W., & Ruiz, J. M. (2002). Psychosocial influences on the development and course of coronary heart disease: current status and implications for research and practice. *Journal of consulting and clinical psychology*, *70*(3), 548.

Somerville, L. H., Heatherton, T. F., & Kelley, W. M. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature neuroscience*, *9*(8), 1007-1008.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). Stai. Manual for the State-Trait Anxiety Inventory (Self Evaluation Questionnaire). *Palo Alto California: Consulting Psychologist*, *22*, 1-24.

Stadler, C., Sterzer, P., Schmeck, K., Krebs, A., Kleinschmidt, A., & Poustka, F. (2007). Reduced anterior cingulate activation in aggressive children and adolescents during affective stimulation: association with temperament traits. *Journal of Psychiatric Research*, 41(5), 410-417.

Stanford, M. S., Houston, R. J., Mathias, C. W., Villemarette-Pittman, N. R., Helfritz, L. E., & Conklin, S. M. (2003). Characterizing aggressive behavior. *Assessment*, *10*(2), 183-190.

Steele, V. R., Anderson, N. E., Claus, E. D., Bernat, E. M., Rao, V., Assaf, M., ... & Kiehl, K. A. (2016). Neuroimaging measures of error-processing: Extracting reliable signals from event-related potentials and functional magnetic resonance imaging. *NeuroImage*, *132*, 247-260.

Sterzer, P., Stadler, C., Krebs, A., Kleinschmidt, A., & Poustka, F. (2005). Abnormal neural responses to emotional visual stimuli in adolescents with conduct disorder. *Biological Psychiatry, 57*(1), 7-15.

Stieben, J., Lewis, M. D., Granic, I., Zelazo, P. D., Segalowitz, S., & Pepler, D. (2007). Neurophysiological mechanisms of emotion regulation for subtypes of externalizing children. *Development and Psychopathology*, *19*(2), 455-480.

Sun, S., & Yu, R. (2014). The feedback related negativity encodes both social rejection and explicit social expectancy violation. *Frontiers in human neuroscience*, *8*, 556.

Suris, A., Lind, L., Emmett, G., Borman, P. D., Kashner, M., & Barratt, E. S. (2004). Measures of aggressive behavior: Overview of clinical and research instruments. *Aggression and Violent Behavior*, 9(2), 165-227.

Swaab, T. Y., Ledoux, K., Camblin, C. C., & Boudewyn, M. A. (2012). Language-related ERP components. *Oxford handbook of event-related potential components*, 397-440.

Swann, A. C. (2003). Neuroreceptor mechanisms of aggression and its treatment. *The Journal of Clinical Psychiatry*, *64*, 26-35.

Talmi, D., Atkinson, R. & El-Deredy, W. (2013). The Feedback-Related Negativity signals salience prediction errors, not reward prediction errors (in-press). *Journal of Neuroscience*, 33(19), 8264-8269.

Taylor, S.P. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality*, 35(2), 297-310.

Titone, D. A., & Salisbury, D. F. (2004). Contextual modulation of N400 amplitude to lexically ambiguous words. *Brain and Cognition*, *55*(3), 470-478.

Tremblay, R. E. (2000). The development of agressive behavior during childhood: What have we learned in the past century?. *International journal of behavioral development*, *24*(2), 129-141.

Tremblay, P. F., & Belchevski, M. (2004). Did the instigator intend to provoke? A key moderator in the relation between trait aggression and aggressive behavior. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, *30*(5), 409-424.

Tremblay, R. E., Nagin, D. S., Séguin, J. R., Zoccolillo, M., Zelazo, P. D., Boivin, M., ... & Japel, C. (2004). Physical aggression during early childhood: Trajectories and predictors. *Pediatrics*, *114*(1), e43-e50.

Tremblay, R., & Nagin, D. S. (2005). The Developmental Origins of Physical Aggression in Humans. I: Richard Tremblay, William W. *Developmental origins of aggression. New York: Guilford Press.(Kap. 5)*.

Tuente, S. K., Bogaerts, S., & Veling, W. (2019). Hostile attribution bias and aggression in adults-a systematic review. *Aggression and Violent Behavior*.

Twenge, J.M., Baumeister, R.F., Tice, D.M., & Stucke, T.S. (2001). If you can't join them, beat them: Effects of social exclusion on aggressive behavior. *Journal of Personality and Social Psychology, 81*(6), 1058-1069.

Ullsperger, M., & Von Cramon, D. Y. (2003). Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, *23*(10), 4308-4314.

Van der Gronde, T., Kempes, M., van El, C., Rinne, T., & Pieters, T. (2014). Neurobiological correlates in forensic assessment: a systematic review. *PLoS ONE, 9*(10), 1-15.

Van der Meij, L., Klauke, F., Moore, H. L., Ludwig, Y. S., Almela, M., & van Lange, P. A. (2015). Football fan aggression: The importance of low basal cortisol and a fair referee. *PloS one*, *10*(4), e0120103.

Van der Molen, M. J., Poppelaars, E. S., Van Hartingsveldt, C. T., Harrewijn, A., Gunther Moor, B., & Westenberg, P. M. (2014). Fear of negative evaluation modulates electrocortical and behavioral responses when anticipating social evaluative feedback. *Frontiers in human neuroscience*, *7*, 936.

Van der Molen, M. J. W., Dekkers, L. M., Westenberg, P. M., van der Veen, F. M., & Van der Molen, M. W. (2017). Why don't you like me? Midfrontal theta power in response to unexpected peer rejection feedback. *NeuroImage*, *146*, 474-483.

Van der Veen, F. M., van der Molen, M. J., Van der Molen, M. W., & Franken, I. H. A. (2016). Thumbs up or thumbs down? Effects of neuroticism and depressive symptoms on psychophysiological responses to social evaluation in healthy students. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(5), 836-847.

Van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & behavior*, *77*(4-5), 477-482.

Vassilopoulos, S. P., Brouzos, A., & Andreou, E. (2015). A multi-session attribution modification program for children with aggressive behavior : Changes in attributions, emotional reaction estimates, and self-report aggression. *Behavioral and cognitive psychotherapy*, *43*(5), 538-548.

Verhoef, R. E., Alsem, S. C., Verhulp, E. E., & De Castro, B. O. (2019). Hostile Intent Attribution and Aggressive Behavior in Children Revisited: A Meta-Analysis. *Child development*, *90*(5), e525-e547.

Verhulst, F. C., Root, H. M., Gullickson, T., & Ramser, P. (1996). The Epidemiology of Child and Adolescent Psychopathology. *Psyccritiques*, *41*(11), 1150.

Ventura, P. (2014). Let's face it: reading acquisition, face and word processing. *Frontiers in psychology*, *5*, 787.

Vila-Ballo, A., Cunillera, T., Rostan, C., Hdez-Lafuente, P., Fuentemilla, L., & Rodriguez-Fornells, A. (2015). Neurophysiological correlates of cognitive flexibility and feedback processing in violent juvenile offenders. *Brain research*, *1610*, 98-109.

Von Borries, A. K. L., Brazil, I. A., Bulten, B. H., Buitelaar, J. K., Verkes, R. J., & De Bruijn, E. R. A. (2010). Neural correlates of error-related learning deficits in individuals with psychopathy. *Psychological medicine*, *40*(9), 1559-1568.

Vredenburg, K., Flett, G. L., & Krames, L. (1993). Analogue versus clinical depression: a critical reappraisal. *Psychological Bulletin*, *113*(2), 327.

Waechter, S., & Stolz, J. A. (2015). Trait anxiety, state anxiety, and attentional bias to threat: Assessing the psychometric properties of response time measures. *Cognitive Therapy and Research*, *39*(4), 441-458.

Waters, H., Hyder, A. A., Rajkotia, Y., Basu, S., Rehwinkel, J. A., Butchart, A., & World Health Organization. (2004). The economic dimensions of interpersonal violence.

White, K. R., Crites Jr, S. L., Taylor, J. H., & Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. *Social Cognitive and Affective Neuroscience*, *4*(2), 191-198.

Wilkowski, B. M., & Robinson, M. D. (2010). The anatomy of anger: An integrative cognitive model of trait anger and reactive aggression. *Journal of personality*, *78*(1), 9-38.

Weidler, C., Habel, U., Hüpen, P., Akkoc, D., Schneider, F., Blendy, J. A., & Wagels, L. (2019). On the complexity of aggressive behavior: Contextual and individual factors in the Taylor Aggression Paradigm. *Frontiers in psychiatry*, *10*, 521.

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of social psychology*, *26*(4), 557-580.

Williamson, S., Harpur, T. J., & Hare, R. D. (1991). Abnormal processing of affective words by psychopaths. *Psychophysiology*, *28*(3), 260-273.

Wilson, S. J., & Lipsey, M. W. (2006). The effects of school-based social information processing interventions on aggressive behavior, Part I: Universal programs. *Campbell Systematic Reviews*, *2*(1), 1-42.

Wiswede, D., Taubner, S., Münte, T. F., Roth, G., Strüber, D., Wahl, K., & Krämer, U. M. (2011). Neurophysiological correlates of laboratory-induced aggression in young men with and without a history of violence. *PloS one*, *6*(7).

World Health Organization. (1992). The ICD-10 classification of mental and behavioral disorders: clinical descriptions and diagnostic guidelines. *Weekly Epidemiological Record= Relevé épidémiologique hebdomadaire*, *67*(30), 227-227.

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, *24*(28), 6258-6264.

Yi, F., Chen, H., Wang, X., Shi, H., Yi, J., Zhu, X., & Yao, S. (2012). Amplitude and latency of feedback-related negativity: aging and sex differences. *Neuroreport*, *23*(16), 963-969.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition*, *63*(3), 271-313.

Yu, Q., Zhuang, Q., Wang, B., Liu, X., Zhao, G., & Zhang, M. (2018). The effect of anxiety on emotional recognition: evidence from an ERP study. *Scientific reports*, *8*(1), 16146.

Zelli, A., Dodge, K. A., Lochman, J. E., & Laird, R. D. (1999). The distinction between beliefs legitimizing aggression and deviant processing of social cues: Testing measurement validity and the hypothesis that biased processing mediates the effects of beliefs on aggression. *Journal of personality and social psychology*, *77*(1), 150.

Zumkley, H., & Frączek, A. (1992). *Socialization and aggression*. Springer-Verlag.

NEDERLANDSE SAMENVATTING

Inleiding

In het huidige proefschrift is geprobeerd om eerdere studies te repliceren en de associaties tussen de vijandige attributiebias (HAB – '*hostile attributional bias*') en agressie te testen. Hoewel het bekend is dat agressieve individuen HAB vertonen, is er relatief weinig bekend over de neurale processen die gerelateerd zijn aan deze bias. Daarom kunnen functionele beeldvormingstechnieken van de hersenen, zoals het elektro-encefalogram (EEG) nuttig zijn bij het bepalen van de onderliggende neurale processen die gerelateerd zijn aan HAB. Het belangrijkste doel van dit proefschrift was om met behulp van *Event-Related brain Potentials* (ERP's) agressie bij gezonde mannelijke volwassenen te onderzoeken om zo inzicht te krijgen in de vroege stadia van informatieverwerking om te onderzoeken welke factoren (i.e. feedback) veranderingen in de informatieverwerking bij gezonde mannelijke volwassenen veroorzaken.

Overzicht van de huidige proefschriftstudies

Studie 1- Kunnen we met elkaar overweg? De relatie tussen feedback gerelateerde negativiteit en reactieve agressie tijdens de sociale feedbackverwerking bij gezonde personen

Het primaire doel van deze studie was om te onderzoeken of de neuronale reflectie van de verwerking van feedback (*Feedback Related Negativity* - FRN) een associatie heeft met zelfrapportage-indicatoren van agressie. Deze studie was gebaseerd op de hypothese dat agressief gedrag blijft bestaan wanneer personen een probleem hebben met leren van feedback, dat met de FRN gemeten kan worden. In de huidige studie is daarom gebruikgemaakt van een sociaal feedbackparadigma (*The Island Getaway*) om dit ERP-proces te meten. *The Island Getaway* (IGT) is een gecomputeriseerde sociale feedbacktaak die gebaseerd is op het televisieprogramma *"Survivor"* en een *peer*-afwijzingstaak. De huidige studie veronderstelde dat de deelnemers eerder geneigd zouden zijn om medespelers af te wijzen die hen eerder hadden afgewezen en dat de relatieve negativiteit zou worden waargenomen in het ERP voor sociale afwijzing in vergelijking met de sociale acceptatie feedback (i.e. de FRN). Concreet verwachtten we dat gezonde deelnemers die hoog scoren op zelfgerapporteerde agressie minder gevoelig zouden zijn voor negatieve sociale feedback en daarom een kleinere FRN-amplitude zouden hebben in de afwijzingsfeedback.

De resultaten van deze studie toonden aan dat de FRN-amplituden niet anders waren voor negatieve feedback dan voor positieve feedback. Bovendien was er geen verband tussen de FRN-amplitude en de zelfgerapporteerde agressie. Daarnaast werd het stemgedrag van de deelnemers niet geassocieerd met reactieve agressie of FRN. Ten slotte concludeerde deze studie dat de huidige versie van de IGT niet gevoelig genoeg is voor het vinden van associaties tussen agressie en leren van feedback in een gezonde populatie. Daarnaast kan ook de conclusie getrokken worden dat de hersenreacties in de IGT, gebruikt als een instrument om gedrag en hersenreacties op sociale feedback te testen in een volwassen gezonde mannelijke populatie, afhankelijk zijn van het uiterlijk van de feedbackstimuli.

Studie 2- Emotionele modulatie van de N400: Manipuleren van de emotionele betekenis van homoniemen

Het primaire doel van deze studie was om te onderzoeken of de N400 een nuttige maat is in andere contexten, met name in de context van agressie. Daarom hebben we onder strikte voorwaarden onderzocht of de negatieve emotionele inhoud van een woord een grotere N400 effect kan uitlokken dan de neutrale emotionele inhoud van een woord. We waren in het bijzonder geïnteresseerd of homoniemen woorden met ofwel geen emotionele betekenis of een negatieve emotionele betekenis verschilden ten opzichte van de N400 respons. Een ander doel van deze studie was onderzoeken of de valentie (positief of negatief) en het angstniveau (dispositie of toestand) van de deelnemer een effect had op de amplitude van de N400, aangezien in eerdere studies een invloed van de stemming op de N400 werd gevonden. In deze studie werd gebruik gemaakt van de homoniemen taak waarbij de emotionele woorden die we gebruikten identiek waren aan de neutrale woorden in vorm en articulatie, maar alleen verschilden in semantische inhoud, die we afhankelijk maakten van de context waarin het woord werd gepresenteerd. Voor wat betreft de indicatoren van de hersenen verwachtte we in de huidige studie een significante toename van het N400-effect voor het homoniem binnen een emotionele context vergeleken met het homoniem dat ingebed is in een neutrale context. Specifiek verwachtten we in de ze studie dat de N400 uitsluitend zou worden ontlokt door de emotionele inhoud van het woord. Daarnaast verwachtten we een verband te vinden tussen affect en angst aan de ene kant en de N400 aan de andere kant.

De resultaten van deze studie toonden een significant hoofdeffect van emotie, wat suggereert dat de negatieve emotionele inhoud van een woord een groter N400-effect ontlokt, zelfs onder zeer strenge voorwaarden (met behulp van homoniemen), vergeleken met de woorden met een neutrale emotionele inhoud. Echter, is deze studie er niet in geslaagd om de correlaties te vinden tussen affect en angstgevoelens aan de ene kant en de N400-amplituden aan de andere kant. Ten slotte, de belangrijkste conclusie die uit deze studie kan worden getrokken is dat de N400 modulatie wordt beïnvloed door de emotionele context van het woord zelf en niet door hun fysieke verschijning kan worden verklaard, iets wat eerdere studies niet konden uitsluiten.

Studie 3- Geen effect van zelfrapportage-agressie-indicatoren op N400 in een lexicale beslissingstaak met bijbehorende woorden in een gewelddadige context

Het primaire doel van deze studie was het onderzoeken van gezonde studentenpopulaties met impliciete indicatoren van bias (vooringenomenheid) om te zien of hogere scores op zelfgerapporteerde agressie geassocieerd werden met een meer aan agressie gerelateerde interpretatie bias *(‘aggression related interpretation bias’* - ARIB). De ARIB is onderzocht met behulp van de zogenaamde Agressieve Interpretatietaak (AIT). De AIT is een lexicale beslissingstaak waarbij het eerste woord van een woordpaar werd gepresenteerd en de deelnemers werd gevraagd om zo snel mogelijk te beslissen of het tweede woord (*target*) in het woordpaar een bestaand Nederlands woord is of een logatoom. De AIT werd verdeeld in zes soorten woordparen (*prime-target*): 1- Ambigue *prime*, *target* geassocieerd in gewelddadige context; 2- Ambigue *prime*, *target* geassocieerd in neutrale context; 3- Ambigue *prime*, *target* niet geassocieerd; 4- Niet ambigue *prime*, *target* geassocieerd in neutrale context; 5- Niet ambigue *prime*, *target* niet geassocieerd; 6- Niet ambigue *prime*, *target* logatoom. De huidige studie verwachtte dat hogere scores van zelfrapportage-agressie-indicatoren gepaard zouden gaan met een snellere reactietijd (RT) op gewelddadige contextdoelen, maar niet voor neutrale context targets. Voor de hersenindicatoren verwachtte de huidige studie dat hogere scores van zelfrapportage-agressie-indicatoren alleen geassocieerd zouden zijn met kleinere N400-amplitudes bij gewelddadige context targets, maar niet bij neutrale context targets.

De resultaten van deze studie toonden aan dat de geassocieerde woorden leidden tot een snellere reactietijd (RT) in vergelijking met logatomen. Daarnaast werden logatomen ook geassocieerd met een grotere N400-amplitude vergeleken met bestaande Nederlandse woorden. Target type woorden die in een gewelddadige context werden geassocieerd, hadden geen invloed op de reactietijd en de N400-indicatoren. Bovendien was er geen verband tussen reactietijden, N400, en zelfrapportage-indicatoren van agressie. Aan de andere kant vond de huidige studie een significant grotere N400 voor logatomen in vergelijking met woorden. De huidige studie vond echter alleen een significant hogere N400 voor niet-geassocieerde woorden in vergelijking met geassocieerde woorden op de Fz-elektrode. Daarnaast werd de N400 in deze studie niet in verband gebracht met zelfrapportage-indicatoren van agressie en woede.

De belangrijkste conclusie die uit dit onderzoek kan worden getrokken is ten slotte dat de AIT alleen de associaties kan vinden tussen de ARIB in termen van RT, N400 en zelfrapportage-agressie-indicatoren in een forensische populatie en misschien niet gevoelig genoeg is om op betrouwbare wijze meer subtiele associaties te detecteren die mogelijk aanwezig zijn in gezonde populaties. Daarom zijn verdere studies nodig bij zeer agressieve (bijvoorbeeld forensische) deelnemers om de relatie tussen agressie en ARIB meer in detail te begrijpen.

Studie 4- Cognitieve gedragsverandering om negatieve en positieve attributies te moduleren bij individuen die hoog scoren op reactieve agressie.

Het primaire doel van deze studie was om te onderzoeken of zelfgerapporteerde agressie en woede gerelateerd zijn aan FRN en ERN tijdens een attributietraining voor gezichtsuitdrukkingen en om te bepalen of FRN en/of ERN gerelateerd zijn aan veranderingen in HAB van pre-tot post-training en om een beoordeling te geven van de relevantie van het FT-paradigma om de neurale correlaties van HAB in agressie te bepalen. De HAB is onderzocht met behulp van het experimentele paradigma *Face Task* (FT), dat is ontwikkeld door Penton-Voak et al. (2013). In de huidige studie werd verwacht dat hogere scores van zelfrapportage-agressie-indicatoren gepaard zouden gaan met kleinere ERN- en FRN-reacties. Daarnaast werd in de huidige studie verwacht dat grotere FRN- en ERN-responsen gepaard zouden gaan met een grotere verandering in HAB gedurende de hele training. Daarnaast verwachtte het huidige onderzoek dat er bij deelnemers met hogere scores op reactieve agressie vaker woedegezichtsuitdrukkingen zouden worden waargenomen dan bij deelnemers met lage scores tijdens de gezichtsuitdrukkingstraining. Ten slotte verwachtte het huidige onderzoek dat de deelnemers hun perceptie van ambigue emotionele uitingen tijdens de training zouden veranderen en er meer positieve gezichtsuitdrukkingen zouden worden waargenomen.

Het resultaat van deze studie toonde aan dat de training succesvol was in het verhogen van de positieve attributies van ambigue emotionele gezichtsuitdrukkingen. Bovendien werden lagere zelfgerapporteerde agressiescores geassocieerd met het classificeren van meer gezichten als gelukkige gezichten. Echter, de correlaties tussen HAB voor de training en reactieve agressie en woede waren niet significant. Daarnaast was de huidige studie ook succesvol in het vinden van de verwachte grotere FRN als reactie op negatieve feedback vergeleken met positieve feedback, en waren de ERN-amplituden groter voor incorrecte trials in vergelijking met correcte trials. Zowel FRN- als ERN-amplituden werden echter niet geassocieerd met zelfgerapporteerde indicatoren van agressie.

De belangrijkste conclusie die uit dit onderzoek kan worden getrokken is dat onze resultaten aantonen dat de FRN- en ERN-amplituden niet gevoelig genoeg zijn voor de effectiviteit van de cognitieve bias modificatie van de interpretatie (CBM-I) procedure bij een steekproef van individuen uit de gezonde populatie die laag scoren op zelfgerapporteerde agressie. Daarnaast is de aanpak die in de huidige studie wordt gehanteerd mogelijk beter voor klinische populaties en waarschijnlijk met een provocatieve manipulatie. Verder wordt er in deze studie geconcludeerd dat het FT-paradigma mogelijk alleen gevoelig genoeg is om betekenisvolle individuele verschillen in HAB te meten tussen (groepen) deelnemers die in extremere mate van elkaar verschillen op het gebied van agressie-indicatoren, wat niet het geval is binnen de gezonde populaties.

Algemene discussie en conclusie

Er moet erkend worden dat het huidige proefschrift mindere succesvol was in het vinden van een associatie tussen EEG- en prestatiemetingen (i.e. N400, ERN, en FRN) en zelfgerapporteerde indicatoren van agressie in diverse populaties van gezonde individuen. Deze resultaten zijn mogelijk te verklaren door het gebruik van gezonde individuen om agressief gedrag te onderzoeken, evenals het niet provoceren van deelnemers tijdens de experimenten. Om die reden suggereert het huidige proefschrift dat agressief gedrag als een vorm van psychopathologie een uitzondering zou kunnen zijn bij het bestuderen van de associatie met HAB. Dat wil zeggen, alleen binnen zeer agressieve populaties waarin het (eigen) agressief gedrag boven klinisch significante niveaus ligt, kan er HAB worden gevonden. Toekomstige studies moeten ook populaties includeren waarin het (eigen) agressief gedrag boven klinisch significante niveaus ligt. Daarnaast moeten toekomstige studies de deelnemers tijdens de experimenten provoceren om het verband te vinden tussen EEG- en prestatiemetingen (i.e. N400, ERN en FRN) en zelfrapportage-indicatoren van agressie te kunnen vinden. Ten slotte moet er worden opgemerkt dat het voor de onderzoekers onduidelijk blijft of de reactie op provocatie universeel is bij verschillende vormen van agressie en of gezonde personen verschillend reageren op dergelijke signalen.

Ten slotte, hoewel het huidige proefschrift geen associaties kon vinden tussen gedrags- en psychofysiologische reflecties van HAB aan de ene kant en zelf-rapporterende indicatoren van agressie aan de andere kant in gezonde mannelijke volwassen proefpersonen, repliceert dit proefschrift nog steeds eerdere bevindingen die aantoonden dat negatieve emotionele inhoud van een woord een groter N400-effect kan ontlokken. Bovendien is het huidige proefschrift een replicatie van eerdere bevindingen waaruit bleek dat de N400 welke ontlokt is door één woord een meer anterieure verspreiding lijkt te hebben met een maximum over frontale of centrale gebieden in het brein. Daarnaast is dit proefschrift erin geslaagd om eerdere bevindingen te repliceren die aantonen

S

dat de CBM-training succesvol is in het aanpassen van het beoogde bias op een positieve manier. Daarnaast was dit proefschrift succesvol in de replicatie van de eerdere bevindingen die grotere FRN-amplitudes laten zien als reactie op negatieve feedback vergeleken met positieve feedback, en grotere ERN-amplituden voor incorrecte trials in vergelijking met correcte trials. In het algemeen vult dit proefschrift een gat in de HAB-literatuur en biedt het belangrijke informatie voor de ontwikkeling van preventie inspanningen en toekomstige interventiestudies en onderstreept het de belangrijke rol van de feedbackverwerking in de ontwikkeling en het voortbestaan van HAB.

ACKNOWLEDGEMENTS

I would like to express my appreciation to Dr. Freddy van der Veen, Dr. Jorg Huijding, and Prof.Dr. Ingmar Franken for their support, encouragement, and guidance throughout this project. Special thanks to the Erasmus Behavioral Lab, Christiaan Tieman, Marcel, Boom, Michel Looije, and, Gerrit Jan Bie for their help in coding and programming the experiments, and for their valuable feedback during the past five years. I would like also to extend my gratitude to the people at the Department of Psychology at Erasmus University Rotterdam, especially all my Ph.D. friends for their helpful feedback and their beautiful friendship. Also, I would like to thank my parents, my brothers, and sisters for their love and support as I pursue my passion. Special thanks also go to my older sister (Fatimah) for setting me on a path that ultimately led to where I am today.

Finally, I would like to thank my wife (Fatimah) and my daughter (Shomookh) who were so patiently suffered from my busy schedules and inconsistencies during the writing of this thesis. Lastly, very special thanks go also to my wife for transferring some of her ideas into this thesis cover.

CURRICULUM VITAE

Ibrahim Qasem Hakami was born on October 30, 1984, in Jazan, Saudi Arabia. He obtained a bachelor's degree in Psychology from Imam Mohammad Ibn Saud Islamic University (Imam), in Riyadh 2007. Following his bachelor's degree, Ibrahim Hakami worked as a teaching assistant in which he taught basic psychological sciences including Clinical Psychology, Counseling Psychology, Psychotherapy, and Abnormal Psychology to bachelor students at Imam university. Following this, Ibrahim went on to complete his postgraduate studies in Clinical and Counseling Psychology at Chestnut Hill College, Pennsylvania, USA, and he was awarded a Master's of Science degree in 2012. During his Master's degree, Ibrahim took practical healthcare internship at TODAY, Inc in the USA in which he specialized in providing comprehensive mental health care services (Prevention, Intervention, Treatment, and Continuing Care) for individuals, families, and communities who are affected by substance abuse or any co-occurring mental health concerns. Immediately after obtaining his degree, Ibrahim started working as Lecturer at the faculty of the social science department of psychology, Imam University. Furthermore, Ibrahim did an internship as a psychologist at Al-Amal Hospital in Riyadh, Saudi Arabia in which he specialized in providing mental health care and drug addiction services. In 2015, Ibrahim granted funding from Imam University to start his Ph.D. program at Erasmus University Rotterdam, the Netherlands under the supervision of Prof. dr. Ingmar Franken. dr Freddy van der Veen, dr. Jorg Huijding. During his Ph.D., Ibrahim focused on employing electroencephalogram (EEG), to identify the underlying neural processes associated with Hostile Attributional Bias. Specifically, his work focused on examining aggression through the use of brain activity Event-Related Potential (ERPs) in order to shed light on the early stages of information processing, resulting in the present thesis. Currently, Ibrahim works as an assistant professor in the department of psychology at Imam university.

Submitted manuscripts

Hakami, I., F.M. van der Veen, Franken., I.H.A. & Huijding, J. Can We Get Along? The relationship between Feedback Related Negativity and Reactive Aggression in Healthy Individuals During Social Feedback Processing.

Hakami, I., F.M. van der Veen., Yu Yuen, Chak., Wieser, M. J., van Strien, J. W., & Franken, I.H.A. Emotional modulation of the N400: Manipulating the emotional meaning of homonyms.

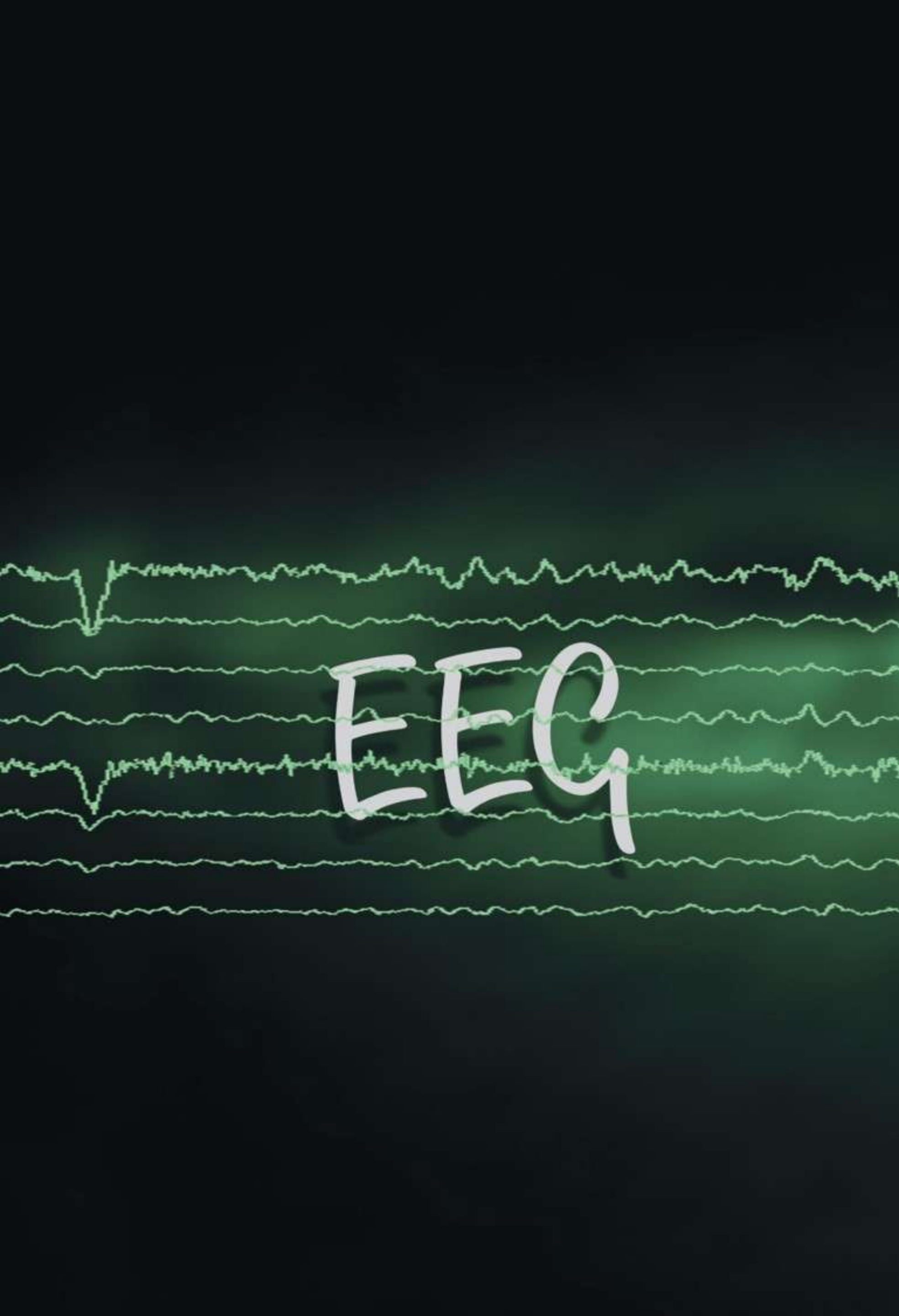
Hakami, I., F.M. van der Veen, Franken., I.H.A. & Huijding, J. No Effect of Self-Report Aggression Measures on The N400 in A Lexical Decision Task with Associated Words in a Violent Context.

Hakami, I., F.M. van der Veen, Franken., I.H.A. & Huijding, J. Cognitive Behavioral Modification to modulating Negative and Positive Attributions in Individuals scoring High on Reactive Aggression Measures.

Conference presentation

Hakami, Ibrahim, Frederik M. van der Veen, Jorg Huijding, Ingmar.H.A. Franken (2018). Can we get along? Feedback Related Negativity for Reactive Aggression individuals during social feedback possessing. The 58th Annual Meeting of the Society for Psychophysiological Research (SPR), Quebec City Convention, Canada.

Hakami, Ibrahim, Frederik M. van der Veen, Jorg Huijding, Ingmar.H.A. Franken (2019). Are you for real? : Interpretation Bias in Individuals with high reactive Aggression. The Graduate Research Day 2019, Erasmus University Rotterdam, the Netherlands.


EEG