The neural correlates of attention bias and interpretation bias in aggression

Rebecca Crago

Thesis submitted for the degree of Doctor of Philosophy.

University of East Anglia
School of Psychology

May 2019

Abstract

Preferentially allocating attention towards hostile stimuli, and attributing hostile intent towards ambiguous stimuli, is thought to contribute to the aetiology of aggression. Using behavioural and ERP methodology, across five studies, this thesis investigated the neural correlates of attention and interpretation bias within aggression. The first four studies explored attention bias towards angry, happy and neutral stimuli across two stimulus types; words and faces. Behavioural results showed a significant correlation between aggression and increased reaction time to probes replacing hostile words and angry faces. However, this effect was not replicated in the follow up studies for either modality. Overall, the ERP results showed significant effects of congruency (evoked P1/P300 amplitudes differed between probe positions, following the simultaneous presentation of two stimuli) across all studies. However, these effects did not always interact with aggression. Nevertheless, study three indicated that low aggression participants differentiated between angry and neutral faces, whereas, high aggression participants had relatively stable amplitudes. Interestingly, results showed differences in ERP patterns when participants responded to different modalities of stimuli. The findings suggest that angry faces are subject to automatic processing and therefore demand attentional resources. However, hostile words may be subject to slower processing and may not grab attention in the same way as angry faces. The final study used a recognition task to investigate neural correlates of interpretation bias. Behavioural results revealed between-group differences suggesting that aggressive individuals had an increased hostility-related interpretation bias. Largely, the interpretation bias ERP results mirrored those found across the attention bias studies, although processes relating to interpretation bias influence the later LPP component. I believe the original design of the studies presented in this thesis, and the subsequent findings, contribute to the understanding of attention and interpretation biases in aggression. Based on previous results, attention and interpretation theories, and current findings, I consider how cognitive biases may contribute to the maintenance of aggression and make recommendations for future work.

Table of contents

List of Tables

List of Figures

Acknowledgements

I would firstly like to thank my supervisors, Dr Laura Biggart, Dr Louis Renoult, Dr Gavin Nobes, and Dr Jennifer Bowler. I am very grateful for the support, advice and encouragement shown to me over the last three years. A special thank you to Jennifer for her help with programming and data collection. Without this vital contribution this thesis would not have been possible.

I would like to thank Tamara Satmarean for her assistance with data collection, and all the participants that gave up time to take part in the studies. I am also grateful for all the support shown to me by the staff within the School of Psychology at UEA, and colleagues turned friends that I have shared the highs and lows with during this journey.

A would like to thank my family for their endless support and constant encouragement, especially my parents who have always provided me with the best advice and given me the confidence to achieve my goals. Finally thanks to Lee for always believing in me – your words of strength, love and kindness have proved invaluable.

Author's declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the School of Psychology Ethics Committee at the University of East Anglia.

Name: Rebecca Crago

1 Introduction

Aggression generally refers to behaviour which causes harm or distress to another (De Castro, Veerman, Koops, Bosch, & Monshouwer, 2002). Aggressive behaviour plays an important role in many, if not all, violent crimes. However, a relatively small number of individuals are responsible for these crimes (Brooks-Crozier, 2011). Identifying individuals with a predisposition for aggressive behaviour could have important implications for dealing with violent offenders within the criminal justice system. Furthermore, understanding mechanisms that underlie aggressive behaviour and the cognitive processes that contribute to a violent and potentially criminal offence is crucial in designing prevention, intervention and rehabilitation policy and practice.

Cognitive biases are likely to play a significant role in the aetiology of aggression (e.g. (Dodge & Frame, 1982; Smith & Waterman, 2003). Cognition is a general term which describes the many processing stages which occur between stimulus presentation and response (Pashler & Sutherland, 1998). These can include selective attention, interpretation, memory and judgement (Weems, Costa, Watts, Taylor, & Cannon, 2007). Cognitive biases refer to differential processing at any of these stages. In particular this thesis focuses on attention bias and interpretation bias. Attention bias is the process in which individuals preferentially allocate attention towards hostile or threatening stimuli relative to neutral stimuli (MacLeod, Mathews, & Tata, 1986). A negative interpretation bias is the process in which individuals interpret ambiguous or mildly aversive actions as more negative and dangerous, and attribute a greater level of hostile intent to these scenarios (Waters, Craske, Bergman, & Treanor, 2008a), for example, interpreting an accidental push in a crowd as an act of provocation with hurtful intent. Cognitive biases that occur in response to inappropriate situations or environments may result in maladaptive behaviours such as aggression.

The thesis starts with a detailed overview of the literature which examines the current literature on three main topics; attention bias, interpretation bias and aggression. EEG techniques are then introduced before reviewing why this methodology may be useful for understanding neural processes associated with cognitive biases. To my knowledge very few studies have explored neural correlates of cognitive biases in aggression, therefore the literature review includes some relevant studies on anxiety, particularly when evaluating the methods used for measuring attention bias.

Across five studies, novel neurological methods were used to investigate neural correlates of cognitive biases in aggression. Attention bias to angry stimuli was investigated across two sets of studies: the first including word stimuli, and the second including faces. The first study investigated responses to a selective attention task which included angry and neutral word stimuli. Building upon possible limitations that emotionality and anger is confounded in Study 1, and to further explore the attentional bias effects in response to stimuli of different valence, the follow up study (Study 2) included angry-neutral word pairs, along with happy-neutral and angry-happy word pairs. The aim was to investigate the differences in attention bias, and the associated ERP correlates, between stimulus modalities (words and faces). Therefore the second set of studies (Study 3 and 4) were methodologically identical to the first set of studies (1 and 2). Attentional bias effects in response to angry-neutral face pairs (Study 2) and in response to happy-neutral and angry-happy face pairs (Study 4) were investigated. The final study (Study 5) investigated responses to a recognition task, with the aim of better understanding interpretation bias in aggression. For all five studies EEG was recorded during task completion and therefore between-group differences in behavioural data (reaction time) and ERP data (evoked amplitude) within an aggressive sample was explored.

I believe that identifying the distinct mechanisms that contribute to cognitive biases may help explain why environmental stimuli provoke aggressive responses in some individuals. Understanding these processes can inform rehabilitation programmes; if cognitive biases contribute to a behavioural response, change in cognitions may result in a change in behaviour.

2 Literature Review

2.1 Attention Bias

2.1.1 Introduction to attention bias

Cognitive biases are not a direct cause of social behaviour, however they act as mediating processes that connect biological, environmental and situational inputs to behavioural outputs (Huesmann, 1998). Attention bias is a cognitive bias described as a process in which individuals show differential allocation of attention towards hostile or threatening stimuli relative to neutral stimuli (MacLeod et al., 1986). It is proposed that there are three operations when attending to a new stimulus (e.g., Posner, 1980; Posner & Petersen, 1990). The first is an initial transient shift of attention to the stimulus, the second is engagement of attention and the third is disengaging attention from the stimulus. Attention bias refers to differential or maladaptive cognitive processing at one or more of these three stages and can include facilitated engagement, difficulty in disengagement and attentional avoidance (see Cisler & Koster, 2010; Koster, Crombez, Verschuere, Van Damme, & Wiersema, 2006). Facilitated engagement refers to the process in which threat-related stimuli are detected faster than neutral stimuli. Facilitated processing is a mechanism by which threatening information is prioritized by the preferential orientating of spatial attention in its location (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & Van Ijzendoorn, 2007). Difficulty in disengagement refers to the difficulty in allocating resources away from threat-related stimuli once it has been engaged. Attentional avoidance refers to the avoidance of allocating attention towards threat-related stimuli and directing attention to stimuli located in opposite locations (Koster et al., 2006).

Within the literature there are a number of terms used when referring to negatively biased attention. These include; threat-related attention bias (usually used within the anxiety literature), and hostile-, aggression-, and anger-related attention bias. These all imply preferential allocation of attention to aversive stimuli compared to neutral stimuli; however the type of aversive stimuli used may vary slightly across studies. Although these terms will be referred to throughout the literature review when discussing previous work, when presenting the studies within this thesis the term aggression-related attention bias will be used.

2.1.2 Theories of attention bias

Within a social environment, fast and efficient detection of threat is critical for survival. According to Darwinian evolution theory, the neurocognitive mechanisms involved with attending to threat stimuli may reflect an adaptive advantage (Darwin & Darwin, 2009). Adaptive heuristics may reflect the tendency for individuals to be prepared for the worst, therefore these biases may result from relying on automatic processing rather than rationality (Gilbert, 1998). This may be particularly relevant to individuals that demonstrate an attention bias towards hostile stimuli as they are hypervigiliant when detecting threat (Nesse, 1998; Nesse, 1994). Hyper-vigilance for threat can be maladaptive as consistently allocating attentional resources to threatening and fear-inducing stimuli can reinforce psychopathological symptoms. The presence of habitual attention biases can contribute to the maintenance of behavioural outcomes such as anxiety (Bar-Haim et al., 2007; Muris & Field, 2008), depression (Kovacs & Beck, 1978; Mogg, Bradley, & Williams, 1995b), schizophrenia (Green, Williams, & Davidson, 2003), and aggression (Smith & Waterman, 2005; Wilkowski & Robinson, 2008b).

Basic premises of current models of attention (e.g., Cisler & Koster, 2010; Mathews & Mackintosh, 1998) support this evolutionary approach. These models suggest that attention to threat is determined by both task demands and stimulus input. Attention to current tasks and ongoing behaviour is interrupted when stimulus input exceeds a certain threshold and is appraised as highly threatening (Mathews & Mackintosh, 1998; Mogg & Bradley, 1998). Individuals with maladaptive attention bias mechanisms have overly sensitive threat appraisal systems that more readily evaluate incoming stimuli as hostile (Koster et al., 2006).

2.1.3 Assessment of methods for measuring attention bias

Most commonly attention bias research has been concerned with the association between attention bias to threat and anxiety. However, similar methods have been adopted from the anxiety literature in order to investigate aggression-related attention biases. This section of the literature review will give a summary of the research on attention biases in anxiety with the aim of evaluating the validity of methods used to measure attention bias.

There is substantial evidence to suggest that anxious individuals consistently show an attention bias towards threatening stimuli. For example, a meta-analysis by (Bar-Haim et al., 2007) examined threat-related attention biases in anxiety and found that across multiple paradigms and under different experimental conditions, anxious individuals consistently show significantly faster reaction times when responding to threat-related stimuli compared to neutral stimuli, suggesting a heightened vigilance for threat. It is suggested that an increased attention to such stimuli within the environment may cause a heightened perception of danger, which in turn reinforces anxious feelings (van Honk, Tuiten, de Haan, vann de Hout, & Stam, 2001a).

Measurement of these attention biases is usually based on reaction time during behavioural cognitive tasks such as the Emotional Stroop task or dot-probe tasks. The Emotional Stroop task is adapted from the original Stroop task in which participants are required to define the colour of written colour names (e.g. red, green or blue) (Stroop, 1935). Some of the colours are printed in their true colour; however others are printed in a colour ink which is different to the name of that colour (for example, the word ‘green’ printed in red ink). Naming the colour of words that are printed in a different colour ink takes longer and is more prone to errors compared to naming the colour of words that are printed in their true colour. The emotional Stroop is an adaptation of the original paradigm; instead of colour names being presented, words of positive and negative emotional valence are displayed (for example, Gotlib & McCann, 1984; Mathews & MacLeod, 1985; Watts, McKenna, Sharrock, & Trezise, 1986).

Mathews and MacLeod (1985) conducted one of the first studies to show that anxious participants take longer to name the colour of a threatening word in comparison to a neutral word, suggesting that there is greater interference of the meaning of the threatening words and an underlying attention bias. Other work has shown this consistent effect using the modified Stroop task (Mathews & MacLeod, 1994; Mogg, Mathews, & Weinman, 1989; Richards & French, 1990; Richards & Millwood, 1989). However, there are a number of criticisms of this method; firstly it is unclear if an increased delay in colour naming of emotional words reflects an attention bias for attending to threat stimulus, or a delay in response generation (Mogg, Millar & Bradley, 2000); secondly, reaction times may not reflect a true measure of selective attentional processes (Fox, 1993); and finally this task, along with other adapted versions such as the emotional Stroop task, predominantly use words as stimuli.

Another task used to measure selective attentional processes is the visual dot-probe task (MacLeod et al., 1986). Compared with tasks that present stimuli centrally one at a time (Bishop, 2008; MacLeod et al., 1986), this task allows for a more direct assessment of competition models of attentional selectivity (e.g., (Desimone & Duncan, 1995) by presenting aversive and benign items simultaneously. Studies investigating attentional processes in relation to anxiety have predominantly included one threatening and one neutral stimuli. Participants are required to respond to probes (targets) which appear in place of previously presented stimuli. A faster response to probes that appear in place of threatening stimuli, in comparison to probes that appear in the place of neutral stimuli is thought to reflect a vigilance to threat (Koster, Crombez, Verschuere, & De Houwer, 2004).

MacLeod et al. (1986) completed one of the first studies to implement the dot-probe paradigm. The sample consisted of 16 participants referred by a general practitioner for anxiety management training, and 16 undergraduate controls. Participants completed a standard dot-probe paradigm consisting of threatening and neutral words. (MacLeod et al., 1986) found that clinically anxious participants had reduced latencies when detecting probes appearing in the prior location of threat words compared with controls. It was concluded that anxious participants have a consistent attentional shift towards threat words. This finding was replicated by Mogg, Mathews, and Eysenck (1992), and Mogg, Philippot, and Bradley (2004) who demonstrated that participants with social phobia had increased vigilance for angry faces compared to happy and neutral faces during a visual probe task, in comparison to a non-clinical control group.

Heightened vigilance towards threat stimuli has been evidenced in clinical (Bradley, Mogg, White, Groom, & Bono, 1999; Mogg et al., 2004), and non-clinical samples (Bradley, Mogg, Falla & Hamilton, 1998; Mogg, Bradley, De Bono, & Painter, 1997). Furthermore, Mogg et al. (1995) found that clinically anxious individuals showed an attention bias towards probes presented in the position of previously presented negative words under both supraliminal and subliminal conditions. This shows evidence that attention biases are evident at both pre-conscious (automatic), and post-conscious stages of cognitive processing. More recent literature supports the fairly consistent relationship between anxiety and threat-related attention bias (for example, Koster et al., 2004; Roy et al., 2008; Telzer et al., 2008; Waters, Mogg, Bradley, & Pine, 2008b). Furthermore, a review by Mogg and Bradley (2005) demonstrated that attention bias effects in individuals experiencing GAD were consistent across both the Stroop and visual dot-probe tasks.

2.1.4 Attention bias and aggression

Trait anxiety and trait anger are considered to be consistent and stable aspects of an individual's temperament (Spielberger, Jacobs, Russell, & Crane, 1983). It has been shown that attention bias is a relatively consistent characteristic of trait anxiety and therefore attention bias associated with anger may also be an important component contributing to anger-related behaviours. This section of the literature review will outline and evaluate different definitions of aggression before reviewing the theoretical explanations for the relationship between attention bias and aggression.

2.1.4.1 The operationalisation of aggression

Across studies aggression is generally defined as a behaviour which causes harm or hindrance to another (De Castro et al., 2002). Similarly, further explanations describe human aggression as a behaviour directed toward another individual that is carried out with the belief that it will harm another individual and the victim will be motivated to avoid the consequences of the behaviour (Anderson & Bushman, 2002; Baron & Richardson, 2004; Berkowitz, 1993; Bushman & Anderson, 2001; Geen, 2001). The most commonly used operationalisation of aggression is the Aggression Questionnaire (Buss & Perry, 1992), which includes items such as ‘I have threatened people I know’, ‘Given enough provocation, I may hit another person’ and ‘Sometimes I fly off the handle for no good reason’. This measure records the likelihood of an individual’s participation in an aggressive act.

Anger is conceptualised as the disposition to experience intense feelings of irritation or rage frequently and for long periods of time (Spielberger et al., 1983). Aggression is the behavioural expression of the anger emotion. Findings suggest that anger is related to both physical and verbal aggression (Hazaleus & Deffenbacher, 1986; Parrott & Zeichner, 2002). Aggressive behaviour has been divided into two main subtypes; impulsive, sometimes known as reactive aggression, and premeditated proactive aggression (Houston, Stanford, Villemarette Pittman, Conklin, & Helfritz, 2004). Reactive aggression refers to angry, emotional or affective aggression which is generally expressed in a physical behavioural response following provocation; this is compared to proactive aggression which is more often premeditated and is motivated by a desire for dominance (Dodge & Coie, 1987). Each of these types of aggression may have particular effects on cognition and attentional processes.

There is evidence to suggest that males and females may demonstrate different types of aggression in response to feelings of anger. Bjorkqvist, Osterman, and Lagerspetz (1994) suggested females are more likely to adopt an indirect expression of aggression. Similarly, Archer (2004) conducted a meta-analytic review of sex differences in aggression in real life settings. Results suggested that the greatest sex differences were in physical aggression, with males having increased levels of physical aggression compared with females. Differences in verbal aggression were smaller but males still had increased levels compared with females. Interestingly there were no sex differences in levels of anger; this suggests that both males and females experience feelings of anger, but males are more likely to express them in a physical way. This is a relatively consistent finding across the literature with another meta-analysis (Knight, Fabes, & Higgins, 1996) and a longitudinal study (Cairns, Cairns, Neckerman, Ferguson, & Gariepy, 1989) also finding evidence of this. The majority of studies investigating aggression-related attention biases have utilised mixed gender samples, and do not take into consideration these potential differences in anger expression. As males are more likely to express anger with a physically aggressive response, understanding cognition in relation to male aggression may have greater implications for intervention and rehabilitative methods as they can be adapted to the particular needs of male aggressors.

2.1.4.2 Theoretical explanations for attention bias and aggression

Literature suggests that similar attentional processes can result in feelings of anxiety or anger depending on the appraisal of presented stimuli. Dimberg and Öhman (1996) stated that the relationship between the sender and receiver of information is essential for the appraisal of stimuli, suggesting that angry faces can be met with anger or anxiety. Anger is a response to perceived provocation and therefore the receiver is motivated to aggressively confront and remove the threat, whereas anxiety is a response to perceived fear and therefore the receiver is motivated to avoid the threat (Smith, McHugo, & Kappas, 1996). This research suggests that attentional processes inform interpretation processes and these determine whether an anxious or aggressive response is provoked. This is consistent with the work of Crick and Dodge (1994) whose social information processing model highlights six steps in encoding, evaluating and responding to the environment.

The social information processing model aims to explain the cognitive process which occur between stimulus presentation and response (Pashler & Sutherland, 1998). Cognitive processing of the environment has an impact on the subsequent behaviour enacted in a particular situation. These behaviours then become the foundation of social adjustment evaluations made by others (Ladd & Mize, 1983; Rubin & Krasnor, 1986). Crick and Dodge's (1994) social processing model (Figure 1) provides a description of how individuals perceive and understand their surroundings. In order to engage and react to social situations appropriately, Crick and Dodge (1994) propose that a number of steps must be followed.

During stage one it is proposed that individuals selectively attend to both internal and external cues, with each stimulus having to be correctly encoded. Stage two which involves forming an accurate representation of the stimuli, occurs either immediately after or during stimulus encoding. As encoding and interpretation are integrated processes, each informs the other. Interpretation of cues may depend on a number of independent processes, including, accessing mental representations stored in long term memory, event and goal analysis specific to the situation, perspective taking, evaluation of past experience, and inferring the meaning of the situation. This stage (two) can be influenced by scripts, schemata and social knowledge, previously stored in memory. During stage three the individual must select a desired outcome. Next, a number of possible responses are generated, these may be new behaviours formed in response to a novel situation or they may be accessed from memory. At stage five a process of evaluation of each response for possible consequences and outcomes is applied. Finally, it is hypothesized that at stage six, the chosen response is regulated and enacted using protocols and scripts.

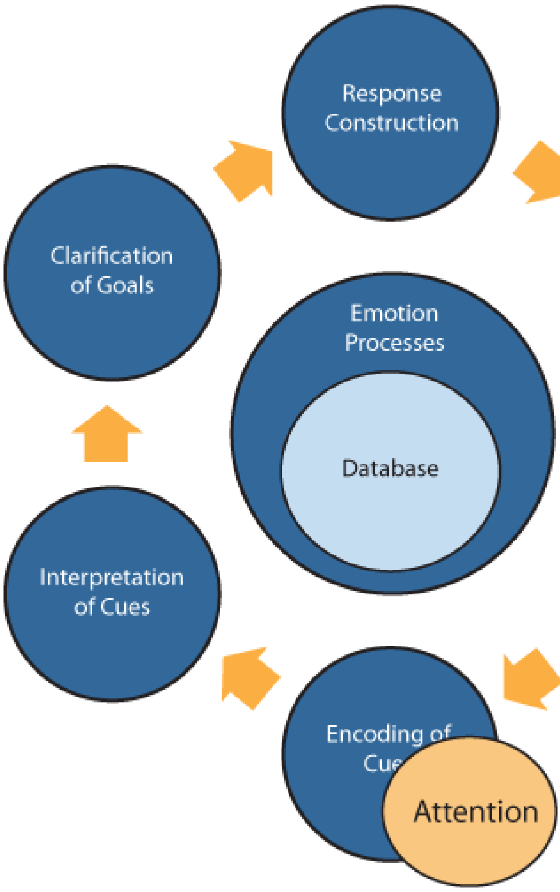

*Figure 1: S*ocial information processing model (Crick & Dodge, 1994).

The reformulated model addresses criticisms made by connectionist theorists who propose that processing stages occur simultaneously along a number of parallel paths and does not follow a rigid sequential structure (Feldman & Ballard, 1982).

Crick and Dodge (1994) suggest that maladjusted behaviour is a result of maladaptive processing at any of these stages. More specifically attention bias refers to a maladaptive process in which individuals engage in greater processing of aggression-related cues over non-aggressive-related cues. Cognitive theories of aggression and antisocial behaviour propose that these biases are evident in aggressive populations (e.g. Coccaro, Noblett, & McCloskey, 2009; Crick & Dodge, 1994; Dodge & Frame, 1982; Dodge, Pettit, McClaskey, Brown, & Gottman, 1986; Smith & Waterman, 2003). Attention bias may be a particularly important component of aggressive behaviour as it is the first stage of information processing and has consequences for all following cognitive processes.

The trait-congruency hypothesis (Blaney, 1986; Bower, 1981; Miranda & Persons, 1988) suggests that personality traits have a direct impact on the cognitive processes used when attending to the environment. This hypothesis states that affective traits such as anxiety and anger are linked to activation of the relevant emotion networks. Therefore trait-related cognitive biases increase the likelihood of the experience of a particular emotion. For example, high trait anger individuals may bias attention towards anger-related stimuli, have increased recall of anger-related information and process hostile cues more efficiently. This theory suggests that facilitative biases allow for quicker processing of emotional stimuli which are congruent with internal traits.

There are also suggestions in the literature that threat and non-threat stimuli are processed differently and this can influence reaction time results (visual attention bias). The amygdala is centrally involved in activating an approach (fight) or avoidance (flight) strategy in response to threat stimuli (Schulkin, 2003). This fight-flight response is a fast and effective system used to process stimuli in the environment which may be harmful. In contrast, non-threat stimuli are assessed using a slower system which passes the pre-frontal cortex and are evaluated before action (LeDoux, 2003). Aggressive individuals who demonstrate an attention bias towards hostile stimuli may be particularly sensitive in detecting threat (Nesse, 1998; Nesse, 1994) and consistently rely on the rapid 'fight' response when processing such stimuli. This quick and efficient mechanism could contribute to a hostile-related attention bias and may be aided by commonly used schemas.

A characteristic of attention bias to threat is difficulty in disengaging from potentially harmful stimuli (see Cisler & Koster, 2010). Poor attempts to regulate attentional control can result in fixations on threat-related stimuli and poor task performance. Eysenck and colleagues (Eysenck, Derakshan, Santos, & Calvo, 2007) stated that attentional control theory can be used to explain attention bias in anxiety. He suggested that poor attentional control contributes to attention bias as it disrupts two executive functions. The first of these is inhibition; this refers to the ability to regulate, or inhibit when necessary, dominant automatic processes. The second function is shifting, which refers to the ability to shift attention successfully between tasks contexts or operations. Although this theory was based on anxiety-induced attention biases, similar processes are involved in aggression-induced attention biases and therefore this theory can be helpful in understanding how attentional control may influence levels of hostility-related biases in aggression, particularly as aggression has also been linked with poor attentional control (e.g. Meesters, Muris, & van Rooijen, 2007). Eysenck et al. (2007) would argue that aggression disrupts the balance between stimulus-driven and goal-driven processes (e.g. Corbetta & Shulman, 2002; Posner & Petersen, 1990) by impairing inhibition and therefore weakening top-down regulatory control.

A number of theorists have proposed explanations for the relationship between hostility-related attention biases and aggressive behaviour. Generally these theories suggest that hostility-related selective attention, which drives aggression, is the product of increased stimulus-driven attentional capture by angry cues (somewhat aided by existing schema), combined with suboptimal effortful regulatory control (e.g., Strack & Deutsch, 2004; Wilkowski & Robinson, 2010). Biased selectivity in aggression is particularly associated with later stages of attention when ruminative processes and difficulties in disengaging from hostile stimuli can influence attentional capture (Wilkowski & Robinson, 2010). Wilkowski and Robinson (2008a) conducted a review of current proposed cognitive models of trait anger and attention bias. They integrated the findings and concluded that anger reactivity can be primarily explained by not only ruminative attention and effortful control, but also automatic hostile interpretations. This shows that vigilance to negative stimuli contributes to aggressive behaviour but other cognitive processes should be considered. This is consistent with the Social Information Processing model proposed by Crick and Dodge (1996) who state six key steps in processing information encountered in the environment, in which attention and interpretation are two essential phases in response formation.

2.1.5 Attention bias to word stimuli

2.1.5.1 Attention bias to angry words

Within the attention bias literature there is little consideration of how distinct stimulus types may differentially affect attentional processes. The aim of this section of the literature review is to outline and evaluate the relationship between aggression and negative attention bias by reviewing studies which have included word stimuli within experimental paradigms to measure attention bias. The research investigating attention bias to face stimuli will be reviewed in the next section (Section 2.1.6).

Smith and Waterman (2005) investigated processing biases to an Emotional Stroop task in undergraduate males and females categorised according to their self-reported aggression score. They found that males had a significantly delayed response to colour naming ‘direct aggression’ words, showing an attention bias towards such stimuli. Females showed some delay in colour naming ‘indirect aggression’ words, although this did not reach significance. Interestingly, the results also indicated that physical aggression was the best predictor for hostile-related attention biases in both males and females. This research makes a number of contributions: first, it shows that attention biases are evident in non-clinical normative samples; second, it illustrates that processing biases may also be essential in understanding female aggression along with male aggression; and finally, it suggests that attention biases may be particularly salient in individuals with high levels of physical aggression.

van Honk, et al. (2001b) also administered the Emotional Stroop task, which included threatening and neutral words, to male and female participants categorised into high and low trait anger groups based on self-reports of trait anger. Participants completed the task in both masked and unmasked conditions. The unmasked task consisted of a fixation cross followed by presentation of the target word in one of four colours. During the masked conditions, following the appearance of the target word, a mask showing a random string of rotated and reversed letters, in the same colour as the target stimuli, was presented on screen. Participants were instructed to name the colour of the target word as quickly as possible. Results showed differences in responses between the high and low trait anger groups for the unmasked task only. High trait anger participants took relatively longer to colour name the threatening words in comparison to the neutral words, whereas low trait anger participants were quicker to name the colour of the threatening words. The authors propose that these findings can be attributed to interference of meaning of the threatening word in the high anger group, and to facilitation in the low anger group. Attentional facilitation refers to the allocation of resources used in detecting threat-related information. The predominant role of fear is to facilitate detection of threat by allocation of attentional resources (LeDoux, 1996). Attentional interference refers to difficulty in disengaging from threat-relevant information which then restricts processing resources needed for another task (Fox, Russo, Bowles, & Dutton, 2001; Fox, Russo, & Dutton, 2002). This suggests that processing of threat-related stimuli in aggressive individuals interferes with their ability to complete the task (e.g. naming the colour of the word). This evidence indicates that attention is directed towards aggressive words and that such stimuli are processed and evaluated to a greater extent, suggesting difficulties with disengaging from aggression-related stimuli. It is the combination of these factors which contribute to an attention bias across tasks.

Hostility-related attention biases are evident in both forensic (offender) and non-forensic (undergraduate) samples. (Smith & Waterman, 2003) assessed attention bias towards violently themed stimuli across these samples using both an emotional Stroop and dot-probe task. Aggression was defined by index offence in the offending population and by self-reported anger (measured using anger subscale of The Aggression Questionnaire (Buss & Perry, 1992) in the undergraduate population. In both samples, and across both tasks, aggressive individuals showed a response bias for the violently themed words, when compared with non-aggressive controls. Aggressive participants responded more quickly to probes that replaced the aggressively themed words in the dot-probe task, and showed greater interference in colour-naming the aggression words compared with neutral words in the Stroop task. This study is consistent with previous findings that violent stimuli may be particularly salient to aggressive individuals. It also shows that aggression-related attention biases are not only observable in forensic populations but in individuals from normative samples with relatively higher anger scores.

Further evidence by Chan, Raine, and Lee (2010) shows that male batterers may allocate more attentional resources to aggressive words which may have consequences for attending to the environment. They used an emotional Stroop task to measure reaction times to colour naming of aggressive and neutral words. They found that batterers had longer reaction times when naming the colour of negative words when compared to neutral words. This effect was not displayed in the control group. This suggests that physically aggressive males show an attention bias to aggressive words. Chan et al. (2010) also found that batterers scored particularly high on reactive aggression. This finding is consistent with theoretical accounts of aggression, based on the frustration-aggression model (Berkowitz, 1993; Dollard, Miller, Doob, Mowrer, & Sears, 1939).

Although research has shown a significant attention bias effect for violent words amongst aggressive populations, other research suggests that these biases are only evident under provoked circumstances (see Cohen, Eckhardt, & Schagat, 1998; Eckhardt & Cohen, 1997). Eckhardt and Cohen, (1997) investigated attention biases towards mood-congruent stimuli in high and low trait anger individuals following an insult. Participants scoring in the upper and lower third of the Trait Anger Scale (Spielberger et al., 1983) were categorised into high and low trait groups respectively. Within each group half of the participants were allocated to an insult group, whereas the other half received no insult. A modified version of the emotional Stroop task including anger, positive emotion and neutral words, was administered. Subjects were shown a target colour before each trial in which the words were presented in colour ink. They were required to indicate if the colour of the ink was the same as the previously presented target colour. Those participants in the insult condition were called an offensive name, by an accomplice of the experimenter, while on route to the laboratory. Results showed that high trait anger subjects took longer to colour name the anger words in comparison to both positive and neutral words, but this effect was only evident under the *insult* conditions. Low trait anger participants showed no attention bias effects in either the *insult* or *no insult* condition. This shows evidence for the mood-congruency hypothesis (Miranda & Persons, 1988) and suggests that attention biases towards anger-related stimuli are only evident under provoked situations where levels of both trait and state anger are high.

Studies investigating attention biases in aggression have used varying methods of conceptualising aggression. Some have used trait anger (Eckhardt & Cohen, 1997; van Honk et al., 2001a) or an anger subscale (Smith & Waterman, 2003), while others have used physical and verbal aggression subscales (Smith & Waterman, 2005). Primarily trait anger is measured as this is considered a consistent internal characteristic, however this relates to feelings of anger and does not necessarily imply an aggressive response. Therefore, it is not clear whether levels of aggression can be inferred from trait anger scores.

The literature assessing hostility-related attention biases to words in aggressive populations have revealed a number of key findings. Firstly there is evidence to suggest that cognitive biases are evident in normative and forensic samples, meaning that implications can target individuals with high trait aggression as well as violent offenders. Secondly it has been proposed that vigilance for hostile-related stimuli is particularly heightened in individuals with increased levels of physical aggression. Research suggests an association between trait anger and attention bias, however further evidence suggests that processing biases are only observable when both trait and state anger are high.

2.1.5.2 Attention bias to positive words

The research reviewed so far demonstrates that aggressive individuals preferentially attend to hostility-related words compared to neutral words, however it is not clear if these biases remain during experimental paradigms which include positively valenced stimuli. There is very little research which investigates cognitive processes involved with selectively attending to happy stimuli in aggressive samples. Findings from Smith and Waterman (2003) (reviewed in Section 2.1.5.1) show that during a Stroop task in which aggression themed, positive emotion, negative emotion, colour, or neutral words were presented, aggressive groups were slower to name the colour of the aggression-themed word compared to the neutral word. However, the results showed no significant differences in colour naming positive emotion words between groups, suggesting that levels of aggression do not influence patterns of attention to positive emotion words. Although this study includes both positive and negative emotion words, it does not compare the differences in reaction times between these trial types as bias scores were calculated by subtracting the reaction time to the neutral word from each of the other word types. Also, due to the singular presentation of word stimuli during the Stroop task, this does not allow for the measurement of selective attentional processes.

Brugman et al. (2015) explored the predictive value of attention bias on reactive and proactive aggressive behaviour in a non-clinical sample. Participants were required to complete an emotional Stroop task in which neutral, negative, positive and aggression-related words were presented in different colours. Self-rated aggression was measured using the Reactive Proactive Aggression Questionnaire (Raine et al., 2006), and aggressive behaviour was measured using the Taylor Aggression Paradigm (TAP; Taylor, 1967). The study found contrasting attention bias effects for self-reported proactive aggression and reactive aggression on the TAP. Slower colour naming of aggression words, suggesting increased interference, was predictive of increased reactive behaviour on the TAP, whereas faster colour naming of aggressive words in comparison so negative words resulted in a higher level of proactive aggression. These contrasting findings suggest that processes contributing to attention biases in aggression may vary depending on the form of aggression studied. Brugman et al. (2015) suggest that individuals with high levels of proactive aggression may not find aggression words emotionally disturbing and therefore are not allocated any greater attentional resources compared to different word types. While this study included positive word types, the results regarding the association between aggression types and attention bias to happy words were not reported. This suggests that the association between aggression and reaction time to colour naming positive words yielded a null result.

Although knowledge on attention bias to positive words in aggression is somewhat limited, there are a number of studies which have used anxious and non-clinical samples that can be drawn upon. Firstly, Pishyar, Harris, and Menzies (2004) investigated attention bias in self-rated anxiety using a dot-probe task that consisted of negative and neutral pairs and positive and neutral pairs. Participants completed two tasks, one with word stimuli and one with visual stimuli. During the word task, the results showed no significant differences across either stimuli pairing. This study utilised a non-clinical sample, and therefore suggests that attention biases may not be evident in low levels of anxiety. Martin, Williams, and Clark (1991) conducted four experiments which investigated attention biases in anxiety using the Stroop task. Participants with generalized anxiety disorder were slower to name both threat-related and positive words, compared to non-anxious controls. This shows that anxious individuals show an attention bias towards both negative and positive emotion, suggesting greater processing of such stimuli. This suggests that threat and emotion may have been confounded in previous studies, and the authors propose that future work is needed explore whether attention bias to threat in anxiety represents an attention bias for emotionally provoking stimuli. A limitation of this study is that the Stroop task does not allow for measurement of selective attention as stimuli are presented singularly.

Further work by Sutton and Altarriba (2011) investigated attention bias to negative and positive emotion words in a non-clinical sample. Across two experiments the researchers explored effects of attentional processing of positive-neutral and negative-neutral word pairs under masked and non-masked conditions during a modified dot-probe task. Results of experiment one, in which the word pairs were presented unmasked, showed that participants responded faster to probes that appeared in place of negative words compared to neutral words on negative-neutral trials. However on positive-neutral trials there were no significant differences in reaction time. The results of experiment two, in which the word pairs were masked, showed identical results to experiment one. These findings suggest that negative words have a unique effect on the attention system in which they are detected quickly and demand attentional resources. Emotional words with a positive valence do not have the same effect

To my knowledge, no studies investigate selective attentional processes associated with attending to positive stimuli during a dot-probe task, specifically in relation to increased levels of aggression. Including positively valenced stimuli would potentially provide useful knowledge on the complexities of attentional processes relating to different stimuli in the environment. The evidence presented here suggests that individuals (from clinical and non-clinical samples) robustly show an attention bias to threat-related or angry-related stimuli compared to neutral stimuli, however there seems to be very little differences in attentional processes associated with attending to happy and neutral words.

2.1.6 Attention bias to face stimuli

2.1.6.1 Attention bias to angry faces

The literature review thus far has given an overview of attention bias, assessed the experimental paradigms used to measure this phenomenon, and given an outline of the previous literature investigating attention bias to aggression-related and positive words in aggression. This following section will review the literature on attention bias to different facial expressions in aggression.

Only a small number of studies have used faces when researching aggression-related attention biases. Images of facial expressions present an immediate and realistic sense of threat and could pose real life hostility, for example a face expressing an angry expression is a direct sign of aggression (Bradley et al., 1999). Perception of human facial expressions is central to human interaction; due to the social cues it conveys and the messages it communicates (Argyle, 1994). Although words can be threatening in nature, they may be deemed fairly arbitrary as they do not usually pose a direct threat or require a behavioural response. Also words are dependent on the participant's vocabulary knowledge whereas facial expressions are generally universally recognised (although there is some variation globally). Therefore, it is proposed that faces have increased ecological validity compared to words in the context of attention biases in aggression.

Attention bias toward angry faces has been well demonstrated in the anxiety literature (e.g. (Bar-Haim et al., 2007)), but has also found to be evident in healthy populations. (Santesso et al., 2008) used a dot-probe task to investigate attention bias towards angry faces in non-anxious undergraduate students. They found that across this sample participants showed a facilitated attentional response to angry faces, indexed by a quicker reaction time to cued probes following the presentation of an angry face compared to a neutral face.

Although attention bias to threat has consistently been found to be associated with high levels of anxiety, little is known about attentional orienting to anger in high aggression samples. Maoz et al. (2017) conducted one of the few studies that investigated attention bias towards angry faces in a sample of high trait anger participants using the dot-probe task. Participants completed the State-Trait Anger Expression Inventory (Spielberger et al., 1983) and a visual probe task in which angry and neutral face pairs were simultaneously presented. They found that increased trait anger was associated with an attention bias for angry faces such that they had speedier reaction times to probes that replaced angry faces compared to neutral faces. Maoz et al. (2017) suggest that negatively biased attention patterns facilitate increased processing of hostile stimuli which in turn amplifies anger.

Compared with the dot-probe task, the Stroop task is more commonly used to investigate attention bias. van Honk et al. (2001a) conducted a pictorial emotional Stroop task in which participants were asked to colour name images of both neutral and angry facial expressions during unmasked and masked conditions. In the unmasked condition, trials consisted of a fixation cross followed by presentation of the target word in one of four colours. During the masked condition, following the appearance of the target word, a mask consisting of a random letter string was presented. Participants were categorised based on anger scores and results illustrated that during both masked and unmasked tasks, participants with high trait anger showed delayed colour naming of angry faces compared to neutral facial expressions, suggesting an attention bias towards the angry face. These results indicate that hostility-related attention biases for angry faces in high trait anger participants were present, even at the preconscious level.

Putman, Hermans, and van Honk (2004) further investigated the attentional processes associated with high trait anger using very similar methods to that of van

Honk et al. (2001a). Participants completed a pictorial emotional Stroop task including neutral, angry and happy faces under both masked and unmasked conditions. It was hypothesised that attentional interference would result in longer latencies when colour naming the threatening faces when compared to neutral or happy faces. They found support for this hypothesis but only under the masked conditions. This study shows that Stroop performance is potentially affected by conscious control of cognitive-emotional processes.

Ford, Tamir, Gagnon, Taylor, and Brunyé (2012) investigated the relationship between trait anger and selective visual attention to rewarding visual stimuli. They tested both a valence-based account and motivation-based account to assess attention biases in individuals with high levels of trait anger (Trait anger scale; Spielberger et al., 1983) and trait aggression (measured using a total score from The Aggression Questionnaire; Buss & Perry, 1992). Ninety-six male participants completed a selective attention task in which rewarding (e.g. erotic couples, hang gliding), threatening (e.g. people holding weapons), and control (e.g. jet planes) images were presented. Stimuli appeared in pairs and participants were asked to state whether a particular image had appeared on screen after each trial. Participants' eye movements during the selective attention task was recorded using an eye-tracker. They hypothesised that if trait anger is associated with an attention bias towards threatening stimuli then this could be explained using a valence-based account. However, if there is a relationship between attention bias for rewarding stimuli and trait anger, these biases would be explained using a motivation-based account. Results suggested that individuals with increased levels of trait anger tended to fixate more on rewarding images compared with threatening images. This study showed support for a motivation-based approach and suggested that people who experience high levels of approach-orientated emotions such as anger, attended to more approach-related stimuli in the environment (rewards). The study found no evidence of a relationship between trait anger and attention bias towards threatening stimuli. Although this study is useful in understanding possible mechanisms associated with attention bias it may lack ecological validity; in real life settings it is rare that rewards are simultaneously presented with threat stimuli.

2.1.6.2 Attention bias to happy faces

The research outlined has demonstrated that aggressive individuals generally preferentially attend to angry faces, compared to neutral faces. However, to my knowledge there are very few studies that have investigated attentional orienting to happy versus neutral or angry faces in aggression. Ciucci et al. (2018) explored the relationship between callous-unemotional traits, aggressive behaviour, and attentional orienting towards emotional stimuli using the dot-probe task in school aged children. Children aged between 11 and 15 completed a dot-probe task in which angry faces (threat), sad and fearful faces (negative but not threat), and happy faces (positive) were each presented alongside a neutral face. Callous-Unemotional (CU) traits was self-reported by the children, whereas aggressive behaviour was measured by determining classmates perceptions of peers aggression. Results showed that irrespective of a child's level of CU traits, participants nominated as more aggressive by their peers showed increased attentional orienting to angry faces. However participants with low attentional orienting to angry faces were only nominated as aggression if they also reported high levels of CU traits. There were no effects of attentional orienting to happy, sad or fearful faces which suggests that attentional facilitation in aggression is unique for angry faces. These results are consistent with previous work that consistently shows that aggressive individuals preferentially attend to angry faces (e.g. Maoz et al., 2017; Putman et al., 2004; van Honk et al., 2001a).

A number of further studies have investigated possible biases relating to attentional processing of positive emotional stimuli in clinical samples with increased levels of anxiety and depression, as well as in healthy control groups. Bantin, Stevens, Gerlach, and Hermann (2016) conducted a recent systematic review to explore selective attention to faces in social anxiety using the dot-probe task. They outlined overall effects on negative-neutral trials and positive-neutral trials. They found on negative-neutral trials socially anxious participants responded faster to probes appearing in place of negative compared to neutral stimuli On positive-neutral trials there were no significant differences in response to probes appearing in place of positive or neutral stimuli. These results suggest that in anxious populations attention bias is specific to angry faces only. These findings are consistent with those found by Salum et al. (2013). Salum et al. (2013) investigated attention bias to threat and happy faces in a fear-disordered group (specific phobia), distress-disorder group (general anxiety disorder, depression), behavioural-disorder group (ADHD, conduct disorder), and no-disorder group. Each participant, from a large school based sample, completed a dot-probe task in which angry-neutral, happy-neutral and neutral-neutral face pairs were presented. Across all groups, there was no evidence of an attention bias on happy-neutral trials; results showed significant effects on angry-neutral trials only. Children with no psychiatric disorder showed increased attention bias for angry faces. This effect was also found in the distress disorder group; participants with higher symptoms had increased vigilance for threat. However in contrast, children with fear-related disorders, those with higher symptoms showed attention bias away from threat. No significant results were found in the behavioural-disorder group. Based on previous findings it would be predicted that individuals with high levels of conduct disorder, who display violent or disruptive behaviour, would show an attention bias to angry faces; however there was no evidence of this. These results suggest that attention bias may contribute to separate psychiatric disorders differently and further research into the unique association between psychiatric symptoms and attention bias is needed.

In contrast to the results from the systematic review conducted by Bantin et al. (2016), Fox et al. (2002) showed evidence for an attention bias towards angry and happy faces, relative to neutral faces in a self-rated anxious sample of undergraduates. Participants completed a cueing task in which the cue was either an ‘angry’, ‘happy’, or ‘neutral’ facial expression. The task included valid trials (target appears in same location as face) and invalid trials (target appears in different location to the face). Results showed a significant effect of valence cue on response time on invalid trials only. Participants had increased reaction times to targets when an emotionally valenced face (angry or happy) appeared in an invalidly cued location, relative to when the face cue had been emotionally neutral. These findings suggest that attention bias in anxiety is associated with difficulties in disengaging from threat-related and emotional stimuli.

There is further evidence to suggest that attention bias towards happy faces is also evident in non-anxious individuals. Waters, Henry, Mogg, Bradley, and Pine (2010) investigated selective attention to faces during a visual probe task in which angry/neutral and happy/neutral face pairs were presented. Results showed that severe anxiety was related to an attention bias to angry faces. The findings also demonstrated that across participants, including non-anxious controls, there was an attention bias towards happy faces relative to neutral ones. This suggests individuals may selectively attend to happy stimuli, regardless of anxious symptoms. This is supported by Pishyar et al. (2004) who found that individuals with low levels of anxiety preferentially attended towards happy faces (compared to neutral faces) and away from threatening faces (compared to neutral faces). Furthermore, Bradley et al. (1997) found a non-significant tendency for healthy control subjects to show vigilance for happy faces compared to neutral faces. It was hypothesised that these findings may reflect the phenomenon of mood regulation, this suggests that attention bias patterns maintain current mood, therefore participants that are happy (non-dysphoric) will attend to happy faces to maintain a happy mood.

In contrast to evidence which suggests attention bias towards happy faces in normative samples, Cooper and Langton (2006) found that early attentional resources are allocated to the location of the relatively threatening face in each pair during a dot-probe task. Undergraduate students not assessed for anxiety levels completed one of two dot-probe tasks in which angry-neutral and happy-neutral face pairs were presented. Happy-neutral face pairs were included to ensure that attention bias effects for angry faces could be attributed to the aggressive nature of the stimuli and not emotionality. In the first condition the face pairs were presented onscreen for 100ms only, whereas in the second condition, face pairs appeared for 500ms. The face pairs were then followed by a horizontal or vertical probe. Participants were required to identify which probe was presented. In the 100ms condition, participants showed vigilance for angry faces compared to neutral and avoidance of happy faces compared to neutral faces. This is in contrast to the 500ms condition where participants showed significant avoidance of angry faces and vigilance for happy faces. These results suggest that deployment of attention occurs as early as 100ms. The authors propose that for both trial types individuals initially attend to the relatively threatening face at 100ms (the angry face on angry-neutral trials and the neutral face on happy-neutral trials) and then shift to the opposing face at 500ms. These findings suggest that when using 500ms presentation during the dot-probe task, reaction times may reflect attentional vigilance, avoidance or both. One possible method for providing more accurate conclusions regarding each of these mechanisms is to include a neutral-neutral baseline condition (Koster et al., 2004).

Research suggests that non-dysmorphic controls generally show an attention bias towards angry faces when they are paired with a neutral face. However the findings regarding attention to happy-neutral face pairs is mixed. Across the literature reviewed here, Ciucci et al. (2018), Bantin et al. (2016), and Salum et al. (2013) showed no evidence of differences between attentional processes associated with happy and neutral faces, whereas Fox et al. (2002), Waters et al. (2010), and Pishyar et al. (2004) report vigilance for happy versus neutral faces. Finally, Cooper and Langton (2006) found an overall vigilance for neutral faces compared to happy. Due to the differences in samples, it may be hard to draw comparisons between these studies however this evidence suggests that further work is needed to untangle the attentional processes associated with different psychiatric symptoms.

Previous work has focused on emotional and neutral word pairs; however, little is known about selective attentional processes involved with attending to emotional stimuli if presented simultaneously with other emotional distracter stimuli. A few studies have investigated the role of emotionally valenced distracter items. For example, across three studies (Hansen & Hansen, 1988) found asymmetry in the processing of emotionally angry faces compared to emotionally happy or neutral faces. In particular, in study one participants were faster to detect threatening targets in friendly crowds than vice versa (Hansen & Hansen, 1988). They suggest that threatening faces perhaps ‘pop out’ in crowds.

This work was built upon by Öhman, Lundqvist, and Esteves (2001) who used a visual search paradigm to test the hypothesis that individuals preferentially orient attention toward threat. Participants were asked to search for the differing (odd-one-out) face among a matrix of otherwise identical distracters. Across the five experiments results showed faster and more accurate detection of threatening faces compared to friendly faces when they were among both neutral and emotional distracters. Also participants were more efficient in locating threatening faces compared to sad or scheming faces which suggest that this effect is specific to threat faces and not dependent on other characteristics of the face, for example, valence or uniqueness. However, the visual search task used schematic faces instead of facial images, these stimuli may be less ecologically valid as they do not demonstrate any real potential threat.

Pineles and Mineka (2005) investigated selective attention to different emotional faces using a dot-probe task. To my knowledge this is the only study to include an angry versus happy trial type in which attention orienting between such stimuli has been explored. The study was designed to investigate whether individuals with high social anxiety show an attention bias for cues of either external (threatening faces) or internal sources (heart-rate information) of potential threat. To assess attention to external threat, participants completed a dot-probe in which reaction time for pairings of different combinations of facial expressions (threatening, happy and neutral) was measured. It was hypothesised that individuals experiencing high levels of social anxiety would show a greater bias towards angry faces compared to both happy and neutral faces, compared to the participants experiencing low levels of social anxiety. The authors found no main effect of anxiety, stimulus pairing or probe position and no interactions. This suggests that reaction times across the different stimulus types and across anxiety groups were relatively stable. However, based on previous anxiety literature they conducted further analyses based on a bias score computed from the two face pairings that included threat faces (threat-happy and threat-neutral). They found no difference in bias scores for threat-happy face pairs between anxiety groups, however, they found that there was a trend level effect of group for the threat-neutral face pairs, such that high social anxiety group had a greater bias towards threat faces than neutral, compared with the low social anxiety group. Pineles and Mineka (2005) recognise that these results should be interpreted with caution. Nevertheless, this study contributes to the knowledge of attentional orienting to threat-happy faces. Further research is needed to replicate these findings and suggest why there is potentially no difference in reaction time to probes replacing threat and happy faces if they are simultaneously presented.

2.2 Interpretation Bias

The literature reviewed thus far focuses on attention biases; however attentional orientating is not a singular cognitive process and influences and is influenced by other simultaneous processes. The social information processing model (Crick & Dodge, 1994) describes six processing stages important for response formation; attention and interpretation are two important cognitive phases. White, Suway, Pine, Bar-Haim, and Fox (2011) suggest that attention and interpretation biases should be simultaneously studied as they are not distinct processes. The next section of the literature review will provide a summary of the research investigating interpretation biases in aggressive behaviour.

2.2.1 Introduction to interpretation bias

A negative interpretation bias is a type of cognitive bias which influences the encoding and interpretation of stimuli presented within the environment. This type of bias is defined as interpreting ambiguous or mildly aversive scenarios as more negative and dangerous, overestimating danger and underestimating the ability to cope (Waters et al., 2008a). Interpretation biases have been considered to contribute to the maintenance of maladaptive behaviours such as anxiety and aggression. For example, interpreting a benign situation as provoking or hostile has subsequent implications for the formation of an aggressive behavioural response (Crick & Dodge, 1994). Hostile Attribution Bias (HAB), termed by Nasby, Hayden, and DePaulo (1980), is a form of cognitive bias which relates to attributing negative, hostile or angry intentions to the behaviour of individuals in the environment. Individuals with increased levels of HAB tend to evaluate both benign and ambiguous stimuli as negative.

Cognitive theories of aggression and antisocial behaviour highlight that aggressive individuals have increased attention to, and engage in greater processing of, aggression-related cues over non-aggressive-related cues, interpret others' actions with more hostility and generate proportionately more aggressive responses to ambiguous behaviour (e.g. Coccaro et al., 2009; Crick & Dodge, 1994; Dodge & Frame, 1982; Dodge et al., 1986; Smith & Waterman, 2003). These findings suggest that aggressive individuals have cognitive biases at several stages of processing that contribute to an aggressive response.

2.2.2 Theoretical explanations of interpretation bias and aggression

Early work suggested that aggressive boys are more likely than non-aggressive boys to attribute hostile rather than accidental behaviour to their peers after an ambiguous provoking event (Dodge, 1980; Dodge & Frame, 1982; Dodge & Newman, 1981; Steinberg & Dodge, 1983). This work was essential in forming the influential social information processing theory. Crick and Dodge's (1994) social information processing theory (described in Section 2.1.4.2) suggests that hostile attribution bias is particularly relevant to social interactions and situations. Crick and Dodge (1994) propose that aggressive individuals demonstrate poor identification of stimuli during the encoding and interpretation stages of processing and subsequently attribute hostile intent to social situations. Individuals demonstrating an interpretation bias towards hostile stimuli for example, interpret benign stimuli as more negative, which in turn is more likely to result in an unkind or negative behaviour. This cognitive model proposes that all behaviour is the consequence of cognitions, suggesting that changing maladaptive thinking patterns and teaching individuals to use adaptive and constructive strategies could impact positively on behavioural outcomes.

When assessing interpretation it is important to consider the role of social cues and schemata. Schemata are cognitive heuristics used to quickly sort information (Bem, 1985), and therefore are cognitively efficient. However reliance on schemata can result in an ineffective interpretation which can lead to an inappropriate social response. It is hypothesised that overreliance on aggressive or negative schema can have detrimental results on a child's social adjustment. Dodge and Coie (1987) investigated this by presenting children with hypothetical prevocational situations. With reference to the described situation, children were asked to describe the intent of their peer. It was aimed to explore how much children rely on the information provided in the scenario or general mental structures based on experience, to attribute intent. Results demonstrated that aggressive children were more likely to make interpretations of intent based on schemata compared with non-aggressive children. It is concluded that maladjusted children show greater biases towards negative social cues and have well-developed schemata that interfere with their ability to interpret the social environment effectively. This shows how cognitive processing may aid the understanding of problematic behaviours. However correlational cross sectional studies, such as this one, cannot inform cause and effect relationships. It may be that, maladaptive cognitions cause behavioural responses, or negative cognitive biases could be a result of repeated reinforcement of undesirable behaviour. However, it is most likely that cognition and behaviour co-occur with each influencing the other.

Further to this, attributions of causality are important aspects of interpretation. Causal attributions refer to the inferences made about the reasons why things occur in our social environment and usually relate to judging the motivations of other individuals behaviour (Weiner & Graham, 1984). They are therefore thought to play a significant role in goal construction. It is considered that socially adjusted children make casual attributions related to positive self-evaluations (Aydin & Markova, 1979). However there has been mixed evidence for the relationships between causality attributions and aggressive behaviour (Crick & Ladd, 1993; Goetz & Dweck, 1980). In terms of understanding children's aggressive behaviour responses to social situations, attributions of intent have been of particular importance (Dodge, 1985). It is hypothesised that hostile attribution bias has a significant impact on behavioural outcomes.

2.2.3 Assessments of methods for measuring interpretation bias

There are a number of different methodologies for measuring interpretation bias, for example ambiguous story completion task (Dill, Anderson, Anderson, & Deuser, 1997; Rule, Taylor, & Dobbs, 1987), rated responses to ambiguous scenarios displayed by text or video (Dill et al., 1997; Epps & Kendall, 1995), or recognition tasks (Mathews & Mackintosh, 2000; Micco, Henin, & Hirshfeld-Becker, 2014). During an ambiguous story completion task participants are asked to complete a story-stem by outlining what events may have happened next and what the main character might think and do. These open responses are then coded for negativity or aggressiveness (Dill et al., 1997). A further commonly used method for assessing interpretation bias is the presentation of ambiguous scenarios. These scenarios can be explained via text or displayed by actors in a video. Participants are shown each scenario and then asked to describe the behaviour of each of the actors in the video. In one of the studies conducted by Dill et al. (1997), participants were shown dyadic interactions which varied in aggressive content and were then asked to rate the degree to which 28 adjectives described the behaviour of each of the actors. Individuals with increased levels of negative interpretation bias would rate the aggression-related adjectives as better describing the behaviour of the actor.

The story stem completion task and ambiguous scenario tasks can be easily modified (e.g. type of story or scenario presented, and open or closed questions regarding each scenario presented). Many similar tasks have been used across different fields of literature, however, the examples outlined have specifically explored hostility-related interpretation biases in aggression. These methods, along with questionnaire measures, evaluate conscious interpretations; the participants are explicitly asked to attribute intent to a protagonist in an ambiguous scenario. In contrast, the recognition task is a more complex task which aims to measure more implicit biases. The recognition task involves the presentation of ambiguous scenarios, followed by positively and negatively valenced statements. Participants are asked to rate the similarity between the ambiguous scenarios and the valanced statements. It is predicted that individuals making more negative interpretations of the ambiguous scenarios will rate the negative statements as more similar to their perceived outcome of the scenario (Mathews & Mackintosh, 2000). A version of the recognition task is more commonly used as a manipulation check in cognitive bias modification (CBM-I) research (Micco et al., 2014). However the task has been validated as an appropriate measure of interpretation bias following training (Salemink & van den Hout, 2010). When the recognition task is used as a manipulation check as part of CBM-I techniques, the valence of the scenario is ambiguous until the final word. The final word presentation forces a positive interpretation and is displayed as a word fragment in which participants have to complete. This is to ensure participants are attending to the scenario and the positive interpretation.

These different tasks have been used across the interpretation bias literature; however they are all comparable in that they ask participants to attribute thoughts and feelings to unfamiliar situations. Aggressive behaviour is hypothesised to be associated with making a greater number of hostile attributions.

2.2.4 Interpretation bias and aggression

Within the interpretation bias literature, research has been primarily conducted to investigate the influence of maladaptive interpretation biases in both aggression and anxiety. An interpretation bias towards hostile stimuli is evident in aggressive samples, whereas anxious individuals show an interpretation bias towards threatening stimuli. This could be attributed to the difference in fight and flight responses to threat stimuli in aggressive and anxious individuals (see Serin, 1991). The tendency for anxious individuals to interpret social situations in a negative or threatening way is a relatively stable phenomenon. Hadwin, Frost, French, and Richards (1997) examined whether self-reported levels of trait anxiety in children was associated with their interpretation of ambiguous stimuli. Participants were asked to interpret ambiguous pictorial homophones which could be either rated as threatening or neutral. Results demonstrated that individuals with increased levels of anxiety rated homophones as more threatening compared to less anxious individuals, suggesting an interpretation bias in high trait anxiety children.

Although the nature of interpretation biases may be specific to different behaviours, similar methodologies have been used to investigate interpretation bias in both anxiety and aggression. Studies have demonstrated a significant relationship between aggression and a negative interpretation bias (for example, Dill et al., 1997; Dodge & Frame, 1982; Dodge & Newman, 1981; Dodge, Price, Bachorowski, & Newman, 1990; Epps & Kendall, 1995; Hall & Davidson, 1996; Sancilio, Plumert, & Hartup, 1989; Steinberg & Dodge, 1983). Consistently these studies show that aggressive traits are associated with hostile attribution bias, such that intent is perceived as aggressive in nature. Dill et al. (1997) used Structural Equation Modeling to investigate the effects of aggressive personality on hostile interpretations of social interactions in a normative young adult sample. The first of two studies examined whether aggressiveness was assocated with the amount of rated hostilty in imagined outcomes of ambigously aggressive story stems. Participants aggression was measured using the Buss and Perry Aggression Questionnaire (Buss & Perry, 1992). They were then required to complete three ambigously aggressive story stems in which they had to indicate what the main character in the story might think, do or feel. The structural equation modelling revealed all four subsclaes of aggression loaded onto an aggressive personality factor. It was found that aggressive personality predicted aggressive thoughts of the main character in the story stems. Study two investigated whether aggressive personality would predict the amount of aggression percieved in the behaviour of actors in three vidoetaped interactions. The three scenarios consisted of one nonaggressive, one ambigously aggressive and one highly aggressive interaction. Participants were asked to rate the degree to which, a list of 28 adjectives, described the behaviour of both the actors in the video using a 7-point Likert scale from ‘not at all’ to ‘extremely’. The results showed that aggresive personality predicted a hostile perception bias in resposne to the actors in both the ambiguous and aggressive videotape interactions, however aggression predicted a perception bais to a much lesser extent in clearly non-aggressive settings. This article suggests that aggressive traits are positively related to hostile interpretations of ambiguous and aggressive hypothetical scenarios. It was hypothesised that schemas influence the perceptions and expectations of social interactions even when individuals are not personally involved.

Interpreting hostile intent to peers has been robustly linked to aggressive behaviour in children (Dodge & Coie, 1987; Dodge & Frame, 1982; Dodge & Newman, 1981; Dodge et al., 1990; Dodge & Tomlin, 1987; Fitzgerald & Asher, 1987; Guerra & Slaby, 1989; Quiggle, Garber, Panak, & Dodge, 1992; Sancilio et al., 1989; Steinberg & Dodge, 1983) and adults (Dill et al., 1997; Epps & Kendall, 1995; Hall & Davidson, 1996). Work in this area initially focused on cognitive biases in children with the aim to understand the development and maintenance of aggressive behaviour. Further work followed with adult samples; subsequent conclusions suggest that biases in cognitive processing, especially attributing hostile intent, are robust and enduring.

Early work by Dodge and Frame (1982) was influential in demonstrating, across three studies, a hostile interpretation bias in young boys. The first study showed that aggressive boys over attribute hostility to peers only when they are a recipient of an outcome, and not when they are observers of an event or behaviour that was directed at someone else. The second study showed that selective recall of hostile cues preceded biased attribution judgements, and the third study showed that boys who initiated acts of verbal or physical aggression were more likely to be the targets for peers' acts of aggression. This research revealed a number of key findings which were important for understanding possible mechanisms of hostile interpretation bias. It shows that the direction of the intended behaviour influences the interpretation of such behaviour, that attentional processes involved with attending to hostile cues influence subsequent judgements and interpretations, and finally that the environment and experiences of aggressive individuals may contribute to the maintenance of hostile interpretation biases.

Dodge et al. (1990) continued to explore the relationship between interpretation bias and aggression. The study investigated hostile attribution bias in 128 juvenile offenders aged between 14 and 19 years. The main aim of this study was to examine such biases in children with severe aggressive conduct disorder. During the experiment, participants were shown a video containing three different types of vignette (ambiguous, prosocial and accidental), during which they were asked to imagine they were the protagonist in the story. They were then asked to attribute intent using a multiple choice format (to be mean, it was an accident, to be helpful, it is unclear). An interpretation bias was positively correlated with under socialised aggressive conduct disorder, reactive aggression and number of violent crimes. This study suggests that, within clinical samples, attributing hostile intent may contribute to interpersonal reactive aggression that involves anger and violence.

Hostile attribution bias has been demonstrated in a number of studies which have recruited clinical and non-clinical child samples (for example, Dodge et al., 1990; Milich & Dodge, 1984; Slaby & Guerra, 1988), however the correlation between attributed hostility and aggression/anger has also been consistent across non-clinical adult populations (e.g. Dill et al., 1997; Epps & Kendall, 1995; Hall & Davidson, 1996). Epps and Kendall (1995) investigated interpretation of hostile, benign and ambiguous scenarios in high self-rated anger (Spielberger Trait Anger Expression Inventory; Spielberger et al., 1983) and hostility (Buss-Durkee Hostility Inventory; Buss & Durkee, 1957). Participants were asked to give scaled responses to unfamiliar situations which outlined an interpersonal interaction. As predicted participants scoring high on anger gave more negative interpretations of scenarios. Hostile attribution bias was evident for both hostile and ambiguous scenarios; however this relationship was less robust for benign scenarios. The results suggest that aggressive individuals are sensitive to hostile environmental cues; therefore they may disproportionately attend to a small number of such cues, even in the presence of dominant non-hostile cues.

The relationship between hostile attribution bias and behaviour has particularly focused on reactive aggression (Crick & Dodge, 1996; Dodge & Coie, 1987; Dodge et al., 1990). Reactive aggression refers to angry, emotional or affective aggression which is usually expressed in a physical response after provocation; this is compared to proactive aggression which is more often premeditated and is motivated by a desire for dominance (Dodge & Coie, 1987). Dodge and Coie (1987) conducted four studies to explore the influences of proactive and reactive aggression on school children’s behaviour. During study three, four groups of socially rejected boys (reactive aggressive, proactive aggressive, reactive-proactive aggressive, and nonaggressive) and a control group of average boys were required to interpret the intentions of a provocateur in a number of video recorded vignettes which displayed provoking scenarios involving peers. Results showed that the reactive aggression and reactive-proactive group both gave negative interpretations on the scenarios, whereas no biases were shown in the proactive or nonaggressive group. This research suggests that making hostile attributions of intent may be particularly salient in individuals who report high levels of reactive aggression. Hostile attribution bias for instrumental situations have also been associated with physical aggression (Dodge, 1980; Dodge & Somberg, 1987); this association is perhaps not surprising as attributing hostile intent usually precedes aggressive behaviour.

Helfritz-Sinville and Stanford (2014) conducted more recent work into the association between hostile attribution bias and subtypes of aggression. They compared interpretation bias in impulsive aggressors, premeditated aggressors, and a non-aggressive control group. Participants were required to rate the intentionality and hostility of 24 vignettes which described intentional, ambiguous, and unintentional everyday conflict scenarios. They were also asked how angry the situation would make them and whether they would have responded aggressively in each given scenario. The results showed no evidence of hostile attribution bias; however premeditated aggressors reported a greater likelihood of being rude in ambiguous situations, even if they did not significantly rate the situations as more intentional or hostile. This suggests that premeditated aggressors are able to successfully interpret the situation but they are still motivated to assert their dominance in potentially provoking situations.

Similarly, Lobbestael, Cima, and Arntz (2013) explored the association between hostile interpretation bias and reactive and proactive aggression in a sample of male patients with mixed diagnoses. Participants were asked to respond to eight vignettes which depicted ambiguous provocative scenarios. To each scenario participants gave an open response explaining what happened in the described situation, and also ranked the likelihood of four given answers (hostile, negative, positive, and neutral). The open responses were coded and categorised as hostile, negative, positive, or neutral. Increased frequency of hostile responses reflected a hostile interpretation bias. Both forms of aggression were measured using the Reactive and Proactive Aggression Questionnaire (Raine et al., 2006). Results showed that reactive aggression was predicted by hostile interpretation bias, however proactive aggression did not. These findings suggest that the nature of aggression may be different across subtypes. Considering this mixed evidence, and the different effects between various forms of aggression, there is little research which has explicitly investigated negative interpretation biases in a physically aggressive sample.

2.3 Electroencephalography

To this point the literature review has provided a summary of the research on attention bias to word and face stimuli in aggression. In outlining and evaluating the methods used to measure attention bias, additional studies on attention bias and anxiety have also been considered. All research included so far has relied on behavioural measures, most commonly reaction time. This next section of the literature review will describe and evaluate the advantages of applying novel neurological methods to cognitive bias research. To do this EEG will be described and relevant research will be discussed.

2.3.1 Why use ERP methodology?

Attention bias is predominantly measured using behavioural analysis, such as self-report and reaction time measurements. Although cognitive biases are relatively automatic processes that operate outside conscious awareness (MacLeod & Rutherford, 1992), behavioural methods are commonly used to identify bias. For example, reaction time in the dot-probe task is thought to be a direct indicator of visual attention allocation (Mogg & Bradley, 1999c). Reaction time measures are a valid resource within psychological research, however they do not only represent the cognitive processes of interest but a combination of processes including evaluation, decision-making, and motor processes (Donders, 1969; Sternberg, 1969).

Poldrack et al. (2017) suggest that current neuro-imaging methods have been influential in understanding the biological basis of human behaviour. These methods can therefore be used to identify neural predictors of violent behaviour. Specifically, EEG is a type of neuro-imaging measure, with accurate temporal resolution, which can be used to understand the neural correlates of cognitive processes. EEG detects automatic attentional processes by recording event-related potentials (ERPs) directly from the scalp. ERPs are recorded evoked amplitudes time locked to a specific event or point of interest (e.g. stimulus response). EEG can capture changes in brain processes between milliseconds (O'Toole & Dennis, 2012) and therefore the ERP technique provides a direct measure of neural activity and allows partial isolation of distinct cognitive processing stages (reviewed in Luck, 2005). Neural activity is measured by a change in amplitude. Amplitude refers to the difference between pre-stimulus baseline voltage and the largest voltage evoked by an event of interest within a given time window (Polich, 2007).

The P300 wave, sometimes referred to as the late positivity potential (LPP), late positive complex (LPC) or P3b, has been one of the most commonly investigated components in ERP research (for review see Polich, 2007). This component appears as a positive deflection at posterior parietal sites between 300 and 800ms after stimulus onset (Coles, Smid, Scheffers, & Otten, 1995). Generally, P300 reflects the allocation of neural resources for information processing tasks, including the distribution of attentional resources, categorization of stimuli, and updating of working memory (Polich, 2007). The P300 component is consistently evoked in response to the oddball paradigm in which attended events are surprising (Pritchard, 1981). During the oddball paradigm, participants are required to respond to an infrequent target that occurs in a background of frequent non-target stimuli. Infrequent targets elicit an increased positive potential compared to non-targets (Polich & Criado, 2006). The P300 is therefore particularly sensitive to differential processing of stimuli in relation to their task relevance and can be used as an index for measuring selective attention (Coles et al., 1995; Donchin & Coles, 1988; Oliver-Rodríguez, Guan, & Johnston, 1999; Polich, 2007).

The temporo-parietal attentional network appears to be a crucial generator of the P300 ERP component (Knight, Scabini, Woods, & Clayworth, 1989; Verleger, Heide, Butt, & Kömpf, 1994; Yamaguchi & Knight, 1992). These studies have demonstrated that participants with lesions in the temporo-parietal junction (TPJ) have reduced P300 amplitude. Specifically, (Verleger et al., 1994) showed that during an auditory oddball task, participants with TPJ lesions had reduced P300 in response to targets, and during a visual oddball task the same participants had attenuated P300 in response to all standard stimuli. The TPJ is located at the intersection of the posterior end of the superior temporal sulcus, the inferior parietal lobule, and the lateral occipital cortex (Krall et al., 2015). The TPJ located in the right hemisphere has been associated with distinct cognitive processes (Decety & Lamm, 2007) and has found to be involved with the orienting of attention and theory of mind (Corbetta, Patel, & Shulman, 2008).

While P300 latency is considered to measure stimulus evaluation time, P300 amplitude is thought to reflect neural resources available to process stimuli (Hillyard & Kutas, 1983). Therefore the P300 ERP component is an index of elaborative stimulus processing and can be a useful tool to assess use of neural resources associated with attention allocation to different stimuli. The P300 is considered a relatively robust measure of emotional processing and information processing biases in anxiety (Moser, Hajcak, Huppert, Foa, & Simons, 2008). The temporal resolution of EEG allows for the identification of neuro-cognitive processes related to physical aggression at different stages and is useful in investigating when processing stages occur after stimulus presentation.

The P300 component reflects later more elaborative stages of attentional processing, whereas the P1 component reflects spatial attentional at earlier stages of processing (e.g. Hillyard & Anllo-Vento, 1998; Woldorff et al., 2002). Therefore both components may be useful when measuring neural correlates of attention bias. The P1 is the earliest ERP marker of visual attention and appears as an increased positive deflection between 80 and 130 milliseconds following stimulus presentation, maximal in the occipital cortex (Hillyard, Mangun, Woldorff & Luck, 1995). ERP results show that P1 amplitude increases when stimuli are presented in a pre-attended location (Woldorff et al., 2002). Participants completed a task in which two chequerboard arrays were presented in the left and right lower visual field quadrants. Participants were instructed which quadrant to attend to, or were told to passively view the two stimuli. Stimuli in the attended to quadrant evoked a larger P1 amplitude.

2.3.2 ERP correlates of attention bias in anxiety and depression

Attention bias in anxious populations has been studied to a great extent. The anxiety and aggression literature have used similar methodology to investigate attention biases towards threatening or hostile stimuli respectively. More recently, a number of studies investigating attention bias towards threat in anxious individuals using ERP analysis have been published (e.g. (Eldar, Yankelevitch, Lamy, & Bar-Haim, 2010; Fox, Derakshan, & Shoker, 2008; Moser, Huppert, Duval, & Simons, 2008; Mueller et al., 2009). Although the P300 component is most commonly investigated in relation to attentional processes, differences in the P1 component have also been found across the attention bias literature.

The P300 component has been used as an index for social information processing bias in socially anxious individuals (Moser et al., 2008b). ERPs were recorded during completion of a modified version of the Erikson flanker task in which negative and positive facial expressions were displayed. For each trial a threatening or reassuring face was presented flankered by two opposing stimulus Participants were required to categorise the emotion of the central facial expression. Behavioural results showed that generally participants were quicker on trials when reassuring faces were the target compared to threatening faces, and on congruent compared to incongruent trials. There were no significant effects of group. The ERP results showed an effect of target such that P300 amplitude was significantly larger for threatening target faces, than for reassuring target faces. The interaction with anxiety group showed that the low anxiety participants showed no significant difference in amplitude between the two target faces, whereas high anxious participants showed enhanced P300 to threatening target faces. The authors propose that socially anxious individuals demonstrate a negative bias during elaborative stimulus processing stages (Moser et al., 2008b).

A number of different tasks have been to assess the neural correlates of attention bias, however the dot-probe task is a less commonly used paradigm. Nevertheless there are a handful of studies which have utilised this method for measuring the processes associated with attention bias in anxiety and depression. For example, Mueller et al. (2009) investigated the neural correlates of attention bias to threat in anxiety. They used a go/no-go version of the dot-probe task to explore differences in P1 amplitude between different face pairs in participants with social anxiety disorder. Results showed that anxious participants had increased P1 potential to the presentation of angry-neutral face pairs compared to happy-neutral face pairs. These findings suggest individuals with increased levels of anxiety show an electrophysiological response to threatening stimuli, which could provide a neural marker for attention bias which is known risk factor for anxiety.

Mingtian, Xiongzhao, Jinyao, Shuqiao and Atchley (2011) explored attention bias to differently valenced pictures using behavioural and EEG data extracted during a dot-probe task. Patients with major depressive disorder and never depressed control patients completed a dot-probe task in which negative-neutral and positive-neutral picture pairs were presented. Pictures depicted images of nature, sport, buildings, and household objects etc., only images of faces were excluded. The probe was either presented at 100ms or 500ms post stimulus pair presentation. Behavioural results suggest that at 500ms depressed patients failed to avoid attending to the negative stimuli relative to the control participants. The ERP results demonstrated that control participants showed significantly larger P1 amplitudes to valid compared to invalid trials when presented with positive-neutral stimulus pairs. The depressed group did not show this effect. As P1 amplitude is generally increased when stimuli appear in a pre-observed location, these results suggest that control participants attended to the positive pictures compared to neutral pictures. Together the results suggest that depressed individuals avoid attending to positive stimuli and instead preferentially attend to negative information in their environment. The main effects observed for behavioural and ERP data were only found at 500ms probe presentation which suggests that attention bias in depression appear later and at more elaborative stages of processing.

Similarly to results presented by Mingitan et al. (2011), Hu et al. (2017) found that participants with major depressive disorder (MDD) showed biases when attending to negatively sad information. Depressed individuals and healthy controls completed a dot-probe task in which fear-neutral, sad-neutral, and happy-neutral face pairs were presented. Behavioural results showed that MDD participants had shorter reaction times on sad-neutral trials when the probe appeared in place of the sad face, suggesting vigilance for sad emotion. The ERP results showed that depressed individuals had increased P300 amplitude in response to sad-congruent trials compared to happy-congruent and fear-congruent trials. In contrast, the healthy controls showed no significant differences between types of emotion. Taken together the findings by Mingitan et al. (2011) and Hu et al. (2017) suggest that biases in attention can be reflected in differences in P1 and P300 amplitude, showing that ERP patterns evoked by stimuli presented during the dot-probe paradigm may be sensitive to early and late attentional processing.

Along with work exploring the neural correlates of attentional orienting to threat in anxiety and depression, ERP analysis has also been used to assess the effectiveness of attention bias modification (ABM). Specifically, (Eldar & Bar-Haim, 2010) examined changes in attention processing after ABM. Success of the training programme was measured by assessing the change in electrophysiological responses. During the study an anxious and non-anxious control group completed a modified dot-probe task in which angry and neutral faces were presented. Half of each group completed a training condition and the other half completed a placebo task. The behavioural results showed that anxious participants, trained to avoid threat showed a gradual reduction in reaction time to neutral targets as training progressed. Trained anxious participants also showed differences in ERP results; after training they showed decreases in P300 amplitude in response to face pair presentations compared with pre-training amplitudes. After training, anxious individuals showed P300 patterns that were similar to those shown in the non-anxious participants.

O'Toole and Dennis (2012) conducted a similar study in which participants completed a modified dot-probe task aimed to train toward or away from threat stimuli. The results showed that changes in amplitude between pre- and post-training conditions were significant for P1 only in the non-anxious group. Before training non-anxious participants showed greater P1 amplitude to non-threatening versus threatening face cues. After taking part in the *train away* AMB task, participants showed reductions in P1 amplitudes to all cues. These results suggest that training towards non-threat stimuli may reduce early, automatic capture of attention of face cues even in a normative sample. Furthermore, Sass, Evans, Xiong, Mirghassemi and Tran (2017) used the dot-probe to assess the effectiveness of attention training in anxious populations. Participants were assigned to a training or placebo group and presented with threat-neutral, pleasant-neutral, threat-pleasant, and neutral-neutral word pairs. As expected, those participants assigned to the training group reported significantly less symptoms of anxiety post intervention, whereas there were no significant changes in the placebo group. Attention training to pleasant stimuli was also associated with greater P100 amplitude in response to neutral stimuli within threat-neutral word pairs from pre-to-post training. However P100 or later P300 amplitude did not reflect increased processing of pleasant stimuli on pleasant-threat trials. This suggests that attention training may only be effective if the stimuli used to 'train towards' is rated as lower in arousal compared to the other stimuli presented within each pair. This is an important consideration for future work as it suggests that emotional arousal of stimuli influences attention bias effects, especially if both stimuli are presented within a high arousal context.

These studies show that attention bias to threat stimuli may be characterised by a distinct neural pattern. ERPs have been used with a number of behavioural paradigms, including the modified dot-probe task, to measure attentional selectivity for threat stimuli in psychological disorders such as anxiety and depression. Therefore, similar methods can also be used to investigate biases associated with other maladaptive behaviours such as aggression.

2.3.3 ERP correlates of aggression

Although there are a greater number of studies exploring threat-related biases in anxious populations, there is some evidence to suggest that attenuations in P300 amplitude are associated with hostile-related attention bias in aggressive populations. However it is unclear whether these variations in amplitude are consistent across different anger or aggressive styles. It is suggested that reduction in P300 amplitude in response to hostile stimulus is a particularly dominant effect in impulsive aggression (Barratt, Stanford, Kent, & Alan, 1997; Gerstle, Mathias, & Stanford, 1998; Harmon-Jones, Barratt, & Wigg, 1997; Mathias & Stanford, 1999).

Evidence suggests that variation in P300 amplitude is associated with antisocial behaviour. A meta-analysis of 38 studies (Gao & Raine, 2009) reviewed findings relating to aggression, antisocial personality disorder, conduct problems or psychopathy. Included studies employed an experimental design specifically intended to target the P300 ERP component. Results indicated that antisocial individuals had significantly smaller P300 amplitudes and longer P300 latencies. The authors proposed that individuals with generic anti-social behaviour show inefficient deployment of neural resources in processing cognitive task-relevant information. These findings were found across standard oddball, more complex non-oddball, and Stroop tasks. These findings should be interpreted with caution as the more complex non-oddball tasks include a variety of different tasks and therefore it may not be possible to make comparisons across these. Also, although these findings were significant, the effect sizes were small. This research shows that anti-social individuals may have different patterns of P300 amplitude which reflects the ability to process task relevant events, but these studies do not distinguish between stimulus types.

The literature suggests that attenuations in P300 amplitude may be particularly salient in violent anti-social individuals. Bernat, Hall, Steffen, and Patrick (2007) investigated the relationship between P300 amplitude and both violent and non-violent criminal offenders. One-hundred and thirty eight adult inmates completed a standard visual oddball task in which they were asked to ignore frequent non-target stimuli. Participants were categorised based on their convicted offence. Violent offences included murder, robbery, assault and sexual offences, whereas examples of non-violent offences were theft, drug-related crimes and fraud. Prisoners convicted of violent offences were found to have a reduced P300 in response to target stimuli. There was no significant relationship between P300 amplitude and response to target stimuli during the oddball task in participants convicted of non-violent offences.

Most research on P300 impairments and aggressive behaviour has recruited participants in young adulthood and used cross sectional designs (e.g. Bernat et al., 2007; Mathias & Stanford, 1999). Gao, Raine, Venables, and Mednick (2013) used a longitudinal design to discover whether there are neurological markers which highlight increased risk for antisocial behaviour. They studied whether P300 amplitude and antisocial behaviour at age 11 was associated with criminal behaviour at age 23. At age 11, P300 was measured over the temporal-parietal junction whilst a continuous performance task was administered. During this task numerals one to nine were presented, with number five being the target number and presented at a lower frequency compared to the other numbers. Numbers were presented randomly and subjects were required to respond as quickly as possible to targets and ignore all other stimuli. Anti-social behaviour was measured using The Child Behaviour Checklist (Achenbach, McConaughy, & Howell, 1987), which was completed by parents and measures Aggression, Non-Aggressive Antisociality, and Total Antisociality subscales. Official court records for offences including property, drug, violence, and serious driving offences were searched when the participants were aged 23 years to construct a measure of criminality. Reduced P300 amplitude was found to be associated with antisocial behaviour at age 11 and criminal behaviour at age 23. These findings highlight that targeting youth antisocial behaviour may influence later outcomes and that cognitive processes should be considered when implementing interventions. This study suggests neural markers for antisocial behaviour, however it uses arbitrary stimuli and therefore does not provide information regarding selective attentional processes. Therefore further research would be beneficial to investigate if neural markers differ depending on the type of stimuli presented.

2.3.4 Theoretical explanations of P300 effects in aggression

The literature suggests that the P300 component may be a neural correlate of attention deficits and re-orientating. A reduced positive P300 amplitude may be associated with cognitive deficits. Although the current research does not measure valence-specific attentional processes, there are a number of theories to explain why aggression was associated with a reduced P300 amplitude when responding to negative stimuli. As P300 amplitude is thought to represent the allocation of cognitive resources, individuals with increased aggression may utilise fewer resources when attending to hostile-related stimuli. Reliance on schemas could allow for efficient and quick processing of such stimuli. Schemas are defined as building blocks of cognitive knowledge which enable individuals to form mental representations of the world (Piaget & Cook, 1953). Wadsworth (1996) suggested that these schemas provide information on how to react to incoming stimuli or information. They therefore provide pre-defined ‘scripts’ which means that few cognitive resources are employed when attending to stimuli relating to these schemata. Aggressive individuals may have developed and retained strong schemas for threat (Todorov & Bargh, 2002). These aggressive schemas are likely to influence a bias towards hostile-related stimuli, as the schema provides a default response to all stimuli.

P300 amplitude is thought to reflect processing relating to categorization of stimuli and updating of working memory models. It is sensitive to infrequent task events and social expectancy violations elicit larger P300 event related positivity (Bartholow, Fabiani, Gratton, & Bettencourt, 2001; Duncan-Johnson & Donchin, 1977). Change in P300 amplitude therefore reflects the process of updating cognitive models based on stimuli that are being attended to (Donchin & Coles, 1988). Aggressive individuals are more likely to expect hostile stimuli in their environment and therefore have cognitive models which fit with expectancy outcomes, resulting in a relatively stable P300 amplitude (Fanning, Berman, & Long, 2014). In contrast, non-aggressive individuals are less likely to expect to perceive hostile stimuli within their environment and attending to such stimuli may trigger an increased P300 response. In summary, high aggressive individuals may only require few neural resources to update cognitive models as presented stimuli fit with existing models.

2.3.5 ERP effects of attention bias to words

Due to the different neural processes involved with attending to words and faces, and to retain clarity, previous research investigating neural correlates of attention bias to words and faces will be reviewed separately. There are relatively consistent findings showing a hostility-related attention bias to threat words in aggressive populations, however very little is known about the neural correlates of this attention bias. EEG, only in more recent empirical work, has been used in conjunction with behavioural measures to explore the social cognition which contributes to psychological disorders. Although there are a number of studies that have used the dot-probe paradigm and simultaneous EEG recording, these studies have explored psychological disorders such as anxiety and depression. To my knowledge these two techniques have not been used collaboratively to understand attention biases specifically in aggression. The modified dot-probe task allows for two types of analysis; between-group analysis of data time locked to the presentation of each stimulus pair, and within-group analysis of data time locked to the presentation of the probe. This task therefore gives a more complex overview of the processes associated with attention bias.

Helfritz-Sinville and Stanford (2015) conducted one of the few studies that investigated neural correlates of attention bias using ERP's. They used a modified oddball task to assess the P300 component of event-related potential across electrodes Fz, Cz and Pz, in relation to attention biases in the processing of threat stimuli. They investigated how two major subtypes of aggressive individuals, reactive (impulsive) and premeditated, process social and physical threat words compared with non-aggressive individuals. During the task, the all-male sample ($N = 58$) were asked to respond to neutral targets which appeared among physical threat distracters, social threat distracters and neutral distracters. They found that non-agggressives showed increased P300 amplitude when presented with both social and physical threat words compared to neutral words. This enhanced processing was not demonstrated in the aggressive samples. Impulsive and premeditated aggressors had P300 amplitude that was relatively stable across responses to social and physical threat words and neutral words. (Helfritz-Sinville & Stanford, 2015) concluded that aggressive individuals perceive threatening words in a similar way to the neutral words and this may be explained by desensitization of hostile stimuli and resulting emotional processing deficits. If P300 amplitude is a reflection of neural resources attributed to the processing of stimuli, an alternative explanation could be that high aggression participants attribute less resources to the processing of aggression-related stimuli. Cognitive processing shortcuts, such as schemas, could allow for the rapid and efficient evaluation of such stimuli (Piaget & Cook, 1953).

ERP analysis has been conducted with healthy (non-aggressive) subjects to demonstrate how P300 may be crucial in understanding how threat information is processed (Thomas, Johnstone, & Gonsalvez, 2007). A small sample of 22 undergraduates completed two versions of the emotional Stroop task. In the first they were asked to colour name (*blue* or *green*) a set of words they had previously rated as ‘personally disturbing’ which were presented randomly among other neutral words. Identical stimuli were used during the second Stroop task, however participants were asked to identify if each of the words was *threat* or *non-threat*. Behavioural results revealed no significant reaction time effects, which is perhaps not surprising given the normative sample used. ERP results suggested that during both tasks, participants showed a larger P300 amplitude to threat words compared with neutral words. This effect was particularly dominant in the word-relevant task. This evidence shows that healthy individuals demonstrate increased P300 in response to threat stimuli, suggesting that threat and neutral information is processed differently. This finding is important as, firstly, it reveals that varying stimulus types are processed differently. Secondly, understanding social cognition in a normative healthy population can help identify differences in non-normal or forensic samples which may reflect abnormal processing This enables more effective work examining possible differences in P300 in response to hostility-related and neutral stimuli. Further work is needed to develop a greater understanding of why processing of aggressive and neutral stimuli may recruit different levels of P300 amplitude, and to examine the behavioural outcomes of these varying levels of processing resources.

Surguy and Bond (2006) investigated P300 abnormalities in a sample with less severe aggression. A healthy sample of 32 volunteers was divided using a median-split into high and low aggression groups based on responses to the Buss-Durkee Hostility Inventory (Buss & Durkee, 1957). Participants completed a novel modified oddball task in which they had to respond to rare food words (targets) only. Aggressive words were also presented with the same frequency as these targets. Both targets and aggressive words appeared randomly among neutral words. ERP's were recorded across frontal, central and parietal midline sites. The results suggested there were no significant overall differences in amplitudes across aggression groups. This is inconsistent with work by Helfritz-Sinville and Stanford (2015) who found a significant effect of word type in the low aggression group. However, Surguy and Bond (2006) findings showed a significant interaction between group and electrode sites in response to non-target aggressive probes. High and low aggression groups showed a different pattern of amplitudes across the three electrode sites when responding to aggressive words. The difference in amplitude between Fz and Cz, and between Fz and Pz, showed a significant relationship with group. High aggression participants, compared with the low aggression group, had lower P300 amplitude in response to randomly occurring aggressive words at Fz compared with Cz and Pz. The authors suggest that individuals who report higher levels of aggression have less efficient processing of aggressive stimuli. However, there are a number of criticisms of this work: firstly it only presents findings based on a very small number of mid-line electrodes; and secondly conclusions are based on subtractions of amplitude between two electrode sites, and so it is not clear what this tells us about the cognitive processing of aggression-related words.

Stewart et al. (2010) investigated the neural correlates of approach and withdrawal anger styles and suggest that different anger styles may influence attentional processes relating to negative and positive valenced information. Approach and withdrawal motivational systems play a crucial role in the expression of emotions. Anger is expressed under circumstances of unfairness, provocation or mistreatment. Expression of this emotion and response formation depends on the context of the situation and determines approach or withdrawal mechanisms. The State Trait Anger Expression Inventory (Spielberger, 1991, 1999) conceptualises approach anger styles (anger out) as verbal or physical behaviour directed towards another person or object. Withdrawal anger styles (anger in) are conceptualised as the repression or inhibition of outward signs of anger. Stewart et al. (2010) used ERP's to examine the relationship between anger styles and attention bias to negative, positive and neutral stimuli during an emotion-word Stroop task in which positive, negative and neutral words were presented. Results suggested that individuals with higher anger-out scores showed increased P300 amplitude in response to the negative words compared to both positive and neutral words. This finding is inconsistent with other work which demonstrates reduced P300 amplitude in aggressive individuals (Barratt et al., 1997; Fanning et al., 2014; Helfritz-Sinville & Stanford, 2015; Surguy & Bond, 2006). However Stewart et al. (2010) propose that this increased positive amplitude reflects greater cognitive effort in overriding attention to negative information.

Stormark, Nordby, and Hugdahl (1995) also investigated the attentional processes involved with attending to negative emotional stimuli using behavioural and ERP methodology within a normative sample. They used a spatial orienting task in which a cue was presented to indicate the most likely location of each target. There were three conditions; the valid-cue condition in which the target appeared in the same location as the cue, the invalid condition in which the target appeared in the opposite location to the cue, and a no-cue condition. The stimuli consisted of eight negative emotion cue words and eight neutral cue words. As expected, reaction time data showed a faster response to the validly cued targets compared to invalidly cued targets, but only when the emotion word served as the cue. ERP data was analysed in response to the cue words and the target. In response to cues, participants showed enhanced P300 amplitude when an emotional cue was presented compared to a neutral cue. In response to targets, participants showed an increased P1 and P3 amplitude on invalid trials but only following an emotional cue word. The authors propose that increased P1 and P3 amplitude on invalidly cued trials may reflect enhanced attentional resources involved in disengaging from the emotionally cued location.

There is some mixed evidence regarding the ERP correlates of attention bias to aggression-related words in aggression. Some studies suggest that P300 amplitude is reduced in aggressive individuals (Gao et al., 2013; Helfritz-Sinville

& Stanford, 2015; Surguy & Bond, 2006), whereas Stewart et al. (2010) and Stormark et al. (1995) found enhanced amplitude to negative stimuli. These studies show that P300 (and P1; Stormark et al., 1995) may be sensitive to attentional processing differences between negative and neutral words, however there is no evidence of the ERP correlates associated with attending to positive words in aggression. Previous literature suggests very little behavioural differences in attentional processes associated with attending to happy and neutral words (e.g. Pishyar et al., 2004; Sutton & Altarriba, 2011). Further work investigating any potential ERP differences evoked by positive words would be beneficial in understanding how the emotional valence of stimuli may influence attention biases in aggression. Furthermore, these studies present stimuli singularly and therefore conclusions are based on the differences between evoked ERPs in response to single stimuli presentation. To my knowledge there are no studies investigating attention bias in aggression which have measured evoked ERPs when two emotional words are presented simultaneously.

2.3.6 ERP effects of attention bias to faces

There are a number of studies that have conducted the dot-probe task with simultaneous EEG recording to explore the electrophysiological processes associated with selective attention. Some of these studies are outlined in a recent meta-analysis conducted by (Torrence & Troup, 2018). The studies within this meta-analysis investigate attention bias in populations with disorders such as social anxiety, trait anxiety and panic disorder; however the majority of studies have recruited a general normative sample. The meta-analysis highlights that there are many inconsistencies in current research, such as, stimulus delay time, delay SOA, target type and type of response, which makes comparing results between studies difficult (Torrence & Troup, 2018). Although useful, this meta-analysis only includes studies which utilise a dot-probe task in which stimuli are presented horizontally; therefore excluding a number of studies published in this area.

To my knowledge no studies have explicitly investigated ERP correlates of attention bias in aggression using the dot-probe paradigm in which emotional faces are presented. Nevertheless, results across the general population show that attention bias towards angry and fearful faces can be seen in early ERPs time locked to the onset of the face stimulus. For example, Holmes, Bradley, Nielsen, and Mogg (2009) found that reaction times to probes replacing emotional faces (angry and happy) during a dot-probe task were faster compared to reaction times to probes replacing neutral faces. The ERP results revealed that on angry-neutral trials, angry-congruent trials evoked an increase in N2pc and late N2pc. On happy-neutral trials, happy congruent trials evoked an increased late N2pc only. These results are consistent with models of attention which suggest facilitated orienting towards emotional information. They also suggest that angry faces capture attention faster than happy faces and that they sustain attention once captured.

There is also further evidence to suggest that orient to threatening faces, compared to neutral is characterised by an early increased N2pc response, under conditions of high cognitive load. Holmes, Mogg, de Fockert, Nielsen, and Bradley (2014) studied attention bias to angry facial expressions under conditions of high cognitive load. Participants were required to complete a dot-probe task in which angry and neutral faces were presented, whilst simultaneously holding a sequence of digits in working memory. Reaction time data showed that participants were quicker to respond to probes on trials where the probe replaced the angry faces, and this effect was not influenced by the working memory manipulation. The ERP results showed that there was increased attentional prioritisation for angry faces under conditions of higher cognitive load. This was characterised by an increased N2pc and late N2pc following the onset of face pairs. These results suggest that capture of threat stimuli is enhanced when executive control resources are depleted by additional task demands.

Santesso et al. (2008) has used behavioural and ERP techniques to explore neural correlates of involuntary orienting to emotional faces in a healthy adult sample. The sample consisted of 16 undergraduate students and they were required to complete a dot-probe task which included angry-neutral and happy-neutral face pairs, while EEG was simultaneously recorded. Face pairs were presented for 100ms only in order to investigate involuntary orientating (this is likely not enough time for participants to shift gaze between the two simultaneously presented stimuli). Behavioural results showed that on angry-neutral trials, participants were faster to respond to probes when it appeared in place of angry faces compared to neutral, but on happy-neutral trials participants had speedier reactions in response to probes that appeared in place of neutral faces compared to happy. The ERP analysis revealed that on angry-neutral trials the evoked P1 amplitude was significantly larger when participants responded to the probe that appeared in place of the angry face compared to when it appeared in place of the neutral face. Santesso et al. (2008) suggest that healthy individuals orient attention towards the most threatening facial expression of each pairing and therefore will respond quicker to probes that replace angry when paired with neutral, but will respond quicker to probes that replace neutral when paired with happy. The authors concluded that P1 is the earliest electrophysiological index of spatial attention and that threat cues can modulate these attentional processes.

These findings are consistent with Thomas et al. (2007) who studied later latencies of attentional orientating and found that P300 amplitude was larger in response to threat words in a healthy undergraduate sample using the emotional Stroop task. Although Thomas et al. (2007) and Santesso et al. (2008) have used different modalities of threat stimuli and implemented different tasks to measure attention biases, they both suggest that healthy adults will show an increased positive amplitude in response to aggression-related stimuli at early and later latencies of attentional processing.

In addition to the dot-probe which is an index for selective allocation of attention, simple face presentation tasks have been used to understand the differences in electrophysiological responses to single face presentation. Schupp et al. (2004b) investigated the neural processing of facial expressions in a healthy undergraduate sample not classified by any condition. EEG was recorded while participants viewed happy, angry and neutral faces. Participants had no specific task. Results showed that individuals had increased late positive potential (LPP) to threat faces relative to both friendly and neutral faces. Similarly, Leppänen, Moulson, Vogel-Farley, and Nelson (2007) found that fearful faces evoke an increased N170 at lateral electrodes compared to neutral and happy faces during a simple face presentation task in a normative population. However, in contrast to findings reported by Holmes et al. (2009) the evoked amplitude in response to happy and neutral facial expressions did not significantly differ. These results show that attentional vigilance for angry faces may influence early (N170) and later (LPP) stages of processing.

Bertsch, Böhnke, Kruk and Naumann (2009) investigated the processing of facial expressions in a sample of healthy participants experimentally provoked for aggressive behaviour. They measured ERPs evoked during an emotional Stroop task in which happy, angry, fearful and neutral faces were presented. Aggression was provoked using the TAP (Taylor, 1967) and anger was measured using the subscale from the Buss and Perry Aggression Questionnaire (Buss & Perry, 1992). Participants were assigned to a provoked or non-provoked control condition. The behavioural data showed that provoked participants were slower to name the colour of all emotion expressions compared to neutral faces. The ERP results showed significant differences in P2 and P3 amplitude between the provoked and unprovoked group such that provoked participants generally showed greater positivity compared to non-provoked participants. The P2 amplitude was increased in response to all facial expressions but was greatest for fearful and angry expressions. This is consistent with previous work by Carretie et al. (2001) which reported increased posterior P2 amplitude for negative pictures. The P3 amplitude was increased for happy compared to neutral, and for neutral compared to angry. This is in contrast to previous work by Thomas et al. (2007) which found increased P3 amplitude to threat-related stimuli. These results suggest that provocation affect the processing of facial expressions, perhaps because threat-related faces become motivationally significant in a provoking situation.

Another study to compare evoked amplitude in response to positive and negative pictures is that by Smith, Cacioppo, Larsen, and Chartrand (2003). Although this study used affective pictures instead of facial expressions, results showed that evoked P1 amplitude was larger in response to negative stimuli compared to positive stimuli. These results suggest that P1 is an early marker of attention allocation and that the valence of stimuli influences the amount of attention received during the initial stage of information processing.

The findings regarding angry-neutral trials in normative samples are relatively consistent; there is generally increased amplitude to angry faces compared to neutral, however, the findings regarding happy-neutral trials is somewhat mixed. Holmes et al. (2009) reported increased late N2pc on happy congruent trials and Carretie et al. (2001) reported increased P3 in response to happy versus neutral faces; whereas Schupp et al. (2004b) and Leppänen et al. (2007) report no differences in amplitude between happy and neutral faces across N170 or LPP components. The literature has used a number of different paradigms to investigate attention biases. This may explain that results have provided evidence of processing differences of emotional stimuli across a large number of ERP components, namely P1, N170, P2, P3 and the LPP. Studies recruiting healthy individuals and can aid the understanding of normative attentional orienting. However, understanding a-typical attention biases in aggressive populations has greater implications for understanding real life behaviour and subsequent interventions.

2.3.7 ERP effects of interpretation bias

Although the association between aggression and hostile interpretation bias is fairly robust, little is known about the neural processes involved in such biases, or the time-course in which these occur. Experimental methods in current research have relied on participant response and therefore it is suggested that more modern methods used in cognitive neuroscience, such as Electroencephalogram (EEG), may be useful in determining the underlying processes associated with interpreting hostile stimuli. This may aid the understanding of how biases contribute to aggressive behaviour**.**

Within the attention bias literature, studies that have used EEG methodology show differences in processing associated with attentional orienting across multiple components; for example, N170 (Leppänen et al., 2007), N2pc (Holmes et al., 2009), LPP (Schupp et al., 2004b) and P300 (Helfritz-Sinville & Stanford, 2015). These effects appear between 80ms (P1) and approximately 300-400ms (P300) after stimulus onset and therefore reflect relatively early attentional processing. In terms of cognitive processing, interpretation is a more elaborative stage of processing compared with attention. Therefore, when investigating interpretation bias, later ERP components such as the LPP, may be useful for investigating differences in processing associated with making hostile attributions. The late positive potential (LPP) is commonly used to refer to P300-P600 effects in the context of emotion-related ERP studies. The LPP is a widely distributed positive potential that occurs in the central parietal region between 300 and 800ms after stimulus onset. The LPP is similar to the P300 component, but the increased potential can be sustained for a longer latency (Hajcak & Olvet, 2008). The LPP is evoked during the evaluation of pleasant and unpleasant stimuli in comparison to neutral stimuli (e.g. Foti & Hajcak, 2008). The LPP is particularly sensitive to sentence processing tasks and is thought to reflect cognitive processes involved with expectancy violations, specifically semantic and thematic violations (Van Herten, Kolk, & Chwilla, 2005). Therefore, LPP is increased in response to unlikely and more salient information. Coulson (1998) suggest that increased potential in response to expectancy violations may reflect engagement of attention and updating of memory when individuals evaluate and interpret an expected event. Therefore the LPP and similar components such as the P300/P600 may be an appropriate measure of processes associated with hostility-related interpretation bias.

The N400 is also sensitive to violations of expectancy models and therefore differences in evoked amplitude of this ERP component have been used as a neural marker for hostile attribution bias (Gagnon et al., 2016; Gagnon et al., 2017). The N400 is a negative potential in the ERP waveform that reaches its maximum at approximately 400ms post stimulus onset in central-parietal electrodes (Kutas & Federmeier, 2011). This component is evoked by social expectancy violations, for example Moreno and Vázquez (2011) presented participants with positive and negative sentence stems which were randomly displayed with their emotionally matched expected outcome, or with emotionally mismatched outcome, or with nonsense. They found that nonsense elicited a large N400 amplitude regardless of the valence of the sentence stem. Individuals therefore must use knowledge stored in long term memory to make predictions about the upcoming outcome of a presented sentence (Kutas & Federmeier, 2000). These findings suggest that N400 may provide a neural marker for negative interpretation bias. It could be predicted that individuals with a hostile attribution bias would show an increased N400 in response to positive interpretations of ambiguous scenarios as they would expect a negative resolution.

There have only been a small number of studies which have used EEG methodology to examine interpretation bias. For example, Moser et al. (2008a) investigated interpretation bias in social anxiety using ERPs. They aimed to explore possible psycho-physiological correlates of interpretation bias associated with social anxiety. A low and high socially anxious group were screened and then recruited based on their score on the Social Phobia Inventory (Connor et al., 2000). EEG was recorded while participants viewed 120 ambiguous sentences that were resolved with a positive or negative final word. The final word was either grammatical or non-grammatical; participants were required to determine the type of resolution word for each sentence. Reaction times to the word resolutions were analysed for all correct trials. EEG recordings were taken from three locations along the midline. Average amplitude between 500 and 700 ms post stimulus onset was analyzed. The reaction time results did not yield any conclusive findings, however, the ERP analyses revealed a significant main effect in the low anxiety group but no significant effects in the high anxiety group. Individuals scoring low on social anxiety were characterized by larger P600 in response to negative sentence resolutions compared to positive, suggesting that negative endings were relatively unexpected and therefore they have a positive bias. High socially anxious individuals showed similar P600 in response to both negative and positive sentence resolutions. This suggests that anxious individuals expect negative outcomes and therefore their expectations were not violated (and thus P600 was not increased). Moser, Huppert, Foa, and Simons (2012) replicated and extended their previous work and found consistent results. The authors hypthesise that non-anxious individuals have a positive bias in which social situations are generally interpretted positively (therefore negtative resolutions evoke a peak in P600 amplitude). However in anxious samples there is no evidence of this positivity bias and therefore negative sentence resolutions do not evoke an increased P600 response. The results from these studies fit with expectency models of the P600 (LPP) component and contribute to the understanding of cognitive processes involved with interpreting the environment in social anxiety.

To my knowledge there are only very few studies that have used EEG methodology to investigate interpretation bias in aggression. Godleski, Ostrov, Houston, and Schlienz (2010) explored the variation in P300 amplitude in relational aggression and hostile attribution bias. To measure hostile attribution bias participants had to indicate a reason for provocation for a number of hypothetical vignettes of socially ambiguous relational and instrumental scenarios. Elicited P300 was measured using an auditory perseveration task in which participants were required to respond to high and low pitched tones along with white noise bursts. The findings suggest that relational aggression was associated with a hostile attribution bias and increased P300. An increased late positive potential is thought to reflect a greater allocation of neural cognitive resources, therefore suggesting that individuals with increased levels of relational aggression are overly sensitive to provoking cues. However, Godleski et al. (2010) used two separate tasks to measure hostile attribution bias and evoked P300 amplitude. Therefore conclusions cannot be drawn regarding the brain processes involved with the interpretation of hypothetical ambiguous scenarios.

Gagnon and colleagues have further assessed the neural correlates of interpretation bias in aggression (Gagnon et al., 2016; Gagnon et al., 2017). Gagnon et al. (2016) aimed to identify the neural mechanisms associated with expectations of hostile or non-hostile intent. Fifty non-aggressive participants were presented with 80 scenarios that included hostile and non-hostile situations. Each scenario included three sentences, the first sentence established the nature of the scenario; either hostile or non-hostile, the second sentence described an ambiguous social provocation that was directed at the reader, and the final sentence included a final target word that disambiguated the intention of a character in the scenario as hostile versus non-hostile. There were therefore four conditions; a hostile situation with a non-hostile resolution (mismatch), a hostile situation with a hostile resolution (match), a non-hostile situation with a hostile resolution (mismatch), and a non-hostile situation with a non-hostile resolution (match). Participants had no specific task but were asked to imagine the thoughts and feelings of the character in the scenario. ERPS in response to the target word of each scenario were recorded. The results showed that N400 was increased in response to mismatch resolutions compared to matched resolutions, therefore when the intention of the target word was not expected, a larger N400 was elicited. This effect was particularly salient for non-hostile target words that violated the expectations of hostile scenarios. These findings show that non-aggressive individuals rapidly evaluate the hostile intent behind the ambiguous behaviours of characters in a social context. Consistent with a typical N400 effect, violation of expected outcomes elicits increased amplitude.

Gagnon et al. (2017) extended this work using an aggressive sample; methods were identical to those implemented by Gagnon et al. (2016). They found that in the aggressive group there was an increased N400 effect in response to non-hostile words that violated the hostile expectancy of the scenario. There was also an enhanced late positive potential-like component in response to hostile words that violate the non-hostile intention expectations in response to non-hostile scenarios. These findings provide further evidence that the N400 is a useful component for investigating interpretation bias. It also suggests that increased LPP may reflect the difficulty in integrating non-hostile social cues and therefore may play a role in the attribution of hostile intent.

2.4 Overview of literature

Within the attention bias literature, there is a focus on threat-related attention bias in anxiety. Although studied to a lesser extent, there is also evidence to suggest hostile-related biases in aggression. These studies indicate that aggressive individuals preferentially attend to aggression-related stimuli compared to neutral stimuli across a number of different tasks.

However, there are a number of gaps in the literature which I will identify and aim to build upon. Firstly, there is a lack of studies which use selective attention tasks, such as the dot-probe task, to explore attention biases in aggression. Secondly, studies have predominantly included threat words as stimuli, rather than angry or threat faces, and to my knowledge no studies have directly compared the attentional processes involved with selectively attending to words and faces and whether there are marked differences between modalities. Finally, studies have mainly relied on behavioural methods, such as reaction time and recall, to draw conclusions on attention bias in aggression. More recently, neuro-psychological methods have been used to explore cognitive processes such as attention. However, studies focusing on aggression are somewhat limited. The ERP studies on attention bias and aggression have a number of methodological limitations, for example the tasks do not allow for conclusions to be drawn regarding selective attention, and they analyse only very few midline electrodes. I aimed to address these methodological issues, as well as advancing knowledge on the neural correlates of attention bias.

The review of the interpretation bias literature revealed that there is a fairly robust association between negative interpretation bias and aggressive behaviour. Aggressive individuals interpret ambiguous scenarios as more hostile in nature compared to non-aggressive controls (Epps & Kendall, 1995). They are also more likely to attribute hostile intent to a protagonist in a scenario (Dill et al., 1997). These results have been demonstrated using a number of different experimental tasks, however to my knowledge only very few studies have explored the ERP correlates of interpretation bias in aggression. The aim was to build on these existing studies by assessing the validity of measuring interpretation biases using ERP methods, and also to explore between group differences in ERP patterns in response to making negative interpretations.

I believe that understanding how cognition affects behaviour, particularly attention and interpretation processes, may have rehabilitative value. Literature has shown that cognitions can be influenced and modified by training methods, such as attention bias modification (e.g. Eldar & Bar-Haim, 2010). These training methods have been shown to be successful in reducing anxious symptoms and behaviours. This evidence suggests that modifying cognitions is an appropriate treatment method for changing behaviours. Therefore I suggest that understanding the cognitive processes that contribute to aggressive behaviour may be essential in designing intervention and rehabilitation programmes for aggressive offenders.

2.5 Thesis aims and outline of studies

The overall aim of the thesis is to address each of these gaps in the literature and to increase understanding of how cognitive biases contribute to aggression by identifying neural correlates associated with these biases. More specific aims are:

1. To provide an initial assessment of the validity of the dot-probe paradigm for investigating neural correlates of selective attention bias in aggression
2. To provide an initial assessment of the reliability of a recognition task to investigate the neural correlates of hostile interpretation bias in aggression.
3. To investigate the relationship between aggressive behaviour and attention bias to angry and happy words using behavioural and EEG methods.
4. To investigate the relationship between aggressive behaviour and attention bias to angry and happy faces using behavioural and EEG methods.
5. To investigate the relationship between aggressive behaviour and hostile interpretation bias using behavioural and EEG methods.
6. To establish possible neural correlates associated with negative attention bias and hostile interpretation bias with a view to increasing understanding of cognitive processes underlying aggressive behaviour

This thesis includes five studies which have been designed to address the specific aims outlined above. To assess the validity of the dot-probe paradigm for investigating neural correlates of attention bias in aggression four studies were conducted that used versions of the dot-probe task with simultaneous EEG recording to compare behavioural reaction time results with evoked amplitude in response to differently valenced stimuli. The four studies are made up of two sets of complementary studies, with each set including a different stimulus modality; the first two studies assessed attention bias to words, whereas the second two studies explored attention bias to faces. The first study within each set (Studies 1 and 3) used a simple paradigm which included just one trial type; angry-neutral. During these studies, angry and neutral stimuli (words or faces) were presented simultaneously, and an arrow probe appeared in the position of one of the previously presented stimuli. The difference in reaction time and amplitude in response to congruent (probe replaces angry stimuli) and incongruent (probe replaces neutral stimuli) trials between aggression groups were compared. The second study within each set (Studies 2 and 4) used a more complex dot-probe design which included three trial types; angry-neutral, happy-neutral, and angry-happy. Within each trial type the probe could appear in a congruent or incongruent position. Differences in reaction time and evoked amplitude between trial types and trial congruency between aggression groups were explored. The aim was to explore if attentional processes involved with attending to stimuli during the dot-probe varied between aggression groups, between modalities, and between emotion of the presented stimuli.

The fifth and final study included in this thesis investigated interpretation bias in aggression. A recognition task in conjunction with simultaneous EEG recording was used to identify possible neural correlates of making negative interpretations associated with increased levels of aggression. Due to the novelty of the combined behavioural and EEG methods the first aim was to assess the validity of the recognition task, and secondly to explore differences in interpretation bias scores and associated neural patterns between aggression groups.

3 Study 1 - Attention bias to angry words

3.1 Introduction

This first empirical chapter explores cognitive processes associated with attention bias to angry words in high and low physically aggressive individuals. This study uses an original design including both behavioural and EEG methods, with the aim of identifying neural correlates of attention bias. Attention bias is defined as the preferential allocation of attentional resources to aversive stimuli compared to benign stimuli (MacLeod et al., 1986). I discuss findings in relation to facilitated engagement, which is the process by which threat-related stimuli are detected faster than neutral stimuli (Bar-Haim et al., 2007), and difficulty in disengagement, which is difficulty in allocating resources away from threat-related stimuli once it has been engaged (see Cisler & Koster, 2010; Koster et al., 2006).

The relationship between increased aggression and attention bias is evident in both forensic and non-forensic samples, and across attentional bias paradigms (Smith & Waterman, 2003). Smith and Waterman (2003) assessed attention bias towards violently themed words during two tasks in an offender and undergraduate population. Across both a dot-probe and Stroop task, aggressive participants from both samples showed increased attention facilitation and interference of aggressively themed words. This study shows that violent stimuli may be particularly salient to aggressive individuals. Further evidence from van Honk et al. (2001b) also shows differences in attention bias between high and low trait anger groups using an Emotional Stroop task in which threatening and neutral words were presented. The task was completed under both masked and unmasked conditions. Results showed differences in responses between the high and low trait anger groups for the unmasked task only. High trait anger participants took relatively longer to colour name the threatening words in comparison to the neutral words, which suggests that interference of meaning of the word influenced their ability to complete the task efficiently. Attentional interference refers to difficulty in disengaging from threat-relevant information which then restricts processing resources needed for another task (Fox et al., 2001; Fox et al., 2002). These results suggest that anger may not influence automatic attention biases that are masked from conscious awareness; however, it shows that high trait anger participants have difficulties attending to the colour of the word once they have become consciously aware of the threat word. This evidence indicates that it is a combination of facilitated attention and difficulties with disengaging from aggression-related stimuli that contribute to an attention bias in aggression.

There is some evidence to suggest that aggression-related attention bias may be particularly salient in individuals with increased levels of physical aggression. Smith and Waterman (2005) investigated processing biases to an emotional Stroop task in a non-clinical undergraduate sample categorised according to their self-reported aggression score. With the aim of exploring the effects of different types of aggression, four subscales were studied; hostility, anger, verbal aggression and physical aggression. Results showed physically aggressive males had a significantly delayed response to colour naming words that related to direct acts of aggression, showing an attention bias towards such stimuli. Further evidence from Chan et al. (2010) shows that during an emotional Stroop task, in comparison to a control group, male batterers with increased reactive aggression scores, had longer reaction times when naming the colour of negative words compared to neutral words. This finding is consistent with theoretical accounts of aggression, based on the frustration-aggression model (Berkowitz, 1993; Dollard et al., 1939).

The literature shows a relatively consistent behavioural association between attention bias to aggression-related words and increased aggression. However, reaction time represents a combination of attentional, evaluative, and motor processes (Donders, 1969; Sternberg, 1969). More recently, a small number of studies have employed commonly used behavioural tasks with simultaneous EEG recording, to explore the role of the P300 component in aggression-related attention biases. EEG can capture changes in brain processes between milliseconds

(O'Toole & Dennis, 2012) and therefore provides a direct measure of neural activity evoked by events of interest (reviewed in Luck, 2005). This component appears as a positive deflection at parietal sites between 300 and 800ms after stimulus onset. It is particularly sensitive to selective attention, that is, the differential processing of stimuli in relation to their task relevance (Coles et al., 1995; Polich, 2007). There is evidence to suggest that non-aggressive healthy participants process threat and neutral information differently. Thomas et al. (2007) found that during a Stroop task participants showed an increased P300 amplitude in response to threat words compared to neutral words.

Helfritz-Sinville and Stanford (2015) used a modified oddball task to assess the P300 component in relation to attention biases in the processing of threat stimuli. They investigated how reactive, premeditated, and non aggressive participants process social and physical threat words compared with neutral words. They found that non-aggressive participants showed increased P300 amplitude when presented with both social and physical threat words compared to neutral words. Both reactive and premeditated aggressive participants showed relatively stable P300 amplitude across responses to all word types. These results suggest that aggressive individuals do not differentiate between stimulus types; however it is not clear from this evidence if processing of threat-related or neutral words differs between individuals with high and low levels of aggression. There are two possible interpretations of these findings; in comparison to the non-aggressive participants, aggressive participants perceive neutral words as more similar to the threatening words; or attribute fewer cognitive resources to the processing of aggression-related stimuli compared with neutral stimuli. However, Stewart et al. (2010) found that individuals with higher anger-out scores showed increased P300 amplitude in response to the negative words compared to neutral words during an emotion-word Stroop task. Stewart et al. (2010) propose that this increased positive amplitude reflects greater cognitive effort in overriding attention to negative information.

The literature exploring attention bias to hostile words primarily uses the Stroop task or oddball task to infer preferential attention to aggression-related stimuli. The Stroop task is a measure of interference in attentional processing and can be used to infer attention. A bias on the emotional Stroop task can be attributed to attentional engagement with, or disengagement from the content of aggression-related stimuli (Clarke, Macloed, & Guastella, 2013). The oddball task is frequently employed with EEG methods to examine the P300 component in relation to processing of rare-target and rare-non-target stimuli. Both of these tasks present stimuli singly and therefore a more appropriate method for measuring selective attention is the dot-probe paradigm (MacLeod et al., 1986) which presents aversive and benign items simulataneously. It also allows for two types of EEG analysis; the examination of evoked amplitude in response to stimulus onset (word/face pair) and probe onset.

3.2 Aims and rationale

Collectively, in line with current cognitive models, findings show an aggression-related attention bias in aggressive samples (Wilkowski & Robinson, 2010). However, the published research in this area primarily uses Stroop and oddball tasks. Therefore, using the dot-probe paradigm, the first aim is to test whether findings by Smith and Waterman (2003) would be replicated; that aggressive individuals show a behavioural attention bias towards angry words when they are presented alongside a neutral word.

Previous literature suggests that the ERP component, P300, may act as an electrophysiological marker of selective attention. The findings show that aggression-prone individuals have similar amplitudes across stimulus types when presented with threat-related and neutral words (Helfritz-Sinville & Stanford, 2015). In contrast, within low aggression normative samples it has been shown that there is a pattern of increased P300 amplitude to aggression-related words compared to neutral (Helfritz-Sinville & Stanford, 2015; Thomas et al., 2007). However, further work is needed to examine these differences in greater detail. By comparing evoked amplitude on congruent and incongruent trials in the high and low aggression groups, the aim was to explore whether, compared to less aggressive individuals, aggressive individuals allocate greater cognitive resources to neutral stimuli because they are perceived as hostile, or whether they allocate fewer resources to angry stimuli as they are desensitised to such stimuli and therefore can be processed with greater efficiency. The final aim was therefore to explore the neural correlates of attention bias across high and low aggression groups. In order to investigate the specificity of this bias in greater detail, and draw conclusions as to whether attentional facilitation or difficulty in disengagement contributes to attention bias, the ERP patterns in response to simultaneous angry and neutral word presentation during a selective attention task was analysed. The difference in evoked ERPs following word pair presentation between high and low aggression groups, and the difference in evoked ERPs following probe presentation between congruent and incongruent trials was analysed.

Studies investigating attention biases in aggression have primarily measured trait anger (Eckhardt & Cohen, 1997; van Honk et al., 2001b) as this is considered a consistent internal characteristic. However, this is an implicit form of aggression, relating to feelings of anger, and does not necessarily imply an aggressive reaction to a scenario. The current study investigated neural processing relating to attention bias, specifically in physical aggression. Physical aggression is a measurable explicit behavioural response which is an expression of anger. A male-only sample was recruited because males show higher levels of physical aggression than females (Archer, 2004). Inclusion of these variables allowed for greater comparison with previous work by Smith and Waterman (2005), which found physical aggression to be predictive of hostile attention bias, and Helfritz-Sinville and Stanford (2015), which explored the processing of threat words in impulsive and premeditated physically aggressive men.

3.2.1 Research questions and hypotheses

Overarching research question: Do high aggression participants have an increased attention bias to angry words compared with low aggression participants, and is this reflected in different ERP patterns in response to angry and neutral stimuli between aggression groups?

Hypothesis one: Relative to participants with low levels of physical aggression, participants with increased physical aggression scores will show an increased attention bias to angry words, characterized by a faster reaction time to probes on congruent trials compared to incongruent trials.

Hypothesis two: Increased self-reported attentional control will be correlated with decreased levels of physical aggression and decreased attention bias to angry words.

Hypothesis three: Compared to the low physical aggression participants, the high physical aggression participants will have decreased P300 amplitude in response to the presentation of angry-neutral word pairs.

Hypothesis four: Participants with low levels of physical aggression will show increased P300 amplitude to congruent trials compared to incongruent trials, whereas participants with high levels of physical aggression will show undifferentiated P300 in response to both congruent and incongruent trials.

3.3 Methods

3.3.1 Power Analysis

An a *priori* power calculation was conducted using G*Power 3.1 software (Faul, Erdfelder, Buchner, & Lang, 2009) based on the most complex planned analyses. For repeated measures mixed model ANOVA analyses, based on 40 measurements, 2 groups and a small to medium effect size (f = 0.20), a minimum sample size of 12 participants will be needed to achieve 90% power, when $\alpha = .05$.

3.3.2 Participants

Data were collected from 36 male University of East Anglia (UEA) students and staff, and members of the wider community. In order to take part in the study participants had to be male, aged between 18 and 35, speak English as their first language and have normal or corrected vision. They also were unable to take part if they had been diagnosed with a psychological condition in the last 12 months, were receiving psychological treatment or were taking anabolic steroids. Efforts were made to recruit participants with a wide range of aggression scores by, for example, distributing adverts that included questions such as ‘Do you tend to lose your temper?’ (Appendix A) and ‘Do you frequently get road rage?’ Of the total sample, 51% were students recruited through the university SONA system (University of East Anglia student study sign up system), the remaining 49% were volunteers recruited from across the university using various methods, for example, email and social media advertising, poster campaign, distributing leaflets and word of mouth. The sample ranged in age from 18 to 35 (M = 21.77, SD = 4.55). The majority of the sample was White British (83%), with the other 17% being African, Asian and of mixed ethnicities. The majority of the sample had some university education, ranging from undergraduate to PhD level (54.3%). All other participants had sixth form level education (45.7%).

One participant was ineligible and was therefore excluded from analyses. Three further participants were also excluded from analyses; two due to excessive noise during EEG recording, and one due to a fault in recording. Therefore for all continuous analyses the total sample consisted of 32 participants (*M* = 21.97, *SD* = 4.70). The participants were categorised into high and low aggression groups based on the physical aggression subscale. Two participants had scores that equaled the median and consequently could not be grouped, therefore both behavioural and ERP between-subjects analyses included a sample of 30 participants (15 high physical aggression, 15 low physical aggression).

3.3.3 Self-report measures

3.3.3.1 Demographics

Participants provided some basic information about themselves, for example, age, gender, ethnicity and employment status (Appendix B).

3.3.3.2 Aggression Questionnaire (Buss & Perry, 1992; Appendix C)

The aggression questionnaire involves responding to 29 statements on a 5-point likert scale which ranges from ‘extremely uncharacteristic of me’ to ‘extremely characteristic of me’. There are four subscales which make up the 29 items; nine items measure *physical aggression*, five *verbal aggression*, eight *measure anger* and eight items measure *hostility*. Example items include statements such as ‘I tell my friends openly when I disagree with them’ (verbal aggression), ‘I have threatened people I know’ (physical aggression), ‘I sometimes feel that people are laughing at me behind my back’ (hostility), and ‘Some of my friends think I am a hothead’ (anger). Each item is scored from one to five, with items 4 and 19 being reversed scored. Total scores range between 29 and 145, with higher scores representing a higher level of aggression. Participants completed this questionnaire online via Qualtrics. The Aggression Questionnaire (Buss & Perry, 1992) is a consistently used measure of aggressive attitudes and behaviours (Giancola & Parrott, 2008; Helfritz-Sinville & Stanford, 2015; Smith & Waterman, 2003). Harris (1997) conducted an analysis of the four subscales of the aggression questionnaire and found that they all have moderate to high internal reliability. The analysis also showed that the measure had some degree of construct validity. The physical aggression scale has good reported internal consistency (αs = .85) (Buss & Perry, 1992).

3.3.3.3 Attentional Control Scale (ACS); (Derryberry & Reed, 2002; Appendix D).

Participants are asked to respond to 20 statements on a four point scale, with 1 being ‘almost never’, 2 being ‘sometimes’, 3 being ‘often’ and 4 being ‘always’. They are asked to indicate how much they think the statement applies to them. Nine items of the twenty refer to *attention focusing* and 11 to *attention shifting*. Example items include ‘when I am working hard on something, I still get distracted by events around me’ and ‘I have trouble carrying on two conversations at once’. Eleven items are reverse scored and then all items are totalled to give a final score. Higher scores reflect better attentional control. The ACS has good reported reliability with reported Cronbach’s alpha being between .71 (Verwoerd, de Jong, & Wessel, 2008) and .88 (Derryberry & Reed, 2001).

3.3.3.4 Delinquency Questionnaire (DQ; taken from (Tarry & Emler, 2007; Appendix E).

This questionnaire is used to determine the participant’s delinquent involvement. They are asked to respond to 24 statements, indicating how many times they have behaved in a certain way in the last 12 months. Items include statements such as, ‘purposefully annoyed, insulted, or taunted strangers in the street’, ‘driven a car on the roads without a licence’ and ‘been involved in a group fight’. Responses range from zero to three, with zero being equal to ‘never’, 1 being ‘once or twice’, 2 being ‘a few times’ and 3 being ‘several times’. Scores for the 24 statements are summed to give a total score between 0 and 72, with higher scores representing a higher level of delinquency. The scale has excellent reported reliability (α = .94; Tarry & Emler, 2007).

3.3.3.5 Trait form of the State-Trait Anxiety Inventory (STAI-T) ;(Spielberger & Gorsuch, 1970; Appendix F)

The trait form of the STAI is an established and widely used measure of trait anxiety. For each of its 20 items, participants are required to rate themselves on a 4 point scale representing general perception of stressful situations that may involve danger or threats to the individual (Spielberger et al., 1983). There are 9 positive items, for example ‘I am happy’, reflecting the absence of anxiety, and 11 negative items, for example ‘I feel like a failure’, reflecting the presence of anxiety. Participants are asked to state how they generally feel in relation to each statement on a 4 point scale, with 1 being almost never, 2 sometimes, 3 often and 4 almost always. The positive statements are reverse scored and a composite score is generated by summing the individual items (range 20-80). A higher score reflects a higher level of anxiety. Barnes, Harp, and Jung (2002) examined the reported internal reliability of the STAI in over 50 research articles and concluded that on average the scale had an internal consistency reliability coefficient of 0.91. Further past research has shown this measure to be reliable and internally consistent with Cronbach’s alpha ranging from 0.86 to 0.95 (Spielberger et al., 1983).

3.3.4 Attention bias test

Attention bias was measured using the probe classification version of the dot-probe task, adapted from MacLeod, Rutherford, Campbell, Ebsworthy, and Holker (2002). In comparison to the original dot-probe task in which participants are required to respond as quickly as possible to a single probe, in the classification version, participants have to indicate the type of probe that is displayed for each trial (for example, left facing arrow or right facing arrow). Therefore, participants are required to attend to the probe in greater detail, encouraging more equal monitoring of both areas of the display (Mogg & Bradley, 1999c). The task was programmed using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) and administered in a laboratory. Participants were seated 60cm from a 23-inch monitor, affording a visual angle of approximately 3 degrees between items (cf. see, MacLeod, & Bridle, 2009). There were a total of 96 trials, with each of the 12 word/face pairs being presented eight times. Each trial began with a fixation point (three small crosses) in the centre of the computer screen for varying duration (range 1060 to 1973ms), followed by presentation of the stimulus pair for 500ms in a randomised order (approx. 6 minutes). The word/face pairs were separated by a vertical distance of 3cm above and below the central fixation cross. Next, a left- or right-pointing arrow probe (“<” vs. “>”) appeared in the prior location of the angry or neutral stimulus until response (see Figure 2). Congruent trials are defined as those in which the arrow appears in the prior location of the angry word/face, whereas incongruent trials refer to those in which the arrow appears in the prior location of the neutral word/face.

The direction (left or right) and location (top or bottom) of the arrow probe was equally distributed across trial types and presentation order was randomised throughout the test. Participants were instructed to identify the direction of the on probe using the arrow keys as quickly and accurately as possible. A one-second blink screen followed the target response to minimize ERP artifacts, after which the next trial started immediately. Aggression-related attention bias is characterized by faster reaction times to congruent trials compared with incongruent trials. There were 10 practice trials (where a “Correct!” or “Incorrect” feedback message appeared after the participant had pressed the arrow key). A break occurred halfway through the test (after 48 trials).

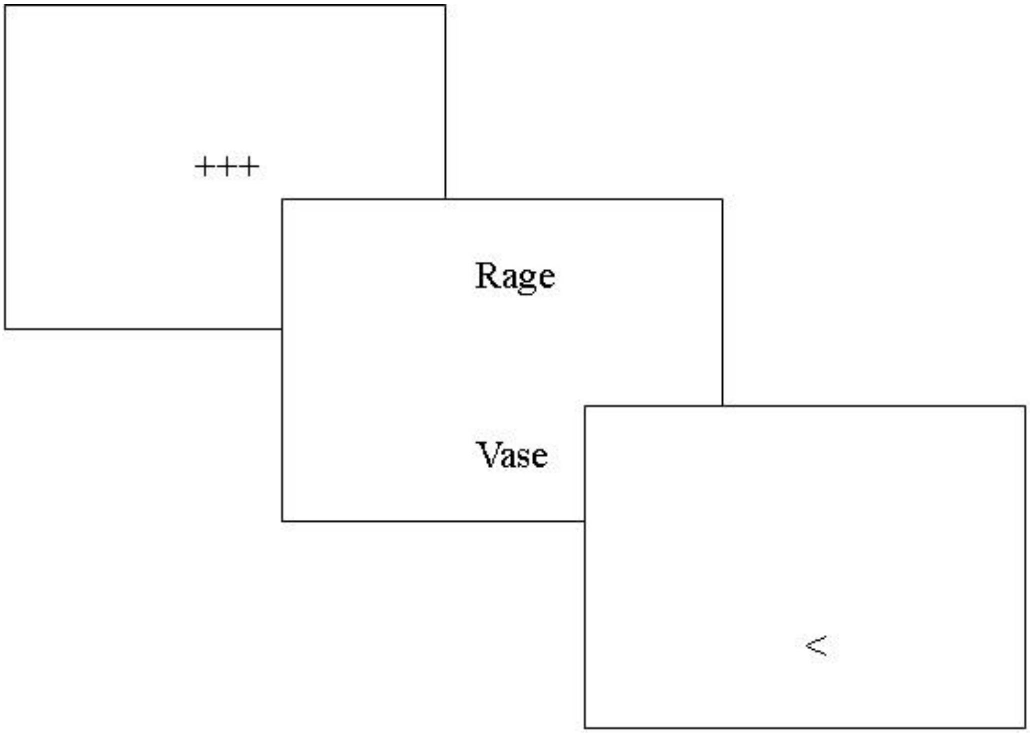

Figure 2: Procedure for the dot-probe task; a) fixation cross is presented in the centre of the screen for a randomized time between 1060 and 1973ms; b) the word pair is presented for 500ms; c) an arrow probe is presented in the prior location of either aggression-related or neutral word and stays on screen until participant response.

3.3.5 Attention bias test stimuli

The stimuli included 12 angry-related words, compared with 12 neutral household-related words (black text on a white background) (see Table 1). Eight of the twelve matched words were used based on prior studies (Faunce, Mapledoram & Job, 2004; Liossi, White & Schoth, 2011), and the further four were developed by the researcher to complement the existing word pairs. The words were matched for length and frequency using the Brysbaert database (Brysbaert & New, 2009). It was decided to use household-related neutral words to control for semantic relatedness and minimise the possible confound of category priming, due to the relatedness of angry-words (Mogg, Bradley, Williams, & Mathews, 1993).

Table 1: Attention bias test stimuli: 12 angry-related words and 12 neutral household-related words.

angry	**neutral**
explosive	framework
hostile	chimney
infuriated	percolator
angry	craft
volatile	verandah
irate	mixer
resentful	appliance
vicious	cutlery
rage	vase
vexed	chair
oppose	tables
aversive	curtains

3.3.6 EEG Acquisition

The School's EEG laboratory protocol (Version 1.1, 24.02.15) was followed throughout to ensure safe and responsible administration of the procedure. EEG was recorded with a 32-channel active electrode system (Brain Products GmbH) embedded in a nylon cap (10/10 system extended). An additional electrode was placed under the left eye in order to monitor vertical eye movements (lower electroculography; EOG). The continuous EEG signal was acquired at a 500 Hz sampling rate using FCz as reference. The impedance was kept below 20 kΩ. ERPs were time-locked to the onset of each stimulus pair. This allowed for the measurement of cognitive processes involved in selective attending following the simultaneous presentation of an angry and neutral word. Faster reaction to the probe is considered to reflect the allocation of attentional resources in the direction of the previously presented word/face prior to the appearance of the probe. Therefore differences in processes associated with competition for attentional resources during the presentation of face/word pairs was explored. Differences in evoked amplitude between congruent and incongruent trials following probe presentation were also explored.

3.3.7 Procedure

Ethical approval was granted from the School of Psychology Ethical Committee. The study was split into two sections; an online questionnaire and a lab session. Initially participants were asked to complete an online questionnaire (Qualtrics); this involved giving consent, reporting demographic information and completing the BPAQ. It also included an eligibility checklist (Appendix G) and gave the participants an opportunity to create their own unique ID (Appendix H) so all data could be matched correctly and stored anonymously. Towards the end of the of the recruitment process, to obtain an even distribution of aggression scores, a number of participants were screened for higher levels of aggression and only participants that gained a total aggression score of 82 or above were invited to attend the laboratory session. Five participants who completed the screening process and scored below 82 did not take part in the second session of the study.

Within one month of completing the online questionnaire, participants then took part in a 90- minute laboratory session. During this session, participants were asked to read an information sheet (Appendix I) and sign a consent form (Appendix J). The true objectives of the study were not revealed until debrief with the aim of minimising demand characteristics. It was then requested that consenting participants complete paper copies of the four questionnaires (ACS, DQ, STAI-T and AIHQ) and three experimental tasks. They completed the dot-probe word task, followed by the dot-probe face task (results of this are reported in Chapter 5). Participants also completed a recognition task to assess interpretation bias (results of this are reported in Chapter 7). Participants were given very basic information when competing the dot-probe task, they were informed that they would see two faces appear on the screen, followed by an arrow. They were asked to respond as quickly and accurately to the arrow as possible. During completion of the experimental tasks, participants wore a nylon cap embedded with 32 electrodes.

One electrode was also placed under the left eye to record eye movements. Participants were instructed to remain as still as possible during the tasks and were asked to try and not to blink during stimulus presentation to reduce the occurrence of muscle or ocular artifacts in the EEG recording. To record accurately, it was necessary to put a water-based gel into the hair under each electrode using blunt syringes. Before being fully debriefed (Appendix K) the participants were given the opportunity to wash their hair. The testing session lasted 1 and a half hours; the experimental tasks lasted approximately 30 minutes of this time. To minimize order effects, completion of the computer-based tasks and questionnaires was counterbalanced. Participants received course credits or shopping vouchers (£10) as compensation.

3.3.8 Data analysis plan

3.3.8.1 Behavioural attention bias data

Median reaction times on congruent (probe replacing angry word/face) and incongruent (probe replacing neutral word/face) trials were extracted as they are not skewed by extreme scores (e.g., Whelan, 2008). An *attention bias index score* was calculated by subtracting the median reaction time on incongruent trials from the median reaction time on congruent trials. Therefore a negative bias score indicates that participants responded more rapidly when probes replaced angry than neutral words.

Both median-split and correlational approaches were used to evaluate how levels of physical aggression in participants were related to attention bias. The association between attention bias index and physical aggression was explored using Pearson’s correlation. A 2 (congruency; congruent, incongruent) x 2 (physical aggression; high, low) ANOVA was conducted to explore differences in reaction times between congruent and incongruent trials in the high and low physical aggression group.

3.3.8.2 EEG data

Offline analyses were conducted using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014), which are open source toolboxes running under Matlab 7.12 (R2013a, The Mathworks). High- and low-pass filter half-amplitude cut-offs were set at 0.1 and 40 Hz, respectively. Before averaging, trials contaminated by excessive artifacts were rejected automatically using a step function (Luck, 2005) with a voltage threshold of ± 100 µV in moving windows of 200ms and with a window step of 100ms. Noisy channels were interpolated using the EEGLAB function eeg_interp (spherical interpolation). The data was not re-referenced offline.

ERP data extracted from the raw EEG data was time-locked to the onset of the face/word pair. Data was segmented into epochs of 1200ms; from -200ms to 1000ms post stimulus (word/face) onset, with -200-0ms pre word/face pair onset as baseline. Mean amplitude between 100-200ms, 200-300ms, 300-400ms, 400-500ms, 500-600ms, 600-700ms, 700-800ms, 800-900ms, and 900-1000ms post stimulus onset were extracted for statistical analyses. Epochs between 100 and 500ms refer to pre-probe presentation, whereas epochs between 500 and 1000ms refer to post-probe presentation. The timing of effects in relation to word/face pair onset and probe onset are important for distinguishing between ERP components. Analysis focused on posterior parietal electrode sites, including CP1/2, CP5/6, P7/8, P3/P4 and TP9/10, where P300 component is considered to be maximal (e.g., Iwaki, Sutani, Kou, & Tonoike, 2007; Polich, 2007).

To explore the main effect of aggression across all trials, one-way ANOVAs were conducted to explore whether high and low aggression groups showed differences in evoked amplitudes in response to the onset of the word/face pair at each electrode. This analysis was conducted for each epoch.

To investigate the effect of trial congruency on amplitude between aggression groups, a mixed model ANOVA was performed on ERP measures for all selected epochs for the region of interest. Driven by the hypotheses it was expected that congruency effects would be evident between 500 and 1000ms, however based on qualitative inspection of the waveforms this analysis was conducted for all epochs (100-1000ms). The ANOVA had the following within-subject factors: trial congruency (congruent versus incongruent), electrode (5 levels) and hemisphere (left versus right). Physical aggression group was added as a between-subject factor. ERP measures were evaluated on correct trials only (3425 out of a total 3456 (99.1%).

Greenhouse-Geisser F tests (Geisser & Greenhouse, 1958) are reported throughout for all repeated measures to avoid violations of the sphericity assumption. Across the results section, some alpha values above $p = .05$ are presented. The decision was made to present p values that were above the conventional significance value, to show significance levels of electrodes across all epochs. With visual reference to the ERP waveforms I present significance values across epochs where there are qualitative differences. Therefore, this transparency in reporting allows for closer examination of the epochs in which electrodes reach significance, and at which epochs electrodes may be outside conventional significance levels. For consistency of reporting, I refer to p values above 0.05 and below 0.1 as ‘approaching significance’.

3.4 Results

3.4.1 Data preparation

3.4.1.1 Missing Items

The DQ and STAI-T had no missing items. The ACS and the BPAQ (from the physical aggression subscale) each had one case of missing data. Missing values were replaced with the mean of the completed items for each questionnaire (method used by Judah, Grant, Mills, & Lechner, 2014). This simple approach was selected as it is considered to make relatively little difference if missing data represent less than 5% of the dataset (Tabachnick & Fidell, 2013).

3.4.1.2 Normality of data

All data (BPAQ and subscales, ACS, and STAI-T) was normally distributed apart from the delinquency questionnaire, which was just outside acceptable limits of ±2 (due to floor effect) (Appendix L).

The two reaction time variables (congruent and incongruent trials) extracted from the dot-probe task were not normally distributed due to skewness and kurtosis calculations (scores divided by the subsequent standard error) that were outside acceptable limits of ±2 (Appendix L), and were therefore analysed using non-parametric test where appropriate. However, the skewness and kurtosis scores for the calculated bias (congruent minus incongruent) were within acceptable limits therefore this data was analysed using parametric tests.

3.4.1.3 Reliability of questionnaires

The BPAQ (α = .92), physical aggression subscale from BPAQ (α = .90), anger subscale from BPAQ (α = .81), hostility subscale from BPAQ (α = .88), DQ (α = .81), and STAI-T (α = .94) demonstrated good internal reliability. The Verbal Aggression subscale from the BPAQ (α = .77) and ACS was only moderately reliable (α = .66).

3.4.2 Descriptive Results

Table 2: Means (SD) for the whole sample, and low and high physical aggression groups for all questionnaire measures.

	Total aggression	Physical aggression	Verbal aggression	Hostility
Whole sample (n = 32)	72.13 *(19.43)*	20.38 *(8.62)*	14.88 *(4.02)*	21.06 *(7.34)*
Low physical aggression (n = 15)	59.95 *(15.10)*	13.01 *(2.67)*	13.07 *(3.65)*	20.80 *(7.30)*
High physical aggression (n = 15)	84.33 *(16.80)*	28.07 *(5.92)*	16.87 *(3.80)*	21.27 *(8.09)*

	Anger	Delinquency	ACS	STAI-T
Whole sample (n = 32)	15.81 *(5.15)*	5.25 *(5.42)*	52.19 *(5.65)*	39.31 *(11.13)*
Low physical aggression (n = 15)	13.07 *(4.38)*	3.00 *(2.90)*	51.75 *(6.55)*	38.47 *(12.28)*
High physical aggression (n = 15)	18.13 *(4.55)*	7.40 *(6.70)*	52.73 *(5.12)*	40.27 *(11.02)*

3.4.2.1 Aggression Questionnaire (Buss & Perry, 1992)

The sample was categorised based on a median split of the *physical aggression* subscale of the Aggression Questionnaire. The high aggression group (M = 28.07, SD = 5.92) significantly differed from the low aggression group (M = 13.01, SD = 2.67); $t(28) = 8.976$, $p < .001$, d = 3.28) (see Table 2 for a closer inspection of the means of low and high aggression groups). There was also a strong positive correlation between physical and total aggression scores (Table 3).

Table 3: Correlations between all subscales of the aggression questionnaire (n = 33).

	Physical Aggression	Verbal aggression	Anger	Hostility
Total Aggression	.788 (<.001)	.826 (<.001)	.858 (<.001)	.667 (<.001)
Physical Aggression		.606 (<.001)	.619 (<.001)	.146 (.424)
Verbal aggression			.657 (<.001)	.465 (.007)
Anger				.482 (.005)

3.4.2.2 Questionnaire variables

The aggression data was explored and this showed that total aggression (and all other subscales) significantly correlated with delinquency; r = .556, p = .001 (one-tailed). However this chapter focuses on the aggression data as the sample was normative and therefore there was a floor effect of delinquency. Total aggression; r = .386, p = .015 (one-tailed), anger; r = .382, p = .016 (one-tailed), and hostility; r = .611, p < .001 (one-tailed) all positively correlated with anxiety. To investigate whether anxiety was a possible covariate, the effect of anxiety on attention bias was investigated. Anxiety did not significantly correlate with an attention bias for angry *words;* r = .197, p = .140 (one-tailed). This was supported by independent samples t-tests conducted on categorical anxiety data created using a median split (Mdn = 38.0). There was no significant difference in attention bias for negative *words* between high (M = 0.00, SD = 23.96) and low anxiety (M = -6.20, SD = 19.12); $t(27)$ = 0.773, p = .446, d = 0.286.

3.4.3 Results relating to hypotheses

Results are presented for physical aggression only. Based on previous research it is suggested that increased attention bias to angry stimuli may be particularly marked in violent or physically aggressive individuals (Smith & Waterman, 2005). The current behavioural results support this; only levels of physical aggression influenced attention bias to words. Also, having explored the effects of the different subscales of aggression on evoked P300 amplitude, it was evident that physical aggression was the main driver of these significant ERP findings. See Appendix M for the significant main effects and interactions with total aggression.

3.4.3.1 Hypothesis one

3.4.3.1.1 *Correlations.*

Results revealed a positive moderate correlation between *physical aggression* and attention bias to negative words; $r = -.442$, $p = .006$ (one-tailed) (Figure 3). This result suggests that those participants who scored higher on the physical aggression subscale of the BPAQ were quicker to respond on congruent trials compared to incongruent trials, showing support for hypothesis one. Total aggression; $r = -.145$, $p = .215$ (one-tailed), verbal aggression; $r = .035$, $p = .424$ (one-tailed), anger; $r = -.104$, $p = .287$ (one-tailed), and hostility; $r = .188$, $p = .151$ (one-tailed) did not correlate with attention bias score.

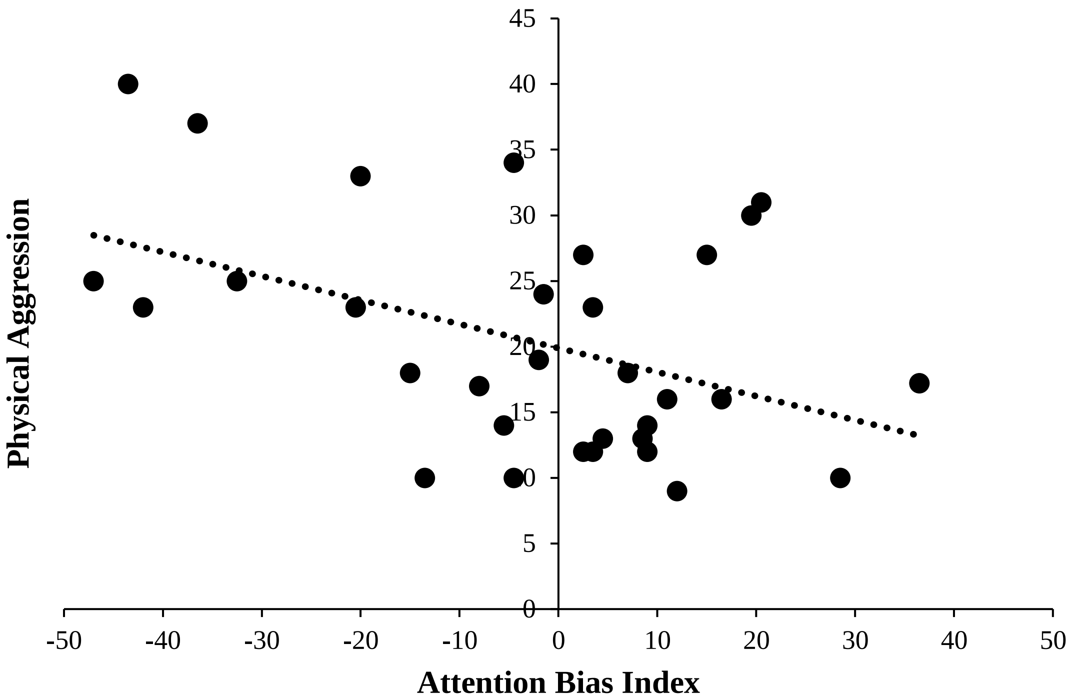

Figure 3: Scatterplot and regression line (r = -.442, p = .006) to show the correlation between physical aggression and attention bias index (n = 32).

3.4.3.1.2 *Median split analysis of group effects.*

An ANOVA was conducted to measure the difference in reaction time on congruent and incongruent trials between those scoring high on physical aggression and those scoring low. The results revealed a significant interaction between trial congruency and physical aggression ($F(2,30) = 8.174$, $p = .008$, $\eta_p^2 = 0.226$) (see Figure 4 below). Participants with high physical aggression ($M = -12.60$, $SD = 23.55$) exhibited a significantly greater attention bias to angry words than those with low physical aggression ($M = 7.33$, $SD = 13.21$); $t(30) = -2.859$, $p = .008$, $d = 1.04$. These findings are consistent with correlational evidence and provide further support for predictions made in hypothesis one.

Table 4: Reaction times (ms) to congruent and incongruent trials in the high and low physical aggression groups (Mean and SD).

	High physical aggression (n = 15)	Low physical aggression (n = 15)	Whole sample (n = 32)	*p*-value
Congruent trials	488.43 (81.19)	484.00 (69.50)	486.33 (73.67)	.838
Incongruent trials	501.03 (88.72)	476.67 (64.18)	489.05 (76.95)	.367
Bias index	-12.60 (23.55)	7.33 (13.21)	-2.72 (20.82)	.008
p-value	.083	.057	.844	/

The Wilcoxon tests (as shown in Table 4 above) suggest that the difference in reaction time between congruent and incongruent trials is approaching significance in both the high and low physical aggression groups. However the high physical aggression group have quicker reaction times to congruent trials, whereas low aggression participants have quicker reaction times to incongruent trials.

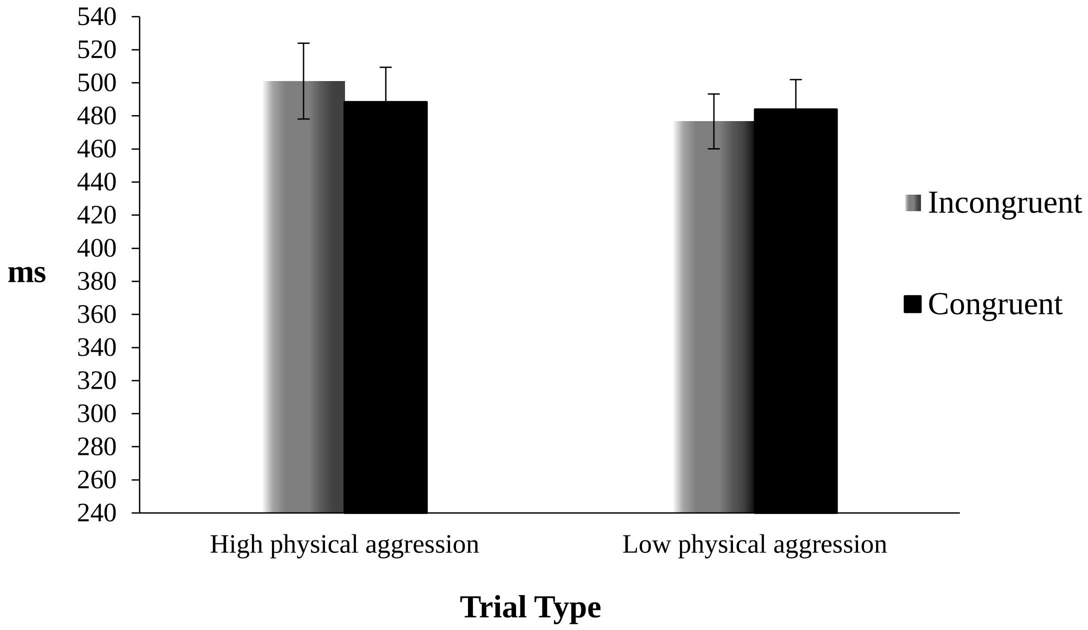

Figure 4: Mean reaction time (ms) on congruent and incongruent trials in the high (n = 15) and low (n = 15) physical aggression groups.

3.4.3.2 Hypothesis two

Attentional control did not significantly correlate with aggression; $r = -.263$, $p = .073$ one-tailed) or attention bias for *words*; $r = .034$, $p = .428$ (one-tailed) showing no support for hypothesis two. Therefore, attentional control was not explored as a possible mediator of the relationship between aggression and attention bias to words.

3.4.3.3 Hypothesis three

For each epoch, one way ANOVAs were conducted to investigate the difference in evoked amplitude in response to angry-neutral word onset between low and high aggression groups in electrodes across the region of interest.

3.4.3.3.1 *Pre-probe differences in aggression group.*

Between 100 and 200ms the effect of aggression was significant at CP6, $F(1,28) = 4.821$, $p = .037$, $\eta_p^2 = .147$; CP2, $F(1,28) = 8.640$, $p = .007$, $\eta_p^2 = .236$;

P8, $F(1,28) = 4.767$, $p = .038$, $\eta_p^2 = .144$; and P4; $F(1,28) = 6.858$, $p = .014$, $\eta_p^2 = .197$. It also approached significance at TP9, $F(1,28) = 3.537$, $p = .070$, $\eta_p^2 = .112$; TP10, $F(1,28) = 3.752$, $p = .063$, $\eta_p^2 = .118$; and CP1, $F(1,28) = 4.098$, $p = .053$, $\eta_p^2 = .128$. Between 200 and 300ms the effect of aggression was significant at CP5, $F(1,28) = 7.201$, $p = .012$, $\eta_p^2 = .205$; CP6, $F(1,28) = 6.564$, $p = .016$, $\eta_p^2 = .190$; CP1, $F(1,28) = 9.257$, $p = .005$, $\eta_p^2 = .248$; CP2, $F(1,28) = 5.870$, $p = .022$, $\eta_p^2 = .173$; and P4, $F(1,28) = 4.604$, $p = .041$, $\eta_p^2 = .141$; and approached significance at TP9, $F(1,28) = 3.379$, $p = .077$, $\eta_p^2 = .108$; TP10, $F(1,28) = 3.592$, $p = .068$, $\eta_p^2 = .114$; P7, $F(1,28) = 3.410$, $p = .075$, $\eta_p^2 = .109$; and P3, $F(1,28) = 3.914$, $p = .058$, $\eta_p^2 = .123$. Between 300 and 400ms there was a significant effect of aggression at CP6, $F(1,28) = 4.424$, $p = .045$, $\eta_p^2 = .136$; and P4, $F(1,28) = 5.079$, $p = .032$, $\eta_p^2 =.154$; and approached significance at CP1, $F(1,28) = 3.877$, $p = .059$, $\eta_p^2 = .122$; and CP2, $F(1,28) = 4.084$, $p = .053$, $\eta_p^2 = .127$. There were no significant effects between 400 and 500ms. In contrast to predictions, the waveform (Figure 5) reveals that high aggression participants have increased positive amplitude in response to angry-neutral word pair onset, compared to low aggression participants. The difference in amplitude is a long lasting effect that is evident for the whole duration of the trial, however the waveform (Figure 5) reveals that effects between 100 and 200ms may reflect the P1 component, 200 and 300ms the N2 component (P1/N2 complex), 300 and 400ms the P300 component.

3.4.3.3.2 *Post-probe differences in aggression group.*

Between 500 and 600ms there were no significant differences between aggression groups. Between 600 and 700ms the effect of aggression was significant at TP9, $F(1,28) = 6.833$, $p = .014$, $\eta_p^2 = .196$; TP10, $F(1,28) = 6.469$, $p = .017$, $\eta_p^2 = .188$; and approached significance at P3, $F(1,28) = 3.351$, $p = .078$, $\eta_p^2 = .107$. Between 700 and 800ms the effect of aggression was significant at TP10, $F(1,28) = 6.601$, $p = .016$, $\eta_p^2 = .191$; and approached significance at TP9, $F(1,28) = 3.209$, $p = .084$, $\eta_p^2 = .103$. Between 800 and 900ms the effect of aggression was significant at TP10, $F(1,28) = 4.824$, $p = .037$, $\eta_p^2 = .147$. Finally, there were no significant effects between 900 and 1000ms. Post probe presentation, the effects of aggression are maximal at TP10. Inspection of the waveform shows that effects between 600 and 900ms may reflect an LPP like component that is later and long lasting. Across the central-parietal and parietal electrodes there is a second slow inclining peak that begins at approximately 750ms (following the first inclining peak which begins at approximately 200ms).

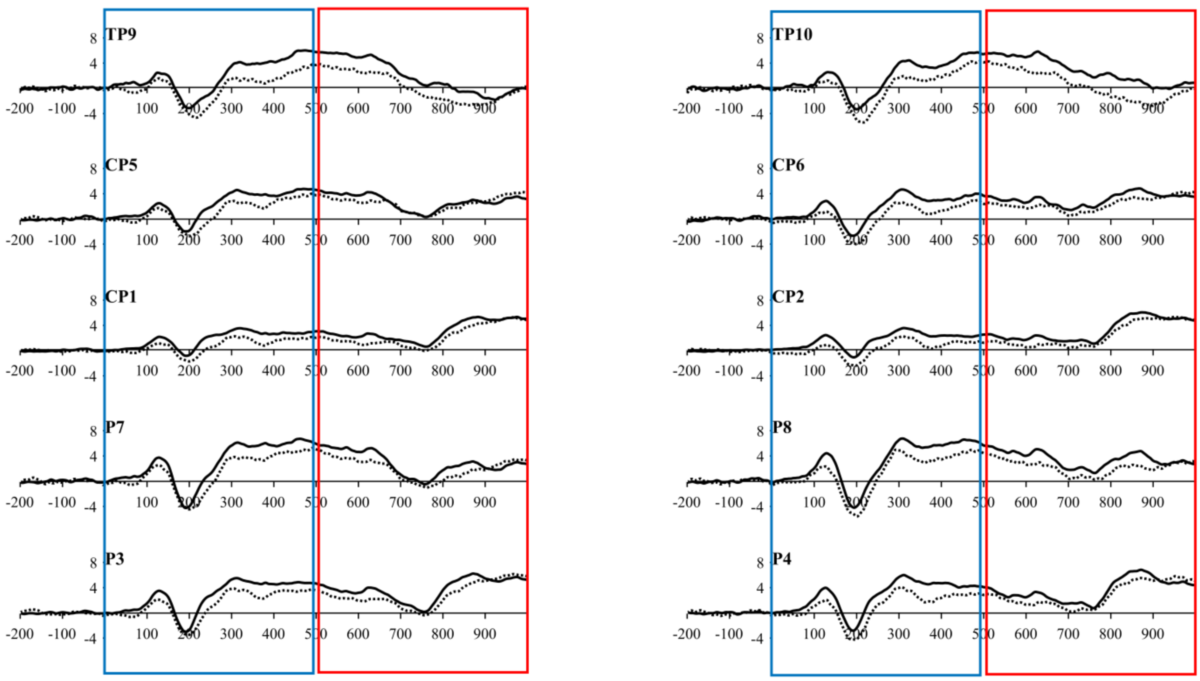

Figure 5: Grand average ERPs for the effect of physical aggression group (high vs. low) across all trials. The high physical aggression group (n = 15; black) is compared with the low physical aggression group (n = 15; dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

3.4.3.4 Hypothesis four

Qualitatively the waveform (Figures 6 and 7) shows that there are potential differences in amplitude between congruent and incongruent trials pre-probe presentation (500ms). This suggests that data time-locked to the probe onset would not have a valid baseline (confirmed by the ERP analyses below). Therefore, the whole trial based on data time-locked to the word onset was evaluated to avoid the possible confound of introduction of post-probe trial type effects created by pre-arrow change in baseline (Poulsen et al., 2005; Mingtian et al., 2011). It was predicted that trial congruency effects would only be evident between 500 and 900ms (post probe presentation). However, the same statistical analysis was conducted across all 100ms epochs (100-1000ms) to confirm the predictions. This also allowed for better conclusions regarding the latency of the effects and investigate attentional processes that occur between word and probe presentation.

3.4.3.4.1 *Post-probe differences in congruency.*

The results revealed no main effect of physical aggression, however there was a significant main effect of congruency between 500 and 600ms, $F(1,28) = 6.114$, $p = .020$, $\eta_p^2 = .179$. P1 amplitude was increased in response to incongruent trials compared to incongruent trials. There was also a significant interaction between congruency and hemisphere between 700 and 800ms, $F(1,28) = 4.424$, $p = .045$, $\eta_p^2 = .136$. Post-hoc tests between 700 and 800ms showed no significant effects of congruency in either hemisphere.

3.4.3.4.2 *Pre-probe differences in congruency.*

Surprisingly, the results yielded significant congruency effects before the presentation of the probe. The results showed a main effect of congruency between 100 and 200ms, $F(1,28) = 11.437$, $p = .002$, $\eta_p^2 = .290$); 200 and 300ms, $F(1,28) = 5.056$, $p = .033$, $\eta_p^2 = .153$; 300 and 400ms, $F(1,28) = 8.149$, $p = .008$, $\eta_p^2 = .225$; and 400 and 500ms, $F(1,28) = 6.158$, $p = .019$, $\eta_p^2 = .180$. There was also a significant interaction between congruency and hemisphere between 300 and 400ms, $F(1,28) = 4.376$, $p = .046$, $\eta_p^2 = .135$. Post-hoc tests showed that the main effect of congruency was significant in the right hemisphere only, $F(1,28) = 11.998$, $p = .002$, $\eta_p^2 = .300$. There was also a significant interaction between congruency, electrode and physical aggression in the right hemisphere, $F(1,28) = 2.787$, $p = .038$, $\eta_p^2 = .091$. Follow up tests showed that the effect of congruency was significant in the low physical aggression group, $F(1,14) = 9.874$, $p = .007$, $\eta_p^2 = .414$, and only approached significance in the high physical aggression group, $F(1,14) = 3.630$, $p = .077$, $\eta_p^2 = .206$. These results show that between 300 and 400ms the effects of congruency are most salient in the low physical aggression group and in the right hemisphere.

Inspection of the waveform indicates that the effect of trial congruency consisted of more positive P1 amplitude between 100 and 200ms, and P300 amplitude between 300 and 500ms for incongruent trials than congruent trials at posterior sites. The P300 effect peaked around 400ms after stimulus onset and was maximal at TP10 (see Figure 6).

There were no main effects of physical aggression, however the ANOVA revealed a close to significant interaction between trial congruency, electrode, hemisphere and aggression, $F(4,112) = 2.622$, $p = .054$, $\eta_p^2 = .086$, within the 300 to 400ms epoch. To further investigate this complex interaction, post-hoc ANOVAs were performed to assess which electrodes the effect of trial congruency was significant in each aggression group. In the low aggression group, the main effect of trial congruency was significant at electrodes TP10, $F(1,14) = 7.129$, $p = .018$, $\eta_p^2 = .337$); CP6, $F(1,14) = 7.557$, $p = .016$, $\eta_p^2 = .351$; and P8, $F(1,14) = 4.961$, $p =.043$, $\eta_p^2 = .262$. In the high aggression sample, the effect of trial congruency was significant at CP2, $F(1,14) = 5.538$, $p = .034$, $\eta_p^2 = .283$. Results indicate that effects of trial congruency may be slightly greater in low aggression participants compared with the high aggression participants, and may be more salient in the right hemisphere compared to the left hemisphere. (Figures 6 and 7).

There is only tentative evidence for hypothesis 4. Although a significant interaction was found, surprisingly this appeared earlier than expected; effects peaked between 300 and 400ms which is pre-probe presentation. High physical aggressive participants showed less differentiation in amplitude when responding to congruent and incongruent trials compared to low aggression participants. Low aggression participants showed increased amplitude in response to incongruent trials compared to congruent trials.

To further explore the interaction between trial type and physical aggression group a number of further tests were conducted to investigate in response to which trial type (congruent or incongruent) and in which physical aggression group (high or low) the differences were evident. For each electrode, for each epoch, and for congruent and incongruent trials, a Pearson correlation (two-tailed) was conducted to assess the association between physical aggression and evoked amplitude. The results show that amplitude on both congruent and incongruent trials positively correlated with physical aggression at multiple electrodes and at multiple epochs (Appendix N). However, due to the number of correlations only the correlation between physical aggression and amplitude at CP2 on incongruent trials between 200 and 300m survived FDR correction. This suggests that participants with increased levels of aggression have increased amplitude on incongruent trials (probe replaces neutral word) at a relatively early stage of processing.

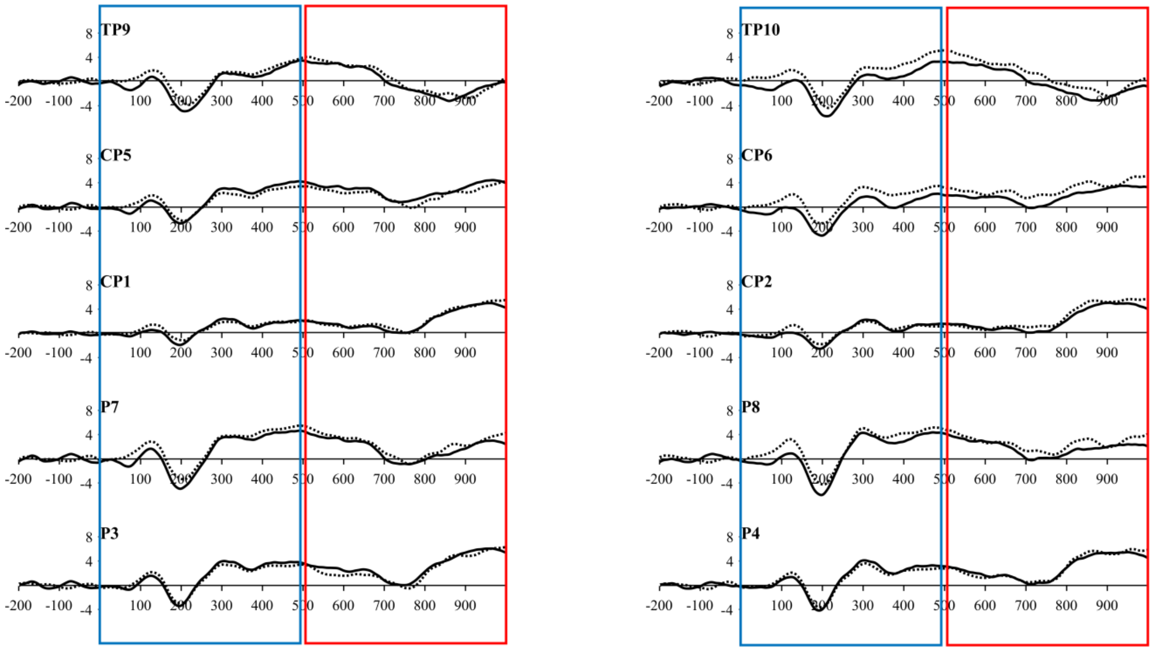

Figure 6: Grand average ERPs (n = 15) for the effect of trial congruency in the low physical aggression group. Mean amplitude to congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

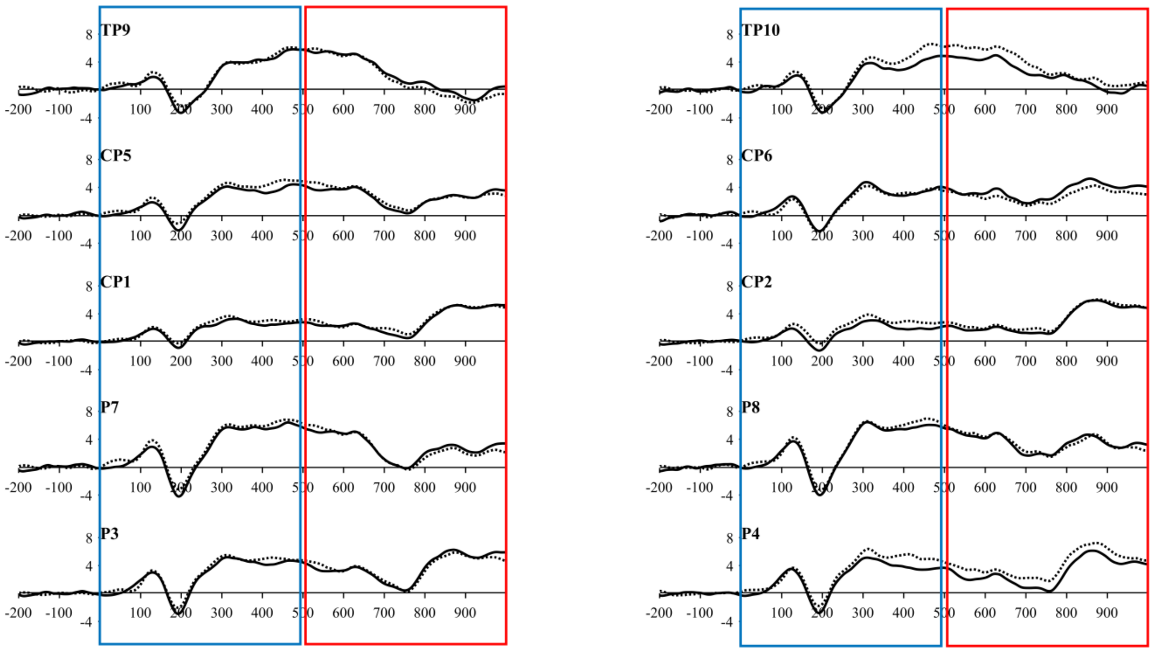

Figure 7: Grand average ERPs (n = 15) for the effect of trial congruency in the high physical aggression group. Mean amplitude to congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

3.5 Discussion

This first study investigated attention bias to aggression-related words in physical aggression using a selective attention task. The dot-probe paradigm was used to explore attentional selectivity when angry and neutral words were simultaneously presented. Concurrent EEG recording gave the opportunity to explore the cognitive processes involved with attention bias in an aggressive sample.

3.5.1 Main findings and interpretations

Hypothesis one predicted that participants scoring high on physical aggression would have an increased attention bias to angry words, relative to participants scoring lower on physical aggression. There was evidence to support this hypothesis. Results showed that higher levels of physical aggression were associated with faster responses to probes replacing angry words compared with probes replacing neutral words, suggesting that aggressive individuals have facilitated orienting towards angry words. This is consistent with Smith and Waterman (2003) whose work demonstrated an attention bias to aggressively themed words in both a dot-probe task and emotional Stroop task. Behavioural findings could be attributed to the trait-congruency hypothesis (Blaney, 1986; Bower, 1981; Miranda & Persons, 1988). This suggests that affective personality traits are linked to the activation of relevant emotion networks. Therefore it is proposed that facilitative biases in aggressive populations allow for attentional orientating towards aggression-related words as they are consistent with internal traits.

Crucially, further evidence shows a significant interaction between trial congruency and physical aggression group. Participants with higher physical aggression scores have a faster reaction time on congruent trials compared to incongruent trials, whereas participants with lower physical aggression scores have a faster reaction time on incongruent trials compared to congruent trials. This finding could be attributed to two different factors. The dot-probe task is a paradigm used to capture both facilitative and disengagement biases (Koster et al., 2004). The most recent accounts of attention suggest that facilitated attention (bottom-up, stimulus driven) and attentional avoidance (top-down, goal driven) both contribute to observed components of attention bias (Cisler & Koster, 2010). Attentional facilitation refers to automatic vigilance for threat (Davis & Whalen, 2001), whereas avoidance refers to strategic cognitive control activated to regulate attention allocation (Derryberry & Reed, 2002; Eysenck et al., 2007). A possible explanation of the findings is that low aggression participants have attentional avoidance of angry words and high aggression participants have attentional facilitation for angry words. High aggression participants may also have delayed reaction time on incongruent trials during the dot-probe task, compared to the low aggression group as that they are slower to disengage from angry words and subsequently take longer to respond to the arrow that appears in place of neutral words. In line with the theory that aggressive individuals are less able to disengage from aggression-related stimuli, I hypothesised that attentional control may play an important role in attention biases in aggression; it was predicted that increased attentional control would be associated with decreased levels of physical aggression and decreased attention bias to angry words. However there was no support found for either of these relationships, therefore this suggests that the attention bias effects cannot be explained by poor attentional control within this sample.

Based on previous evidence which suggests reduced P300 amplitude in response to target stimuli in aggressive populations (Bernat et al., 2007; Gao & Raine, 2009; Gao et al., 2013), hypothesis three predicted that the high physical aggression participants will have decreased positive amplitude in response to angry-neutral word pair presentation compared to the low physical aggression participants. There was no evidence to support this hypothesis as the main effect of aggression between 100 and 400ms and 600 and 900ms showed that the high aggression group had increased amplitude compared to the low aggression group. The effect of aggression seems to affect both the P100 and P300 component, suggesting that levels of aggression influence both early attentional processes relating to spatial attention, and later more elaborative stages of processing such as distribution of resources, categorization and updating memory models (Polich, 2007). Previous literature suggests that anti-social individuals show inefficient deployment of neural resources in processing cognitive task-relevant information and therefore have reduced P300 amplitude when responding to target stimuli (Gao & Raine, 2009). However there is some mixed evidence regarding the ERP correlates of attention bias to aggression-related words in aggression. The current findings are consistent with those found by Stewart et al. (2010); they showed that individuals with higher anger-out scores showed increased P300 amplitude in response to the negative words during an emotional Stroop task.

In the current study, the behavioural results suggest biased attention towards angry words in the high aggression group; therefore the increased amplitude may reflect increased processing of angry words. Although these findings suggest differences between aggression groups following the presentation of a simultaneously presented angry and neutral word pair, I cannot make conclusions regarding which word in the pair was evoking the difference in amplitude or the attentional processes that contribute to these differences. For example, does attentional facilitation of the angry word, or increased attentional resources allocated to disengaging from the angry word, evoke increased amplitude in the high physical aggression group. In order to draw further conclusions an additional control trial would be needed, for example including a happy-neutral, neutral-neutral condition. To further understand the complex processes involved with simultaneously attending to two stimuli, differences in amplitude following probe presentation which resolves the trial as congruent or incongruent was explored.

Hypothesis four predicted that there would be differences in evoked amplitude between congruent and incongruent trials, and that this would interact with aggression. The hypothesis was based on previous findings by Helfritz-

Sinville and Stanford (2015) who reported similar P300 amplitude in response to presentation of both threat and neutral words during a modified oddball task in the aggressive sample, whereas control participants exhibited enhanced amplitude to the threat words (social and physical) compared to neutral words. Previous work by Thomas et al. (2007) also showed that individuals with low levels of aggression had increased P300 amplitude to threat-related stimuli compared to neutral stimuli. It was theorized that aggression-related words are evaluated in terms of possible danger and potentially may require further processing in order to formulate goals and select an appropriate response (Smith et al., 1996). High aggression participants are less likely to show this pattern of processing as they may become desensitized to aggressive stimuli within the environment (Helfritz-Sinville & Stanford, 2015). It was therefore predicted that following probe presentation, high aggression participants would show similarity in P300 amplitude across congruent and incongruent trials, whereas low aggression participants would show increased amplitude on angry-congruent trials.

Results *post-probe* presentation show significant effects of congruency between 500 and 600ms only. The effect of congruency did not significantly interact with aggression, suggesting very few differences between evoked P300 amplitude in response to congruent and incongruent trials in either the high or low physical aggression group. However, post-hoc correlations showed that physical aggression positively correlated with amplitude on congruent trials between 600 and 900ms consistently at electrode CP6. This suggests that following probe presentation attention bias to angry words in the high aggression group, may be reflected by increased amplitude on congruent trials. However, these correlations did not survive FDR correction. Furthermore, inspection of the waveform suggests that high physical aggression participants show different ERP patterns across different electrode sites. For example they show increased P600/LPP in response to congruent trials at CP6, but show increased P600/LPP in response to incongruent trials at P4. This evidence suggests that ERP effects may vary across electrodes and latencies and therefore should be interpreted with caution

Surprisingly, more salient effects of congruency were found *pre-probe* presentation (between 100 and 500ms). The effect of congruency interacted with aggression such that high physical aggression participants show much smaller differences between congruent and incongruent trials relative to low physical aggression participants. In contrast to findings by Helfritz-Sinville and Stanford (2015) and Thomas et al. (2007) results showed that participants scoring low on physical aggression had increased P300 amplitude on incongruent trials compared to congruent trials between 300 and 400ms post word onset, maximal over the right hemisphere. This is in line with the behavioural findings which indicated attentional avoidance of angry words in the low physical aggression group. Low aggression participants may have a positive bias in which they avoid angry words and pay greater attention to neutral words, reflected in quicker reaction times and increased attentional processing (evoked amplitude). However, due to the latency of these, no robust conclusions can be drawn. The results of the current study suggest that trial congruency effects may be evident before the presentation of the probe. Following word pair onset, participants show different evoked P300 amplitude in response to upcoming congruent and incongruent trials. Theoretically, it is not clear why effects would be evident during these epochs, however, a possible explanations is that the early effects of congruency are long lasting effects evident from the previous trial. For example, Hajcak and Olvet (2008) found that increased LPP potential can be increased even after emotional stimuli offset. Early congruent effects may therefore reflect attentional processes that relate to probe presentation (and not word-pair presentation). However due to the length of the trial, the additional one-second blink screen between trials, and the speed of attentional allocation this is somewhat unlikely.

To my knowledge, and not surprisingly, no other studies have used the dot-probe paradigm to explore effects of trial types pre-probe presentation. However there are a few studies that have used a similar methodology; for example Mingtian et al. (2011) explored attention bias in depressed patients using the dot-probe task and EEG methodology. They assessed differences in amplitude between invalid and valid trial types following probe presentation. Qualitatively, Mingtian et al. (2011) provide waveforms that depict the whole trial; inspections of these waveforms show potential pre-probe differences, however this is not confirmed by statistical analysis. I suggest that future dot-probe research investigating congruency effects across different trial types should adopt an analytical approach in which the length of the whole trial is statistically analysed.

3.5.2 Limitations and future work

When evaluating this research there are a number of considerations to take in to account. Firstly, there are questions regarding the ecological validity of using hostile words as sources of anger/threat. In natural settings it is uncommon for single word presentations to provoke an aggressive response. In this sense, facial stimuli may provide more realistic measures of attention bias in response to aggression-related stimuli in the environment as they are important for human interaction; they provide social cues which convey messages important for communication (Argyle, 1994). I suggest that follow up work using faces instead of words may complement these current findings.

Secondly, Martin, Williams, and Clark (1991) found that threat words were more emotional, as well as more threatening, than control words and suggest that previous studies have confounded threat and emotionality. Results showed that anxious participants show a similar bias towards negative emotionally valenced threat words and positive emotion words, suggesting a bias to all emotion stimuli. However, within the current sample, across correlational and between-subject analyses there was no evidence of an association between anxiety and attention bias for angry words. Further research would be needed to compare attention biases to positive and negative emotional words in anxiety and aggression.

Familiarity or subjective frequency of the aggression-related words used may contribute to a bias towards such stimuli. For example Bradley et al. (1997)

suggests that anxious individuals may be primed towards threat words as they are more likely to think about possible threatening events, describe themselves as frightened, fearful or scared. Therefore, these words are used more frequently and become more familiar. This priming effect may contribute to aggressive individuals bias toward anger-related words; words such as ‘hostile’, ‘angry’, and ‘rage’ may be primed in individuals scoring high on physical aggression measures.

The results only provide tentative evidence for the difference in attentional processes between high and low physical aggression groups. Based on the current findings firm conclusions regarding the specificity of the processes relating to attention bias cannot be made. The complex results of this study require replication before solid conclusions can be drawn. A critical aim of follow up work would be to understand why effects of congruency are evident before probe presentation.

3.5.3 Contributions

This study has made a number of contributions to the understanding of cognitive biases associated with aggression. It suggests there are differences in selective attentional processes displayed by individuals with relative low and high levels of physical aggression when responding to angry-neutral trials at word pair onset, and in response to angry-congruent and angry-incongruent trials. Specifically it proposes that physical aggression is associated with an attention bias to angry words, reflected in a speedier reaction time to probes that replaced such words. This bias is characterized by relatively undifferentiated ERPs on congruent and incongruent trials. To my knowledge, this is the first dot-probe study to investigate selective attention processes to angry words in aggression using ERP methodology. However, the latencies of observed effects are unprecedented and currently there are no clear explanations for the early effects of congruency, therefore they will require replication before further conclusions can be drawn.

Although the findings are surprising, they have shown that ERPs are sensitive to attentional processes and make a number of methodological recommendations for future work using the dot-probe task with simultaneous EEG recording. Firstly I suggest using a pre stimuli baseline instead of a pre-probe baseline due to possible pre-probe effects. I also recommend analysing the whole trial length to better understand the latency of effects and contribute to the interpretation of ERP effects in relation to the dot-probe paradigm. There is some confusion in the literature regarding data that is time-locked to the stimuli presentation and data that is time-locked to the probe presentation, and subsequent interpretations of each analytical method. I suggest that transparent and consistent methodology is needed in order to make better comparisons between studies and draw more robust conclusions.

3.5.4 Conclusions

In line with current literature, the first aim of the study was to provide evidence of a behavioural attention bias towards angry words in a physically aggressive sample. Using the dot-probe paradigm in which angry and neutral words were presented simultaneously, selective attentional selectivity for angry words was explored. The results showed clear evidence for an attention bias for angry words in physical aggression. Very little is known about processing biases in aggression, therefore a further aim of the study was to explore the neural correlates of attention bias across high and low physical aggression groups. The novel use of EEG methodology allowed for the exploration of cognitive processes involved with selective attention following the presentation of an angry and neutral word pair.

The ERP results showed that low and high physical aggression participants had different ERP patterns in response to congruent and incongruent trials. Participants scoring low on physical aggression show increased amplitude on incongruent trials compared to congruent trials, whereas participants who scored higher on physical aggression showed much greater similarity in amplitude across trial types. However, due to the early latency of these findings, it is not clear how a negative attentional bias contributes to these differences in ERP pattern.

Using an original design this study provides an initial contribution to the understanding of the cognitive processes involved with attention bias in aggression. I suggest that further analyses will be needed to understand why high physical aggression groups show little differentiation between congruent and incongruent trials, and to explore the complexity of attentional processes involved with increased evoked P300 in response to incongruent trials in low physical aggression groups. Crucially, replication of the pre-probe congruency effects are needed before conclusions based on these findings can be drawn. Due to the complexity of the findings it is suggested that further studies utilising both reaction time and ERP data will further contribute to this field.

4 Study 2 - Attention bias to angry and happy words

4.1 Introduction

The previous chapter outlined the first of the studies within this thesis to investigate attention bias to aggression-related words in physical aggression. Study 2 aimed to explore the attentional processes involved with attending to words of different emotional valence. The main focus was to investigate whether attention bias is specific for aggression-related words or whether aggressive individuals attend to emotionally salient stimuli in general, for example, happy words. By exploring the neural correlates of these biases the aim was to identify any differences in processing involved with selectively attending to angry and happy words.

As described in the previous chapter, research suggests that aggressive individuals preferentially attend to aggression-related stimuli compared to neutral stimuli (e.g. Smith & Waterman, 2005; van Honk et al., 2001b). This fairly robust association was replicated in Study 1 using a dot-probe task. Participants with higher physical aggression scores were quicker to respond to probes that replaced angry words compared to probes that replaced neutral words. It is suggested that further research into attention bias towards differently valenced stimuli is needed to distinguish between an aggression specific bias and a more general emotional bias. To my knowledge there are very few studies which investigate selective attention biases to emotional words (angry or happy) compared with neutral words, specifically in relation to aggression. However, Smith and Waterman (2003) conducted an aggression-emotion themed Stroop task in which aggression themed, positive emotion, negative emotion, colour, or neutral words were presented. They found that aggressive individuals were slower to name the colour of the aggression-themed word compared to the neutral word. However, no significant differences in colour naming positive emotion words and neutral words were found. This suggests that levels of aggression do not influence patterns of attention to positive emotion words. Although this study includes both positive and negative emotion words, due to the single presentation of word stimuli during the Stroop task, firm conclusions based on the measurement of selective attentional processes cannot be drawn.

Although there is limited evidence of attention bias to positive stimuli in aggressive samples, there are some studies exploring attention bias in anxious and healthy samples that can be drawn upon. Martin et al. (1991) found that participants with general anxiety disorder were slower to name both threat-related and positive words, compared to non-anxious controls during a Stroop task. However, Pishyar et al. (2004) investigated attention biases in self-rated anxiety using a dot-probe task that consisted of negative and neutral pairs and positive and neutral pairs. This study yielded no significant attention bias effects for either negative or positive words. Taken together these results suggest that non-specific attention biases to valenced stimuli may only be evident in individuals with severe anxiety. Furthermore, Sutton and Altarriba (2011) investigated attention bias to negative and positive emotion words in a non-clinical sample during a dot-probe task. They found that participants responded faster to probes that appeared in place of negative words compared to neutral words on negative-neutral trials. However on positive-neutral trials there were no significant differences in reaction time. These results suggest that negative words may be detected quickly and have a unique effect on the attention system.

There is mixed evidence regarding attentional processes associated with attending to happy and neutral words. Nevertheless, there is some evidence to suggest that when positive and negative facial expressions are used to explore attentional processes associated with different emotions, healthy samples show an attention bias to both angry and happy facial expressions if presented alongside neutral faces (Waters et al., 2010; Bradley et al., 1997). For example, Waters et al. (2010) found that non-anxious controls showed an attention bias towards happy faces relative to neutral ones during a visual probe task. It is yet unclear if this effect is consistent across word and face stimulus modalities.

Only a small number of studies have used EEG to investigate the neural correlates of attention bias related to aggression and therefore very little is understood about attentional processes associated with attending to different types of stimuli. Previous evidence suggests that non-aggressive individuals show increased P300 amplitude in response to negative (threat) words (Helfritz-Sinville & Stanford, 2015; Thomas et al., 2007), whereas individuals with increased levels of aggressin show relatievly undifferentiated ERPs in response to negative and neutral words (Helfritz-Sinville & Stanford, 2015). Stewart et al. (2010) conducted one of the few studies to include a positive word condition when investigating attention biases in aggression. Results showed that individuals with higher anger-out scores showed increased P300 amplitude in response to the negative words compared to both neutral and positive words. This suggests that negatively valenced information is processed uniquely by the attentional system.

These studies used the Stroop or oddball tasks; behavioural paradigms in which stimuli are presented singly. However, the dot-probe paradigm, in which two stimuli are presented concurrently, has been used with simultenous EEG recording. For example Holmes et al. (2009) investigated attention bias to angry-neutral and happy-neutral face pairs in a normative healthy sample. Participants had faster reaction times to probes replacing emotional faces (angry and happy) compared to probes replacing neutral faces. ERP results showed that on angry-neutral and happy-neutral trials, congruent trials evoked an increased N2pc compared to incongruent trials, suggesting enahanced attentional capture of both angry and happy faces. There is also evidence to suggest differences between attentional processing of negative and positive stimuli in depresssion. For example, Mingitan et al. (2011) found that depressed individuals avoid attending to positive stimuli, reflected in reduced P1 amplitude in response to positive-neutral stimulus pairs. In addition, Hu et al. (2017) found depressed individuals had increased P300 amplitude in response to sad-congruent trials. Due to the different samples used across these studies it is difficult to suggest how results obtained from an aggressive sample may compare. Also, these studies include pictorial or facial stimuli. Images of emotional faces may not be comparable to emotional words; different words can more accurately convey specific concepts relating to physical aggression, whereas an angry face is a more general indication of threat.

These studies show that the dot-probe task may be sensitive to early (P1/N2) and later (P300) stages of attentional processing in healthy samples and psychological disorders such as depression. Therefore this methodological approach could be beneficial for understanding attention biases associated with increased levels of aggressive behaviour. Although the dot-probe paradigm has been used with simultaneous EEG recording, to my knowledge these two techniques have not been used together with the aim of understanding attention biases specifically in aggression. Finally this approach is more commonly used to explore attention to angry versus neutral, and happy versus neutral stimuli, however relatively little is known about attentional processes to simultaneously presented angry-happy stimuli.

There were two overarching aims of this study. Firstly, the aim was to explore whether aggressive individuals show an attention bias to positive and negative words when they are paired with a neutral word distracter. Comparing angry-neutral and happy-neutral word pairs on the dot-probe task would allow for the comparison of attentional processes associated with different emotional stimuli. A critique of previous attention bias research is that aggression and emotionality have been confounded; therefore it is not clear if aggressive individuals show an attention bias to aggressive stimuli or all emotional stimuli. It was predicted that previous findings (e.g. Smith & Waterman, 2005; van Honk et al., 2001b) that physically aggressive participants would have increased attention bias to angry words characterised by faster reaction time on congruent trials (probe replaces angry word) compared to incongruent trials (probe replaces neutral word) would be replicated. I recognise that as the dot-probe task cannot distinguish between processes of engagement and disengagement, theoretically this could be due to faster reaction times to probes replacing angry words or slower responses to probes replacing neutral words. Based on theories of aggression (Wilkowski & Robinson, 2010), the mixed evidence of attention bias to positive words, and the null finding by Smith and Waterman (2003), on happy-neutral trials, it was predicted that there would be no evidence of attention bias to happy words in either the physically aggressive or non-aggressive group.

Secondly, the aim was to discover whether aggressive individuals still show an attention bias towards negative words if they are paired with a positive word distracter. In line with attentional theory, it was suggested that when two stimuli are presented simultaneously both facilitative and disengagement processes contribute to attention bias (Koster et al., 2004). Therefore, selective attention processes involved with attending to angry words when they are presented alongside similarly emotional stimuli were explored. There is evidence to suggest that activation levels for all emotional stimuli take longer to decay than for neutral stimuli (McKenna, 1986); therefore greater attentional resources may be recruited in disengaging with such stimuli. Previous studies suggest that negative stimuli has a unique impact on the attentional system (Sutton & Altarriba, 2011), therefore including a condition in which both stimuli are emotionally salient helps to understand the complex processes involved with selective attention. Due to the novelty of including simultaneously presented angry and happy words, firm predictions regarding this stimuli combination were not made. However, previous evidence suggests that attentional interference on a Stroop task is greater for threatening faces compared to happy faces (Putman et al., 2004). Therefore it was suggested that the high physical aggression group, compared to the low physical aggression group, would show an increased attention bias for angry words.

In Study 1 EEG methodology was used to investigate the neural correlates of attention bias to negative and neutral word stimuli in an aggressive sample. The

ERP results showed a main effect of trial congruency in the low physical aggression, such that they showed increased positive amplitude on incongruent trials compared to congruent trials. The high aggression sample had relatively stable amplitude across trial types. Surprisingly the effects appeared pre-probe presentation. Therefore, the findings require replication before firm conclusions can be drawn and this was part of the rationale for Study 2.

In addition to replicating findings regarding angry words paired with neutral words, attentional processes involved with attending to differently valenced emotional stimuli were also explored. Previous studies have explored ERP correlates of attention bias to the single presentation of positive and negative emotion stimuli across healthy (e.g. Holmes et al., 2009; Schupp et al., 2004b), aggressive (Stewart et al., 2010) and anxious samples (e.g. Fox et al., 2008). In addition studies have used the dot-probe task to measure ERP correlates of attention bais to posiitve-neutral and negative-neutral stimulus pairs in healthy (e.g. Santesso et al., 2008) and depressed individuals (e.g. Mingitan et al., 2011; Hu et al., 2017). Finally, Pineles and Mineka (2005) explored attention bias to threat-happy stimulus pairings during the dot-probe paradigm in participants categorised by high and low social anxiety scores. They found the bias score obtained for angry-happy face pairs did not differ between anxiety groups. I recognise that caution should be taken when interpreting these results as the studies include emotional facial expressions, instead of words. Due to the different effects of faces and words on the attentional system, I believe it is important to distinguish between stimulus modalities and explore attention bias effects to positive and negative words. To my knowledge no studies have used the dot-probe task with simultaneous EEG recording to explore attentional selectivity to happy-neutral and angry-happy word pairs specifically in aggression. Due to these novel stimulus pairings, the predictions are somewhat exploratory in nature.

4.2 Aims and rationale

The literature suggests that aggressive participants show an attention bias towards angry stimuli compared with neutral stimuli. Study 1 implemented this method to shed some light on the attentional processes involved with attending to angry and neutral words in aggression. Due to the novelty of the findings, the first aim was to replicate the behavioural and EEG findings from Study 1.

The second aim is to extend previous findings on attention bias in aggression by investigating how different emotionally salient stimuli may be processed. To reduce the possible confound of emotionality when using negative-neutral word pairs, a happy-neutral word pair was included to test whether aggressive individuals have a bias for positive words (as well as negative), or whether the attention bias effect found in Study 1 is unique for negative words. By comparing the evoked ERPs in response to angry-neutral and happy-neutral word pair presentation it was possible to investigate the processes involved when attending to differently valenced stimuli.

As reaction time on the dot-probe task can reflect attentional facilitation of the target stimuli or difficulties in disengaging from the non-target (distracter) stimuli (Koster et al., 2004), an additional angry-happy trial type was included. By comparing angry-neutral and angry-happy trial types the aim was to explore whether the valence of the distracter stimuli influences the processing of angry words. The dot-probe task is used to measure selective attention processes, with speedier reaction times to congruent trials thought to reflect allocation of attention to the stimuli presented in the same prior location as the probe. Therefore the aim was to test whether aggressive individuals still selectively attend to angry stimuli if paired with happy stimuli. By analysing both behavioural and ERP data, the aim was to further understand the cognitive processes associated with attending to differently valenced emotional words. Based on findings from Study 1, that attentional processing of angry words influences the P100 and P300 component, predictions are made regarding both components.

4.2.1 Research questions and hypotheses

Overarching research questions:

- Is the attention bias effect in high aggression participants specific to angry stimuli, or do they also show an attention bias to positively-valenced happy stimuli?
- Do high aggression participants show differences in evoked P300 amplitude compared to low aggression participants when selectively attending to negative and positive emotionally-valenced stimuli?

4.2.1.1 Behavioural

4.2.1.1.1 *Correlational hypotheses*

Hypothesis one: Physical aggression score will be positively correlated with *angry-neutral* bias score such that those with higher physical aggression will have an increased bias towards angry words.

Hypothesis two: There will be no significant correlation between *happy-neutral* bias score and physical aggression.

Hypothesis three: Physical aggression score will be positively correlated with *angry-happy* bias score such that those with higher physical aggression will have an increased bias towards angry words.

4.2.1.1.2 *Between-subject hypotheses*

Angry-neutral

Hypothesis four; main effect: Participants will be quicker to respond to probes that replace *angry* words compared to probes that replace *neutral* words.

Hypothesis five; low physical aggression: Participants will have a significantly faster reaction time on trials where the probe replaces *angry* words compared to trials where the probe replaces *neutral* words.

Hypothesis six; high physical aggression: Participants will have a significantly faster reaction time on trials where the probe replaces *angry* words compared to trials where the probe replaces *neutral* words. This difference in reaction time between congruent and incongruent trials will be greater in the high physical aggression group compared to the low physical aggression group.

Happy-neutral

Hypothesis seven; main effect: Due to evidence which suggests no interference of positive emotion words on attention bias tasks (Smith & Waterman, 2003), it was predicted that there would be no difference between reaction times on congruent and incongruent trials.

Hypothesis eight; low physical aggression: Participants will show no difference in reaction time between probes that replace *happy* words and probes that replace *neutral* words.

Hypothesis nine; high physical aggression: Participants will show no differences in reaction time between probes that replace *happy* words and probes that replace *neutral* words.

Angry-Happy

Hypothesis 10; main effect: Participants will be quicker to respond to probes that replace *angry* words compared to probes that replace *happy* words.

Hypothesis 11; low physical aggression: Participants will have a significantly faster reaction on trials where the probe replaces *angry* words compared to trials where the probe replaces *happy* words.

Hypothesis 12; high physical aggression: Participants will have a significantly faster reaction on trials where the probe replaces *angry* words compared to trials where the probe replaces *happy* words. This difference in reaction time between congruent and incongruent trials will be greater in the high physical aggression group compared to the low physical aggression group.

4.2.1.2 ERP

4.2.1.2.1 *Main effect of aggression*

Hypothesis 13: Based on results from Study 1, it was predicted that high physical aggression participants will have increased P300 amplitude in response to all trial types (angry-neutral, angry-happy, and happy-neutral) at word pair onset compared to low aggression participants.

4.2.1.2.2 *Main effect of valence*

Hypothesis 14: Angry stimuli will evoke increased amplitude; therefore, angry-happy and angry-neutral trials will evoke increased positive P300 amplitude compared to happy-neutral trials following word pair onset. This effect will be most salient in the high physical aggression group.

4.2.1.2.3 *Effect of congruency*

Angry-neutral

Hypothesis 15: The general task effect across all participants will show increased positive P100/P300 amplitude on incongruent trials compared to congruent trials.

Hypothesis 16: Participants scoring low on physical aggression will show increased P100/P300 amplitude on incongruent trials compared with congruent trials.

Hypothesis 17: Participants scoring high on physical aggression will show similar P100/P300 amplitude in response to congruent and incongruent trials. This will be due to increased P100/P300 amplitude on incongruent trials due to the allocation of resources when attending to the simultaneously presented angry word.

Happy-neutral

Hypothesis 18: Due to previous evidence which suggests very little evidence for neural correlates of attention bias to happy words in aggression it was predicted that there would be no difference in evoked P100/P300 amplitude in response to congruent trials compared to incongruent trials.

Hypothesis 19: Due to previous evidence which suggests very little evidence for neural correlates of attention bias to happy words in aggression it was predicted that the low aggression group would show no difference in evoked P100/P300 amplitude in response to congruent trials compared to incongruent trials.

Hypothesis 20: It is not predicted that cognitive processes related to the allocation of attention to *happy-neutral* trials would be influenced by levels of physical aggression. Therefore it was predicted that the high aggression group would show no differences in evoked P100/P300 amplitude in response to congruent trials compared to incongruent trials.

Happy-angry

Hypothesis 21: The main task effect will show increased P100/P300 amplitude to incongruent trials compared to congruent trials.

Hypothesis 22: Based on findings from Study 1 it was predicted that participants scoring low on physical aggression will show increased P100/P300 amplitude to incongruent trials compared to congruent trials.

Hypothesis 23: Based on findings from study 1, and consistent with hypothesis 17, it was predicted that participants scoring high on physical aggression will show similar P100/P300 amplitude in response to congruent and incongruent trials.

.

4.3 Methods

4.3.1 Power Analysis

An a *priori* power calculation was conducted using G*Power 3.1 software (Faul, Erdfelder, Buchner, & Lang, 2009) based on the most complex planned analyses. For repeated measures mixed model ANOVA analyses, based on 8 measurements, 2 groups and a small to medium effect size (f = 0.20), a minimum sample size of 28 participants will be needed to achieve 90% power, when $\alpha = .05$.

4.3.2 Participants

Data were collected from 56 male University of East Anglia (UEA) students and staff, and members of the wider community. In order to take part in the study participants had to be male, aged between 18 and 35, be right-handed, speak English as their first language and have normal or corrected vision. They also were unable to take part if they had been diagnosed with a psychological condition in the last 12 months, were receiving psychological treatment or were taking anabolic steroids. One left-handed participant was ineligible and therefore subsequently excluded from analysis. Four further participants were excluded from analyses; two due to a fault in recording and two due to excessive EEG noise. Therefore, the final sample consisted of 51 participants (mean age = 21.39, *SD* = 3.49).

Participants were recruited using various methods such as poster adverts, the University SONA system for undergraduate students, paid participant panel and word of mouth. Sixteen participants (31%) were recruited through the University SONA system, the other 35 were recruited through various advertisement methods. Of the total 51 participants, 44 were students, six were in full time work, and one in part-time work. 97.9% of the sample had some university credit (ranging from currently undertaking the first year of an undergraduate degree, to holding a Master’s degree).

4.3.3 Self-report measures

Self-report measures were the same as those used in Study 1. They consisted of The Aggression questionnaire (Buss & Perry, 1992), the Attentional Control Scale (Derryberry & Reed, 2002), The Delinquency Questionnaire (Tarry & Emler, 2007), and the Trait form of the State-Trait Anxiety Inventory (Spielberger & Gorsuch, 1970). See Chapter 3, Section 3.3.3 for further details on each of these.

4.3.4 Attention bias test

Attention bias was measured using the probe classification version of the visual-probe task, adapted from MacLeod et al. (2002), and programmed using E-Prime software (Schneider et al., 2002). Participants were seated 60cm from a 23 inch monitor (black text/colour images on a white background), affording a visual angle of approximately 3 degrees between items (cf. see, MacLeod, & Bridle, 2009). Each trial began with a fixation point (three small crosses) in the centre of the computer screen for varying duration (range 1060 to 1973ms), followed by presentation of the stimulus pair. Each word/face pair was presented for 500ms separated by a vertical distance of 3cm above and below the central fixation cross. Next, a left or right pointing arrow probe (“<” vs. “>”) appeared in the prior location of the stimulus pair with equal probability until response. Participants were instructed to indicate the direction of the arrow probe on screen using the arrow keys as quickly and accurately as possible. A one-second blink screen followed the target response to minimise ERP artefacts, after which the next trial started immediately. Aggression-related attention bias is characterized by faster reaction times to arrow probes located in the congruent (replacing angry) versus incongruent (replacing neutral) position. The test included ten practice trials followed by a further 192 test trials. A break occurred after every 48 trials (three breaks evenly distributed throughout the task).

4.3.5 Attention bias test stimuli

The stimuli consisted of 32 angry, 32 happy, and 32 neutral words. The 32 angry-related words were taken from (Smith & Waterman, 2003) and were matched with 32 neutral and happy words based on length and frequency using the Brysbaert database (Brysbaert & New, 2009); see Table 5). All neutral words were household items to control for category relatedness (e.g. Faunce et al., 2004; Placanica, Faunce, & Soames Job, 2002). There were 32 word pairs in each condition and these were repeated twice across the task. This gave a total of 64 angry-neutral pairs, 64 angry-happy pairs and 64 happy-neutral pairs (each individual angry, happy, and neutral word was presented four times across all trials). The 32 word pairs for each condition were split into 8 blocks of four; the presentation of blocks was randomised throughout the test. The direction (left or right) and location (top or bottom) of the arrow probe was equally distributed across trial types.

Table 5: Attention bias test word stimulus: angry, happy and neutral words matched for length and Brysbaert frequency.

No.	Aggression		Positive		Neutral	
1	strike	3.09	genius	3.07	shower	3.11
2	bloody	2.95	kissed	2.93	stairs	293
3	anger	2.81	hopes	2.80	piano	2.78
4	assault	2.77	beloved	2.69	bottles	2.66
5	kick	3.34	warm	3.27	ball	3.32
6	insult	2.73	laughs	2.78	drawer	2.67
7	hate	3.68	love	3.87	room	3.80
8	shoot	3.51	happy	3.74	floor	3.44
9	stab	2.51	heal	2.62	oven	2.52
10	hurt	3.71	hope	3.79	door	3.72
11	argue	2.91	pride	2.97	plate	2.97
12	temper	2.65	wisdom	2.65	tables	2.67
13	fight	3.57	funny	3.67	light	3.59
14	attack	3.29	dreams	3.24	window	3.39
15	punish	2.60	divine	2.57	fridge	2.59
16	rape	2.59	goal	2.73	lamp	2.59
17	annihilate	1.56	loveliness	1.57	spectacles	1.56
18	prison	3.18	spirit	3.17	camera	3.14
19	cut	3.69	fun	3.67	bed	3.61
20	riot	2.40	lust	2.33	sofa	2.36
21	destroy	3.15	excited	3.23	bedroom	3.10
22	injure	1.65	cuddly	1.66	jigsaw	1.61
23	threaten	2.57	cheerful	2.51	curtain	2.52
24	knife	3.10	faith	3.09	chair	3.19
25	annoyed	2.18	devotion	2.26	chimney	2.15
26	rifle	2.59	adore	2.47	porch	2.53
27	scream	2.94	admire	2.78	closet	2.95
28	intimidate	1.92	affection	1.98	headphones	1.81
29	stare	2.64	charm	2.78	frame	2.72
30	rage	2.61	fond	2.69	shed	2.62
31	shout	2.66	loyal	2.67	towel	2.72
32	kill	3.76	glad	3.63	book	3.51
Mean		**2.82**		**2.85**		**2.82**

4.3.6 EEG acquisition

The Electroencephalogram (EEG) was recorded with a 32-channel active electrode system (Brain Products GmbH) embedded in a nylon cap (10/10 system extended). An additional electrode was placed under the left eye in order to monitor vertical eye movements (lower electroculography; EOG). The continuous EEG signal was acquired at a 1000Hz sampling rate using FCz as reference. The impedance was kept below 20 kΩ. ERPs were time-locked to the onset of each stimulus pair.

4.3.7 Procedure

Ethical approval was obtained from the School of Psychology research ethics committee. All participants were invited to attend a laboratory session on UEA Campus. There was no targeted recruitment or screening procedures used for this study. During the testing session participants were first asked to read a full information sheet (Appendix O) and sign a consent form (Appendix P), before providing demographic information (appendix X). Participants were then fitted with the nylon cap embedded with 32 electrodes and EEG trace was recorded during completion of two *dot-probe tasks* (words and faces). The true objectives of the study were not revealed until after task completion. Participants received minimal instructions for each of the experimental tasks to reduce demand characteristics. They were told that they would see two faces appear on screen, followed by an arrow; their only job was to respond as quickly and accurately to the arrow as possible. During a single experimental session, participants completed the dot-probe word task and the dot-probe face task. Results obtained from the dot-probe word task are reported in this chapter, results obtained from the dot-probe face task are presented in Chapter 6. Participants were also required to complete the four questionnaires (AQ, ACS, DQ and STAI-T); these were presented online via Qualtrics. All participants were provided with both a written and verbal debriefing (Appendix Q). The experimental tasks and questionnaire measures, and the order in which they completed the two dot-probe tasks, were counter balanced. The order of questionnaire completion was also randomised. The testing session took approximately 90 minutes and participants received SONA credits or payment as compensation.

4.3.8 Data analysis plan

4.3.8.1 Behavioural attention bias data

Reaction time data was extracted using E-merge software. Reaction times on correct trials only (accurate identification of arrow orientation) were analysed (98.87% of all trials on the word task). The median reaction times were extracted for each of the three trial types (angry-neutral, angry-happy, happy-neutral) and for congruent and incongruent probe positioning for each trial type. Therefore reaction time was extracted for six conditions. Median reaction times were extracted as they are not skewed by extreme scores (e.g., Whelan, 2008). The six conditions consisted of angry-neutral trials where the probe could appear in place of the angry (congruent) or neutral (incongruent) stimuli, angry-happy trials where the probe could appear in place of the angry (congruent) or happy (incongruent) stimuli, and happy-neutral trials where the probe could appear in place of the happy (congruent) or neutral (incongruent) stimuli (see Figure 8). These were used to analyse the difference between trial congruency for each of the three trial types. An *attention bias index score* was calculated for angry-neutral, angry-happy and happy-neutral trials (AN, AH and HN). These were calculated by subtracting the median reaction time to incongruent trials from median reaction time on congruent trials. Therefore, attention bias on angry-neutral trials were calculated by subtracting median reaction time on trials where the probe replaces the neutral stimuli, from median reaction time on trials where the probe replaces the angry stimuli. Therefore a negative bias score indicates that participants responded more rapidly when probes replaced angry than neutral stimuli.

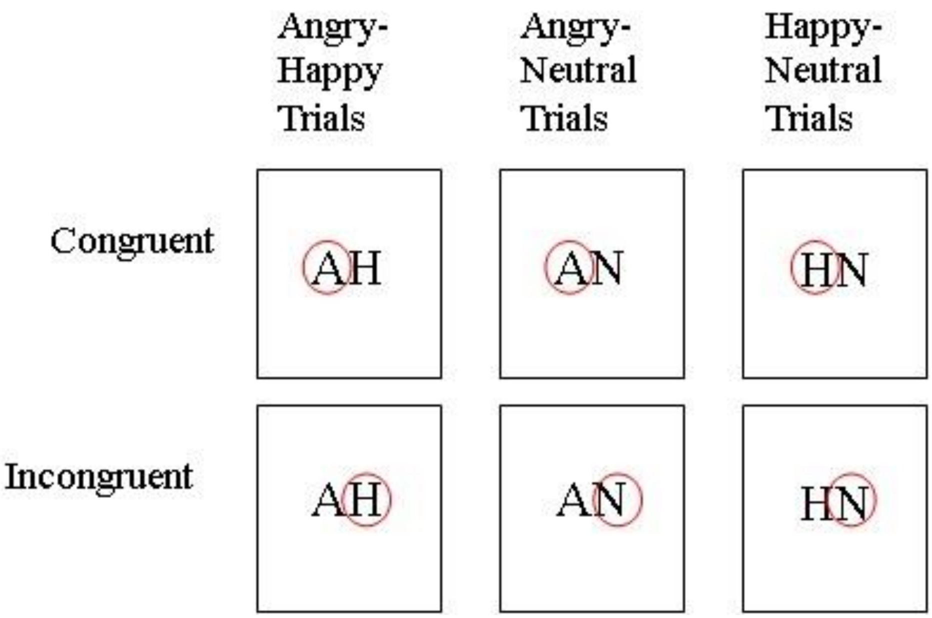

Figure 8: Positioning of the probe (highlighted by the red circle) on congruent and incongruent trials for each trial type.

The relationship between attention bias and aggression was investigated using both correlational and between-subject analyses. Correlations were used to explore the relationship between physical aggression and reaction time data.

Further to the correlations a repeated measures mixed model ANOVA was conducted to investigate the effect of valence. Trial type (3 levels; angry-neutral, happy-neutral, angry-happy), and trial congruency (2 levels; probe appears in congruent or incongruent position) were added as within-subject factors. Physical aggression was included as a between-subject factor (2 levels; high and low physical aggression). Due to the moderate positive correlation between attentional shifting and attention bias on *angry-neutral* trials, attentional shifting was added as covariate (see Section 4.4.2.2). Driven by the individual hypotheses, further post-hoc analyses were conducted to explore the effect of congruency for each trial type.

4.3.8.2 EEG data

Offline analyses were conducted using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014), two open source toolboxes running under Matlab 7.12 (R2013a, The Mathworks). High- and low-pass filter half-amplitude cut-offs were set at 0.1 and 40 Hz, respectively. Before averaging, trials contaminated by excessive artifacts were rejected automatically with a step function (Luck, 2005) with a voltage threshold of ± 100 µV in moving windows of 200ms and with a window step of 100ms. Noisy channels were interpolated using the EEGLAB function eeg_interp (spherical interpolation).

The EEG was segmented into epochs of 1200ms; from -200ms to 1000ms post word/face pair onset. Mean amplitude between 100 and 200ms, 200 and 300ms, 300 and 400ms, 400 and 500ms, 500 and 600, 600 and 700, 700 and 800, 800 and 900ms, and 900 and 1000ms time-locked to the onset of the word/face pair were extracted for statistical analyses with a -200-0ms baseline. Based on results obtained from Study 1, the whole trial epoch time-locked to the word/face onset was extracted instead of using a pre-probe baseline. Therefore, epochs between 100 and 500ms refer to pre-probe presentation, whereas epochs between 500 and 1000ms refer to post-probe presentation. This allowed for the P1 and P300 effects in response to word/face pair presentation and in response to probe presentation to be explored. Using a large number of short epochs allowed for the exploration of where and at which latencies the effects reached the greatest significance. The analyses focused on posterior parietal electrodes (CP1/2, CP5/6, P7/8, P3/P4 and TP9/10), where P300 component is considered to be maximal (e.g. Polich, 2007).

For initial analyses the repetition factor was not included, therefore analyses includes trials presented across the whole task (2 presentations). ERP measures were evaluated on correct trials only (98.16% of all trials on the face task and 98.87% of all trials on the word task).

4.3.8.2.1 *Effect of aggression*

To explore the main effect of aggression one-way ANOVAs were conducted to investigate the differences between high and low physical aggression groups for each trial type across the electrodes of interest for all epochs.

4.3.8.2.2 *Effect of valence*

To explore the differences between amplitude evoked by angry and happy valenced words, a mixed model ANOVA was conducted. The ANOVA consisted of trial type (3 levels; angry-neutral, angry-happy, happy-neutral), hemisphere (2 levels; left and right), and electrode (5 levels) as within-subject factors. Physical aggression (2 levels; high and low) was added as a between-subject factor. Post-hoc tests were conducted to explore the nature of the significant differences in each trial type. This analysis was conducted for epochs between 100 and 500ms to test the effect of valence pre-probe presentation.

4.3.8.2.3 *Effect of trial congruency*

A 3 (trial type; angry-neutral, angry-happy, happy-neutral) x 2 (congruency; congruent, incongruent) x 5 (electrode) x 2 (hemisphere; left, right) x 2 (physical aggression; high, low) mixed model ANOVA was conducted for each epoch. Follow up planned comparisons were conducted to explore the effect of trial type and trial congruency where appropriate. Differences between congruent and incongruent trials post-probe presentation (between 500 and 1000ms) were expected, however qualitatively inspection of the waveform revealed potential congruency effects pre-probe presentation and therefore the same statistical analysis was conducted across all 100ms epochs (100-1000ms) to further investigate these qualitative observations.

Based on results from Study 1 the ERP results presented here, and subsequent post-hoc tests, only include physical aggression. This allows for better comparisons between datasets and subsequent findings. In addition, preliminary analyses were conducted with both total and physical aggression as between-subject factors, results showed that effects were generally more statistically significant when the model included physical aggression.

Results and post-hoc analyses for each of the two dot-probe tasks are presented separately. Results of the dot-probe word task are presented in this chapter, results of the dot-probe faces task are presented in Chapter 6. Greenhouse-Geisser (Geisser & Greenhouse, 1958) *F* test is reported throughout for all repeated measures to ensure there are no violations of the sphericity assumption. Consistent with Study 1, some alpha values above the conventional significance value of $p = .05$ are presented to show significance levels of electrodes across all epochs (based on qualitative inspection of waveform). I refer to p values above 0.05 and below 0.1 as ‘approaching significance’.

4.4 Results

4.4.1 Data preparation

4.4.1.1 Missing data

The Delinquency Questionnaire (DQ) and Attentional Control Scale (ACS) had no missing items. The Buss and Perry Aggression Questionnaire had one piece of missing data from the verbal aggression subscale. The Trait Anxiety Inventory (STAI-T) had two pieces of missing data. Missing values were replaced with the mean of the completed items for each questionnaire (method used by Judah et al., 2014).

4.4.1.2 Normality of data

All data (BPAQ and subscales, ACS, and STAI-T) were normally distributed apart from the delinquency questionnaire (due to floor effect) (Appendix R). The six reaction time variables (angry-neutral congruent, angry-neutral incongruent, happy-neutral congruent, happy-neutral incongruent, angry-happy congruent, angry-happy incongruent) were also assessed for normality; these were not normally distributed due to skewness and kurtosis calculations (scores divided by the corresponding standard error) that were outside acceptable limits of ±2. Although kurtosis calculations were generally within acceptable limits, the data was positively skewed towards lower reaction times. Therefore analysis of reaction time data utilised non-parametric tests. The skewness and kurtosis scores for the calculated biases (congruent minus incongruent) were almost all within acceptable limits (bias score for angry-neutral trials was slightly skewed) (Appendix R). Preliminary analyses were conducted using both parametric and non-parametric tests and these showed that results were comparable using both versions. Therefore, along with the other bias scores, angry-neutral bias was analysed using parametric tests.

Across all six reaction time variables there was one extreme outlier (3 standard deviations above the mean) which was replaced with the next highest score plus one. There were some consistent other outliers (2 standard deviations above the mean) which shows that across all trial types some participants were slower to react to the stimuli. The data for each of these participants was explored and it was decided not to remove or adjust these as they were stable across the data and therefore did not affect the calculated bias scores. For the dot-probe word task angry-neutral bias ranged in score from -32.0 to 56.5, angry-happy bias ranged from -36.0 to 36.5, and happy-neutral bias ranged between -31.0 to 37.5.

4.4.1.3 Reliability of questionnaires

The BPAQ (α = .84), physical aggression subscale from BPAQ (α = .80), anger subscale from BPAQ (α = .81), ACS (α = .82), and STAI-T (α = .90) demonstrated good internal reliability. The verbal aggression subscale from the BPAQ (α = .72), hostility subscale from BPAQ (α = .63), and DQ (α = .73) were only moderately reliable.

4.4.2 Descriptive results

4.4.2.1 Aggression Questionnaire (Buss & Perry, 1992)

For consistency with Study 1, data were categorised based on a median split of physical aggression scores. There were three participants that scored the median and could therefore not be categorised. Between-subjects analysis included 25 participants scoring low on physical aggression and 23 participants scoring high. The high physical aggression group (M = 26.13, SD = 4.81) significantly differed from the low physical aggression group (M = 15.08, SD = 2.94); $t(48) = 9.691$, $p < .001$. d = 2.77) (see Table 6). Table seven shows the relationship between the four subscales of the aggression questionnaire.

Table 6: Means (SD) for the whole sample, and low and high physical aggression groups for all questionnaire measures.

	Total aggression	Physical aggression	Verbal aggression	Hostility	Anger	Delinquency	Attentional focusing	Attentional shifting	ACS	STAI-T
Whole sample (*n* = 51)	73.21 *(14.35)*	20.35 *(6.60)*	15.07 *(3.96)*	22.04 *(5.05)*	15.75 *(5.32)*	3.25 *(3.49)*	16.59 *(4.09)*	12.24 *(2.42)*	51.37 *(8.18)*	43.80 *(9.13)*
Low physical aggression (n = 25)	63.94 *(10.54)*	15.08 *(2.94)*	13.78 *(3.59)*	21.40 *(5.20)*	13.68 *(4.25)*	2.04 *(2.24)*	14.96 *(3.09)*	12.48 *(2.10)*	53.44 *(6.23)*	43.90 *(8.63)*
High physical aggression (*n* = 23)	83.74 *(11.38)*	26.13 *(4.81)*	16.26 *(4.06)*	23.00 *(5.12)*	18.35 *(5.58)*	4.57 *(4.26)*	18.09 *(4.36)*	12.00 *(2.78)*	49.61 *(9.31)*	44.71 *(9.90)*

Table 7: Correlations (one-tailed) between all subscales of the aggression questionnaire (n = 51).

	Total Aggression	Physical Aggression	Verbal aggression	Anger	Hostility
Total Aggression		.810 (<.001)	.612 (<.001)	.715 (<.001)	.549 (<.001)
Physical Aggression			.472 (<.001)	.418 (.001)	.184 (.099)
Verbal aggression				.242 (.044)	.084 (.278)
Anger					.241 (.044)

4.4.2.2 Questionnaire variables

Correlations between all the questionnaire variables were explored to investigate the associations between dependent variables. Delinquency did not correlate with any of the aggression subscales or the total score. This could be attributed to the floor effect of delinquency in this non-forensic sample. There was a strong positive correlation between anxiety and hostility, $r = .626$, $p < .001$ (one-tailed); however, anxiety did not correlate with any reaction time or attention bias index variables.

ACS scores correlated with total aggression score; $r = -.298$, $p = .017$ (one-tailed), physical aggression; $r = -.251$, $p = .038$ (one-tailed), and anxiety score; $r = -.306$, $p = .015$ (one-tailed). To further investigate the correlation between attentional control scale and aggression, Pearson's correlations were conducted between the two attentional control subscales (shifting and focusing) and all aggression subscales. Results showed a moderate positive correlation between attentional focusing and total aggression; $r = .415$, $p = .001$ (one tailed), physical aggression; $r = .344$, $p = .007$ (one-tailed), and anger; $r = .318$, $p = .012$ (one tailed). Attentional shifting did not correlate with any of the aggression subscales ($p \geq .263$). On the dot-probe word task, attentional shifting moderately positively correlated with *angry-neutral* bias only, $r = .337$, $p = .008$ (one-tailed) (Figure 9), such that a negative attention bias was associated with decreased attentional shifting scores. As attentional shifting correlated with the main dependent variable and could be a possible confound, this was included as a covariate in the ANOVA model.

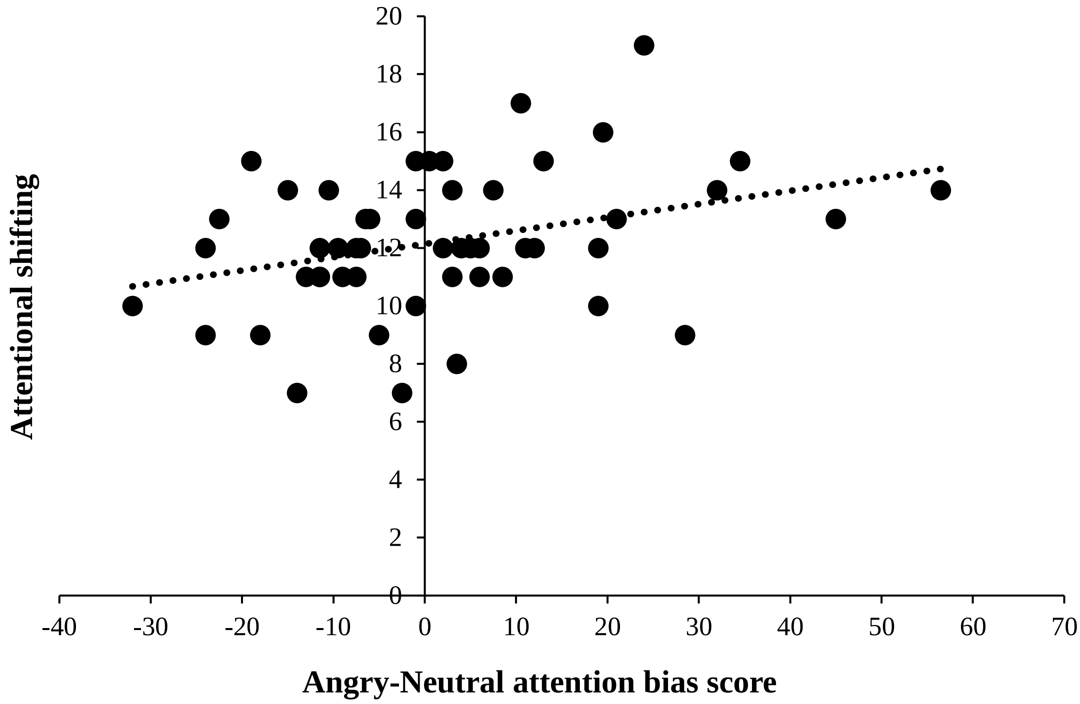

Figure 9: Scatterplot and regression line (r = .337, p = .008) to show the correlation between angry-neutral attention bias index and attentional shifting (n = 51).

Results relating to hypotheses

4.4.3 Behavioural data

4.4.3.1 Effect of aggression.

Table eight gives an overview of the means and standard deviations across all trial types in the high and low physical aggression group. Although inspection of the means shows that generally high physical aggression participants were slower to respond to probes across all trial types, these differences did not reach significance ($p > .565$) (see Figure 10).

Table 8: Mean reaction time (ms) for target stimuli and bias score of each trial type within the total, high and low physical aggression groups (SDs).

Trial type	**AN**			**HN**		
Reaction time variable	**Angry target**	**Neutral target**	**Bias**	**Happy target**	**Neutral target**	**Bias**
Low physical aggression (n = 25)	436.86 (59.36)	433.54 (53.77)	3.32 (16.49)	437.36 (57.58)	433.00 (54.70)	4.36 (17.30)
High physical aggression (n = 23)	439.89 (55.19)	442.41 (52.29)	-2.52 (17.67)	441.80 (55.16)	437.48 (48.45)	4.33 (13.81)
Whole sample (n = 51)	442.58 (58.72)	440.49 (53.38)	2.09 (17.82)	442.97 (57.39)	439.21 (53.60)	3.76 (15.80)

Trial type	**AH**		
Reaction time variable	**Angry target**	**Happy target**	**Bias**
Low physical aggression (n = 25)	434.90 (50.26)	434.94 (54.38)	-0.04 (14.83)
High physical aggression (n = 23)	440.12 (53.18)	439.02 (52.02)	1.09 (14.71)
Whole sample (n = 51)	441.84 (54.20)	441.32 (54.58)	0.52 (15.27)

4.4.3.1.1 *Effect of valence and trial congruency*

Correlational results. There were no significant correlations between physical aggression and *angry-neutral* bias score (p = .156, one-tailed), *angry-happy* bias score (p = .254, one-tailed), or *happy-neutral* bias score (p = .191, one-tailed). These results do not support hypotheses one or three as it was suggested that participants higher on aggression would have an increased bias score characterized by quicker reaction times on angry-congruent trials compared to incongruent trials on both *angry-neutral* and *angry-happy* trial types. The results support hypothesis two as it was predicted that *happy-neutral* bias would not be correlated with aggression, it was suggested that both high and low aggression groups would attend to both stimuli similarly.

Median split analysis of group effects. Between-subject analysis consisted of a 3 (trial type) x 2 (trial congruency) x 2 (physical aggression) omnibus ANCOVA. This was conducted to explore the interaction between trial types and congruency. Attentional shifting was added as a covariate. There were no main effects of trial type which suggests that reaction times across the three trial types were relatively similar. However, the interaction between trial type and trial congruency approached significance; $F(2,90) = 2.952$, $p = .060$, $\eta_p^2 = .062$. The interaction between trial type, trial congruency and attentional shifting also approached significance; $F(2,90) = 2.833$, $p = .067$, $\eta_p^2 = .059$. To explore this interaction, post-hoc tests were conducted to explore in which trial types the effects of congruency was significant. A 2 (congruency; congruent, incongruent) x 2 (physical aggression; high, low) ANOVA was conducted for each trial type. Attentional shifting was included as a covariate.

On *angry-neutral* trials, post-hoc tests showed there was a main effect of trial congruency, $F(1,45) = 6.730$, $p = .013$, $\eta_p^2 = .130$, and also a significant interaction between trial congruency and attentional shifting, $F(1,45) = 7.169$, $p = .010$, $\eta_p^2 = .137$. Without the attentional shifting covariate in the model, the main effect of trial congruency was non-significant ($p = .872$). This suggests that the main effect of trial congruency is only significant when controlling for attentional shifting and therefore attentional shifting may be a possible moderator of attention bias on *angry-neutral* trials. Inspection of the means shows that participants were slightly quicker to respond to incongruent trials compared to congruent trials. This shows no support for hypothesis four. It was expected that angry words would grab attention and therefore participants will be quicker to respond to probes that replace *angry* words compared to probes that replace *neutral* words. Including attentional shifting as a covariate, post-hoc tests also showed no significant interaction between trial congruency and physical aggression for *angry-neutral* trials (p =

.328). This suggests that reaction times were similar across both high and low aggression groups and there was no significant effect of trial congruency across either group. Therefore, there was no evidence for hypothesis five or six.

On *happy-neutral* trials, post-hoc tests showed that there was a close to significant effect of trial congruency, $F(1,46) = 3.656$, $p = .063$, $\eta_p^2 = .074$. This showed that there was a quicker response when the probe appeared in place of the neutral word compared to when it appeared in place of the happy word.
This evidence does not show support for hypothesis seven. Based on the lack of evidence of attention bias effects for happy words, it was predicted that participants would show no difference in reaction time to probes that replace *happy* words and probes that replace *neutral* words. The post-hoc tests also showed no significant interaction between trial congruency and physical aggression on *happy-neutral* trials ($p = .994$). This suggests that the main effect of trial congruency is similar across both aggression groups. It was predicted that low and high aggression groups would show no difference in reaction time to probes that replace *happy* words and probes that replace *neutral* words. However inspection of bar graph (Figure 10) shows that participants have faster responses to probes replacing neutral words compared to probes replaces happy words. Therefore, hypothesis eight and nine are not supported. There were no significant effects of attentional shifting.

On *angry-happy* trials, post-hoc tests showed no significant effects of congruency ($p = .657$). It was predicted that participants would be quicker to respond to probes that replace *angry* words compared to probes that replace *happy* words. However, hypothesis ten was not supported. Post-hoc tests showed no significant interaction between trial congruency and physical aggression on *angry-happy* trials ($p = .827$). This suggests that the main effect of trial congruency is similar across both aggression groups and therefore reaction times for congruent and incongruent trials are similar across participants scoring high and low physical aggression. Therefore, hypothesis 11 and 12 are not supported. There were no significant effects of attentional shifting.

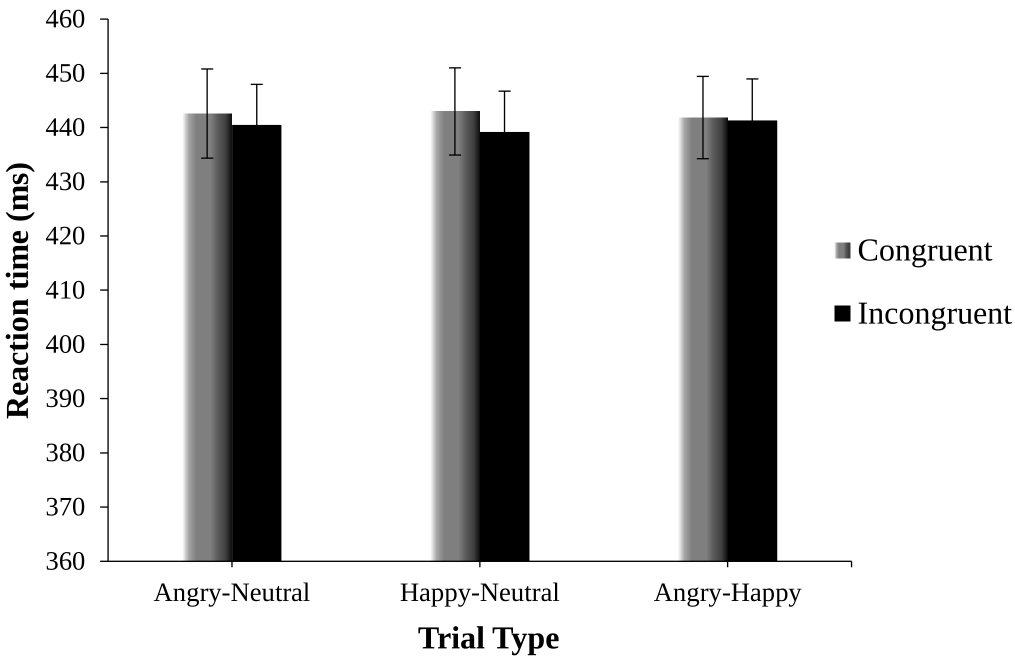

Figure 10: Reaction time for congruent and incongruent trials by trial type (n = 51; error bars = +/- 1 standard error).

4.4.4 ERP data

4.4.4.1 Effect of aggression

The main effect of aggression across angry-neutral, angry-happy and happy-neutral trial types was explored at each electrode for each epoch before probe presentation. On *angry-neutral* trials the main effect of aggression was significant at P3, $F(1,46) = 4.103$, $p = .049$, $\eta_p^2 = .082$; and approached significance at CP5, $F(1,46) = 3.512$, $p = .067$, $\eta_p^2 = .071$; and CP1, $F(1,46) = 3.512$, $p = .067$, $\eta_p^2 = .071$, between 400 and 500ms only. On *angry-happy* trials the main effect of aggression was significant at CP2, $F(1,46) = 6.810$, $p = .012$, $\eta_p^2 = .129$; and approached significance at P3, $F(1,46) = 3.884$, $p = .055$, $\eta_p^2 = .078$, between 400 and 500ms only. Across *angry-neutral* and *angry-happy* trials the high physical aggression group had an increased amplitude compared to the low physical aggression group between 400 and 500ms (Figure 11).

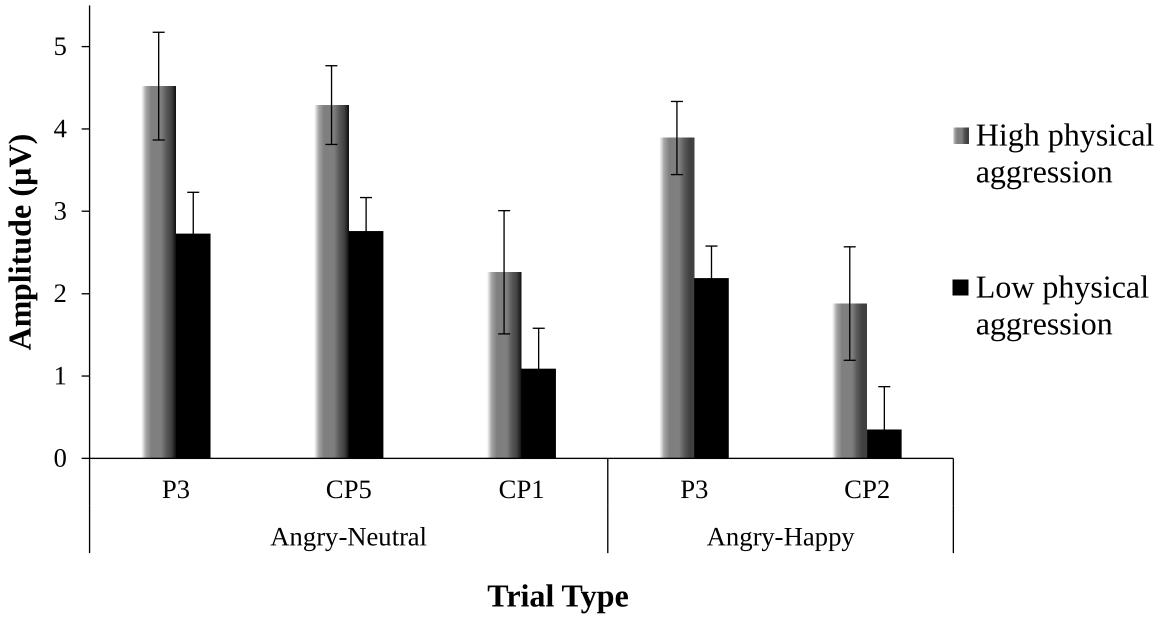

Figure 11: Evoked amplitudes of high and low physical aggression groups on angry-neutral and angry-happy trial types between 400 and 500ms (error bars = +/- 1 standard error).

On *happy-neutral* trials the main effect of aggression was significant at CP2, $F(1,46) = 4.264$, $p = .045$, $\eta_p^2 = .085$, between 200 and 300ms only. These results show that low physical aggression participants show early increased amplitude in response to happy-neutral trials compared to the high physical aggression group.

Overall, there are only minor effects of aggression across all three trial types and these differences were evident pre-probe presentation (Figure 12). However, the findings for angry-neutral and angry-happy trials show support for hypothesis 13 as it was predicted that high physical participants would have increased P300 response to word pairs.

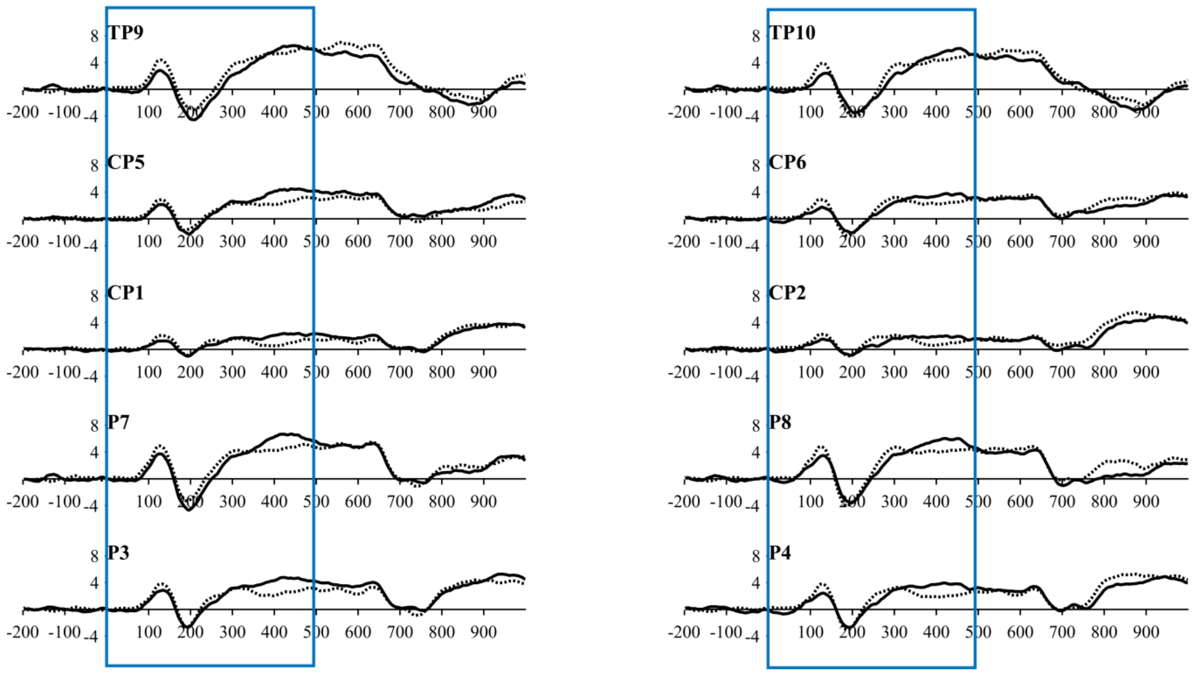

Figure 12: Grand average ERPs for the effect of physical aggression group (high vs. low) across all trials. The high physical aggression group (n = 23; black) is compared with the low physical aggression group (n = 25; dotted) Pre-probe epochs are highlighted by the blue box.

4.4.4.2 Effect of valence

A mixed model omnibus ANOVA was conducted for each epoch pre-probe presentation to explore the effect of valence. Trial type (3 levels; angry-neutral, angry-happy, happy-neutral), electrode (5 levels), and hemisphere (2 levels) were added as within subject factors. Physical aggression was added as a between-subject factor. The ANOVA revealed a significant main effect of trial type between 200 and 300ms, $F(2,92) = 3.860$, $p = .026$, $\eta_p^2 = .077$; 300 and 400ms, $F(2,92) = 4.008$, $p = .024$, $\eta_p^2 = .080$; and 400 and 500ms, $F(2,92) = 4.125$, $p = .019$, $\eta_p^2 = .082$. Figure 13 shows that across all three epochs angry-neutral trials evoke the greatest positive amplitude, followed by happy-neutral trials, and then angry-happy trials. This effect did not interact with physical aggression.

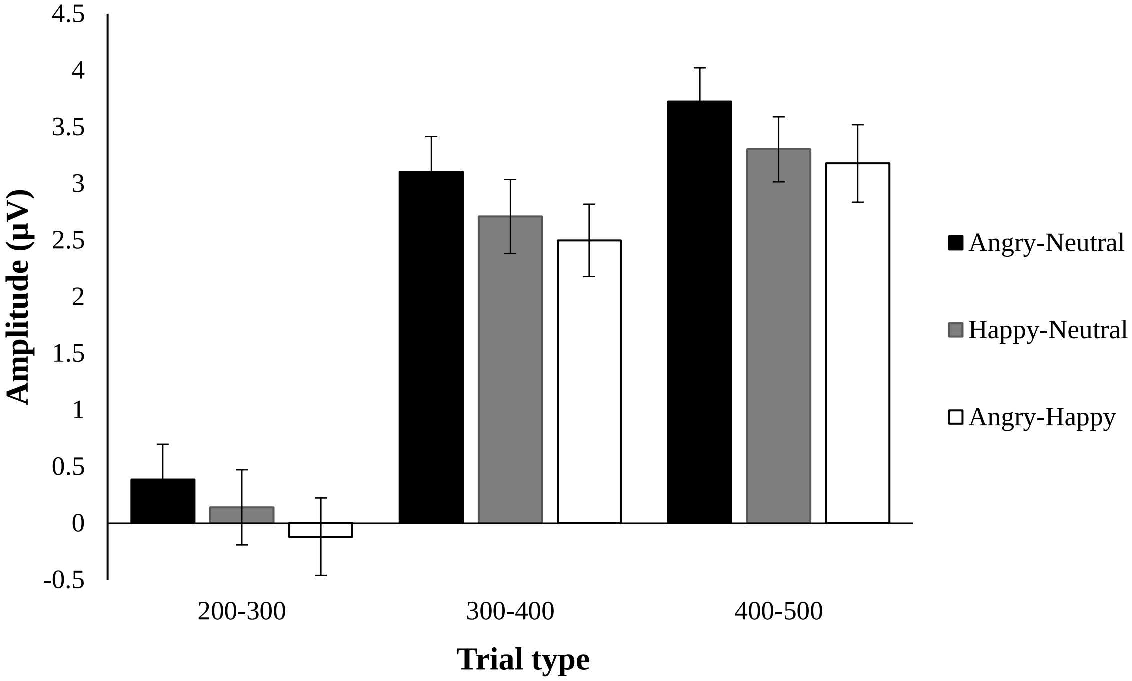

Figure 13: Bar graph to show the differences between trial types in average evoked amplitude of all electrodes in our region of interest across all participants (n = 51) (error bars = +/- 1 standard error).

To explore the main effect of valence, a 2 (trial type) x 5 (electrodes) x 2 (hemisphere) was conducted for each trial type pairing (angry-neutral/angry-happy, angry-neutral/happy-neutral, angry-happy/happy-neutral). Post-hoc effects show that for *angry-neutral and angry-happy* trials there is a main effect of trial type between 200 and 300ms, $F(1,50) = 7.648$, $p = .008$, $\eta_p^2 = .133$; 300 and 400ms, $F(1,50) = 8.216$, $p = .006$, $\eta_p^2 = .141$; and 400 and 500ms, $F(1,50) = 6.684$, $p = .013$, $\eta_p^2 = .118$. There was also an interaction between trial type and electrode between 200 and 300ms, $F(4,200) = 2.919$, $p = .039$, $\eta_p^2 = .055$; 300 and 400ms, $F(4,200) = 5.064$, $p = .003$, $\eta_p^2 = .092$; and 400 and 500ms, $F(4,200) = 2.587$, $p = .048$, $\eta_p^2 = .049$. Results of follow up test for each electrode can be found in Table nine. The results show a consistent significant difference in amplitude between *angry-neutral* trials and *angry-happy* trials across TP9, CP1, P7, and P8. The waveform (Figure 14) shows that in general participants show increased amplitude on angry-neutral trials compared to angry-happy trials.

Table 9: Results of the follow up tests exploring trial type differences between angry-neutral and angry-happy trial types across the electrodes of interest for each epoch.

		Angry-Neutral/Angry-Happy
200-300	TP9	$F(1,50) = 10.545$, $p = .002$, $\eta_p^2 = .174$
	TP10	$F(1,50) = 3.048$, $p = .087$, $\eta_p^2 = .057$
	CP5	$F(1,50) = 3.548$, $p = .065$, $\eta_p^2 = .066$
	CP6	*NS*
	CP1	$F(1,50) = 3.278$, $p = .076$, $\eta_p^2 = .062$
	CP2	$F(1,50) = 4.582$, $p = .037$, $\eta_p^2 = .084$
	P7	$F(1,50) = 5.456$, $p = .024$, $\eta_p^2 = .098$
	P8	$F(1,50) = 13.414$, $p = .001$, $\eta_p^2 = .212$
	P3	*NS*
	P4	*NS*
300-400	TP9	$F(1,50) = 13.198$, $p = .001$, $\eta_p^2 = .209$
	TP10	$F(1,50) = 4.237$, $p = .045$, $\eta_p^2 = .078$
	CP5	*NS*
	CP6	*NS*
	CP1	$F(1,50) = 5.410$, $p = .024$, $\eta_p^2 = .098$
	CP2	$F(1,50) = 2.986$, $p = .090$, $\eta_p^2 = .056$
	P7	$F(1,50) = 6.363$, $p = .015$, $\eta_p^2 = .113$
	P8	$F(1,50) = 9.810$, $p = .003$, $\eta_p^2 = .164$
	P3	$F(1,50) = 4.050$, $p = .050$, $\eta_p^2 = .075$
	P4	*NS*
400-500	TP9	$F(1,50) = 6.77$, $p = .012$, $\eta_p^2 = .120$
	TP10	*NS*
	CP5	*NS*
	CP6	*NS*
	CP1	$F(1,50) = 7.801$, $p = .007$, $\eta_p^2 = .135$
	CP2	*NS*
	P7	$F(1,50) = 5.702$, $p = .021$, $\eta_p^2 = .102$
	P8	$F(1,50) = 12.325$, $p = .001$, $\eta_p^2 = .198$
	P3	$F(1,50) = 4.813$, $p = .033$, $\eta_p^2 = .088$
	P4	*NS*

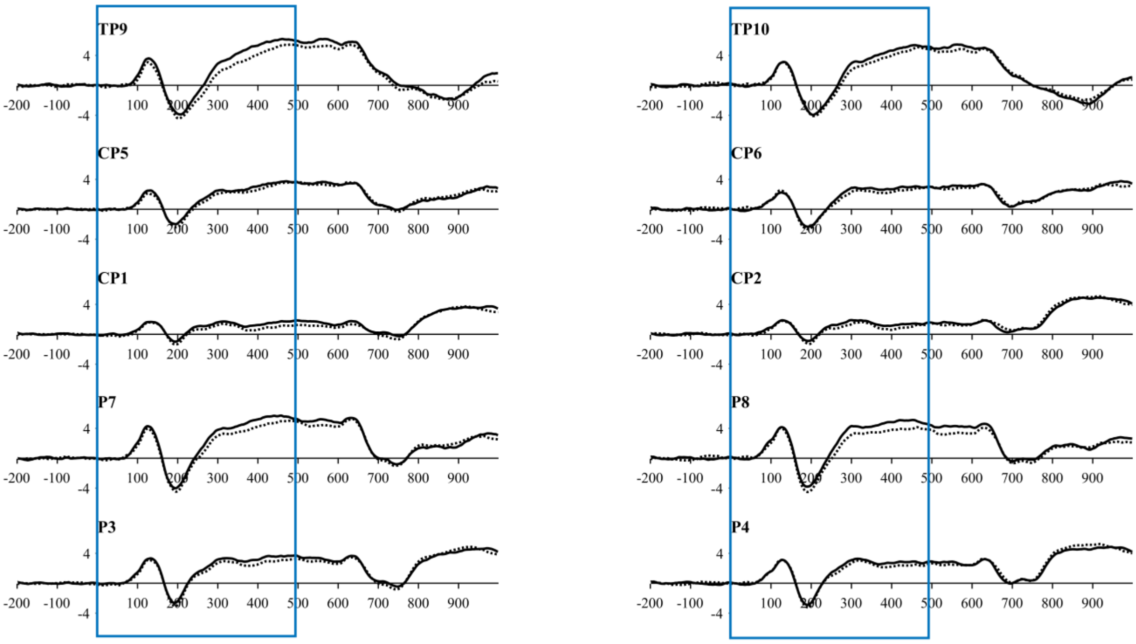

Figure 14: Grand average ERPs (n = 51) for the effect of trial type in all participants. Mean amplitude to angry-neutral trials (black) are compared with mean amplitude to angry-happy trials (dotted). Pre-probe epochs are highlighted by the blue box.

For *angry-neutral and happy-neutral* trials the main effect of trial type was significant between 400 and 500ms, $F(1,50) = 4.482$, $p = .039$, $\eta_p^2 = .082$; and approached significance between 300 and 400ms, $F(1,50) = 3.941$, $p = .053$, $\eta_p^2 = .073$. Between 300 and 400 there was also a significant interaction between trial type, electrode and hemisphere, $F(4,200) = 2.714$, $p = .045$, $\eta_p^2 = .051$. Follow up tests showed that the effect of trial type was significant at TP, $F(1,50) = 4.651$, $p = .036$, $\eta_p^2 = .085$; and CP2, $F(1,50) = 4.544$, $p = .038$, $\eta_p^2 = .083$. The effect also approached significance at CP6, $F(1,50) = 3.207$, $p = .079$, $\eta_p^2 = .066$; and P4, $F(1,50) = 3.698$, $p = .060$, $\eta_p^2 = .069$.These results show increased amplitude to *angry-neutral* trials compared to *happy-neutral* trials between 300 and 400ms (see Figure 15). The findings show there are differences between onset of word pairs across the whole sample, however, there are no significant interactions with aggression suggesting that participants in the low and high physical aggression groups responded similarly to *happy-neutral* and *angry-neutral* word pairs.

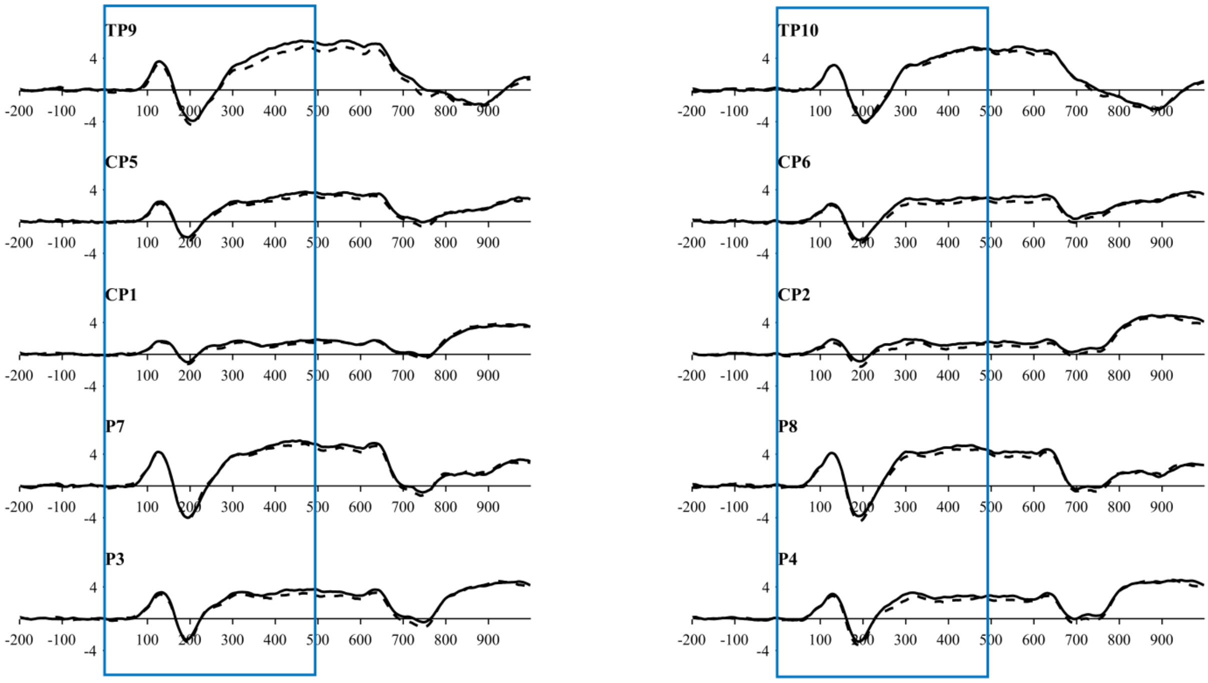

Figure 15: Grand average ERPs (n = 51) for the effect of trial type in all participants. Mean amplitude to angry-neutral trials (black) are compared with mean amplitude to happy-neutral trials (dashed). Pre-probe epochs are highlighted by the blue box.

For *angry-happy and happy-neutral* trials the interaction between trial type and electrode was significant between 200 and 300ms, $F(4,200) = 8.413$, $p < .001$, $\eta_p^2 = .144$; and 300 and 400ms, $F(4,200) = 3.525$, $p = .015$, $\eta_p^2 = .066$. Follow up tests between 200 and 300ms showed that the effect of trial type was significant at P7, $F(1,50) = 9.249$, $p = .004$, $\eta_p^2 = .156$. The effect also approached significance at TP9, $F(1,50) = 3.936$, $p = .053$, $\eta_p^2 = .073$; TP10, $F(1,50) = 3.943$, $p = .053$, $\eta_p^2 = .073$; and P8, $F(1,50) = 3.299$, $p = .073$, $\eta_p^2 = .062$. Follow up tests between 300 and 400ms showed that the effect of trial type was significant at P7, $F(1,50) = 4.724$, $p = .035$, $\eta_p^2 = .086$, only (figure 16).

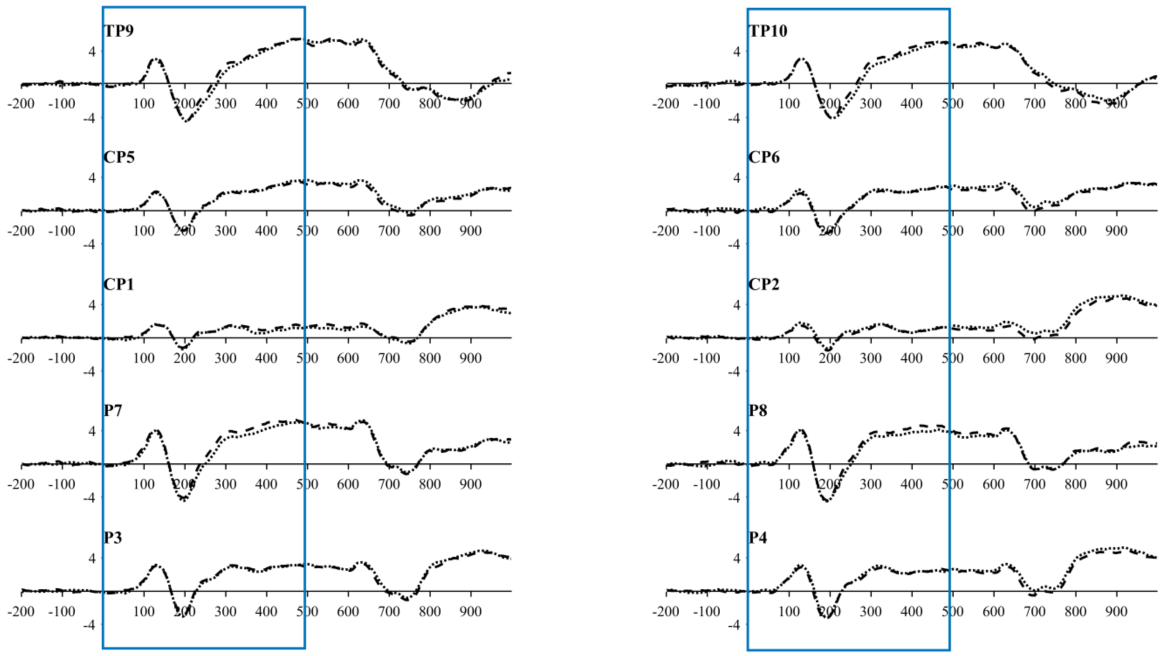

Figure 16: Grand average ERPs (n = 51) for the effect of trial type in all participants. Mean amplitude to angry-happy trials (dotted) are compared with mean amplitude to happy-neutral trials (dashed). Pre-probe epochs are highlighted by the blue box.

Overall the results provide mixed evidence for the hypothesis. It was predicted that *angry-happy* and *angry-neutral* trials would evoke increased amplitude compared to *happy-neutral* trials in the high physical aggression group. Although there were no interactions with physical aggression results showed a general effect of valence in which *angry-neutral* trials evoke increased amplitude compared to *happy-neutral* trials. Although there were only small differences between *angry-happy* and *happy-neutral* trials, surprisingly there was increased amplitude for *happy-neutral* trials. There was also marked differences between *angry-neutral* and *angry-happy* trials in which *angry-neutral* trials evoked an increased amplitude. These results show that *angry-neutral* trials evoke the greatest amplitude (see Figure 13).

4.4.4.3 Effect of trial congruency

4.4.4.3.1 *Post-probe differences in congruency.*

To explore the effect of congruency, a 3 (trial type; angry-neutral, angry-happy, happy-neutral) x 2 (congruency; congruent, incongruent) x 5 (electrode) x 2 (hemisphere; left, right) x 2 (physical aggression; high, low) mixed model omnibus ANOVA was conducted for each epoch. The results show a number of significant interactions with trial type and congruency between 500 and 1000ms. The interaction between trial type and congruency was significant between 600 and 700ms, $F(2,92) = 3.329$, $p = .042$, $\eta_p^2 = .067$; 700 and 800ms, $F(2,92) = 3.788$, $p = .026$, $\eta_p^2 = .076$; 800 and 900ms, $F(2,92) = 4.843$, $p = .011$, $\eta_p^2 = .095$; and 900 and 1000ms, $F(2,92) = 3.899$, $p = .024$, $\eta_p^2 = .078$. The interaction also approached significance between 500 and 600ms, $F(2,92) = 2.950$, $p = .057$, $\eta_p^2 = .060$. Finally the mixed model omnibus ANOVA revealed a significant interaction between congruency, hemisphere and physical aggression between 500 and 600ms, $F(1,46) = 4.687$, $p = .036$, $\eta_p^2 = .092$; and between 700 and 800ms, $F(1,46) = 4.687$, $p = .036$, $\eta_p^2 = .092$. To explore the significant interactions, for each trial type, a 2 (congruency) x 5 (electrode) x 2 (hemisphere) repeated measures ANOVA was conducted to explore for which trial types the effect of congruency was significant. Physical aggression was added as a between-subject factor when conducting post-hoc analyses for the significant interactions between congruency and aggression.

On *angry-neutral* trials, post-hoc analyses showed no effects of congruency across any epoch and therefore hypothesis 15 was not supported. Post-hoc analyses for *angry-neutral* trials also showed no interactions with physical aggression, suggesting that effects of congruency are stable across both groups. As the main effect of congruency was not significant, this suggests that both the high and low aggression group show little difference between evoked amplitude on congruent and incongruent trials. It was hypothesised that participants scoring low on physical aggression would show increased positive amplitude on incongruent trials compared with congruent trials, therefore hypothesis 16 is not supported. It was also hypothesised that participants scoring high on physical aggression will show similar amplitude in response to congruent and incongruent trials therefore hypothesis 17 is supported.

On *happy-neutral* trials, post-hoc tests revealed a significant interaction between congruency and electrode between 500 and 600ms, $F(4,200) = 3.140$, $p = .025$, $\eta_p^2 = .059$. This effect also approached significance between 600 and 700ms, $F(4,200) = 2.393$, $p = .077$, $\eta_p^2 = .046$; and 800 and 900ms, $F(4,200) = 2.510$, $p = .055$, $\eta_p^2 = .048$. Follow up tests showed that the effect of congruency was significant at TP10, $F(1,50) = 6.419$, $p = .014$, $\eta_p^2 = .114$, between 800 and 900ms only. This suggests that there is increased amplitude on congruent trials compared to incongruent trials. However as the effect is only significant at TP10 at one epoch, it suggests that this effect is not very robust. This evidence does not provide support for hypothesis 18.

Further post-hoc analyses showed a significant interaction between congruency, hemisphere and physical aggression for *happy-neutral* trials between 700 and 800ms only, $F(1,46) = 12.855$, $p = .001$, $\eta_p^2 = .218$. Follow up analyses revealed that the effect of congruency was significant in the high aggression group at CP6, $F(1,22) = 8.195$, $p = .009$, $\eta_p^2 = .271$; and P4, $F(1,22) = 4.426$, $p = .047$, $\eta_p^2 = .168$; and approached significance at TP10, $F(1,22) = 3.926$, $p = .060$, $\eta_p^2 = .151$. High physical aggression participants showed an increased positive amplitude on congruent trials compared to incongruent trials (see Figure 17). There were no significant effects in the low aggression group suggesting that trial congruency effects on *happy-neutral* trials may be more salient in the high aggression group. It was predicted that amplitude would be relatively stable across congruent and incongruent trials for both aggression groups, therefore the results show evidence for hypothesis 19 only. Predictions made regarding the high aggression group (hypothesis 20) were not supported.

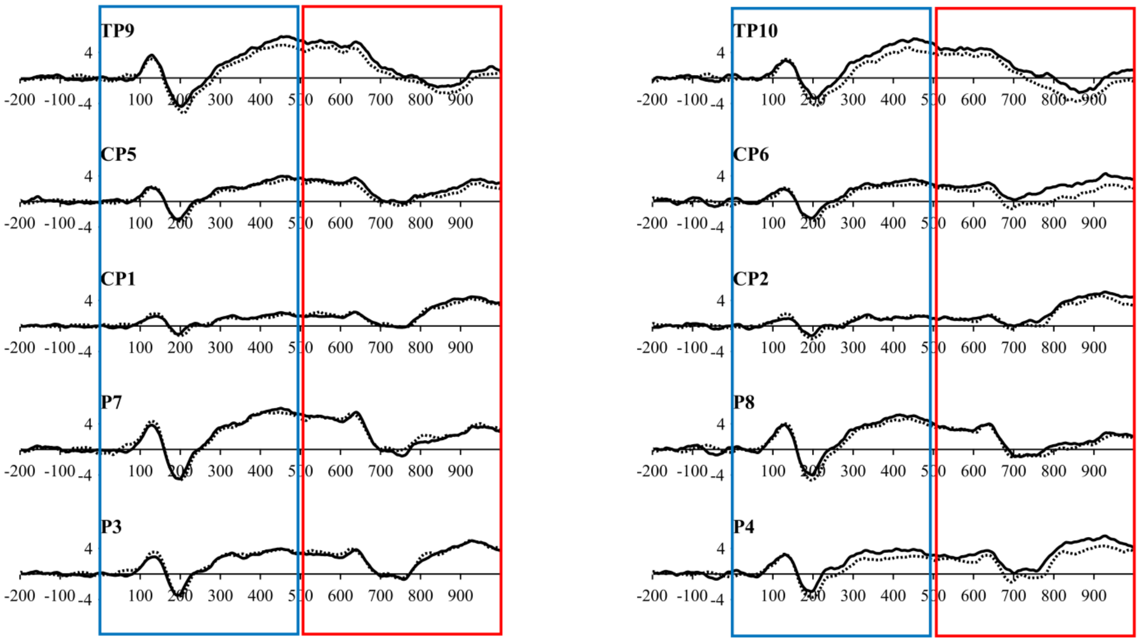

Figure 17: Grand average ERPs for the effect of trial congruency on happy-neutral trials in the high physical aggression group (n = 23). Mean amplitude on congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

On *angry-happy* trials, post-hoc tests showed a main effect of congruency between 500 and 600ms, $F(1,50) = 11.010$, $p = .002$, $\eta_p^2 = .180$; 600 and 700ms, ($F(1,50) = 7.923$, $p = .007$, $\eta_p^2 = .137$; 700 and 800ms, $F(1,50) = 6.410$, $p = .015$, $\eta_p^2 = .114$; and 800 and 900ms, $F(1,50) = 7.457$, $p = .009$, $\eta_p^2 = .130$. The interaction between congruency and electrode also approached significance between 500 and 600ms, $F(4,200) = 2.757$, $p = .058$, $\eta_p^2 = .052$. Follow up tests showed a consistent main effect of congruency at TP9, TP10, CP5, and P3 (see Table 10). The evidence shows support for hypothesis 21. There are consistent differences in amplitude between congruent and incongruent trials on angry-happy trials across multiple electrodes. Results suggest that incongruent trials evoke increased amplitude compared to congruent trials.

Table 10: Significant effects of congruency at each electrode between 500 and 900ms

	500-600ms	600-700ms
TP9	$F(1,50) = 8.717$, $p = .005$, $\eta_p^2 = .148$	$F(1,50) = 4.542$, $p = .038$, $\eta_p^2 = .083$
TP10	$F(1,50) = 9.031$, $p = .004$, $\eta_p^2 = .153$	$F(1,50) = 7.480$, $p = .009$, $\eta_p^2 = .130$
CP2	*NS*	$F(1,50) = 3.932$, $p = .053$, $\eta_p^2 = .073$
CP5	$F(1,50) = 7.183$, $p = .010$, $\eta_p^2 = .126$	$F(1,50) = 4.410$, $p = .041$, $\eta_p^2 = .081$
CP6	$F(1,50) = 12.177$, $p = .001$, $\eta_p^2 = .160$	$F(1,50) = 7.094$ $p = .010$, $\eta_p^2 = .124$
P3	$F(1,50) = 12.496$, $p = .001$, $\eta_p^2 = .200$	$F(1,50) = 11.884$, $p = .001$, $\eta_p^2 = .192$
P4	$F(1,50) = 3.956$, $p = .052$, $\eta_p^2 = .073$	*NS*
P8	*NS*	*NS*

	700-800ms	800-900ms
TP9	$F(1,50) = 5.303$, $p = .025$, $\eta_p^2 = .096$	$F(1,50) = 6.772$, $p = .012$, $\eta_p^2 = .119$
TP10	$F(1,50) = 6.617$, $p = .013$, $\eta_p^2 = .117$	$F(1,50) = 8.573$, $p = .005$, $\eta_p^2 = .146$
CP2	$F(1,50) = 5.125$, $p = .028$, $\eta_p^2 = .093$	*NS*
CP5	$F(1,50) = 4.158$, $p = .047$, $\eta_p^2 = .077$	$F(1,50) = 4.042$, $p = .050$, $\eta_p^2 = .075$
CP6	*NS*	$F(1,50) = 5.187$, $p = .027$, $\eta_p^2 = .094$
P3	$F(1,50) = 7.904$, $p = .007$, $\eta_p^2 = .136$	$F(1,50) = 5.548$, $p = .022$, $\eta_p^2 = .100$
P4	*NS*	*NS*
P8		$F(1,50) = 4.468$, $p = .040$, $\eta_p^2 = .082$

Further post-hoc analyses for *angry-happy* trials showed the interaction between congruency, hemisphere and physical aggression approached significance between 500 and 600ms, $F(1,46) = 3.947$, $p = .053$, $\eta_p^2 = .079$; and 700 and 800ms, $F(1,46) = 2.996$, $p = .092$, $\eta_p^2 = .061$. Follow up tests between 500 and 600ms show an effect of congruency at TP9, $F(1,24) = 3.464$, $p = .075$, $\eta_p^2 = .126$; TP10, $F(1,24) = 4.914$, $p = .006$, $\eta_p^2 = .170$; CP5, $F(1,24) = 4.382$, $p = .047$, $\eta_p^2 = .154$; CP6, $F(1,24) = 13.439$, $p = .001$, $\eta_p^2 = .359$; CP2, $F(1,24) = 3.482$, $p = .074$, $\eta_p^2 =$

.127; and P3, $F(1,24) = 4.861$, $p = .037$, $\eta_p^2 = .168$, in the low aggression group. The effect was also significant at TP9, $F(1,22) = 5.509$, $p = .028$, $\eta_p^2 = .200$; CP5, $F(1,22) = 3.913$, $p = .061$, $\eta_p^2 = .151$; CP6, $F(1,22) = 3.672$, $p = .068$, $\eta_p^2 = .143$; and P3, $F(1,22) = 12.110$, $p = .002$, $\eta_p^2 = .355$, in the high aggression group. Follow up tests between 700 and 800ms show an effect of congruency at TP9, $F(1,24) = 5.439$, $p = .028$, $\eta_p^2 = .185$; TP10, $F(1,24) = 8.576$, $p = .007$, $\eta_p^2 = .263$; CP6, $F(1,24) = 6.730$, $p = .016$, $\eta_p^2 = .219$; CP2, $F(1,24) = 4.329$, $p = .048$, $\eta_p^2 = .153$; and P8, $F(1,24) = 4.079$, $p = .055$, $\eta_p^2 = .145$, in the low aggression group. The effect was also significant at P3, $F(1,22) = 6.541$, $p = .018$, $\eta_p^2 = .229$, in the high aggression group.

In line with the predictions (hypothesis 22) and previous results from Study 1, participants scoring low on physical aggression showed increased positive amplitude to incongruent trials compared to congruent trials (Figure 18). The high aggression participants also showed increased amplitude to incongruent trials compared to congruent trials (Figure 19). This shows no support for hypothesis 23 as it was predicted that high aggression group would show relatively stable amplitude in response to congruent and incongruent trials. Results suggest that both aggression groups show consistent effects of trial congruency. The results demonstrate that effects may be most salient at TP9, CP5, CP6 and P3.

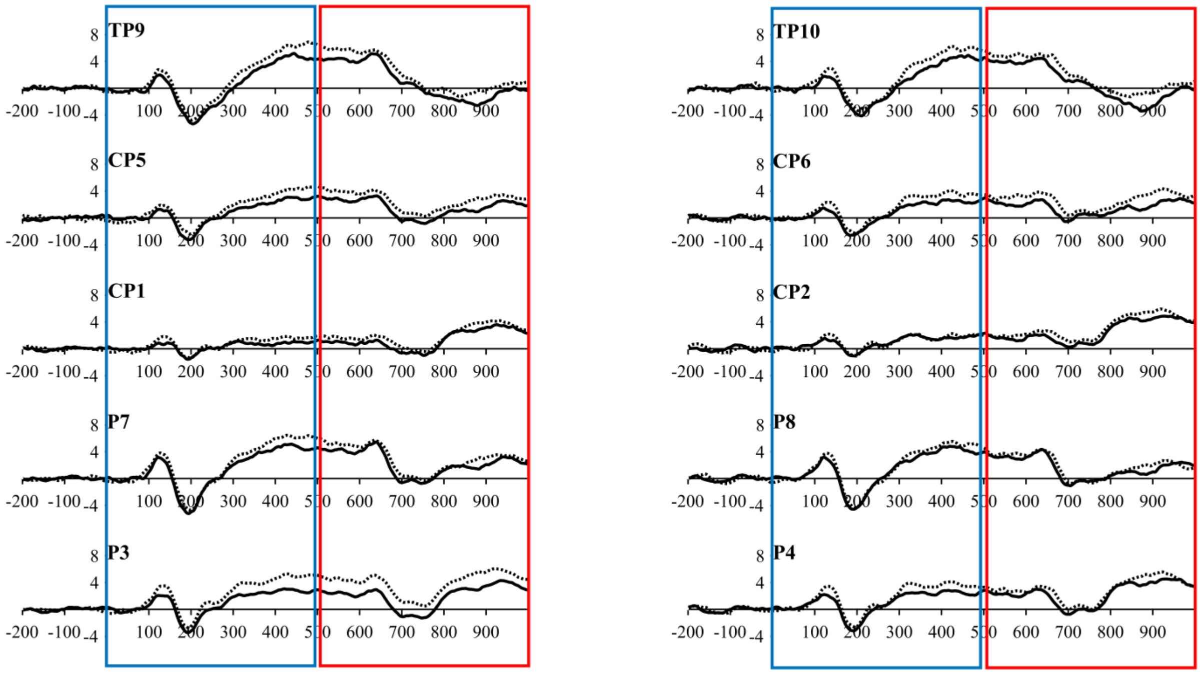

Figure 18: Grand average ERPs for the effect of trial congruency on angry-happy trials in the low physical aggression group (n = 25). Mean amplitude on congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

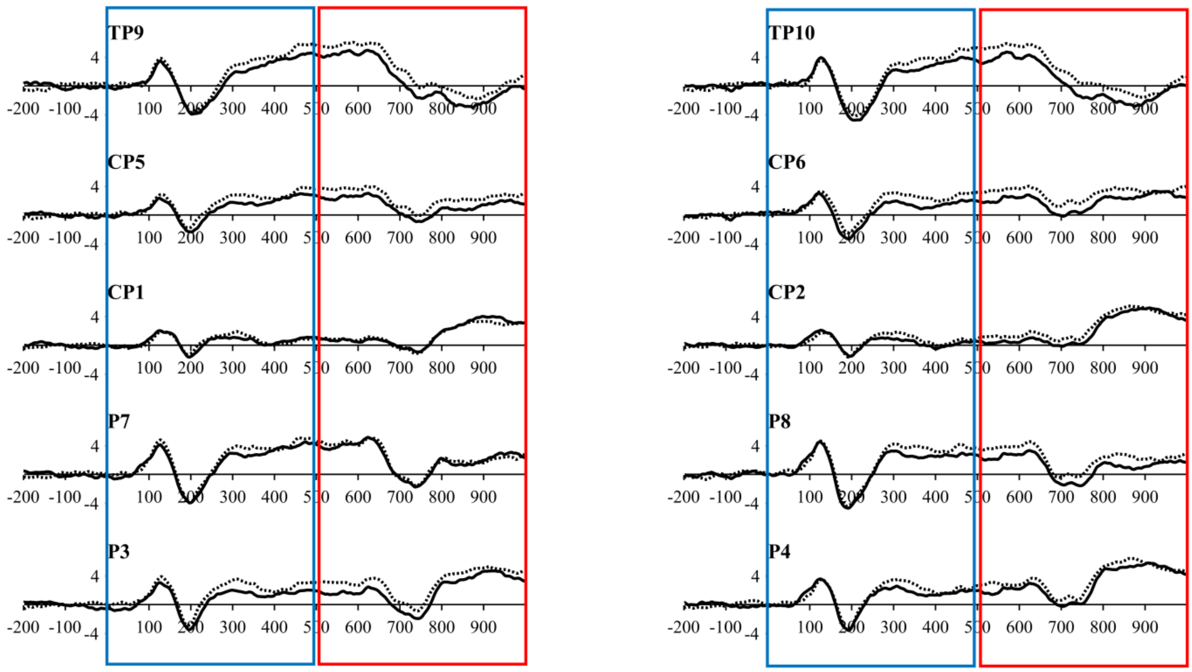

Figure 19: Grand average ERPs for the effect of trial congruency on angry-happy trials in the high physical aggression group (n = 23). Mean amplitude on congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

4.4.4.3.2 *Pre-probe differences in congruency.*

Based on a qualitative inspection of the waveform, effects of congruency at earlier latencies (100-500ms) were also explored. Surprisingly the mixed model ANOVA also revealed significant interactions between trial type and congruency between 300 and 500ms. Between 300 and 400ms, the ANOVA showed a significant interaction between trial type and congruency, $F(2,92) = 4.528$, $p = .014$, $\eta_p^2 = .090$; trial type, congruency and electrode, $F(8,368) = 2.859$, $p = .013$, $\eta_p^2 = .059$; and trial type, congruency, hemisphere and physical aggression, $F(2,92) = 4.102$, $p = .022$, $\eta_p^2 = .082$. Between 400 and 500ms there was a significant interaction between trial type and congruency, $F(2,92) = 3.679$, $p = .029$, $\eta_p^2 = .074$; and between trial type, congruency and electrode, $F(8,368) = 2.543$, $p = .022$, $\eta_p^2 = .052$.

Post-hoc tests between 300 and 400ms showed a main effect of congruency for *angry-neutral* trials, $F(1,46) = 4.481$, $p = .040$, $\eta_p^2 = .08$. The interaction between congruency, hemisphere, and physical aggression also approached significance, $F(1,46) = 3.905$, $p = .054$, $\eta_p^2 = .078$. Follow up tests showed a significant interaction between congruency and hemisphere in the high aggression group, $F(1,22) = 4.704$, $p = .041$, $\eta_p^2 = .176$. Effect of congruency was significant in the right hemisphere only, $F(1,22) = 4.81$, $p = .050$, $\eta_p^2 = .163$, at electrode site P8, $F(1,22) = 5.281$, $p = .031$, $\eta_p^2 = .194$. There were no significant effects in the low aggression group. For *angry-happy* trials there was a main effect of congruency, $F(1,46) = 11.181$, $p = .002$, $\eta_p^2 = .196$; the interaction between congruency and electrode approached significance, $F(4,184) = 2.502$, $p = .062$, $\eta_p^2 = .052$. Effect of congruency was significant at TP9, $F(1,50) = 4.015$, $p = .050$, $\eta_p^2 = .074$; TP10, $F(1,50) = 5.965$, $p = .018$, $\eta_p^2 = .107$; CP5, $F(1,50) = 4.770$, $p = .034$, $\eta_p^2 = .087$); CP6, $F(1,50) = 7.511$, $p = .008$, $\eta_p^2 = .131$; P7, $F(1,50) = 4.393$, $p = .041$, $\eta_p^2 = .081$; P3, $F(1,50) = 11.856$, $p = .001$, $\eta_p^2 = .192$; and P4, $F(1,50) = 9.435$, $p = .003$, $\eta_p^2 = .159$. For *happy-neutral* trials there was a significant interaction between congruency, hemisphere and physical aggression, *F*(1,46) = 4.991, *p* = .030, η_p^2 = .098. However follow up tests revealed no significant effects of congruency in either aggression group.

Post-hoc tests between 400 and 500ms revealed no significant effects on *angry-neutral* trials or *happy-neutral* trials. On *angry-happy* trials, post-hoc tests showed a significant main effect of congruency, *F*(1,46) = 8.541, *p* = .005, η_p^2 = .157; and a significant interaction between congruency and electrode, *F*(4,184) = 2.912, *p* = .032, η_p^2 = .060. The effect of congruency was significant at TP9, *F*(1,50) = 5.416, *p* = .024, η_p^2 = .098); TP10, *F*(1,50) = 6.544, *p* = .014, η_p^2 = .116; CP6, *F*(1,50) = 4.530, *p* = .038, η_p^2 = .083; and P3, *F*(1,50) = 8.372, *p* = .006, η_p^2 = .143.

4.5 Discussion

In this chapter attention bias to angry and happy words in low and high physical aggression groups during a selective attention task was explored. The first aim was to replicate Study 1 and test whether findings were consistent. To do this selective attention to angry words when they were paired with a neutral word during a dot-probe task was explored. By using a modified dot-probe paradigm in which two further trial types, happy-neutral and happy-angry, were added, attentional processes involved with selectively attending to positive and negative valenced words were also explored. Including these trial types it was possible to investigate the attention bias to angry and happy words when paired with a neutral distracter, and also to explore attentional processes involved with selective attention when two emotionally valenced words are presented simultaneously. The study used a unique combination of behavioural and ERP methods to measure attention bias which allows for a more robust assessment of cognitive processes. Due to the complexity of the results the behavioural and ERP results for each trial type will be explained and discussed individually before an overview of the main findings are presented.

4.5.1 Main findings and interpretations

4.5.1.1 Behavioural results

4.5.1.1.1 *Angry - neutral*

The correlation and between-subject results showed no evidence for any of the hypotheses relating to angry-neutral trials (hypotheses one, four, five and six). Based on previous literature and theories of attention it was hypothesised that all participants would show an attention bias towards aggression-related words and that this effect would be most salient in the high aggression group. It is proposed that detection of stimuli of negative valence has an adaptive value from a biological and psychological perspective. Protecting oneself from threat is evolutionarily important for survival and therefore will command attentional resources (Vuilleumier, 2005). The results showed no significant difference in reaction time between angry and neutral trials in either aggression group. This is surprising considering the current literature, although past studies have predominantly used the Stroop task (Chan et al., 2010; Smith & Waterman, 2005; Sutton & Altarriba, 2011; van Honk et al., 2001b). However, Smith and Waterman (2003) investigated attention biases in both the Stroop and dot-probe task and consistently found preferential attention for aggression-related words (delayed colour naming on Stroop task and quickened reaction time on the dot-probe). The results from this study did not show evidence of facilitated attention to angry words in aggression.

These findings are inconsistent with Study 1 where a negative correlation between physical aggression and attention bias was found, such that increased aggression was related to a more negative attention bias. Although between-subject effects in the current study did not reach significance, the means were similar to those found in Study 1; high aggression participants had speedier reaction times when responding to probes that appeared in place of angry words compared to neutral, whereas the low aggression group had faster reaction times to probes replacing neutral words, compared to angry words.

4.5.1.1.2 *Happy - neutral*

There was support for hypothesis two as aggression did not significantly correlate with happy-neutral attention bias. The between-subject effects revealed no support for hypotheses seven, eight and nine. Based on the findings of Sutton and Altarriba (2011), who found no attention bias effects on positive-neutral trials on a dot-probe task in a non-clinical sample, it was predicted that reaction times across congruent and incongruent trials would be relatively similar, and these would not vary across aggression groups. However, the main effect of congruency approached significance, such that participants were quicker to respond to probes that appeared in place of neutral words compared to happy words. There were no differences in bias across aggression groups which suggest that both the high and low aggression groups had similar reaction times to both targets. This is similar to the findings of Smith and Waterman (2003) who reported no significant differences in colour naming positive emotion words between aggression groups during an emotional Stroop task. Attention bias to positive stimuli is a phenomenon which is less frequently studied and subsequently evidenced to a lesser extent.

To my knowledge no studies have investigated selective attentional processes associated with attending to positive stimuli during a dot-probe task, specifically in relation to increased levels of aggression. However, there is mixed evidence of attention bias to happy words in the literature relating to anxiety, Pishyar et al. (2004) found no significant differences across happy and neutral targets, whereas Martin et al. (1991) found an attention bias to positive words in a high anxiety sample. Surprisingly, the attention bias towards neutral stimuli approached significance. Due to the differences in findings, and that previous research has used only a high anxiety sample (and not aggression samples), it is somewhat difficult to draw conclusions from the current finding. This effect also only approached significance and therefore should be interpreted with caution. This finding does not contribute to the understanding of selective attentional processes involved with attending to happy stimuli in aggression; however it does suggest a possible general population bias. Theoretically it is not clear why individuals would selectively attend to neutral words compared to happy words, as both negative and positive emotion words have been found to capture attention (Martin et al., 1991). However, based on the findings of Santesso et al. (2008), I tentatively suggest that participants attend to the most threatening stimulus in each stimuli pairing. Neutral stimuli which are somewhat ambiguous in nature can be perceived as hostile ((Mellentin, Dervisevic, Stenager, Pilegaard, & Kirk, 2015). Due to the novelty of the methodology used and the uniqueness of this finding, the need to replicate this in future work is recognised.

4.5.1.1.3 *Angry - happy*

The results provide no support for any of the hypotheses regarding angry-happy trials. In line with attentional theory that negative stimuli capture attention (e.g. Fox et al., 2000; Hansen & Hansen, 1988), it was predicted that across both groups there would be an attention bias towards angry words compared to the happy words. However, results showed no main effect of congruency which suggests no significant difference in reaction time between congruent and incongruent trials.

There is some evidence to suggest that both negative and positive stimuli capture attention if paired with a neutral distracter (e.g. Martin et al., 1991; Waters et al., 2010), however it is not known if biases are evident when negative and positive words are simultaneously presented. Previous evidence suggests that during a lexical decision task high trait anger individuals were found to have quicker reaction times when responding to the anger-related emotional words, compared with all other emotion words (Parrott et al., 2005). Furthermore, research suggests participants have longer latencies when colour naming threatening faces in comparison to both happy and neutral faces during a Stroop task (Putman et al., 2004). To my knowledge, currently there is no previous evidence concerning selective attention to simultaneously presented angry and happy words in aggression. The findings suggest that participants attend to angry and happy words relatively similarly, therefore either individuals have attentional facilitation to both word types, or attention bias to angry words is less salient when presented alongside an emotional distracter. However, across this study, this sample showed no evidence of a bias towards happy words when they were paired with neural words, or angry words when paired with neutral words. This suggests that the high physical aggression sample either show no evidence of attention bias for angry or happy words, or the behavioural measures are not sensitive to low level differences in processes associated with attention. Overall the absence of an attention bias to angry words is contradictory to previous evidence and therefore interpretations about the mechanisms of attention bias should be made tentatively.

4.5.1.2 ERP results

Findings were most sensitive to levels of physical aggression; it was suggested that differences in amplitude may be particularly salient in those individuals that form a violent reaction in response to a hostile situation. The initial

ERP analysis showed support for hypothesis 13. Based on results from Study 1, it was predicted that high physical aggression participants will have increased amplitude in response to all trial types (angry-neutral, angry-happy, and happy-neutral) at word pair onset compared to low aggression participants. This was supported by results which showed a significant effect of aggression on *angry-neutral* and *angry-happy* trials, such that the high physical aggression group had an increased amplitude compared to the low physical aggression group between 400 and 500ms post word pair onset. This is in contrast to previous literature which suggests reduced P300 amplitude in response to target stimuli in anti-social individuals (Bernat et al., 2007; Gao & Raine, 2009; Gao et al., 2013). It is theorized that aggressive individuals are less able to allocate resources to task-relevant stimuli and therefore do not show enhanced amplitude in response to word onset.

However results of the current study showed enhanced P300 amplitude in the low aggression group on *happy-neutral* trials. This suggests that the results vary across different stimulus pairings and therefore may be valence specific. It was found that high physical aggression participants showed increased amplitude to word pair onset compared to low aggression only when the word pair included an angry word. This is consistent with findings from Study 1 which showed increased amplitude in response to *angry-neutral* word pair presentation in the high physical aggression group. Across all three trial types, the differences between aggression groups was only found at one or two electrode sites and at very few epochs, therefore these effects are perhaps not very robust.

The findings revealed mixed evidence for hypothesis 14. It was predicted that word pairings that included an angry word (*angry-neutral* and *angry-happy*) would evoke an increased amplitude compared to *happy-neutral* word pairs, and that this effect would be most salient in the high physical aggression group. Although a significant interaction with aggression was not found, in support of the hypothesis results showed that *angry-neutral* trials evoke increased P300 amplitude compared to *happy-neutral* trials. This finding is comparable to work by Mueller et al. (2008) which found that participants with social anxiety disorder had increased P1 amplitude in response to angry-neutral face pairs compared to happy-neutral face pairs. Although this sample consisted of anxious participants, and the paradigm used facial images instead of words, these findings suggest that negative stimuli may influence early and later stages of attentional processing. It also suggests that angry words are allocated increased attentional resources. However in contrast to this proposed explanation, and the hypothesis, it was also found that participants showed increased amplitude for *happy-neutral* trials compared to *angry-happy* trials. Surprisingly, there was also a marked differences between *angry-neutral* and *angry-happy* trials in which *angry-neutral* trials evoked an increased amplitude. These results suggest that across all trial types *angry-neutral* trials evoke the greatest amplitude. To my knowledge there are no studies that have directly compared evoked P300 amplitude in response to these trial types. I suggest that angry-neutral trials may evoke increased potential due to the salience of the angry stimuli when presented alongside a neutral distracter. It is a well studied phenomenon that angry words command greater attention compared to neutral words (e.g. Smith & Waterman, 2003; van Honk et al., 2001b). There is also some evidence to suggest that greater resources are allocated to processing happy words (Sass et al., 2014). Due to the competition of resources between angry and happy faces, the effects of attentional facilitation to angry words may be attenuated when presented alongside a happy word. This is reflected in reduced P300 amplitude on angry-happy trials compared to angry-neutral and happy-neutral trials. The results showing differences in evoked P300 amplitude between trial types suggests a general population bias. However, no difference in ERP responses to word pairs between aggression groups were found, therefore conclusions cannot be drawn regarding how aggression may influence attentional processes associated with positive and negative stimuli presentation.

The comparison between trial types allows for conclusions concerning valence effects based on the presentation of negative-neutral, positive-neutral and positive-negative word pairs to be drawn. However, attentional selectivity cannot be inferred and therefore it is unclear which stimulus from the word pair is driving the patterns in evoked ERP potential. It is likely that a combination of attentional facilitation and disengagement contributes to the P300 amplitude for each trial type. Further analyses of ERP patterns in response to probe presentation were conducted to better understand potential differences in attention processes when probes appear in congruent and incongruent locations.

4.5.1.2.1 *Angry - neutral*

For *angry-neutral* there was no main effect of congruency following probe presentation, showing no support for hypothesis 15. Participants showed similar amplitude in response to congruent and incongruent trials. This finding was consistent across aggression groups. Based on Study 1, it was suggested that the low aggression group would show increased amplitude in response to incongruent trials. There were no effects of congruency in the low aggression group showing no support for hypothesis 16. Consistent with Helfritz-Sinville and Stanford (2015) current results showed that physically aggressive males show relatively undifferentiated ERPs in response to probes on congruent and incongruent trials and therefore hypothesis 17 was supported. Overall, the null results for epochs following probe presentation (500-900ms) show that aggressive individuals have relative uniformity when attending to probes following the presentation of angry and neutral words during a selective attention task. However in contrast to predictions the uniformity in ERP amplitude is consistent across both aggression groups. The current results are consistent with previous findings which found no significant differences in amplitudes across aggression groups when participants responded to aggressive words among neutral distracters during a modified oddball task (Surguy & Bond, 2006). The absence of a significant interaction between congruency and aggression in the current study is perhaps explained by the lack of behavioural differences in reaction time. Without a significant behavioural attention bias effect it was not possible to make interpretations of the ERP results in relation to reaction time on the dot-probe.

Surprisingly a main effect of congruency on angry-neutral trials between 300 and 400ms was found. The results suggested that individuals show increased amplitude to incongruent trials compared to congruent trials. These effects are unexpected as they appear before probe presentation at (500ms); however they replicate the early main effect of congruency found in Study 1. In the current study, follow up tests showed that the effect was particularly salient in the high aggression group at P8.

4.5.1.2.2 *Happy - neutral*

Based on the attention bias literature which provides very little evidence of attention bias to happy compared to neutral words (Pishyar et al., 2004; Smith & Waterman, 2003; Sutton & Altarriba, 2011), it was predicted that evoked amplitude on congruent and incongruent trials on *happy-neutral* trials would be relatively stable. However, results showed no support for hypothesis 18 as the main effect of congruency was significant at temporal-parietal electrodes between 800 and 900ms (300 to 400ms following probe presentation). This suggests that participants had increased P300 amplitude on congruent trials compared to incongruent trials. Although the findings did not support the predictions, there is some mixed evidence regarding attention bias to positive words in the literature. Therefore the results are comparable to findings which showed that control participants demonstrated larger evoked P300 amplitude in response to both pleasant and unpleasant words compared to neutral words (Sass et al., 2014). The literature on attention to emotional facial expressions provides much greater consistency in results. For example, Holmes et al. (2009), suggest greater processing of emotionally provoking stimuli (angry and happy faces) in normative healthy samples. The findings, although refer to a different modality of stimuli, are consistent with findings showing increased processing of positive stimuli compared to neutral stimuli.

Further analyses showed that the effect of congruency was greater in the high aggression group compared to the low aggression group. It was expected that no effects of trial congruency would be found in either aggression group. However results suggest that although the main effect of congruency is significant across the whole sample, differences between congruent and incongruent trials are most salient in the high aggression group. This may suggest that high aggression participants have increased processing of happy words compared to neutral words. This finding is inconsistent with findings on angry-neutral trials. It was expected that participants would show increased processing of emotionally salient stimuli when presented alongside neutral stimuli. Results suggest that high aggression participants had relative uniformity when responding to congruent and incongruent trials on angry-neutral trials. This is the first study to investigate ERP correlates of attention bias to happy words in aggression and therefore this finding will require replication.

4.5.1.2.3 *Angry – happy*

On *angry-happy* trials, there was a main effect of congruency between 500 and 900ms post word pair onset across multiple electrodes. Inspection of the waveform shows that there is consistently increased amplitude in response to incongruent trials (probe replaces happy word) compared to congruent trials (probe replaces angry word). Surprisingly, there were no interactions with aggression which suggests that both groups processed congruent and incongruent trials similarly. However, a general population bias in which incongruent trials would evoke increased amplitude was predicted, therefore hypothesis 21 was supported. This finding is somewhat contradictory to the previous literature. For example, Stewart et al. (2010) found that individuals with increased levels of anger had increased P300 amplitude to negative words compared to both positive and neutral words, and Santesso et al. (2008) found that P1 was larger for validly cued angry probes on angry-neutral trials, compared to validly cued happy probes on happy-neutral trials.

It was expected that effects of congruency would be visible from 600ms onwards, however, inspection of the waveform shows that they are evident as early as 300ms post word pair presentation. The later congruency effects seem to be long lasting and affecting several ERP components. At parietal electrodes for example, the waveform reveals; a positive peak a little after 600 (likely the P1) followed by a first negative peak (the N1), possibly reflecting the P1/N1 complex; there is then a very short positive peak (the P2) and another small negative peak (likely an N2) before a positive deflection at approximately 800ms that is likely the P300 component. The finding that congruency effects are evident across multiple components suggests that valence specific attentional processes influence early (such as the P1 and N1) and later (P300) cognitive processing.

Increased amplitude on incongruent trials could be attributed to the competition of attentional allocation between two stimuli high on emotional valence. There is evidence to suggest that both positive and negative emotional stimuli evoke increased P300 amplitude (Sass et al, 2014). The dot-probe task is a paradigm used to capture both facilitative and disengagement biases (Koster et al., 2004) and therefore based on the current analyses conclusions as to whether facilitation or disengagement processes contribute to the differences between congruent and incongruent trials cannot be made. It could be proposed that increased amplitude to incongruent trials may reflect the recruitment of resources needed to down regulate the simultaneously presented angry face distracter in order to complete the task. It is suggested that consistent with an inhibitory account of P300 (Polich, 2007), greater cognitive resources are needed to inhibit attentional facilitation of the angry word. However, neuro-cognitive models of aggression suggest that individuals with increased levels of physical aggression show greater deficits in regulatory control over incoming perceptual stimuli (e.g., Wilkowski & Robinson, 2010). Therefore, biases would be particularly salient in the high aggression group; conversely the effect of congruency was found to be consistent across aggression groups.

Surprisingly and consistent with the qualitative inspection of the waveform, significant effects of congruency were also found between 300 and 400ms. These were in the same direction as later effects and showed increased amplitude to incongruent trials compared to congruent trials. This finding is unexpected as theoretically effects of congruency should not be evident before the presentation of the probe at 500ms. However this effect has been internally replicated; in the current study and Study 1, pre-probe congruency effects were found at 300 and 400ms for angry-neutral trials. This suggests that pre-probe effects require further examination and replication. To my knowledge this is the first study to include an angry-happy trial type when investigating selective attentional processes in aggression. The behavioural results revealed no significant differences in reaction time, therefore conclusions based on complementary behavioural and ERP data is limited. This study has provided original evidence that suggests differences in ERP patterns in response to two simultaneously presented valenced stimuli, however additional work is needed to replicate these results and further understand the complex cognitive processes involved with attention biases.

4.5.2 Limitations

There are a number of possible limitations to consider when evaluating this work. The first of these is the complex nature of the dot-probe task used to measure selective attention. The task involved three different trial types, and two probe locations for each trial type (angry-neutral, angry-happy and happy-neutral, where the probe could appear in either location for each), resulting in six trial combinations. The trials were counterbalanced and randomly distributed across blocks. Therefore participants were presented with multiple stimuli very quickly. Therefore the task perceptual load could be high which means participants became less able to distinguish between different types of words. This could result in overlapping processes for each trial, for example selectively attending to, and processing an angry word when the next word pairing is presented. There is evidence to suggest that increasing the cognitive load during attention tasks may not affect emotion processing per se, but interfere with the ability to complete the task effectively (Berggren, Koster & Derakshan, 2012). Possible differences between behavioural results in Study 1 and 2 may be attributed to the task complexity. In Study 1 only angry and neutral words were presented and therefore the distinction between the two targets may be more salient. However, further evidence suggests that perceptual load and cognitive load have different effects on selective attention processes, whereas increased cognitive load such as working memory or dual task coordination is expected to increase distracter interference, increased perceptual load reduces distracter interference and ultimately improves task performance. Lavie, Hirst, De Fockert and Viding (2004). Compared to Study 1, in the current study, the perceptual load was increased as participants had to respond to three different stimulus pair combinations instead of one. However, the requirements of the task were the same (respond to the direction of the probe), therefore previous research suggests that the complexity of the task should not influence participants' ability to respond to each trial effectively.

Another possible limitation of this study is that the positive and negative words were not matched for valence and arousal. Therefore ERP results which show increased processing on incongruent trials compared to congruent trials on angry-happy trials may be confounded by increased valence or arousal. Although both stimuli were matched based on length and frequency using the Brysbaert database (Brysbaert & New, 2009), studies have shown that the arousal value of the stimuli is associated with attentional facilitation and that valence is a less determining factor when studying attention bias (e.g. Vogt, De Houwer, Crombez, Koster, & Van Damme, 2008). When investigating the effect of positive compared to negative stimuli on attention I would recommend matching the stimuli for arousal ratings.

There is evidence to suggest that attention biases are affected by current mood states and therefore potentially state anger should have been measured along with trait aggression. Smith et al. (2006) provided evidence to suggest that affective context moderates an attention bias towards negative information.

Participants were primed with either negative or positive information before attentional allocation towards negative or positive images was measured using both behavioural (Stroop task) and EEG methodology. They suggested that when participants were primed with the positive information, attention bias to negative stimuli can be eliminated or attenuated. ERP results showed that when participants were primed with negative information, the P1 amplitude was increased in response to negative stimuli in the testing phase, whereas when participants were primed with positive information, P1 amplitude was increased in response to positive stimuli. Results of the Stroop task showed that in the no-prime and negative prime condition participants showed an attention bias to negative targets, reflected in longer reaction times, whereas in the positive prime condition there were no significant differences in reaction time between positive and negative targets. These results show that attention bias to negative information is not evident in positive affective contexts. This suggests that current mood of the participants in studies may affect the attention bias results. This is consistent with Eckhardt and Cohen (1997) who demonstrated that attention biases were only evident in provoked situations when levels of both trait and state anger were increased. The non-significant between-subject effects for aggression in this current study may be explained by the positive context in which the experiment was conducted. Participants were made to feel comfortable during the laboratory session, with the aim of giving the participants a pleasant testing experience. Although the ‘high physical aggression’ participants report that they have the capacity to be physically aggressive, they were not in an aggressive state when responding to the dot-probe task. Another explanation for the null between-subject findings could be the recruitment of a non-clinical sample. The differences in the high and low aggression groups may be quite subtle as the study utilised a median split of scores based on a healthy sample. Although the aggression scores between-groups were significantly different, more extreme differences may be needed to demonstrate more robust differences in evoked amplitude in response to word stimuli.

4.5.3 Future work

Based on the recognised limitations of the current task, there are a number of improvements that could be made for future work. To reduce task complexity, the dot-probe could be simplified to present each stimulus pairing individually, in which case the participants would take part in three dot-probe tasks all with different stimuli; angry-neutral, angry-happy and happy-neutral pairings. An alternative would be to separate the trial types into separate blocks and not have them randomised throughout. Although there could be potential confounds of order effects this would allow for the more accurate analysis of differences in probe position for each pairing.

The lack of between-subject effects may be attributed to the non-clinical sample and therefore replicating this work with a forensically aggressive sample may be beneficial. Another logical next step for future work would be to include a neutral-neutral control condition in which probes consistently replace the neutral word. Currently reaction times and evoked amplitude in response to probes presented for 500ms may reflect attentional vigilance, avoidance or both (Cooper & Langton, 2006). Using a neutral-neutral trial type for which to compare the three experimental conditions would make it possible to draw more accurate conclusions regarding each of these mechanisms (Koster et al., 2004). It would be expected that there would be no behavioural attention bias and no differences in patterns of ERP activity in the control condition. Using a neutral-neutral control condition would provide a ‘baseline’ for which to compare the evoked ERP amplitude to angry and happy targets. This may help to better distinguish the differences in amplitude for each target and therefore make more informed conclusions regarding facilitation and disengagement processes in aggression.

4.5.4 Contributions

This research has made a considerable contribution to the literature as it identifies ERP patterns for selective attentional processes in response to different valenced words. To my knowledge this is one of the very few studies to investigate selective attentional processes to angry, happy and neutral words using EEG methodology. There are no clear conclusions to be drawn from the behavioural results, however the ERP results suggest that there are effects of congruency on happy-neutral and angry-happy trials. This study replicates the early effects of congruency found on angry-neutral trials in Study 1. The study also extends previous literature and research questions addressed in Study 1 by investigating the ERP patterns of attentional orienting to happy words. These results revealed that participants show increased positive amplitude to happy words when they are individually paired with a neutral or angry word. It is suggested that disengagement processes involved with attending to two simultaneously presented stimuli may be crucial for understanding processes involved with attention bias (Koster et al., 2004). However it is not clear from the current data whether increased positive amplitude to probes that replace happy words reflects processes involved with facilitation of attention to happy words, or disengagement of attention to neutral or angry words. Therefore some caution is required when interpreting these results.

The ERP results show differences between congruent and incongruent trials across a number of different components. This suggests that ERPs are sensitive to differences in cognitive processing of varying stimuli and that EEG methodology may be beneficial in understanding attention bias at all stages of attentional processing. I conclude this as ERP differences were evident even in the absence of reaction time differences. Future work using ERP methods to complement current behavioural methods is necessary for understanding the complex mechanisms driving attention biases in aggression.

4.5.5 Conclusions

Using complementary behavioural and ERP methodology, the main aims of this study were firstly to test whether the findings from Study 1 would be replicated by exploring selective attention to angry words when they were paired with a neutral word during a dot-probe task, and secondly investigate the attention bias to angry and happy words when paired with a neutral distracter. Finally, the aim was to explore attentional processes involved with selective attention when two emotionally valenced words are presented simultaneously. The findings relating to angry-neutral trials show effects of congruency only before probe presentation at 500ms. The results show increased amplitude on incongruent trials compared to congruent trials. This effect replicates the novel results found in Study 1. These findings are unexpected in terms of attention theory therefore tentative conclusions are drawn based on this evidence. Due to the absense of behavioural differences in reaction times it was not possible to determine how differences in cognitve processes (evoked P300 postivity) may drive difference in selective attention.

Results on happy-neutral and angry-happy trials show an overall main effect such that participants have significantly increased positive amplitude to probes that replace happy stimuli compared to probes that replace angry stimuli or neutral stimuli. Therefore on all trial types, high aggression participants show increased amplitude to the stimuli with increased positive valence (although effects only appeared pre-probe for *angry-neutral* trials). To conclude, this study has used an original design to explore the cognitive mechanisms associated with attention bias to negative and positive words in aggression. Results provide initial evidence of differences in evoked amplitude across stimulus types.

5 Study 3 - Attention bias to angry faces

5.1 Introduction

The empirical chapters so far have investigated attention bias to angry, happy and neutral words in aggression. The next two chapters explore attention bias to different emotional faces in aggression. This chapter reports a study which aimed to identify differences in neural patterns of attention bias to angry faces in high and low physical aggression groups. Although both words and faces have been used across the attention bias literature, there has been very little consideration of the differences between stimulus modalities and the influence this may have on the attentional system. Attention bias to angry faces may be particularly significant as a hostile facial expression could present an immediate and realistic sense of threat (Bradley et al., 1999). Therefore, it is proposed that faces have increased ecological validity compared to words in the context of attention biases in aggression.

There have been a small number of studies that have used a pictorial emotional Stroop task to explore attention bias in aggression. van Honk et al. (2001a) investigated attention bias towards angry faces in a sample of high trait anger participants. Participants were asked to colour name images of both neutral and angry facial expressions under masked and unmasked conditions. In comparison to participants with low trait anger, participants with high trait anger demonstrated delayed colour naming of angry faces compared to neutral faces during both conditions. This suggests that attentional interference due to processing of the angry face resulted in poorer task performance, and that biases in attention are evident even at the preconscious level. A study by Putman et al. (2004) reinforced these findings. Thirty-four healthy participants completed a pictorial emotional Stroop task that included neutral, angry and happy faces under both masked and unmasked conditions. Results showed that attentional interference resulted in longer latencies when colour naming the threatening faces compared to neutral or happy faces, only under non-conscious masked conditions. This study shows that Stroop performance is potentially affected by conscious control of cognitive-emotional processes. This suggests that attention bias is a relatively automatic cognitive process which can be influenced by conscious attentional control. This is in line with cognitive theories of aggression (e.g., Crick & Dodge, 1994; Wilkowski & Robinson, 2010) which suggest that hostility-related selective attention is characterized by a combination of increased stimulus-driven attentional capture by angry cues and poor effortful regulatory control (e.g., Strack & Deutsch, 2004; Wilkowski & Robinson, 2010).

Although studies by van Honk et al. (2001a) and Putman et al. (2004) show support for hostility-related attention bias in aggression, they have utilised the Stroop task which has been subject to a number of criticisms. Firstly, it is suggested that increased delay in colour naming may reflect biases in response generation and not biases in attention (Mogg et al., 2000); and secondly it is not a true measure of selective attentional processes (Bishop 2008; Fox, 1993; MacLeod et al., 1986). However, the dot-probe task (MacLeod et al., 1986; MacLeod et al., 2002) was designed to assess the relative allocation of attention to simultaneously presented aversive and neutral stimuli. Evidence suggests that during a dot-probe task, violent offenders respond significantly more rapidly to probes replacing violent versus neutral words in comparison with a undergraduate control group (Smith & Waterman, 2003).

Maoz et al. (2017) conducted one of the few studies that investigated attention bias towards angry faces in a sample of high trait anger participants using the dot-probe task. When presented with angry-neutral face pairs, participants with high trait anger had faster reaction times to probes that replaced angry faces, compared to probes that replaced neutral faces. Maoz et al. (2017) suggest that negatively biased attention patterns facilitate increased processing of hostile stimuli which in turn amplifies anger. Consistent with Maoz et al. (2017), Ciucci et al. (2018) found a robust association between aggressive behaviour and attention bias for angry faces. Children aged between 11 and 15 completed a dot-probe task in which angry faces (threat), sad and fearful faces (negative but not threat), and happy faces (positive) were each presented alongside a neutral face. The results showed that children nominated as more aggressive by their peers showed increased attentional orienting to angry faces, reflected in quicker reaction times to probes that replace angry faces compared to neutral faces. There were no effects of attentional orienting to happy, sad or fearful faces which suggests that attentional facilitation in aggression is unique for angry faces.

Although the dot-probe has been infrequently used when studying attention biases in aggression, there have been a few other studies which have used this methodology with a different sample. Evidence from a recent systematic review suggests that during a dot-probe task in which negative and neutral faces were presented, socially anxious participants respond faster on negative-congruent trials compared to incongruent trials (Bantin et al., 2016). These findings are consistent with a study by Salum et al. (2013), which explored attention bias to threat faces during a dot-probe task in disordered children. Children with no psychiatric disorder, and children with a form of distress disorder, such as depression, showed increased attention bias for angry faces. However, children with fear-related disorders showed an attention bias away from threat. Surprisingly there were no significant effects of attention bias in the behavioural-disorder group. These results are in contrast to Ciucci et al. (2018) which found an association between disordered behaviour and attention bias for angry faces in children. However the results suggest that attention bias may contribute to separate psychiatric disorders differently.

The attention bias literature has used a number of different methodological paradigms across varying samples, however, little is understood about the cognitive processes that contribute to such biases. A number of studies have used simple face presentation tasks with the aim of understanding the differences in electrophysiological responses to different facial expressions. For example, results suggest that attention bias for angry faces may influence early (N170) and later

(LPP) stages of processing. Leppänen et al. (2007) found that during a single face presentation, fearful faces evoked an increased N170 compared to neutral and happy faces in a normative population. Additionally, Schupp et al. (2004b) investigated the neural processing of facial expressions in a healthy undergraduate sample. Participants were required to view angry, happy and neutral faces while EEG was recorded. Individuals had increased late positive potential (LPP) to threat faces compared to both friendly and neutral faces. Bertsch et al. (2009) also explored how a sample of healthy participants processed different facial expressions during an emotional Stroop task following a provocation. The behavioural data suggested that following provocation, participants had delayed colour naming of all emotional faces compared to neutral. The ERP results showed that P2 amplitude was greatest for fearful and angry facial expressions. This finding is consistent with work which shows increased amplitude to angry faces in normative samples and suggests that this effect may be particularly salient following a provocation.

Further studies have used the dot-probe task with simultaneous EEG recording to explore the electrophysiological processes associated with selective attention. Some of these studies are outlined in a recent meta-analysis conducted by (Torrence & Troup, 2018). For example, Santesso et al. (2008) used a dot-probe task, which included angry-neutral face pairs, to investigate neural correlates of involuntary orienting to angry faces in a healthy adult sample. Face pairs were presented for 100ms only in order to investigate involuntary orientating (this is likely not enough time for participants to shift gaze between the two simultaneously presented stimuli; Cooper & Langton, 2006). Behavioural results showed that participants were faster to respond to probes that appeared in place of angry faces compared to neutral. The EEG analysis revealed that evoked P1 amplitude was significantly larger when participants responded to the probe that appeared in place of the angry face compared to when it appeared in place of the neutral face. Santesso et al. (2008) suggest that healthy individuals orient attention towards threatening facial expressions. The authors concluded that P1 is the earliest electrophysiological index of spatial attention and that threat cues can modulate these attentional processes.

Previous findings reveal that normative samples generally show increased positive amplitude to angry faces compared to neutral; however, it is yet unknown how ERP patterns may differ between aggression groups. Helfritz-Sinville and Stanford (2015) conducted one of the very few studies to investigate patterns of P300 amplitude in response to aggression-related and neutral words in aggressive populations. Participants were required to complete a modified oddball task, with simultaneous EEG recording, to investigate the attentional processing of social and physical threat words. Helfritz-Sinville and Stanford (2015) found that non aggressive individuals showed relatively similar processing of threat and neutral words, whereas non aggressive individuals showed increased P300 amplitude to both threat words when compared to neutral words. The current study aimed to expand on this work by exploring whether this effect would be consistent using angry faces instead of words. The dot-probe task was used to explore the cognitive processes associated with attending to two simultaneously presented stimulus. Due to there being no clear distinction between stimulus modalities in the previous literature, and the difficulty in comparing modalities across different tasks and samples, the predictions were similar to those made in Study 1 regarding attention bias to angry words.

5.2 Aims and rationale

Collectively, past research suggests a behavioural bias towards angry faces in line with current cognitive models of aggression (Wilkowski & Robinson, 2010). There is evidence to suggest that during an emotional Stroop task, high anger (van Honk et al., 2001a) and healthy (Putman et al., 2004) samples show greater interference when colour naming angry faces. However, the evidence suggesting biased attention to angry faces during a selective attention task, such as the dot-probe is limited (Maoz et al., 2017). I have drawn from the general attention bias literature using healthy and anxious samples and multiple paradigms to make predictions about aggressive populations. Therefore, the first aim of the current study was to test whether previous findings would be replicated by examining whether non-clinical individuals with high trait physical aggression display a visual attention bias towards angry faces using the dot-probe task. By comparing the results of this study to those found in Study 1 it is possible to make comparisons between stimulus modalities. It was predicted that reaction times on the dot-probe task will yield similar effects for both words and faces.

The study also aimed to determine the neural characteristics of attention biases to angry faces by examining ERP correlates of this bias. The P1 (Santesso et al., 2008), P2 (Bertsch et al., 2009), and LPP (Schupp et al., 2004b) have been shown to be increased in response to angry faces across healthy populations during single presentation tasks. Therefore, the aim was to explore if the ERP pattern in response to angry-neutral face pairs will differ across low and high aggression groups. Additionally, the P300 ERP component has been shown to be similar across stimulus types in aggression-prone individuals when presented with hostility-related and neutral words (Helfritz-Sinville & Stanford, 2015). However, to my knowledge no studies have explicitly investigated ERP correlates of attention bias in aggression using the dot-probe paradigm in which emotional faces are presented. Therefore, the current study investigated neural processing relating to attention bias for angry faces, specifically in physical aggression. Based on previous evidence that suggests increased P1 (Santesso et al., 2008) and P300 (Helfritz-Sinville & Stanford, 2015) in response to angry faces and threat words respectively, it was predicted that the low aggression group would show increased P100 and P300 amplitude in response to probes that replaced angry faces, compared to probes that replaced neutral faces. It was suggested that high aggression participants would show less differentiation between stimulus types.

5.2.1 Research questions and hypotheses

Overarching research question: Do high aggression participants have an increased attention bias to angry faces compared with low aggression participants, and is this reflected in different ERP patterns in response to angry and neutral stimuli between aggression groups?

Hypothesis one: Relative to participants with low levels of physical aggression, participants with increased physical aggression scores will show an increased attention bias to angry faces, characterized by a faster reaction time to congruent trials compared to incongruent trials.

Hypothesis two: Increased self-reported attentional control will be correlated with decreased levels of physical aggression and decreased attention bias to angry faces.

Hypothesis three: Based on results from studies one and two it was predicted that, compared to the low physical aggression participants, the high physical aggression participants will have increased P300 amplitude in response to angry-neutral face pairs.

Hypothesis four: Participants with low levels of physical aggression will show increased P1 and P300 amplitude to congruent trials compared to incongruent trials, whereas participants with high levels of physical aggression will show greater similarity in P1 and P300 amplitude in response to congruent and incongruent trials.

5.3 Methods

The majority of the methods used for this current study are identical to the methods outlined in Study 1; however, the task conducted in this study used a different modality of stimuli (faces instead of words). A full description of the methods are outlined in Study 1 (Chapter 3, Section 3.3).

5.3.1 Power Analysis

The a *priori* power calculation based on the most complex planned analyses for this study can be found in Chapter 3, Section 3.3.1.

5.3.2 Participants and procedures

The sample (see Section 3.3.2) and procedures (see Section 3.3.7) were the same as that recruited for the first study which investigated attention bias to angry words in aggression.

5.3.3 Self-report measures

Self-report measures consisted of The Aggression questionnaire (Buss & Perry, 1992), the Attentional Control Scale (Derryberry & Reed, 2002), The Delinquency Questionnaire (Tarry & Emler, 2007), and the Trait form of the State-Trait Anxiety Inventory (Spielberger & Gorsuch, 1970). These were identical to the questionnaire measures collected for Study 1 which explored attention bias to angry words. See Chapter 3, Section 3.3.3 for further details on each of these.

5.3.4 Attention bias test

The dot-probe task was identical to that used in the first word task (Chapter 3, Section 3.3.4), however word pairs were replaced with face pairs (Appendix S).

5.3.5 Attention bias test stimuli

Stimuli consisted of 12 angry and 12 neutral facial expressions. These were colour images obtained from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015) presented against a white background. The same actor displayed the angry and neutral facial expressions in each pair. To select the stimuli, all 598 faces were downloaded from the database. The 93 items labelled ‘white male’ were reviewed. Of these faces, 12 faces were selected at random for the current study. All faces portrayed mouth-closed expressions. Images were chosen from this database as it provides standardized face stimuli for a number of different expressions. For example, photos were taken in controlled conditions with identical light and exposure. The faces were modified in photoshop to ensure that piercings and facial hair were removed. When selecting stimuli from the database no other variables were controlled for. Individual images were cropped to dimensions of 7.9cm by 11.9cm and resized to 50% of originals in Photoshop, such that each face was just under 4cm by 6cm onscreen.

5.3.6 EEG Acquisition

EEG acquisition was identical to the first word attention bias task (Chapter 3, Section 3.3.6)

5.3.7 Data analysis plan

The analysis plan for both behavioural and EEG data extracted from the *face* task followed the same steps as Study 1 in which attention bias to angry *words* was measured (Chapter 3, Section 3.3.8). ERP measures were evaluated on correct trials of the dot-probe face task only (3372 out of 3456 (97.6%)).

5.4 Results

5.4.1 Data preparation

A full description of data preparation can be found in Chapter 3, Section 3.4.1. The approach to missing items and questionnaire reliability are the same as those outlined in the first attention bias to angry words study (Study 1).

All data (BPAQ and subscales, ACS, and STAI-T) was normally distributed apart from the delinquency questionnaire, which was just outside acceptable limits of ±2 (due to floor effect). To assess the normality of the two reaction time variables (congruent and incongruent trials) extracted from the dot-probe task measuring attention bias to angry faces, skewness and kurtosis scores were divided by their respective standard error scores. These were within acceptable limits of ±2. The calculated bias (angry minus neutral) was also within acceptable limits, therefore data was analysed using parametric tests. Skew and kurtosis calculations can be found in Appendix L.

5.4.2 Descriptive Results

The descriptive results for the aggression data and questionnaire variables are the same as those outlined in Study 1 (Chapter 3, Section 3.4.2).

Exploratory analyses were conducted to assess the relationship between attention bias to faces and anxiety. Similar to the attention bias to words results, there was no significant difference in attention bias for *faces* between high (M = -2.71, SD = 20.29) and low anxiety (M = -6.10, SD = 15.11); $t(27) = 0.512$, $p = .613$, $d = .190$. Attention bias for angry faces did not significantly correlate with anxiety, $r = -.127$, $p = .244$ (one-tailed). Therefore anxiety was not included as a covariate in the subsequent analyses.

5.4.3 Results relating to hypotheses

For consistency across stimulus types only the physical aggression results are reported here. Similar to the attention bias to words task, significant results were more marked when investigating physical aggression. Exploratory analyses were also conducted for total aggression (see Appendix T for the significant main effects and interactions).

5.4.3.1 Hypothesis one

5.4.3.1.1 *Correlations*

Pearson correlations showed that the correlation between physical aggression score and attention bias was significant, r = -.341, p = .014 (one-tailed) (Figure 20). This suggests that participants with higher levels of physical aggression were quicker to respond on congruent trials compared to incongruent trials.

Exploratory analyses with the other aggression subscales showed that verbal aggression; r = -.273, p = .065 (one-tailed), anger; r = -.234, p = .099 (one-tailed), and hostility; r = -.191, p = .148 (one-tailed) did not significantly correlate with attention bias index.

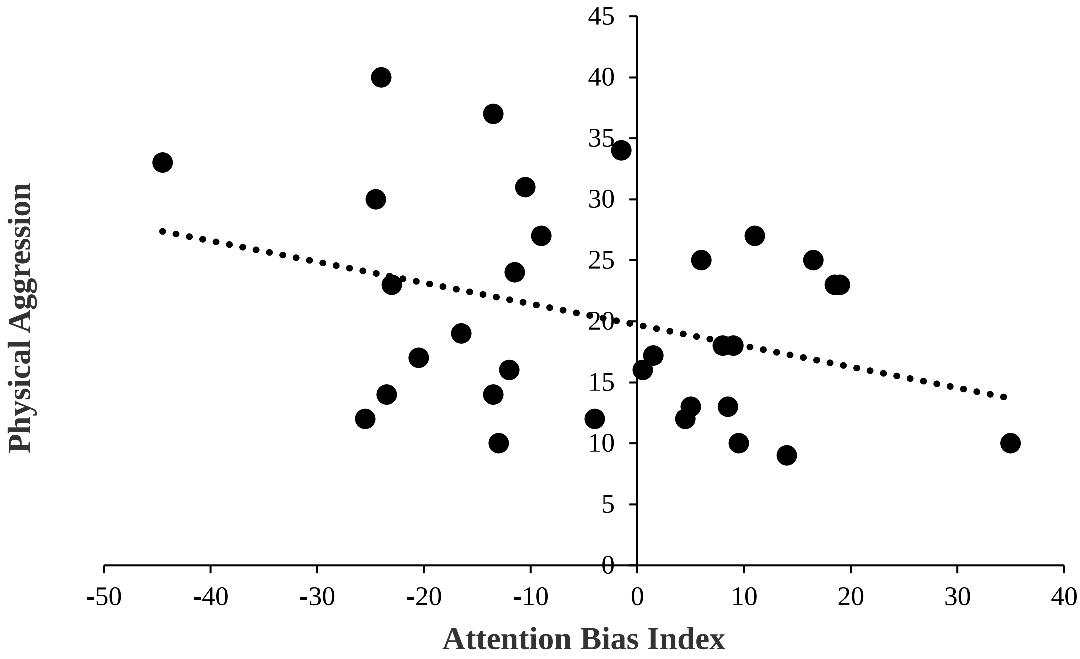

Figure 20: Scatterplot and regression line (r = -.341, p = .014) to show the correlation between physical aggression and attention bias index (n = 32).

5.4.3.1.2 *Median split analysis of group effects*

To further explore the significant correlation between physical aggression and attention bias index, between-subjects analyses were conducted. The ANOVA results showed no significant effects of physical aggression on bias score. The independent samples t-test results also showed no significant differences in bias scores between high and low physical aggression (high physical aggression, *M* = -7.17, *SD* = 18.57; low physical aggression, *M* = -2.23, *SD* = 16.24; *t*(28) = -0.775, *p* = .445, *d* = 0.283).

Table 11: Mean reaction time (ms) to angry and neutral faces in the high and low physical aggression groups (SDs).

	High physical aggression (n = 15)	Low physical aggression (n = 15)	Whole sample (n = 32)	p-value
Congruent trials	580.97 (66.94)	576.30 (56.30)	579.47 (63.11)	.838
Incongruent trials	588.13 (65.29)	578.53 (60.39)	583.34 (64.03)	.679
Bias index	-7.17 (18.57)	-2.23 (16.24)	-3.88 (17.07)	.445
p-value	.157	.603	.208	/

Whilst non-significant, Table 10 shows that means were broadly in the expected direction. Across high and low aggression physical aggression groups, participants were quicker to respond on congruent trials compared to incongruent trials. However, the difference in means is only marginal in the low aggression group, whereas the difference is greater in the high aggression group. The results show tentative support for hypothesis one; the correlation shows a significant association between physical aggression and attention bias to angry faces, however between-subject effects did not reach significance.

5.4.3.2 Hypothesis two

Attentional control did not significantly correlate with aggression; r = -.263, p = .073 (one-tailed) or attention bias for angry *faces*; r = .003, p = .494 (one-tailed). Therefore hypothesis two was not supported and attentional control was not explored as a possible mediator of the relationship between aggression and attention bias to words.

5.4.3.3 Hypothesis three

For each epoch, one way ANOVAs were conducted to investigate the difference in evoked amplitude in response to angry-neutral word onset between low and high aggression groups in electrodes across the region of interest.

5.4.3.3.1 *Pre-probe differences in aggression group.*

Between 100 and 200ms the effect of physical aggression was significant at CP2, $F(1,28) = 4.943$, $p = .034$, $\eta_p^2 = .150$. Between 200 and 300ms the effect of aggression was significant at TP9, $F(1,28) = 4.225$, $p = .049$, $\eta_p^2 = .131$; CP5, $F(1,28) = 4.269$, $p = .048$, $\eta_p^2 = .132$; and CP2, $F(1,28) = 4.515$, $p = .043$, $\eta_p^2 = .139$. Between 300 and 400ms there was a significant effect of aggression at CP2 only, $F(1,28) = 5.315$, $p = .029$, $\eta_p^2 = .160$. There were no significant effects between 400 and 500ms.

5.4.3.3.2 *Post-probe differences in aggression group.*

Between 600 and 700ms the effect of aggression approached significance at P8, $F(1,28) = 3.356$, $p = .078$, $\eta_p^2 = .107$; and P4, $F(1,28) = 4.142$, $p = .051$, $\eta_p^2 = .129$. Between 700 and 800ms the effect of aggression was significant at P4, $F(1,28) = 5.271$, $p = .029$, $\eta_p^2 = .158$; and approached significance at P8, $F(1,28) = 3.555$, $p = .070$, $\eta_p^2 = .113$. There were no significant effects between 500 and 600ms, 800 and 900ms or 900 and 1000ms.

The waveform (Figure 21) reveals that compared to low aggression participants, high aggression participants have increased P2 and P300 amplitude in response to angry-neutral face pair onset showing support for hypothesis three. At P4 and P8 there seems to be a longer lasting effect which is still evident after probe presentation at 500ms.

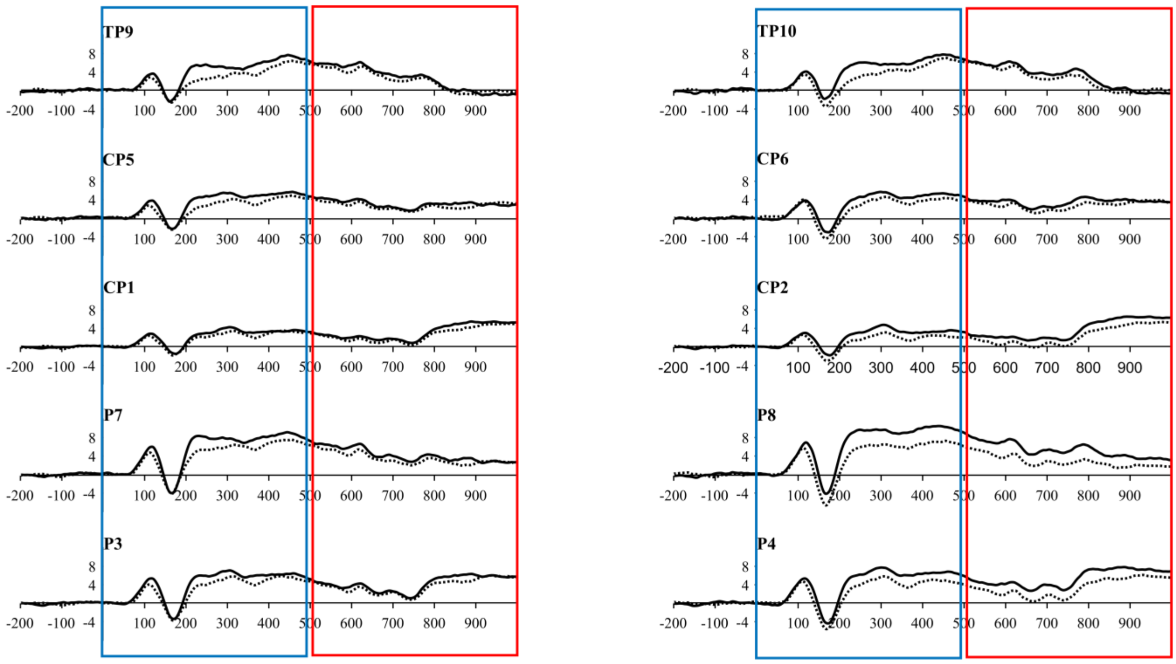

Figure 21: Grand average ERPs for the effect of physical aggression (high vs. low) across all trials. The high physical aggression group (n = 15; black) is compared with the low physical aggression group (n = 15, dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

5.4.3.4 Hypothesis four

A qualitative inspection of the waveform (Figure 22) shows apparent congruency effects from 300ms post face onset. For this reason the length of the whole trial was analysed using a -200 pre-face onset baseline, where 100-500ms refer to pre-probe processes and 500-1000ms refer to post-probe processes. A mixed model ANOVA was conducted to explore the effect of trial congruency within each aggression group.

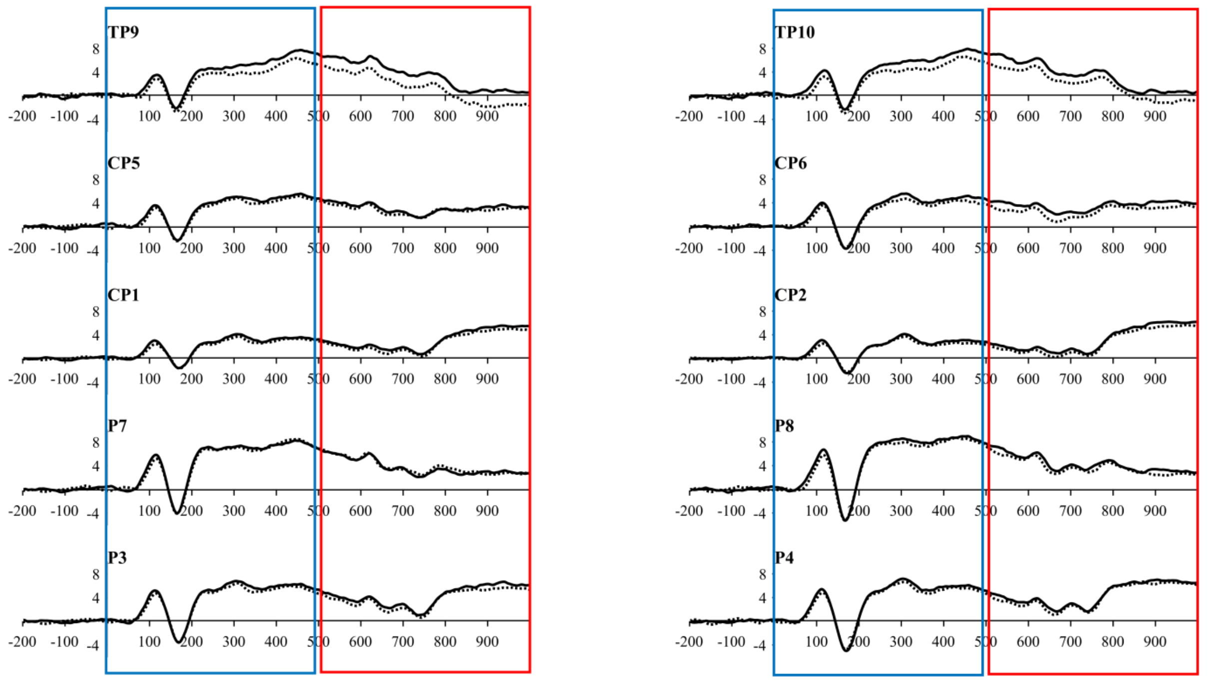

Figure 22: Grand average ERPs (n = 32) for the effect of trial congruency in all participants. Mean amplitude to congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

5.4.3.4.1 *Post-probe differences in congruency.*

The ANOVA results showed the interaction between congruency and electrode was significant between 500 and 600ms, $F(4,112) = 3.495$, $p = .038$, $\eta_p^2 = .111$; and approached significance between 700 and 800ms, $F(4,112) = 3.179$, $p = .058$, $\eta_p^2 = .102$. To explore the trial congruency and electrode interaction and to investigate where effects of trial type were maximal across all recording sites, a further ANOVA was conducted for each of these epochs (500-600 and 700 and 800ms). Each set of electrodes were entered into the ANOVA separately. Between 500ms and 600ms, the effect of trial congruency was significant at TP9, $F(1,31) = 4.622$, $p = .039$, $\eta_p^2 = .130$; and TP10, $F(1,31) = 4.710$, $p = .038$, $\eta_p^2 = .132$. Between 700ms and 800ms the effect of trial congruency was also significant at TP9, $F(1,31) = 4.355$, $p = .045$, $\eta_p^2 = .123$.These results demonstrate that effects of trial congruency were maximal at TP9 and TP10 (Figure 22).

The ANOVA also showed a significant interaction between congruency, hemisphere and physical aggression between 600 and 700ms, $F(1,28) = 5.340$, $p = .028$, $\eta_p^2 = .160$; 800 and 900ms, $F(1,28) = 4.683$, $p = .039$, $\eta_p^2 = .143$; and 900 and 1000ms, $F(1,28) = 5.32$, $p = .029$, ηp2 = .160. This effect also approached significance between 700 and 800ms, $F(1,28) = 3.360$, $p = .077$, $\eta_p^2 = .107$.

Post-hoc tests between 600 and 700ms were conducted to investigate whether the effect of trial congruency is significant in either or both the high and low physical aggression groups. A 5 (electrode) x 2 (hemisphere) x 2 (trial congruency) ANOVA was conducted for each of the aggression groups (high and low). Results showed no effects in the high physical aggression group. In the low physical aggression there was a close to significant effect of congruency, $F(1,14) = 4.541$, $p = .051$, $\eta_p^2 = .245$; and a significant interaction between congruency and hemisphere, $F(1,14) = 5.003$, $p = .042$, $\eta_p^2 = .263$. Follow up tests showed that in the left hemisphere there was a significant interaction between congruency and electrode, $F(4,56) = 3.291$, $p = .037$, $\eta_p^2 = .190$. The effect of congruency was significant at TP9 only, $F(1,14) = 6.149$, $p = .026$, $\eta_p^2 = .305$. In the right hemisphere there was a significant effect of congruency, $F(1,14) = 6.360$, $p = .024$, $\eta_p^2 = .312$. The effect of congruency was significant at P4, $F(1,14) = 4.816$, $p = .046$, $\eta_p^2 = .256$; and approached significance at electrode TP10, $F(1,14) = 4.491$, $p = .052$, $\eta_p^2 = .243$; and CP2, $F(1,14) = 3.604$, $p = .078$, $\eta_p^2 = .205$. The waveforms demonstrate that between 600 and 700ms, the low physical aggression participants show increased amplitude to congruent compared to incongruent trials, whereas the high physical aggression participants show little different in amplitude between congruent and incongruent trials (Figures 23 and 24).

Post-hoc tests between 700 and 800ms showed no significant effects in the high physical aggression group. In the low physical aggression there was a significant interaction between congruency and hemisphere, $F(1,14) = 5.445$, $p = .035$, $\eta_p^2 = .280$. Follow up tests showed that in the left hemisphere there was a significant interaction between congruency and electrode, $F(4,56) = 3.804$, $p = .030$, $\eta_p^2 = .214$. The effect of congruency approached significance at electrode TP9, $F(1,14) = 4.379$, $p = .055$, $\eta_p^2 = .239$.

Post-hoc tests between 800 and 900ms showed there were no significant effects in the high physical aggression group. In the low physical aggression the interaction between congruency and hemisphere approached significance, $F(1,14) = 3.784$, $p = .072$, $\eta_p^2 = .213$. However, follow up tests showed no significant effects of congruency in the left or right hemisphere.

Post-hoc tests between 900 and 1000ms showed no significant effects in the high physical aggression group. In the low physical aggression the effect of congruency, $F(1,14) = 4.411$, $p = .054$, $\eta_p^2 = .240$; the interaction between congruency and electrode, $F(4,56) = 2.959$, $p = .067$, $\eta_p^2 = .174$; and the interaction between congruency and hemisphere, $F(1,14) = 4.031$, $p = .064$, $\eta_p^2 =.224$, all approached significance. Follow up tests showed the effect of congruency was significant at TP9, $F(1,14) = 5.460$, $p = .035$, $\eta_p^2 = .281$; and TP10, $F(1,14) = 5.442$, $p = .035$, $\eta_p^2 = .280$.

These results show support for hypothesis four. As predicted, participants scoring high on physical aggression showed no significant effects of trial congruency, such that they had relatively similar evoked amplitude on congruent and incongruent trials. The results also show that low aggressive participants had increased evoked amplitude in response to congruent trials compared to incongruent trials between 600 and 1000ms, maximal in the right hemisphere, at electrode sites, TP9, TP10, P4 and CP2. The effect of congruency seems to be long-lasting and affecting several ERP components at posterior electrodes. Inspection of the waveform (Figure 23) reveals: a positive peak a little after 600 (likely the P1) followed by a first negative peak (the N1), then a short positive peak (the P2) and another negativity (likely an N2) and finally another positive deflection that is likely to be a P300.

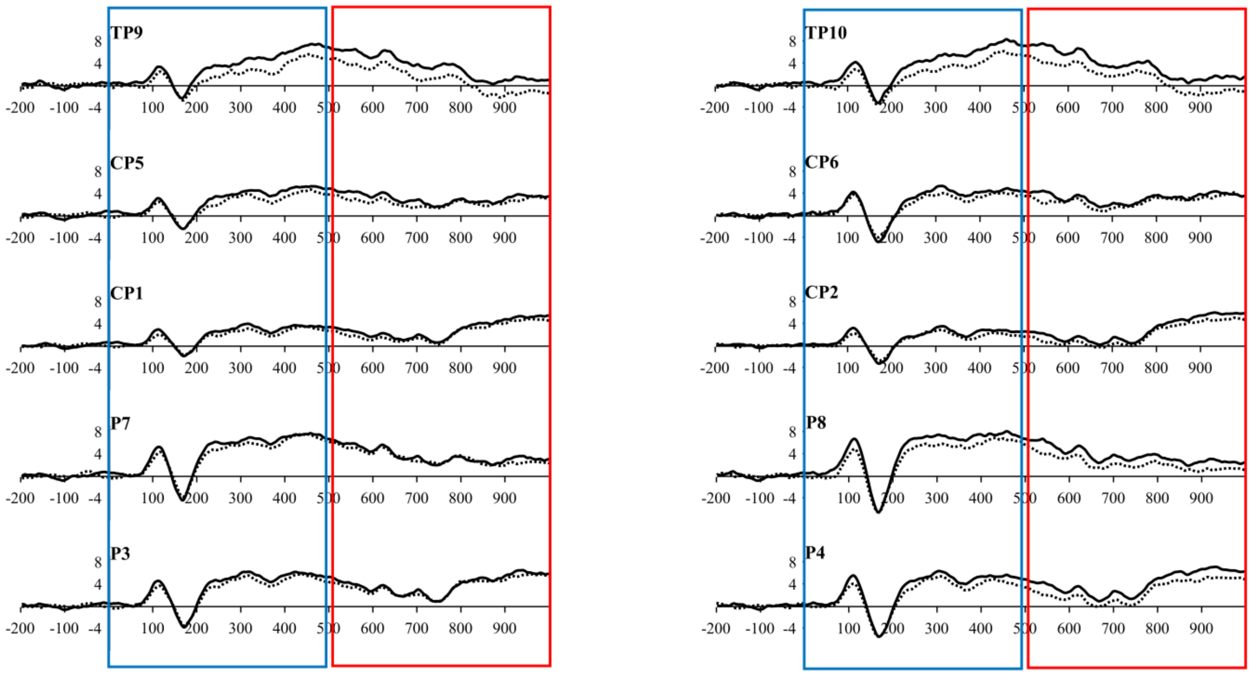

Figure 23: Grand average ERPs (n = 15) for the effect of trial congruency in the low physical aggression group. Mean amplitude to congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

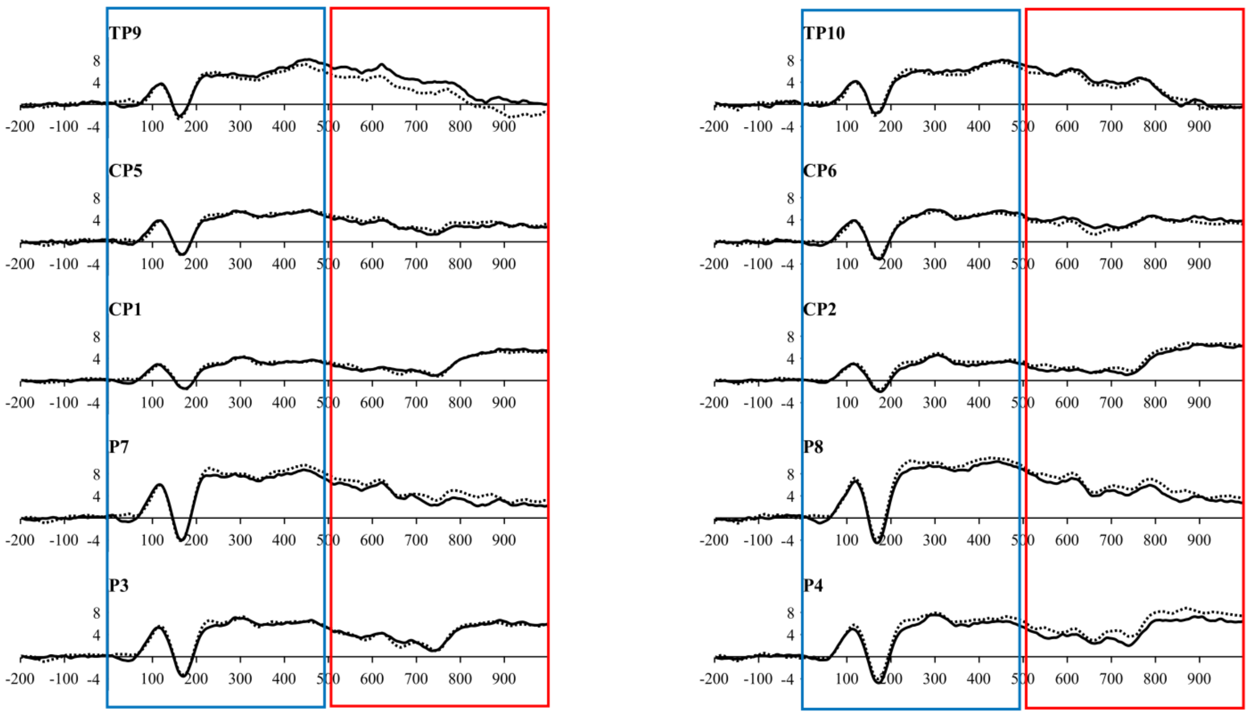

Figure 24: Grand average ERPs (n = 15) for the effect of trial congruency in the high physical aggression group. Mean amplitude to congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

5.4.3.4.2 *Pre-probe differences in congruency.*

Finally, early effects of trial types were also observed before arrow onset (see Figure 23), confirming that a pre-arrow baseline would not have been appropriate (Poulsen et al., 2005; Mingtian et al., 2011). The ANOVA results showed a main effect of congruency between 300 and 400ms, $F(1,28) = 4.482$, $p = .043$, $\eta_p^2 = .138$; and a significant interaction between congruency and electrode, $F(4,112) = 3.503$, $p = .038$, $\eta_p^2 = .111$. Between 300ms and 400ms**,** post-hoc analyses showed a significant effect of trial congruency at electrode sites TP9, $F(1,31) = 6.400$, $p = .017$, $\eta_p^2 =.171$; and TP10, $F(1,31) = 6.606$, $p = .015$, $\eta_p^2 = .176$. The ANOVA also showed a significant interaction between trial congruency and aggression between 300ms and 400ms, $F(1,28) = 4.747$, $p = .038$, $\eta_p^2 = .145$). Post-hoc analyses revealed a significant effect of trial congruency, $F(1,14) = 7.535$, $p =.016$, $\eta_p^2 = .350$; and a significant interaction between trial congruency and electrode, $F(4,15) = 3.323$, $p = .028$, $\eta_p^2 = .192$ in the low physical aggression group. Effect of trial congruency was significant at electrode sites TP9, $F(1,14) = 11.187$, $p = .005$, $\eta_p^2 = .444$; TP10, $F(1,14) = 6.198$, $p = .026$, $\eta_p^2 = .307$; CP5, $F(1,14) = 7.746$, $p = .015$, $\eta_p^2 = .356$; and P8, $F(1,14) = 5.329$, $p = .037$, $\eta_p^2 = .276$, such that amplitude was increased for congruent trials (see Figure 23). There were no significant results in the high physical aggression sample.

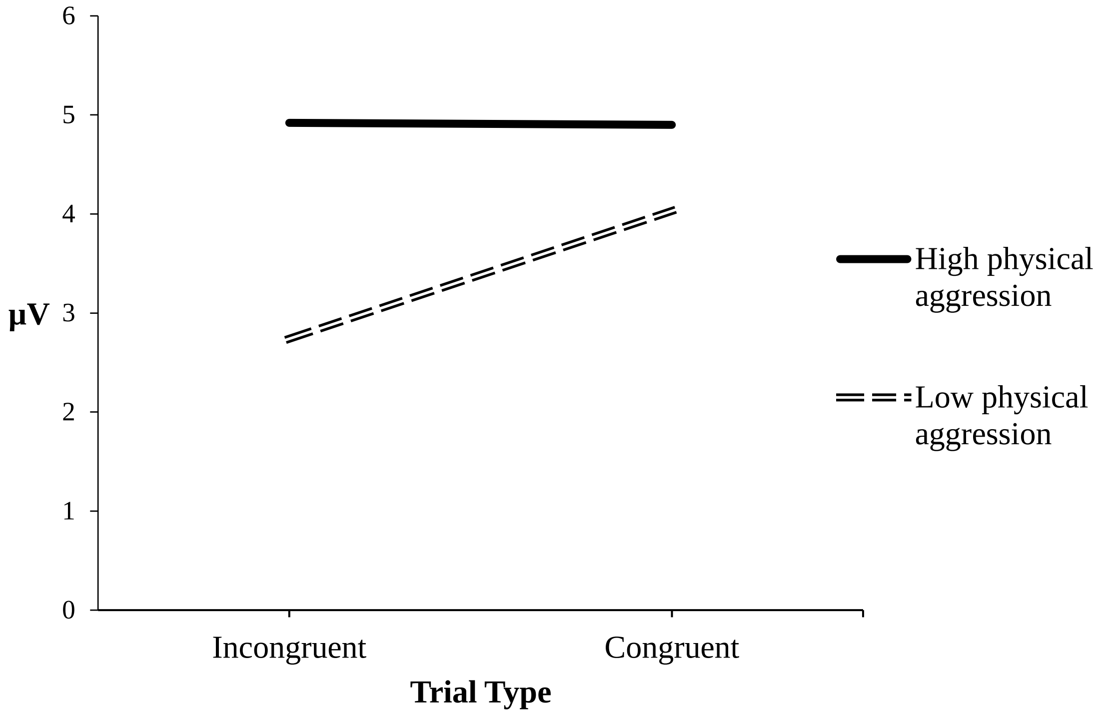

Figure 25: Mean amplitude across all posterior electrodes between 300ms and 400ms on congruent and incongruent trials in the high (n = 15) and low (n = 15) physical aggression groups.

5.4.3.4.3 *Post-hoc tests*

To further explore the interaction between trial type and physical aggression group a number of further tests were investigated to investigate in response to which trial type (congruent or incongruent) and in which physical aggression group (high or low) the differences were evident. For each electrode, for each epoch, and for congruent and incongruent trials, a t-test was conducted to assess the difference in means between the high and low physical aggression groups. The results suggested that across congruent trials, there was no significant difference in aggression groups across any electrode at any epoch. For the incongruent trials there was a number of significant differences between the low and high aggression groups across all epochs, particularly at electrode sites CP2, P4 and P8 which showed significant differences between high and low aggression groups at all epochs excluding 400-500ms (see Appendix U for full results). The

ERP waveforms also show larger differences in P300 amplitude between low and high physical aggression groups on incongruent trials compared with congruent trials (see Figures 26 and 27 below).

This is consistent with correlations which show a significant relationship between physical aggression score and amplitude on incongruent trials at P4 across multiple epochs (see Appendix V). Although this reaches significance this did not survive FDR correction. Physical aggression did not correlate with amplitude on congruent trials.

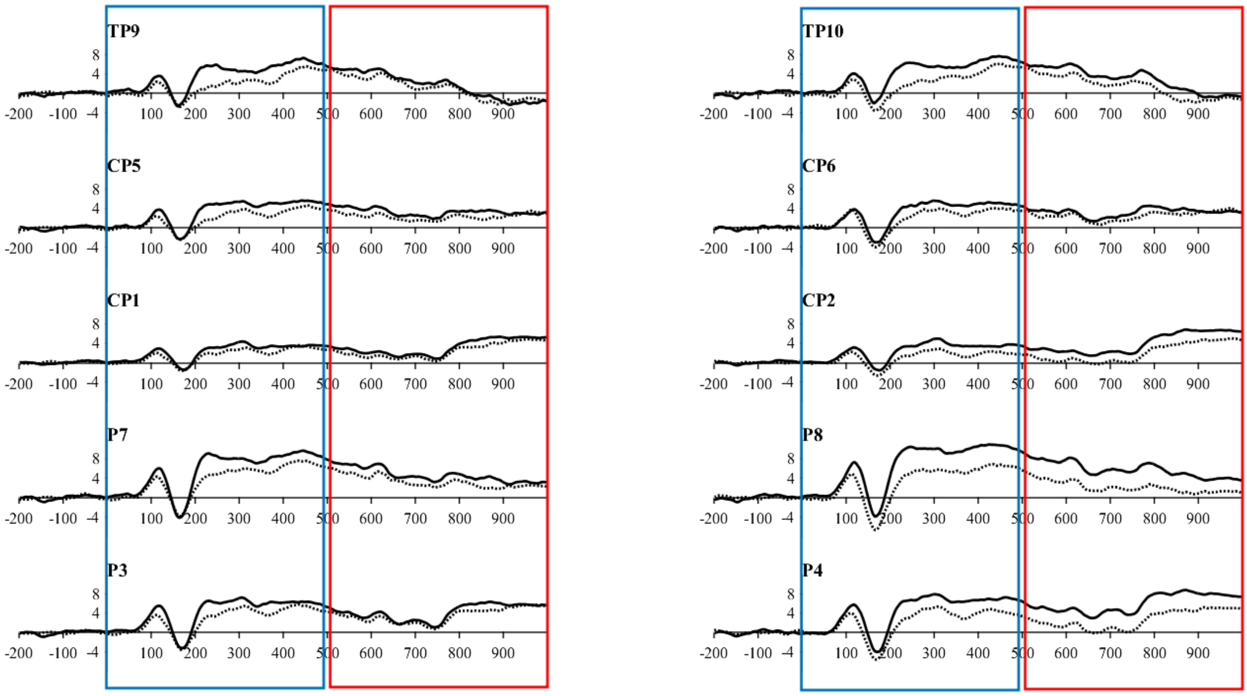

Figure 26: Grand average ERPs for the effect of physical aggression (high vs. low) on incongruent trials only . The high physical aggression group (n = 15; black) is compared with the low physical aggression group (n = 15, dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

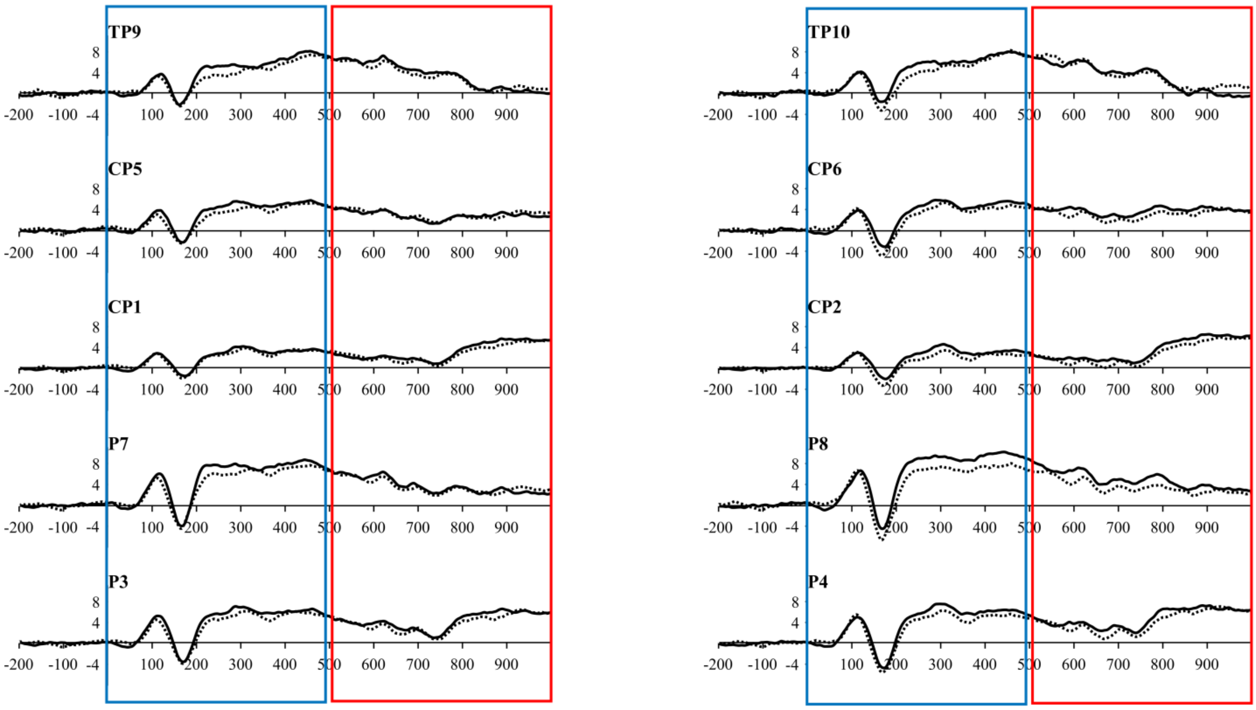

Figure 27: Grand average ERPs for the effect of physical aggression (high vs. low) on congruent trials only . The high physical aggression group (n = 15; black) is compared with the low physical aggression group (n = 15, dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

5.5 Discussion

This study investigated the relationship between physical aggression and selective attentional processes to angry and neutral faces using both reaction time and ERP measurements. The dot-probe paradigm was used to explore attentional selectivity when angry and neutral faces were simultaneously presented. Using simultaneous EEG recording the aim was to investigate how selective attentional processes might contribute to hostile-related attention biases in aggression. Using the same methodology as Study 1 allowed for the comparison with results across studies and explore neural correlates of attention bias for both word and face modalities.

5.5.1 Main findings and interpretations

In line with the first hypothesis, the behavioural results suggest a moderate association between physical aggression and attention bias for angry faces (indicated by shorter reaction time to probes replacing angry faces compared to neutral faces). These findings are consistent with prior dot-probe assessments of attention bias to hostile words in physical aggression (e.g., Smith & Waterman, 2003), and angry faces in high trait anger (Maoz et al., 2017), suggesting that attention bias towards angry stimuli is evident across both word and face modalities. Physical aggression average scores in the present study were comparable with Smith and Waterman (2003). Smith and Waterman (2003) showed that in both an undergraduate and an offender sample, violent individuals display an attention bias towards hostile stimuli. It is proposed that preferential attentional processing of angry faces contributes to the formation of an aggressive response. Aggressive individuals are motivated to confront and remove threat in response to potentially threatening or provoking information in the environment (Smith et al., 1996). The current study provides correlational evidence for the relationship between physical aggression and attention bias for angry faces. However, contrary to expectations, between-subjects effects did not reach significance. It has also been suggested that attentional control may be associated with physical aggression and attention bias to angry faces, however no significant results were found. This suggests that poor attentional control did not contribute to attention biases in participants with increased physical aggression.

Research suggests that between-subject attention bias effects are only evidenced in normative samples differentiated by trait aggression after provocation, where levels of both trait and state anger are high (Cohen et al., 1998; Eckhardt & Cohen, 1997). These studies demonstrated that high trait aggressive individuals only showed an attention bias towards hostile material if the task followed an insult. In the current study, the between-subjects effects may not reach significance due to state anger in the participants being low (only trait anger was measured). However, a more possible explanation is that non-significant between-subjects effect were attributed to the lack of power due to the small sample size.

Hypothesis three explored the differences in amplitude between aggression groups following the face pair presentation across both trial types. In support of the hypothesis, results showed between 200 and 400ms, high physical aggression participants had increased P300 amplitude in response to angry and neutral face pair presentation, compared to low physical aggression participants. This effect was maximal at CP2, but also appeared at TP9 and CP5. This is consistent with findings which suggest that individuals with higher anger scores showed increased P300 in response to negative words, compared to neutral words during an emotional Stroop task (Stewart et al., 2010). In addition, Bertsch et al. (2009) showed that an aggressively provoked group generally had increased P2 and P3 amplitude when responding to all facial expressions during a pictorial emotional Stroop task, compared to the unprovoked group. These findings suggest that increased amplitude to aggression-related stimuli is consistent across word and face stimulus modalities. However these results are in contrast with other work which demonstrates reduced P300 amplitude in aggressive individuals (Barratt et al., 1997; Fanning et al., 2014; Helfritz-Sinville & Stanford, 2015; Surguy & Bond, 2006). The previous studies have utilised different methods such as the oddball and Stroop task and therefore results may not be comparable across studies. To my knowledge the current study is the first to explore attention biases specifically relating to aggression using the dot-probe task. The differences in findings may be attributed to the processes associated with attentional selectivity when attending to two simultaneously presented stimuli.

Hypothesis four concerned the difference in amplitude between congruent and incongruent trials and the interaction between congruency and physical aggression score. It was predicted that attention bias in high physical aggression participants would be characterized by relatively undifferentiated P300 amplitude across both congruent and incongruent trials, whereas participants with low levels of physical aggression would show increased P300 amplitude to congruent compared to incongruent trials. Consistent with Poulsen et al. (2005) data from epochs across the whole length of the trial were analysed time-locked to the face pair onset. Findings are reported for epochs pre- (100-500ms) and post-probe (500-1000ms) separately.

The analysis of post-probe presentation epochs showed that there were no significant effects of trial congruency in the high physical aggression group, such that participants had relatively similar amplitudes in response to congruent and incongruent trials. In the low aggression group, findings showed significant effects of congruency between 600 and 900ms, maximal at TP9, P4 and CP2. These findings suggest that effects of congruency are evident across a number of different ERP components. For example, the effect at 600ms may reflect increased P1 amplitude, whereas the increased amplitude between 800 and 900ms may reflect the P300 component. These findings show support for hypothesis four and are consistent with previous studies that have shown differences in P1 (Santesso et al., 2008) and P300 (Helfritz-Sinville & Stanford, 2015; Thomas et al., 2007) amplitude between negative and neutral stimuli in low aggression samples. For example, Santesso et al. (2008) found enhanced P1 component to cues following angry faces in a non-anxious, non-aggressive undergraduate sample. Schupp et al. (2004b) also investigated the processing of threat, happy and neutral facial expressions in a healthy undergraduate sample not classified by any condition. They showed that individuals had increased late positive potential to threat faces relative to both friendly and neutral faces. This suggests that healthy individuals have increased processing of angry faces.

Overall the results indicated that individuals characterized with high or low levels of physical aggression show different patterns of attention. The current findings are in line with findings obtained by Helfritz-Sinville and Stanford (2015), using a modified oddball task with emotional words. They reported similar P300 amplitude across threat (physical and social) and neutral words in aggressive individuals, whereas control participants exhibited enhanced amplitude to the threat words compared to neutral words. This study has replicated these results using the dot-probe task in which stimuli are presented simultaneously, and using a different modality of aggressive and neutral stimuli (faces instead of words). It includes analyses of a larger number of epochs and electrodes to show the latency of the effect and at which brain region the effect is maximal. Although significant differences between 600 and 900ms were expected, unexpectedly a significant interaction between congruency and physical aggression was also found between 300 and 400ms. Effects of congruency were significant in the low aggression group, but not in the high. This effect appears before the onset of the probe at 500ms and the congruency of the trial has been revealed. The results replicate the pre-probe effects found in studies one and two. These findings could reflect possible priming effects if participants were able to predict where the probe is due to appear, however, this should not be possible due to the counterbalancing and randomisation of trial types. Therefore, a convincing explanation for these early effects cannot be provided and it is suggested that future work using the dot-probe task should analyse the whole trial with the aim of replicating these results and understanding why trial congruency effects are salient pre-probe presentation.

Results indicate that processing of angry and neutral stimuli may occur using a similar contribution of cognitive resources in high physically aggressive participants. It is suggested that aggressive individuals can more readily access aggression-related schemata (belief structures held in long-term memory) (Todorov & Bargh, 2002) and therefore may categorise hostile information with ease. Change in P300 amplitude is thought to reflect processing relating to categorization of stimuli and updating of working memory models based on stimuli that are being attended to (Donchin & Coles, 1988). The low-physically-aggressive individuals may not expect to perceive the angry face and therefore this may trigger a P300 response, whereas the physically aggressive participants are more likely to expect hostile stimuli in their environment and therefore have cognitive models which fit with expectancy outcomes. This may explain why highly aggressive individuals show an attention bias towards hostile stimuli according to reaction time, however efficient systems allow the stimuli to be categorised with little processing. The similarity in neural effort to process both negative and neutral stimuli may affect how individuals perceive, interpret and respond to social cues and may contribute to an aggressive response (General Aggression Model; GAM; Anderson & Bushman, 2002).

Another possible explanation of the findings is that increased aggression is linked to a tendency to perceive hostility in both angry and neutral faces. A recent systematic review by Mellentin et al. (2015) investigated how aggressive individuals perceive facial expressions of different valence. Anger-prone individuals were found to perceive hostility in ambiguous and non-ambiguous non-hostile expressions. This explanation may contribute to the similar processing across faces in the aggressive sample. However, the behavioural results, and previous findings by Smith and Waterman (2003) and Maoz et al. (2017), suggest that aggressive males differentiate to some extent between aggressive and neutral faces, that is, they have visual attentional selectivity for angry faces indicated by quicker reaction time to probes replacing angry faces compared with probes replacing neutral faces. Therefore, it is proposed that other explanations, such as the impaired disengagement hypothesis, should be considered.

Given the results obtained from testing hypotheses three and four, a number of post-hoc tests were conducted to further explore the relationship between aggression and amplitude to congruent and incongruent trials. The aims of these analyses were to further aid the understanding of how simultaneously presented neutral and angry stimuli are processed in relation to aggression. Significant positive correlations revealed that increased physical aggression levels were related to increased P300 amplitude evoked by incongruent trials, while amplitude to congruent trials did not correlate with aggression. The significant correlations were predominantly found at earlier epochs (between 200ms and 300ms, and 300ms and 400ms) and at typical P300 electrode sites, for example P3/P4, CP2 (where the maximal correlation was observed) or CP5. Correlations were also found at more lateral parietal sites such as P7/P8 and temporo-parietal sites such as TP9/10, at which trial-congruency effects (i.e., significant differences between congruent and incongruent trials) were also found. The significant correlations found at these epochs are unexpected as trial congruency effects should only be evident after the presentation of the probe at 500ms. However, significant correlations between amplitude on incongruent trials and physical aggression were also found at expected latencies at P4 (600-900ms) and P8 (800-900ms). An increased P300 response to incongruent trials may explain why individuals with elevated aggression scores showed smaller differences in amplitude between congruent and incongruent trials. The increased P300 amplitude in response to incongruent trials in aggressive individuals may suggest that they are less able to distinguish between angry and neutral faces.

I suggest that relative uniformity in ERP amplitudes across stimulus types in the high aggression group could be attributed to the recruitment of enhanced cognitive processes on neutral trials (reflected in increased P300 amplitude) needed to down regulate the simultaneously presented angry face, consistent with an inhibitory account of P300 (Polich, 2007). On neutral trials it is suggested that participants reporting high levels of physical aggression assign greater cognitive resources (reflected in the increased P300 amplitude) to inhibit the response to the angry face distracter. This pattern of results is in keeping with neurocognitive models of aggression that suggest deficits in regulatory control over incoming perceptual stimuli contribute to visual attention bias (e.g., Wilkowski & Robinson, 2010) in physical aggression, with physically aggressive behaviour being characterized by poor emotion regulation and response inhibition (e.g., Patrick, 2008). These findings suggest that individuals with increased levels of aggression are hypervigilant to threat stimuli in the environment and therefore will attend to these stimuli quicker than other stimuli (reflected in quicker reaction times to probes replacing angry faces compared to neutral faces). Additionally, findings suggest that once engaged with these stimuli individuals with increased aggression find it hard to draw attention away and attend to an alternative stimuli. Therefore, they are required to recruit greater levels of cognitive resources to allocate attention elsewhere (reflected in increased amplitude when probe appeared in place of neutral stimuli in the current task). Theoretically, neural abnormalities in face processing could affect perceptual, cognitive and emotional integration of social cues and contribute to an aggressive response.

Overall, although there are potential explanations for the congruency effects found between 600 and 900ms; the congruency effects between 300 and 400ms are unexpected and currently unexplained. During 300-400ms post face onset, both an angry and neutral face is present on screen, therefore the congruency of the trial is not yet revealed. It is suggested that future work using the dot-probe task will be needed to replicate this work and further understand these effects. However, this work shows that using a pre-probe baseline may not be reliable when investigating attention bias effects using the dot-probe task due to the processing of stimuli between stimuli and probe presentation which potentially invalidates the ERP measurements taken post-probe. This is in line with Poulsen et al. (2005) who claim that using a pre-probe baseline creates a mid-trial change in baseline (between cue and probe presentation) which could introduce post-probe condition differences. To avoid this possible artefact, it is suggested that a pre-cue baseline is used to reference the whole trial epoch.

5.5.2 Limitations and future work

Behavioural analyses revealed only correlational evidence to support hypothesis one. These findings may be explained by the lack of statistical power in analyses using between-subjects designs based on dichotomisation of a continuous variable (Gignac & Szodorai, 2016). Future work including more extreme groups (e.g., very aggressive versus not at all aggressive) would be expected to yield bigger effects sizes. However, ERP analyses showed significant between-subject effects based on a median split of physical aggression score, and correlational evidence to support the current hypotheses. Key findings which show differences between aggression groups and between congruent and incongruent trials, show moderate effect sizes, suggesting that valence-driven attention biases are evident in more aggressive individuals within a normative sample.

Within aggressive samples, neutral facial expressions can be perceived as hostile (Mellentin et al., 2015) and therefore neutral faces may not be an optimal control stimulus for assessing attention bias in aggression. The results reported here suggest that an attention bias for angry faces is characterized by relatively undifferentiated ERPs. However, it is not clear whether these findings reflect a general negative attention bias or whether this ERP pattern is distinct for attention bias to aggressive stimuli. It is also important to consider whether the results reflect an attention bias to angry faces or a more general emotional bias. Anger and emotionality may be confounded in this study. Therefore, future research could explore selective attentional processes involved when attending to angry faces paired with other emotional stimulus types, for example, happy, sad, or frightened faces. This would enable researchers to investigate the specificity of attention bias in aggression.

Attention is a cognitive process which interacts with a number of other processes. In particular, White et al. (2011) highlight the need to investigate attention bias along with interpretation bias. They found that preferential allocation of attention had effects on how ambiguous information was interpreted. Bowler et al. (2017) also used cognitive bias modification techniques in anxious individuals to investigate whether implementing positive interpretation or attention training also had positive effects on the untrained cognitive domain. They found that attention bias training resulted in a reduced threat-related attention bias and an increase in positive interpretation bias. These results demonstrate the need for further work investigating the cognitive mechanisms which underlie both attention and interpretation processes. Particularly within the aggression literature, it has been evidenced that high aggression individuals show a hostile attribution bias, but very little is known about neural processes relating to hostile interpretation of stimuli.

5.5.3 Contributions

This study addresses some of the methodological issues with using the dot-probe task with simultaneous EEG assessment. It has demonstrated that there are a number of ways to analyse and interpret the data and these can have important implications for the conclusions drawn. By presenting whole epoch data time-locked to the face pair onset, this study demonstrates the need for clarity in future dot-probe work. It also suggests that using a pre-probe presentation baseline may be unreliable due to early effects of attention which may be present between face and probe presentation. A large number of epochs and electrodes are presented to gain a wider picture of the attentional processes associated with attending to angry faces across the whole length of the trial, rather than a small 100ms epoch selected from the middle of a trial. This also allows for more specific conclusions regarding the latency and location of these effects to be drawn.

The current findings have important therapeutic implications. For example, attention bias is considered a valid therapeutic target across a range of disorders including aggression (Brugman et al., 2016) and anxiety (e.g., Bar-Haim, 2010), and can be targeted using explicit (cognitive behavioural therapy; e.g., Dehghani, Sharpe, & Nicholas, 2003; Mogg, Bradley, Millar, & White, 1995) and implicit (cognitive bias modification (CBM); Bar-Haim, 2010 for review) techniques. Attention bias modification (ABM), which uses computer-based techniques to implicitly modify threat-related attention bias, has yielded highly successful results in clinically anxious populations (Bar-Haim, 2010; Mogoaşe, David, & Koster, 2014). Present ERP results suggest that, not only is EEG an effective method in measuring attentional processes in physical aggression and therefore could be used alongside current CBM techniques, but also suggest that P300 could be an index used to measure the success of interventions.

5.5.4 Conclusions

Overall, findings indicated that high physical aggression participants show enhanced P300 amplitude in response to angry-neutral face pair presentation, compared with low physical aggression participants. Secondly, individuals with high physical aggression scores show faster reaction times to congruent trials (compared with incongruent trials) but undifferentiated P300 amplitudes across trial types. In contrast, individuals with low aggression scores exhibit increased amplitude to congruent trials compared with incongruent trials. The similarity of P300 amplitude across congruent and incongruent trials and increased susceptibility to selective visual processing in high physical aggression individuals suggests processing abnormalities in valence-driven attentional selectivity among this population. Physical aggression was correlated with amplitude on incongruent trials only; this suggests that differences in attentional processing between the two samples during this task resulted from differing patterns of neural activity in response to incongruent trials. I suggest that individuals with high physical aggression may recruit greater cognitive resources in inhibiting the response to angry face distracters on incongruent trials. The results predominantly support predictions, however a moderate to strong effect of congruency at earlier latencies was also found, which was not predicted. Therefore, conclusions are made tentatively and I acknowledge replication is required.

Relatively little is known about processing biases in aggressive individuals. To my knowledge this work is the first study to investigate selective attentional processes to angry faces in aggression using both behavioural and ERP methodology. Findings shed new light on the cognitive foundations of aggression, and could inform the development of novel therapeutic strategies for modifying visual attention bias in physical aggression.

6 Study 4 - Attention bias to angry and happy faces

6.1 Introduction

Study 3 investigated the attentional processes involved with attending to angry and neutral faces in physical aggression. Results revealed that high and low physical aggression groups have different ERP patterns in response to angry and neutral trials. Based on the results, and a possible limitation that anger and emotionality may have been confounded, a further study was designed which aimed to confirm conclusions drawn from Study 3 and contribute to the understanding of the attentional processing of different facial expressions in aggression.

When assessing aggression-related attention bias, the majority of studies have compared attentional processing of angry faces with neutral faces. Neutral faces are used as a baseline with which to compare reaction time to hostile faces. However, models of attention (Cisler & Koster, 2010; Posner, 1980; Posner & Petersen, 1990; Wilkowski & Robinson, 2010) suggest that processing of non-targets (presented in dot-probe or visual search tasks) have an important role in attentional processes and permit the measurement of selective attention. Attention bias is usually the result of a unique combination of facilitated engagement, difficulty in disengagement and attentional avoidance (e.g. Cisler & Koster, 2010; Koster et al., 2006). Difficulty in disengagement refers to the inferior ability to draw attention away from aggression-related stimuli once it has been engaged. Consistent with the theory that a disengagement processes contribute to attention biases during selective attention tasks, findings from Study 3 suggest that high aggressive individuals showed increased processing on incongruent trials compared to individuals with low aggression. I theorised that this is attributed to impaired disengagement and subsequent processing of the simultaneously presented angry face (non-target distracter). Therefore, the aim of the current study was to explore the cognitive processes involved with selectively attending to differently valenced emotional faces, and consider the role of a valenced distracter in hostility-related attention bias.

The evidence provided in the previous chapter suggests that attention bias to angry faces has been evident in anxious (Bar-Haim et al., 2007), healthy (Santesso et al., 2008) and aggressive populations (Maoz et al., 2017). However, much less is known about cognitive processes associated with attention bias to happy faces. Ciucci et al. (2018) has conducted one of the few studies to explore attentional orienting to emotional stimuli in aggressive children. They used a dot-probe task to measure attention bias to angry, sad, and happy faces, each paired with a neutral face. Aggressive behaviour was measured by asking classmates to report on their perceptions of peers aggression. Results showed that participants rated as more aggressive by their peers showed increased reaction times to probes replacing angry faces, compared to neutral. There was no significant difference in attentional orienting to happy versus neutral faces. Bantin et al. (2016) and Salum et al. (2013) also found that during a dot-probe task there were significant attention bias effects for angry faces on angry-neutral trials, but no significant differences on happy-neutral trials. These results suggest that angry faces may have a specific influence on the attentional system, and that attention bias effects are not due to emotionality of the facial stimuli. In contrast to these findings, there is further evidence to suggest that healthy controls show an attention bias to both angry and happy facial expressions if presented alongside neutral stimuli (Bradley et al., 1997; Pishyar et al., 2004; Waters et al., 2010). For example, Pishyar et al. (2004) found that participants with low levels of anxiety preferentially attended towards happy faces (compared to neutral faces) and away from threatening faces (compared to neutral faces).

Further contradictory evidence comes from Santesso et al. (2008); during a dot-probe task in which happy-neutral face pairs were presented, healthy participants taken from an undergraduate sample, had speedier reaction times in response to probes that appeared in place of neutral faces compared to happy. The authors propose that individuals attend to the most threatening facial expression within each pair. This is consistent with earlier work which showed that at short latencies (100ms) participants initially attend to the relatively threatening face of each pair (the angry face on angry-neutral trials and the neutral face on happy-neutral trials), and then late shift attention to the opposing face (Cooper & Langton, 2006).

There is mixed evidence of attention bias to happy faces when presented with a neutral face. However, these studies have used different paradigms and recruited healthy, anxious and aggressive samples; it is therefore currently unknown if attention biases for happy faces are evident in aggression populations. Using a dot-probe task in which angry-neutral and happy-neutral word pairs were presented the aim was to better understand whether aggressive individual attend to angry faces only, or both emotional faces. On angry-neutral trials it was predicted that previous findings of increased reaction time to probes that replace angry faces, compared to probes that replace neutral faces would be replicated. On happy-neutral trials, it was also predicted that consistent with the literature (Bradley et al, 1997; Fox et al., 2002; Pishyar et al., 2004; Waters et al., 2010), participants would have an increased attention bias to happy compared to neutral faces.

In order to explore the role of the distracter stimuli when two emotional stimuli are presented simultaneously an angry-happy trial type was also included. To my knowledge only one study has included an angry versus happy trial type in which attention orienting between such stimuli has been explored (Pineles & Mineka, 2005). Participants were required to complete a dot-probe in which threatening-neutral, happy-neutral and threatening-happy face pairs were presented. The results showed no differences in reaction time between probe positions on any trial type. These results suggest that attention bias to angry faces are not evident when they are presented alongside an equally emotional happy face. However, the null findings across all stimulus pairings suggest that this effect needs replicating before more concrete conclusions can be drawn. This current study therefore used a similar design to investigate the influence of a happy face distracter on the processing of angry faces within a physically aggressive sample. Further research suggests that threatening faces are detected faster amongst crowds of neutral and friendly distracter stimuli, compared to neutral or friendly stimuli in a crowd of angry distracters (Hansen & Hansen, 1988; Öhman et al., 2001). Therefore, it was predicted that aggressive participants would have speedier reaction times to probes that replaced angry faces, compared to probes that replaced neutral or happy faces.

Attention bias studies have used a number of different methodologies to explore the cognitive processes involved with attention biases. More frequently EEG is used to identify the neural correlates of attention bias. Schupp et al. (2004b) investigated processing of threat faces, compared to friendly and neutral faces in a healthy undergraduate sample not classified by any condition. Participants were required to attend to different facial expressions while EEG was recorded. Results showed that individuals had increased late positive potential to threat faces relative to both friendly and neutral faces, suggesting that processing of threat faces may be reflected by a distinct ERP pattern. However, this study does not contribute to the understanding of cognitive processing of threat faces when there is competition for attentional resources (more than one stimulus presented at one time). Bertsch et al. (2009) used an emotional Stroop task to investigate attention bias to different facial expressions in a healthy sample provoked for aggressive behaviour. The results showed that provoked participants had delayed colour naming of all emotional facial expressions compared to neutral expressions. Generally across all trials, provoked participants showed an enhanced P2 and P3 amplitude compared to unprovoked participants. Although amplitude was enhanced across all trials, this effect was most salient for fearful and angry expressions. This suggests that participants experiencing higher levels of state anger, process negative emotions more elaborately.

Further research has used the dot-probe paradigm and EEG methodology to investigate attentional orienting to both happy and angry faces versus neutral faces in healthy (Santesso et al., 2008) and anxious samples (Holmes et al., 2009; Mueller et al., 2009). Mueller et al. (2009) investigated evoked P1 amplitude in response to face pair presentation. They found that participants with increased levels of social anxiety disorder had enhanced P1 potential to the presentation of angry-neutral face pairs compared to happy-neutral face pairs, providing support for an early neural marker for the automatic detection of threat. Further research has investigated the differences in evoked amplitude following probe presentation. Santesso et al. (2008) found that duirng a dot-probe task in which angry-neutral and happy-neutral face pairs were presented, participants had signficantly larger P1 amplitude when responding to probes that appeared in place of the angry face compared probes that appeared in place of the neutral face. On happy-neutral trials ERP analyses revealed no significant differences in P1 amplitude between probes that appeared in place of happy and neutral faces. These findings suggest that P1 amplitude is increased in response to angry faces only. In contrast, Holmes et al. (2009) found that congruent trials on both angry-neutral and happy-neutral trial types evoked an increased late N2pc. Finally, Carretie et al. (2001) reported increased P3 amplitude in response to happy compared to neutral pictures. These results are consistent with models of attention which suggest facilitated orienting towards emotional information

Previous evidence suggests that selective attentional processes may differ between emotional stimuli (angry or happy) and neutral stimuli. During the dot-probe task it can be predicted that when an angry face is simultaneously presented alongside a neutral face, individuals will orient attention towards the angry face, and that angry faces evoke increased amplitude across a number of different components. However, the evidence regarding attention to happy versus neutral faces is less robust. Holmes et al. (2009) reported increased late N2pc on happy congruent trials and Carretie et al. (2001) reported increased P3 in response to happy versus neutral faces; whereas Santesso et al. (2008) and Schupp et al. (2004b) report no differences in amplitude between happy and neutral faces across P1 and LPP components. These studies tested healthy or anxious samples and therefore predictions regarding aggressive populations are made tentatively. However, findings suggest that attention biases for angry and happy faces may influence early (P1, P2) and later (P3, LPP) stages of attentional processing. To my knowledge no studies have used the dot-probe paradigm with simultaneous EEG recording to explore attention biases to angry and happy faces in aggression.

6.2 Aims and rationale

The literature suggests that aggressive participants show an attention bias towards angry faces compared with neutral stimuli (Putman et al., 2004), however little is understood about seletcive attentional processes in aggression. In Study 3 a dot-probe task with simultanous EEG recording was used to explore attention bias to angry faces in aggressive populations. As well as a significant behavioural attention bias effect, the ERP patterns in response to angry and neutral trials differed across aggression groups. Low aggression participants had increased evoked amplitude in response to probes that replace angry words compared to probes that replace neutral words, whereas high aggression participants had relatively undifferentiated ERPs. The primary aim of Study 4 was to replicate the findings found in Study 3.

By including two other trial types, happy-neutral, and angry-happy, the aim was to explore attention bias to different emotional faces. Firstly, a happy-neutral trial type was included to investigate attention bias to positive facial expressions in aggression. Previous literature suggests there is mixed evidence of attention bias to happy faces, however attentional orienting in aggressive populations using a dot-probe and EEG recording has yet to be studied. Secondly, an angry-happy trial type was included to explore the role of positive distracter stimuli in hostility-related attention bias. Study 3 results indicated the importance of the distracter stimuli during selective attention tasks as these can influence disengagement processes which contribute to attention bias.

EEG methodology has not yet been applied to study the attentional processes involved with attending to different emotional expressions (happy, angry, sad etc.) in aggression. Understanding how aggressive individuals attend to stimuli within their environment could identify neural markers for aggressive behaviour and subsequently inform interventions. Using an original design, and complementary behavioural and ERP methods, the aim was to contribute to the understanding of selective attentional processes involved with attending to angry and happy faces in a psychically aggressive sample. Due to the evidence showing ERP effects of attention bias to emotional face across a number of different components, a number of predictions relating to early (P1) and later (P3/LPP) attentional stages are made.

6.2.1 Research questions and hypotheses

Overarching research questions:

- Is the attention bias effect in high aggression participants specific to angry stimuli, or do they also show an attention bias to positively-valenced happy stimuli?
- Do high aggression participants show different P1/P3 ERP patterns from low aggression participants when selectively attending to negative and positive emotionally-valenced stimuli?

6.2.1.1 Behavioural

6.2.1.1.1 *Correlational hypotheses*

Hypothesis one: Physical aggression score will be positively correlated with *angry-neutral* bias score such that those with higher physical aggression will have an increased bias towards angry faces.

Hypothesis two: There will be no significant correlation between *happy-neutral* bias score and physical aggression because both low and high aggression participants will respond quicker to probes that replace happy faces.

Hypothesis three: Physical aggression score will be positively correlated with *angry-happy* bias score such that those with higher physical aggression will have an increased bias towards angry faces.

6.2.1.1.2 *Between-subject hypotheses*

Angry-neutral

Hypothesis four; main effect: Participants will be quicker to respond to probes that replace *angry* faces compared to probes that replace *neutral* faces.

Hypothesis five; low physical aggression: Participants will have a significantly faster reaction on trials where the probe replaces *angry* faces compared to trials where the probe replaces *neutral* faces.

Hypothesis six; high physical aggression: Participants will have a significantly faster reaction on trials where the probe replaces *angry* faces compared to trials where the probe replaces *neutral* faces. This difference in reaction time between congruent and incongruent trials will be greater in the high physical aggression group compared to the low physical aggression group.

Happy-neutral

Hypothesis seven; main effect: Participants will be quicker to respond to probes that replace *happy* faces compared to probes that replace *neutral* faces.

Hypothesis eight; low physical aggression: Participants will show a significantly faster reaction time to probes that replace *happy* faces compared to probes that replace *neutral* faces.

Hypothesis nine; high physical aggression: Participants will show a significantly faster reaction time to probes that replace *happy* faces compared to probes that replace *neutral* faces.

Angry-happy

Hypothesis 10; main effect: Participants will be quicker to respond to probes that replace *angry* faces.

Hypothesis 11; low physical aggression: Participants will have a significantly faster reaction on trials where the probe replaces *angry* faces compared to trials where the probe replaces *happy* faces.

Hypothesis 12; high physical aggression: Participants will have a significantly faster reaction on trials where the probe replaces *angry* faces compared to trials where the probe replaces *happy* faces. This effect will be greater in the high physical aggression group compared to the low physical aggression group.

6.2.1.2 ERP

6.2.1.2.1 *Main effect of aggression*

Hypothesis 13: High physical aggression participants will have increased P300 amplitude in response to all trial types (angry-neutral, angry-happy, and happy-neutral) at face pair onset compared to low aggression participants.

6.2.1.2.2 *Main effect of valence*

Hypothesis 14: Angry stimuli will evoke increased amplitude; therefore, angry-happy and angry-neutral trials will evoke increased positive P300 amplitude compared to happy-neutral trials following face pair onset. This effect will be most salient in the high physical aggression group.

6.2.1.2.3 *Effect of congruency*

Angry-neutral

Hypothesis 15: The general task effect across all participants will show increased positive amplitude on congruent trials compared to incongruent trials.

Hypothesis 16: Participants scoring low on physical aggression will show increased positive amplitude on congruent trials compared with incongruent trials.

Hypothesis 17: Participants scoring high on physical aggression will show similar amplitude in response to congruent and incongruent trials. This will be due to increased positive amplitude on incongruent trials due to the allocation of resources when attending to the simultaneously presented angry word.

Happy-neutral

Hypothesis 18: It was predicted that the main task effect will show increased positive amplitude on congruent trials compared to incongruent trials.

Hypothesis 19: Participants scoring low on physical aggression will show increased positive amplitude on congruent trials compared to incongruent trials.
Hypothesis 20: Due to the lack of evidence of attention bias to happy faces in aggression, specific hypotheses were not made for the high aggression group.

Angry-happy

Hypothesis 21: The main task effect will show increased positive amplitude to congruent trials compared to incongruent trials.
Hypothesis 22: Participants scoring low on physical aggression will show increased positive amplitude to congruent trials compared to incongruent trials.
Hypothesis 23: Participants scoring high on physical aggression will show an increased positive amplitude to congruent trials compared to incongruent trials, and this effect will be greater in the high physical aggression group compared to the low physical aggression group.

6.3 Methods

6.3.1 Participants and procedures

The sample and procedures for Study 4 were identical to those outlined for the second word task study (Study 2 - Chapter 4, Section 4.3.2 and 4.3.7).

6.3.2 Self-report measures

Self-report measures consisted of The Aggression questionnaire (BPAQ; (Buss & Perry, 1992), the Attentional Control Scale (ACS; (Derryberry & Reed, 2002), The Delinquency Questionnaire (DQ; (Tarry & Emler, 2007), and the Trait form of the State-Trait Anxiety Inventory (STAI-T; (Spielberger & Gorsuch, 1970). These were identical across all studies and are fully outlined in Study 1 (Chapter 3, Section 3.3.3).

6.3.3 Attention bias test

The experimental task used for Study 4 (faces) was identical to that used for Study 2 (words). However, the experimental task included pictorial stimuli instead of verbal stimuli. An in-depth description of the task is found in Chapter 4, Section 4.3.4.

6.3.4 Attention bias test stimulus

The image task consisted of thirty-two angry, happy and neutral facial expressions (colour images) which were obtained from the Chicago Face Database (Ma et al., 2015). All faces portrayed Caucasian male actors (32 in total; the same actor displayed the angry/happy/neutral facial expression in each pair) against a white background (Appendix S). The 32 actors were selected from the 35 white male individuals who had corresponding angry, happy and neutral facial expressions itemised on the database. Norming data for the Chicago Face database (Ma et al., 2015) provides an average score from 1 (not at all) to 7 (extremely) across a number of different factors based on independent ratings of each actors neutral expression. Only mouth closed expressions were chosen. Individual images were cropped to dimensions of 7.9 by 11.9cm and then and resized to 50% of originals in Photoshop, such that each face was just under 4 by 6cm onscreen.

6.3.5 EEG acquisition

EEG acquisition was identical to the second word task (Study 2 - see Chapter 4, Section 4.3.6)

6.3.6 Data analysis plan

Data extraction and preparation was consistent across data from both word and face tasks (Studies two and four) The details are found in Chapter 4, Section 4.4.1) However, analysis of correct trials on the image task consisted of 98.16% of all trials.

6.4 Results

6.4.1 Data preparation

A full description of data preparation can be found in Chapter 4, Section 4.4.1. The approach to missing items and questionnaire reliability are the same as those outlined in Study 2.

All data (BPAQ and subscales, ACS, and STAI-T) was normally distributed apart from the delinquency questionnaire (due to floor effect). The six reaction time variables extracted from the dot-probe task were also assessed for normality (see Appendix R for skew and kurtosis calculations). Within the reaction time data for the six trial/congruency combinations (angry-neutral, happy-neutral, and angry-happy) there were four extreme outliers (3 standard deviations above the mean). These were replaced with the next highest score plus one. There were some other consistent outliers (2 standard deviations above the mean) which shows that across all trial types some participants were slower to react to the stimuli. These were not removed or adjusted as they were stable across the data and do not affect the calculated bias scores.

The six reaction time variables (AN, NA, AH, HA, HN, NH) were not normally distributed due to skewness and kurtosis calculations (each divided by their subsequent standard error scores) being outside acceptable limits of ±2. Generally, data was skewed towards lower reaction time scores. Therefore analysis of reaction time data utilised non-parametric tests. The calculated bias scores (congruent minus incongruent for angry-neutral, happy-neutral and angry-happy trial types) were normally distributed and therefore parametric tests were used to investigate these variables. Angry-neutral bias ranged in scores from -37.0 to 56.0, angry-happy bias ranged from -39.0 to 33.5, and happy-neutral bias ranged from -39.5 to 57.0.

6.4.2 Descriptive results

The descriptive results for the aggression data and questionnaire variables are the same as those outlined in Study 2 (Chapter 4, Section 4.4.2).

Exploratory analyses were conducted to assess the relationship between attention bias to faces and attentional control/anxiety. Results showed that attentional control (total score and both subscales) and anxiety did not significantly correlate with any of the reaction time measures or attention bias scores on the image dot-probe task ($rs < 0.224$, $ps > .113$). These variables were therefore not possible confounds and were not included as covariates in subsequent analyses.

Results relating to hypotheses

6.4.3 Behavioural data

6.4.3.1 Effect of aggression

Based on the findings from Study 3, data was analysed using the physical aggression subscale. This also retains consistency across all attention bias studies. The high and low groups were categorised based on a median split. Table 11 gives an overview of the means and standard deviations across all trial types in each physical aggression group. Although inspection of the means (Table 11) shows that generally high physical aggression participants are slower to respond to probes across all trial types , these differences did not reach significance ($p > .355$).

Table 12: Mean reaction time (ms) for target stimuli and bias score of each trial type within the total, high and low physical aggression groups (SDs).

Trial type	**AN**			**HN**		
Reaction time variable	**Angry target**	**Neutral target**	**Bias**	**Happy target**	**Neutral target**	**Bias**
Low physical aggression (n = 25)	482.52 (69.56)	482.18 (65.39)	0.34 (17.29)	481.66 (66.83)	482.32 (69.54)	-0.66 (18.55)
High physical aggression (n = 23)	497.8 (70.43)	494.22 (74.70)	3.59 (19.90)	493.93 (70.02)	492.87 (75.82)	1.07 (20.36)
Whole sample (n = 51)	493.51 (70.34)	491.53 (70.36)	1.98 (17.97)	490.04 (67.24)	490.08 (71.58)	-0.04 (18.82)

Trial type	**AH**		
Reaction time variable	**Angry target**	**Happy target**	**Bias**
Low physical aggression (n = 25)	478.70 (62.77)	484.96 (62.85)	-6.26 (15.96)
High physical aggression (n = 23)	497.57 (76.82)	499.76 (74.27)	-2.20 (17.26)
Whole sample (n = 51)	491.80 (70.62)	495.87 (69.23)	-4.07 (16.14)

6.4.3.2 Effect of valence and trial congruency

Correlational results. There were no significant correlations between physical aggression and *angry-neutral* attention bias score (p = .741), *happy-neutral* attention bias score (p = .907) or *angry-happy* attention bias core score (p = .999). These results do not support hypotheses one or two as it was suggested that participants higher on aggression would have an increased bias score characterized by quicker reaction times on angry-congruent trials compared to incongruent trials. However, the results support hypotheses three as it was predicted that *happy-*

neutral bias would not be correlated with aggression, it was suggested that both high and low aggression groups would attend to both stimuli similarly.

Median split analysis of group effects. Between-subject analysis consisted of a 3 (trial type) x (2 trial congruency) x 2 (physical aggression) omnibus ANOVA was conducted to explore the interaction between trial type and trial congruency. The results revealed no significant results. Therefore there was no main effect of trial type or interactions with trial type, suggesting that amplitude was relatively stable across angry-neutral, angry-happy, and happy-neutral trials. In line with each specific hypothesis, further planned analyses were conducted to explore the effect of trial congruency within each trial type. These planned analyses consisted of a 2 (congruency; congruent, incongruent) x 2 (physical aggression; high, low) ANOVA for each trial type.

On *angry-neutral* trials, planned analysis revealed no significant effect of congruency (p = .468). It was predicted that participants would be quicker to respond on congruent trials compared to incongruent trials, however there was no evidence of this and therefore hypothesis four was not supported. The planned analyses also revealed no significant interaction between trial congruency and physical aggression for *angry-neutral* trials (p = .548). This suggests the pattern of results were similar to the general task effect across both high and low aggression groups. It was predicted that participants in both the high and low physical aggression group would have significantly faster reaction on congruent trials compared to incongruent trials and that this difference would be more salient in the high physical aggression group compared to the low physical aggression group. There was no significant difference in reaction times between trial congruency for *angry-neutral* trials in either aggression group. There is therefore no support for hypothesis five or six.

On *happy-neutral* trials, planned analyses showed no significant effect of congruency (p = .943). It was predicted that participants would be quicker to respond to probes that replace *happy* faces compared to probes that replace *neutral* faces, however there was no evidence of this and therefore hypothesis seven was not supported. The mixed model ANOVA also showed no interaction between trial congruency and physical aggression (p = .760). This suggests that effects of congruency across both high and low aggression groups are in line with the main effect. The main effect of congruency revealed no significant differences between reaction time on congruent and incongruent trials; therefore, hypothesis eight and nine are not supported. The bar chart (Figure 28) show that the mean reaction time is relatively consistent across both congruent and incongruent trials for both aggression groups.

On *angry-happy* trials, planned analyses showed the effect of congruency approached significance, $F(1,46) = 3.11$, $p = .063$, $\eta p^2 = .063$. Participants were quicker to respond on congruent trials (probe appears in place of angry face) compared to incongruent trials (probe appears in place of happy face) (Figure 28). This main effect shows support for hypothesis ten. This finding could be explained by individuals orienting rapidly to angry stimuli but could also be attributed to difficulties in disengaging from the simultaneously presented angry face when responding to the probe which replaces the happy face. The planned analyses also revealed no interaction between trial congruency and physical aggression for *angry-happy* trials (p = .401). This suggests that effect of trial congruency in the high and low aggression group is in line with the main effect. The main effect suggests that participants are quicker to respond on congruent trials compared to incongruent trials, therefore there is tentative support for hypothesis 11 and 12. However it was predicted that the effect of congruency would be more salient in the high aggression group; the lack of interaction between congruency shows no evidence of this.

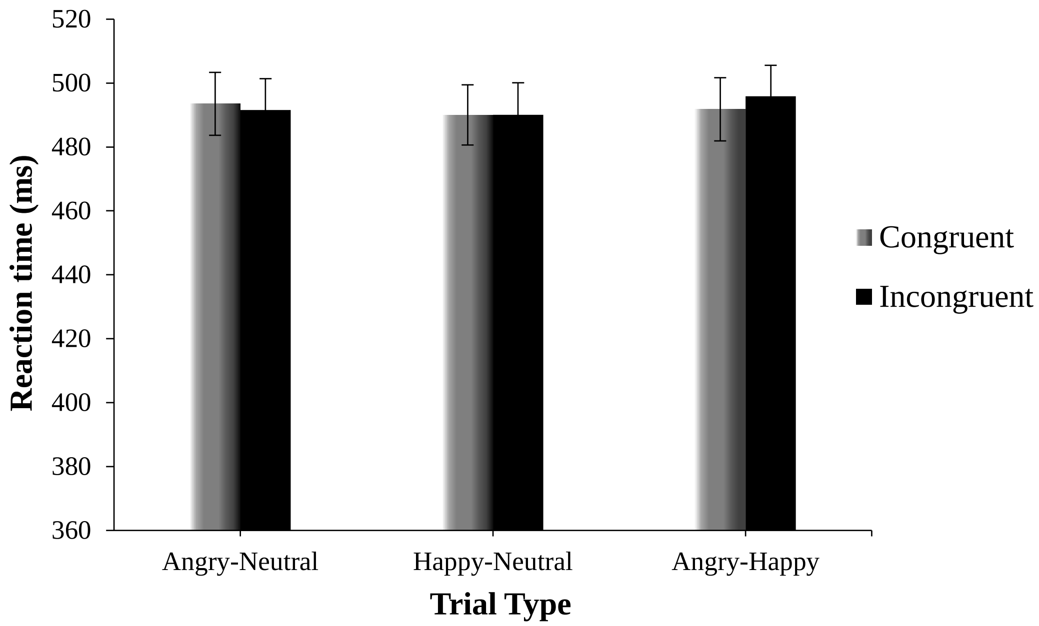

Figure 28: Bar graph to show the reaction time for congruent and incongruent trials for all three trial types (n = 51; error bars = +/- 1 standard error).

6.4.4 ERP data

6.4.4.1 Effect of aggression

The main effect of aggression across angry-neutral, angry-happy and happy-neutral trial types was explored at each electrode for each epoch. On *angry-neutral trials*, the one-way ANOVAs revealed a significant difference in evoked amplitude between aggression groups between 200 and 300ms at CP1, $F(1,46) = 4.236$, $p = .045$, $\eta_p^2 = .084$; and CP2, $F(1,46) = 5.413$, $p = .024$, $\eta_p^2 = .105$. The difference in means also approached significance at P3, $F(1,46) = 3.063$, $p = .087$, $\eta_p^2 = .062$; and P4, $F(1,46) = 3.809$, $p = .057$, $\eta_p^2 = .076$. Inspection of the waveform (Figure 29) shows that low physical aggression show an increased P1 and P2 amplitude in response to angry-neutral face pair presentation, compared to the high physical aggression group.

On *angry-happy* trials there was a close to significant difference between aggression groups at P8 between 800 and 900ms; $F(1,46) = 3.961$, $p = .053$, η_p^2 =

.079. On *happy-neutral* trials the difference in evoked amplitude in response to face pair onset between aggression groups did not reach significance at any epoch. The bar chart (Figure 30) shows that participants in the low physical aggression group have increased positive P2 amplitude in response to angry-neutral face pair presentation, compared to participants with high physical aggression. This shows no support for hypothesis 13 as it was predicted that high aggression would have increased amplitude in response to face-pair presentation.

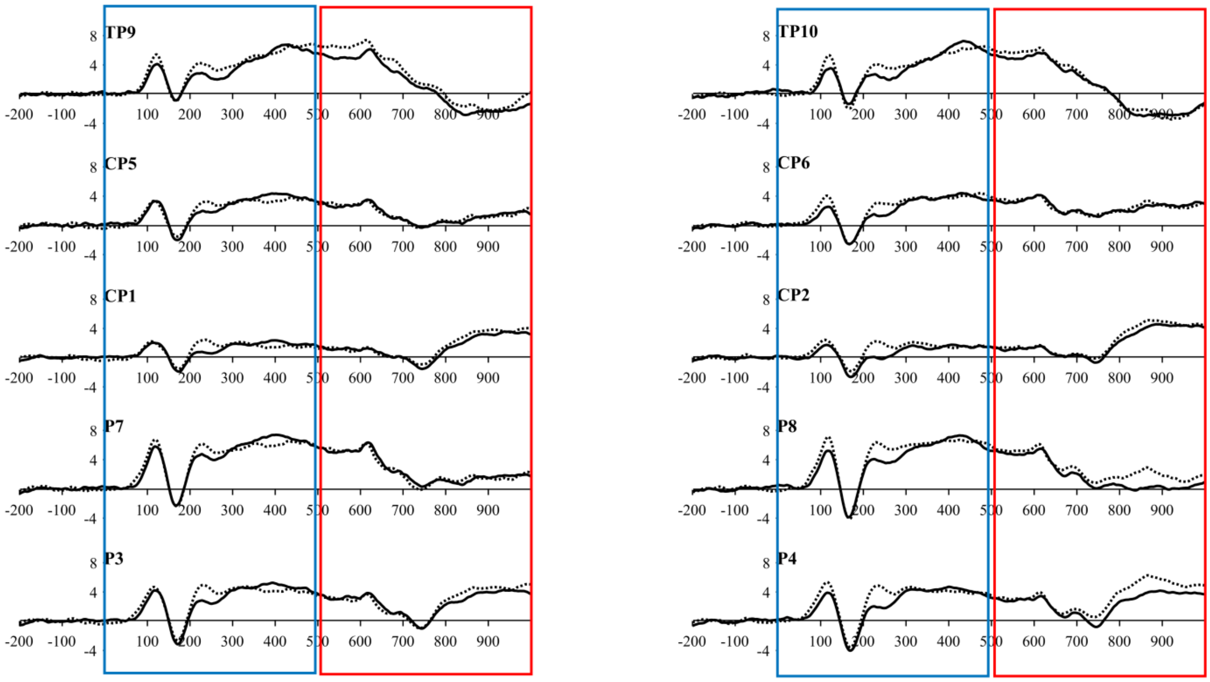

Figure 29: Grand average ERPs for the effect of physical aggression group (high vs. low) on angry-neutral trials. The high physical aggression group (n = 23; black) is compared with the low physical aggression group (n = 25; dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

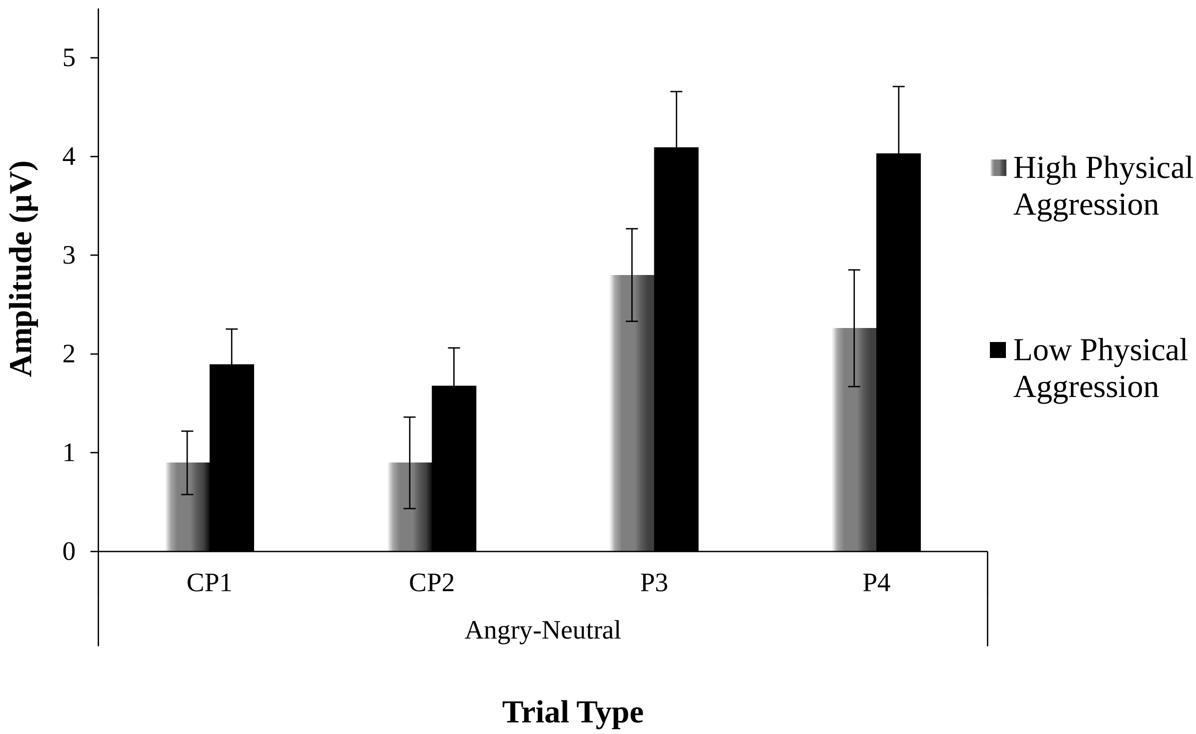

Figure 30: Bar graph to show the significant differences in evoked amplitude between high (n = 23) and low (n = 25) physical aggression groups on angry-neutral trial between 200 and 300ms (error bars = +/- 1 standard error).

6.4.4.2 Effect of valence

A mixed model omnibus ANOVA was conducted for each epoch to explore the effect of valence. Trial type (3 levels; angry-neutral, angry-happy, happy-neutral), electrode (5 levels), and hemisphere (2 levels) were added as within subject factors. Physical aggression was added as a between-subject factor. The ANOVA showed no effect of valence. This suggests that overall amplitude in response to face pair onset on angry-neutral, angry-happy and happy neutral trials was relatively stable (Figure 31). This shows no evidence for hypothesis 14 as it was predicted that *angry-neutral* and *angry-happy* trials would evoke an increased amplitude to *happy-neutral* trials and that this would be particularly salient in the high aggression group.

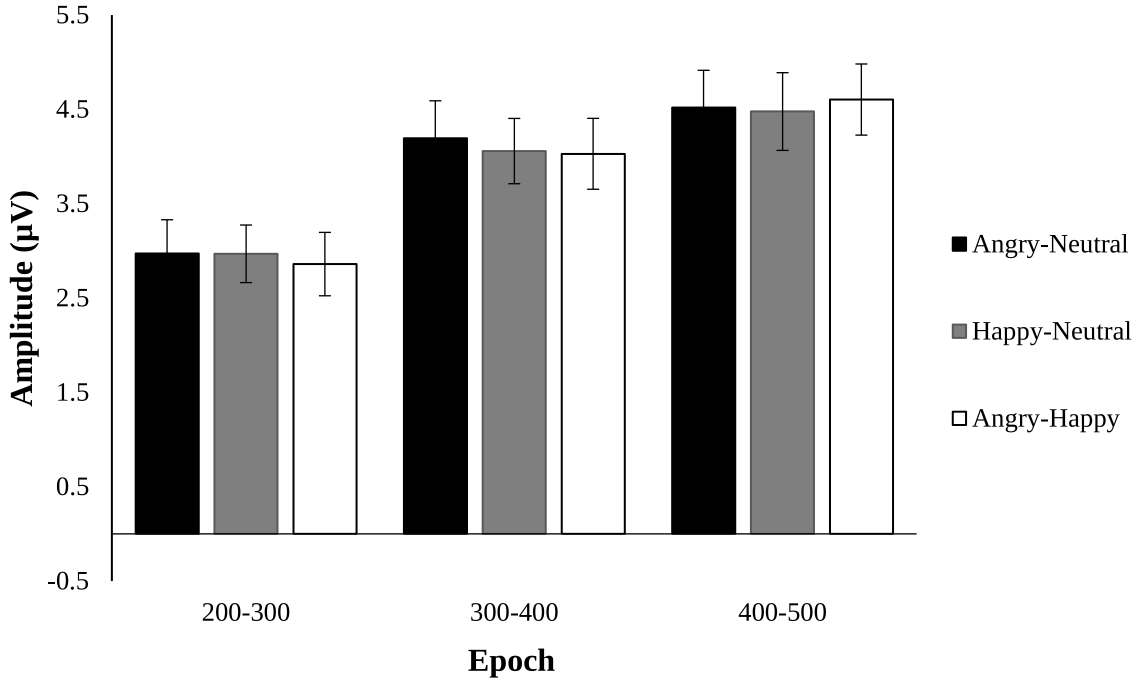

Figure 31: Bar graph to show the differences between trial types in average evoked amplitude of all electrodes in our region of interest across all participants (n = 51) (error bars = +/- 1 standard error).

6.4.4.3 Effect of trial congruency

6.4.4.3.1 *Post-probe differences in congruency.*

To explore the effect of congruency, a 3 (trial type; angry-neutral, angry-happy, happy-neutral) x 2 (congruency; congruent, incongruent) x 5 (electrode) x 2 (hemisphere; left, right) x 2 (physical aggression; high, low) mixed model omnibus ANOVA was conducted for each epoch. The results show a number of significant interactions with trial type and congruency between 600 and 800ms. Results revealed a significant interaction between trial type and congruency between 600 and 700m, $F(2,92) = 3.232$, $p = .045$, $\eta_p^2 = .066$; the effect of congruency also approached significance between 700 and 800ms, $F(1,46) = 3.256$, $p = .078$, $\eta_p^2 = .066$ (Figure 32). To explore these effects and investigate the hypotheses for each trial type, planned analyses were conducted to study the effects of congruency within each trial type. These planned comparisons consisted of a 2 (congruency) x 5 (electrode) x 2 (hemisphere) repeated measures ANOVA.

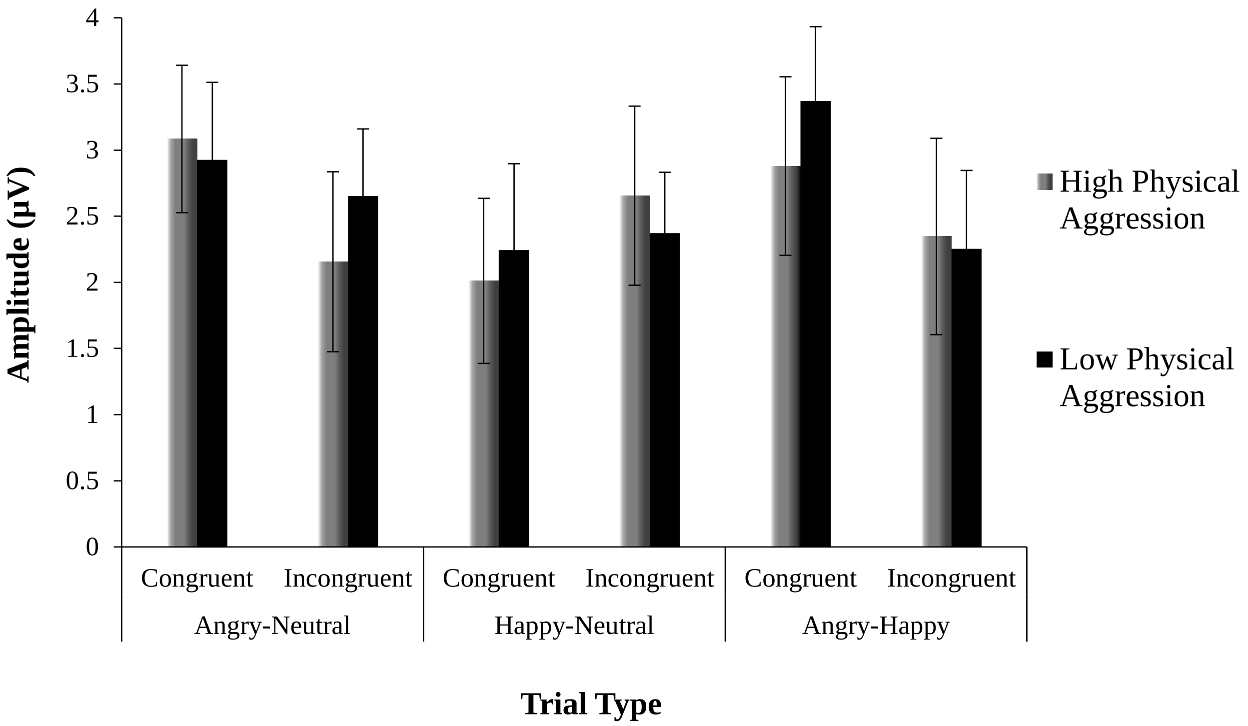

Figure 32: Bar graph to show the differences in congruency for each trial type. The graph shows the averaged evoked amplitude of all electrodes in our region of interest between 600 and 700ms in high (n = 23) and low (n = 25) physical aggression groups (error bars = +/- 1 standard error).

On *angry-neutral* trials, the interaction between congruency and hemisphere approached significance $F(1,46) = 3.136$, $p = .083$, $\eta_p^2 = .064$. Follow up tests showed a significant effect of congruency in the right hemisphere only, $F(1,46) = 7.024$, $p = .011$, $\eta_p^2 = .132$. Effect of congruency was significant at TP10, $F(1,50) = 5.454$, $p = .024$, $\eta_p^2 = .098$; P4, $F(1,50) = 5.411$, $p = .024$, $\eta_p^2 = .098$; and P8, $F(1,50) = 6.814$, $p = .012$, $\eta_p^2 = .120$. Evidence shows support for hypothesis 15. The waveform (Figure 33) shows on *angry-neutral* trials there is an increased positive amplitude for congruent trials compared to incongruent trials. This may reflect a P1-like component in response to probe presentation. However, the waveform suggests that the effect may be a long lasting effect which begins pre-probe presentation. There was no significant interaction between trial congruency and physical aggression, which suggests that effect of congruency is stable across both aggression groups. This suggests evidence for hypothesis 16 as effect of trial congruency on in the low aggression group is similar to that of the overall trial congruency effect. There is no support for hypothesis 17 as it was hypothesised that high aggression participants would not differentiate between congruent and incongruent trials. However there is no interaction with physical aggression which suggests that effect of congruency is stable across aggression groups and therefore suggesting that high aggression participants show increased positive amplitude on congruent trials (see bar chart (Figure 32). On happy-neutral trials there was no significant effects of congruency (Figure 34) suggesting that participants did not differentiate between happy and neutral faces. Hypotheses 18 and 19 were not supported as it was predicted that amplitude would be increased in response to probes that replace happy faces.

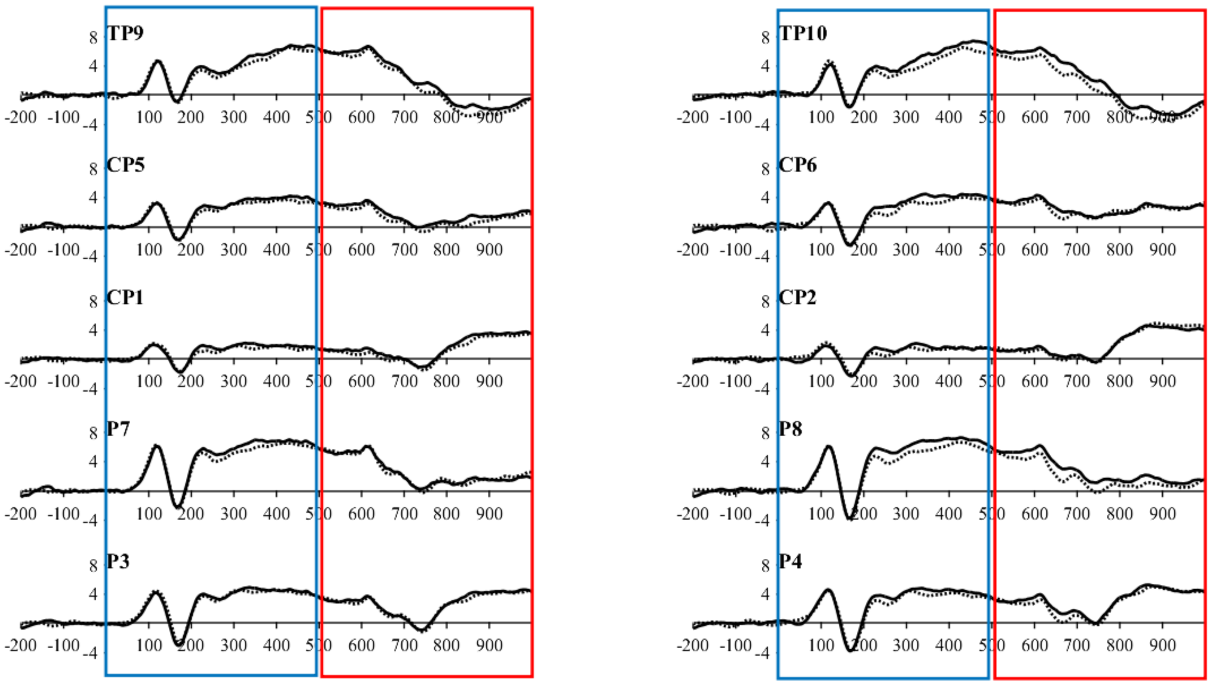

Figure 33: Grand average ERPs for the effect of trial congruency on angry-neutral trials across all participants (n = 51). Mean amplitude on congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

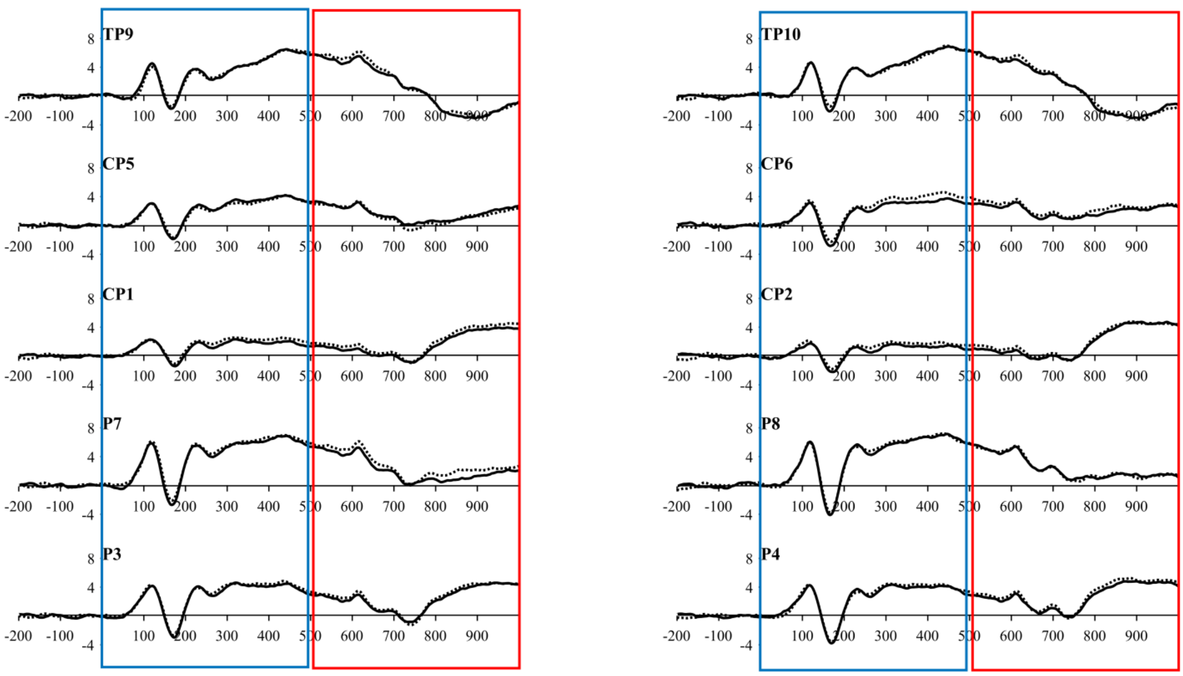

Figure 34: Grand average ERPs for the effect of trial congruency on happy-neutral trials across all participants (n = 51). Mean amplitude on congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

On angry-happy trials, the effect of congruency approached significance between 600 and 700ms, $F(1,46) = 3.828$, p =.056, $\eta_p^2 = .077$; and was significant between 700 and 800ms, $F(1,46) = 5.421$, p =.024, $\eta_p^2 = .105$. Inspection of the waveform (Figure 35) revealed that on angry-happy trials the amplitude is larger on congruent trials compared to incongruent trials. This shows evidence for hypothesis 21. The waveform reveals that following probe presentation, on congruent trials there is enhanced P1 and P2 amplitude, compared to incongruent trials. Qualitative evaluation of the waveform also reveals that the effect may influence the P300 component, although this did not reach significance. There were no significant interactions with aggression, suggesting that in line with the main effect of congruency, both groups show increased positive amplitude to congruent trials compared to incongruent trials. Therefore there are tentative results to show evidence for hypothesis 22 and 23.

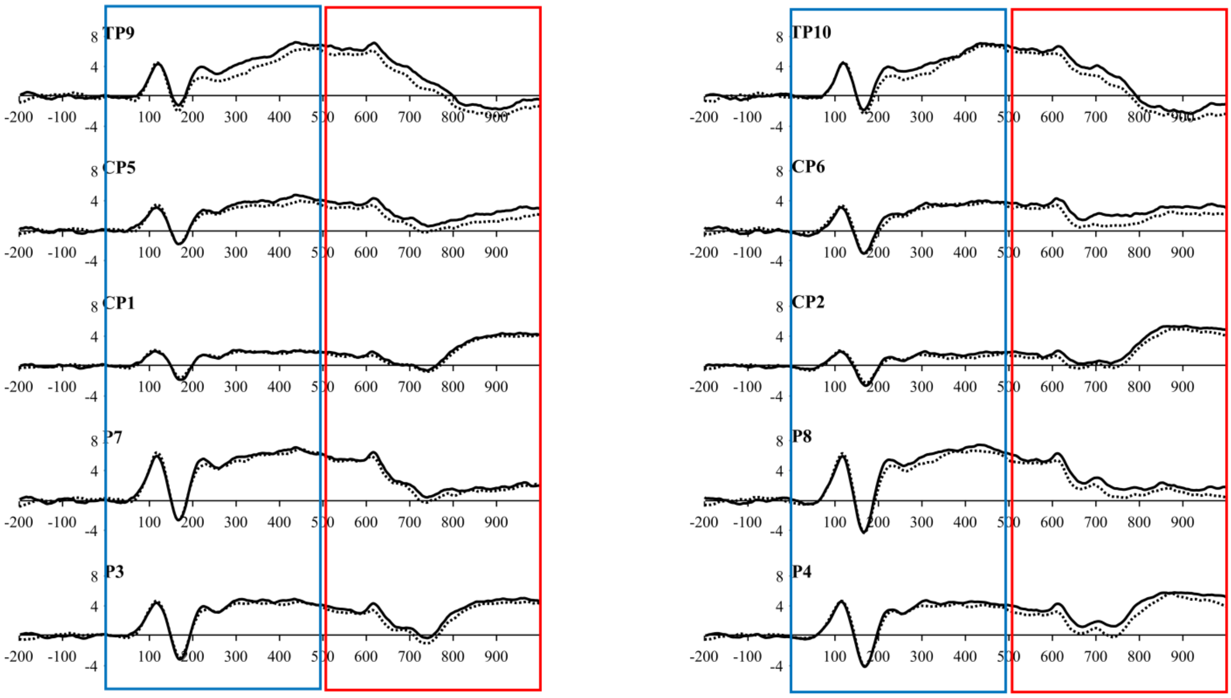

Figure 35: Grand average ERPs for the effect of trial congruency on angry-happy trials across all participants (n = 51). Mean amplitude on congruent trials (black) are compared with mean amplitude to incongruent trials (dotted). Pre-probe (blue) and post-probe (red) epochs are highlighted.

6.4.4.3.2 *Pre-probe differences in congruency.*

Based on a qualitative inspection of the waveform, the effects of congruency at earlier latencies (100-500ms) were also explored. Surprisingly, effects of congruency were found between 200 and 400ms post face onset. The main effect of congruency approached significance between 200 and 300ms, $F(1,46) = 3.326$, $p = .075$, $\eta_p^2 = .067$; and between 300 and 400ms, $F(1,46) = 3.224$, $p = .079$, $\eta_p^2 = .065$. There was also a significant congruency by electrode interaction between 200 and 300ms, $F(4,184) = 3.857$, $p = .015$, $\eta_p^2 = .067$; and 300 and 400ms, $F(4,184) = 3.139$, $p = .028$, $\eta_p^2 = .064$. Finally, there was also evidence of an interaction between congruency and physical aggression pre-probe presentation. Between 200 and 300ms there was a significant interaction between trial type, congruency, hemisphere and physical aggression, $F(2,92) = 3.426$, $p = .038$, $\eta_p^2 = .069$. Between 300 and 400ms there was a significant congruency by hemisphere by physical aggression interaction, $F(1,46) = 5.289$, $p = .026$, $\eta_p^2 = .103$.

Post-hoc tests for *angry-neutral* trials showed there was a main effect of congruency between 200 and 300ms, $F(1,46) = 4.078$, $p = .049$, $\eta_p^2 = .081$; and 300 and 400ms, $F(1,46) = 4.130$, $p = .048$, $\eta_p^2 = .082$. The interaction between congruency and hemisphere approached significance between 200 and 300ms, $F(1,46) = 3.217$, $p = .079$, $\eta_p^2 = .065$; and 300 and 400ms, $F(1,46) = 3.151$, $p = .082$, $\eta_p^2 = .064$. Follow up tests showed a significant effect of congruency in the right hemisphere between 200 and 300ms, $F(1,46) = 6.472$, $p = .014$, $\eta_p^2 = .123$; and 300 and 400ms, $F(1,46) = 6.548$, $p = .014$, $\eta_p^2 = .125$. There were no significant in the left hemisphere. Between 200 and 300ms, the effect of congruency was significant at CP6, $F(1,50) = 6.048$, $p = .017$, $\eta_p^2 = .108$; P4, $F(1,50) = 4.055$, $p = .049$, $\eta_p^2 = .075$; and P8, $F(1,50) = 9.803$, $p = .003$, $\eta_p^2 = .164$. Between 300 and 400ms the effect of congruency was significant at CP6, $F(1,50) = 5.003$, $p = .030$, $\eta_p^2 = .091$; and P8, $F(1,50) = 7.474$, $p = .009$, $\eta_p^2 = .130$. The effect also approached significance at TP10, $F(1,50) = 3.765$, $p = .058$, $\eta_p^2 = .070$. Inspection of the waveform (Figure 33) shows there are early effects of congruency in which congruent trials evoke an increased P2/P3 amplitude compared to incongruent trials.

For *angry-happy* trials there was a significant congruency by electrode interaction between 200 and 300ms, $F(4,184) = 4.792$, $p = .005$, $\eta_p^2 = .094$. There was also a close to significant congruency by hemisphere by physical aggression interaction between 200 and 300ms, $F(1,46) = 3.982$, $p = .052$, $\eta_p^2 = .080$; and 300 and 400ms, $F(1,46) = 4.517$, $p = .039$, $\eta_p^2 = .089$. However, follow up tests showed that in both epochs there were no significant effects in either the high or low physical aggression group.

6.5 Discussion

This chapter investigated selective attentional processes involved with attending to negative and positive facial expressions (namely, angry and happy) in a low and high physical aggression sample. The primary aims of the study were two-fold. Firstly, the aim was to replicate the findings from Study 3 by exploring attention bias to angry faces during a dot-probe task in which they were simultaneously presented alongside a neutral face. Secondly, by including two other trial types, happy-neutral, and angry-happy, the aim was to explore attention bias to different emotional faces. The study explored whether physically aggressive individuals attend to happy faces (as well as angry), when paired with a neutral face distracter. An angry-happy trial type was included to investigate selective attentional processes involved with attending to angry faces when they are presented alongside an emotional distracter. Complimentary reaction time and EEG data was used to make better informed conclusions regarding cognitive processes involved with attention bias in aggression. Due to the complexity of the results the behavioural and ERP results for each trial type will be explained and discussed individually before an overview of the main findings are presented.

6.5.1 Main findings and interpretations

6.5.1.1 Behavioural results

6.5.1.1.1 *Angry-neutral*

There was no evidence for any of the hypotheses relating to angry-neutral trials. It was hypothesised that there would be a significant effect of trial congruency on reaction time across the whole sample but this effect would be particularly salient in the high aggression group. Due to facilitated engagement of threat stimuli, it was predicted that generally participants would have a quicker reaction time on angry trials compared to neutral trials. Fox et al. (2000) suggests that healthy individuals (normal controls) should still show a bias towards angry faces as individuals are evolutionally primed to detect threat in their environment. Fox et al. (2000) propose that detection of angry faces is fast and efficient; although they claimed, it does not have a traditional ‘pop out’ effect. There was no evidence of attention bias in the current sample. There were also no significant effects found across either group suggesting that both high and low aggression groups respond similarly to when the probe replaces angry faces and when the probe replaces neutral faces. Given the literature suggests quite a robust link between aggression and attention bias to angry stimuli (Ciucci et al., 2018; Maoz et al., 2017; Smith & Waterman, 2003; van Honk et al., 2001), this is a somewhat surprising finding. These findings are also inconsistent with Study 3 where an increased attention bias towards angry faces in the high aggression sample was found (significant correlations but no between-subject effects).

6.5.1.1.2 *Happy-neutral*

There was no evidence of attention bias on happy-neutral trials. There was no main effect of trial congruency in either aggression group and no significant correlations. This suggests that there are no significant differences in reaction time when participants respond to probes that appear in place of happy faces and probes that appear in place of neutral faces. It was predicted that the low aggression group would show an attention bias for happy faces compared to neutral, whereas due to perceived hostility in neutral expressions (Mellentin et al., 2015) the high aggression group may show an attention bias for neutral faces. However neither of these predictions were supported. These findings are consistent with Ciucci et al. (2018), Bantin et al. (2016), and Salum et al. (2013) which showed that in aggressive, anxious and disordered participants respectively, there was no evidence of an attention bias to happy faces during a dot-probe task in which happy-neutral face pairs were presented. However, there is some contradictory literature which suggest that healthy controls show a bias to happy stimuli compared to neutral stimuli (Pishyar et al., 2004; Waters et al., 2010). Waters et al. (2010) conducted a visual probe task in which angry-neutral and happy-neutral face pairs were presented. Results demonstrated that anxious individuals showed an attention bias to angry faces compared to neutral faces, whereas the non-anxious controls showed an attention bias to happy faces relative to neutral faces. Given the mixed evidence for attentional selectivity of happy faces, the current findings suggest that increased levels of physical aggression do not influence attentional allocation to happy faces, as both low and high aggression groups responded similarly across probes that replaced happy faces and probes that replaced neutral faces.

6.5.1.1.3 *Angry-happy*

There was evidence for a main effect of trial congruency on angry-happy trials (hypothesis 10). This suggests that in general participants were quicker to respond on congruent trials (probe appears in place of angry face) compared to incongruent trials (probe appears in place of happy face). It was predicted that this difference in reaction time between congruent and incongruent trials would be greater in the high physical aggression group compared to the low physical aggression group. However, no evidence of an interaction with aggression was found, suggesting that the effect of trial congruency is consistent across both low and high aggression groups. Both facilitated attention and poor attentional disengagement can contribute to attention bias (Cisler & Koster, 2010). Therefore differences in reaction times between congruent and incongruent trials could be due to speedier reaction times on congruent trials attributed to facilitated attention to angry stimuli; or delayed reaction time on incongruent trials, due to difficulties in disengaging from the simultaneously presented angry face when responding to the probe which replaces the happy face. These findings suggest that there may be complex attentional processes activated when participants are required to selectively attend to two emotional stimuli.

The current finding that participants are generally quicker to respond to probes replacing angry faces, compared to happy faces is consistent with evidence that suggests that angry faces are easier to detect in a matrix of happy faces, compared to happy faces in a matrix of angry faces (Hansen & Hansen, 1988). This suggests that potentially angry faces are detecting more quickly by the attentional system and command greater levels of processing.

Previous evidence suggests that individuals show an attention bias to angry faces when paired with a neutral distracter (Bar-Haim et al., 2007; Maoz et al., 2017; Santesso et al., 2008). This current study has built on previous work by demonstrating that individuals are attentive to an angry stimulus when paired with a happy distracter. Therefore, individuals may preferentially attend to angry faces regardless of the distracter stimuli (neutral or happy). Interestingly, an attention bias effect to angry stimuli was found when paired with happy, but not neutral facial expressions. To my knowledge this is the first study to include an angry-happy trial type when investigating selective attention in aggression. Therefore, future research will be needed to replicate these results and contribute to the understanding of cognitive processes employed when aggressive individuals are presented with two differently valenced facial expressions.

6.5.1.2 ERP results

The ERP results showed a main effect of aggression such that on angry-neutral trials, the low aggression group had increased P2 amplitude in response to face pair presentation compared to the high aggression group. This is consistent with previous research which shows reduced amplitude in response to task relevant information in aggressive individuals (Bernat et al., 2007; Gao & Raine, 2009; Gao et al., 2013). Gao and Raine (2009) suggested that antisocial behaviour is related to the inefficient deployment of neural resources and therefore participants show reduced processing of stimuli presented during cognitive tasks. However, these conclusions were drawn from studies using standard oddball, more complex non-oddball, and Stroop tasks, therefore these may not be comparable with the dot-probe task used in the current study.

Although these findings are consistent with some previous literature, the observed effect is in contrast the results found in Study 1, 2 and 3. The previous studies presented in this thesis showed that overall high physical aggression participants showed an evoked amplitude that was increased in response to both word pair (Study 1 and 2) and face pair (Study 3) presentation, compared to low physical aggression participants. Due to the similarity of the task used across the studies, it is surprising that a reverse effect would be found in one of the four studies. However, qualitative inspection of the waveform suggests that although low aggression show enhanced P2 amplitude compared to the high aggression group, the high aggression group show some evidence of increased P300 compared to the low aggression group, although this did not reach statistical significance. These results suggest that processing of stimuli may influence stages of attentional processing differently depending on levels of aggression.

The results revealed no differences in evoked amplitude between trial types for any ERP component. This shows that in response to angry-neutral, happy-neutral, and angry-happy face pair presentation, there were no significant differences in attentional processing of emotions. Finally, analysis was conducted to explore the evoked amplitude in response to congruency for each trial type. The results for each trial type are discussed in turn below.

6.5.1.2.1 *Angry-neutral*

The evidence showed a main effect of trial congruency across the whole sample, such that there was increased P1amplitude for probes which appear in place of angry faces compared to probes which appear in place of neutral faces. This finding replicates previous work by Santesso et al. (2008) which showed that during a dot-probe task in which angry-neutral face pair were presented, angry-congruent trials evoked an increased P1 amplitude compared to angry-incongruent trials within a general population sample. This effect is also consistent with the low aggression group in Study 3 (and similar to low aggression groups in current literature; Thomas et al., 2007). The increased amplitude on angry-congruent trials may reflect the increased allocation of resources to process stimuli (Hillyard & Kutas, 1983), or increased salience (e.g. Sass et al., 2010).

Unexpectedly, effects of congruency were evident pre-probe presentation. Results suggest that congruent trials evoke increased P2/P3 amplitude compared to incongruent trials. These findings replicate the early effects of congruency found in previous studies outlined in this thesis; however I acknowledge that theoretically it is not possible to measure congruency effects before the probe has appeared on screen (pre 500ms). Therefore these results will require replication.

Based on previous research (Helfritz-Sinville & Stanford, 2015) and results from Study 3 it was hypothesised that the effect of trial congruency would be salient in the low physical aggression group, however, the high physical aggression group would show relatively stable amplitude across both congruent and incongruent trials. Helfritz-Sinville and Stanford (2015) found that in response to a modified oddball task in which threat and neutral words were presented, both reactive and premeditated aggressive participants showed relatively stable P300 amplitude across responses to social and physical threat words and neutral words. Study 3 found similar results to these using a dot-probe task and therefore it was expected that these results would be replicated in the current study. However, results from this study revealed no significant interaction between trial congruency and aggression group, this suggests that the both low and high aggression groups show an increased P1 amplitude to angry congruent trials.

These findings suggest that high aggression participants show differentiations in ERP patterns in response to angry-congruent and angry-incongruent trials. However, patterns of P1 amplitude were relatively consistent across both aggression groups. Therefore, due to the recruitment of a non-forensic sample, it is proposed that perhaps the groups were not different enough (more than likely down to the high aggression group not experiencing extreme/clinical levels of aggression) to reflect differing attentional processes. As the results are consistent with effects shown in low aggression groups across the literature this could be a valid explanation. However, the high and low physical aggression samples recruited for this current study were comparable to the samples used in Study 3.

6.5.1.2.2 *Happy-neutral*

Due to mixed evidence it was hypothesised that participants would show increased positive amplitude on congruent trials compared to incongruent trials, or there would be no difference in amplitude between congruent and incongruent trials. The results show no significant differences in P1 or P300 amplitude between congruent and incongruent trials across the whole sample and no interaction between trial congruency and aggression group. This suggests that individuals show similar processing of both happy and neutral stimuli regardless of self-reported aggression levels. This finding is consistent with a study conducted by Santesso et al. (2008) which found no significant differences in P1 amplitude between probes that appeared in place of happy and neutral faces. Leppänen et al. (2007) also showed that when participants were presnted with happy and nuetral faces, there were no significant differences in evoked N170 amplityde. These findinsg suggest that happy and neutral faces are processed similarly by the attentional system. However, there is contradictory evidence which suggests healthy individuals show increased amplitude in response to happy congruent trials. Holmes et al. (2009) found that during a dot-probe task in which happy and neutral faces are presented, happy faces evoke increased N2pc amplitude. This mixed evidence suggests that the N2pc may be particularly sensitive to attentional allocation to happy faces. However, in the current study no differences between congruent and incongruent trials were found across any ERP component.

The absence of congruency effects on happy-neutral trials could be explained by the differing valence across stimuli. Happy and neutral faces may be closer in emotional valence compared to other stimuli pairs. There may be greater visible differences between angry and neutral faces and angry and happy faces; for example, an angry face usually has features such as, frowning brows, staring eyes and a shut mouth (Ekman & Friesen, 2003), whereas a happy face is often characterized by a U shape mouth. However, this explanation may not be suitable for explaining biases within aggressive population, as research suggests that aggressive individuals tend to perceive hostility in neutral facial expressions, as well as angry (Mellentin et al., 2015).

6.5.1.2.3 *Angry-happy*

In regards to angry-happy trials, there was an overall task affect in which participants showed increased P1/P2 amplitude on congruent trials compared to incongruent trials (hypothesis 21 supported). However, this did not interact with aggression, suggesting that both high and low aggression groups showed similar evoked ERP patterns. These findings are in line with a previous study by Smith et al. (2003) which demonstrated that participants showed enhanced P1 amplitude in response to negative affective pictures, compared to positive. Furthermore, Schupp et al. (2004b) found that during a simple task in which participants viewed different facial expressions, threat faces evoked increased LPP amplitude compared to both neutral and happy faces. These studies use different paradigms to measure attentional processing of emotional stimuli, however together they demonstrate that angry faces command greater resources at early and later stages of attentional processing, compared to happy faces.

These findings suggest that generally participants show greater processing of angry faces in the environment. The in depth processing of such stimuli may be in preparation for response formation. Happy faces do not usually require a behavioural response, whereas a potentially threatening face may demand an act of self protection (Brosch, Sander, Pourtois, & Scherer, 2008). This combination of stimuli in selective attention tasks is relatively unique and subsequent processing of such stimuli has yet to be studied in the literature. Across both behavioural and ERP results there seems to be something particularly interesting about attentional processes involved with attending to angry faces when they are paired with another emotional face. The P1 component increases when stimuli are presented in a pre-attended location (Woldorff et al., 2002) and therefore reflects spatial attentional at earlier stages of processing (e.g. Hillyard & Anllo-Vento, 1998; Woldorff et al., 2002). Current findings therefore suggest that angry stimuli attract attention and subsequently participants have faster reaction times and increased P1 amplitude in response to probes that appear in place of angry faces, compared to happy faces. This is consistent with previous evidence by Hansen and Hansen (1988) which showed that threatening faces are detected faster amongst crowds of neutral and friendly distracter stimuli, suggesting they more readily attract attention. There is very limited evidence of selective attentional processes associated with attending to angry faces when paired with a happy face; therefore these novel findings of the current study contribute to the understanding of attention processing of simultaneously presented positive and negative emotional faces.

6.5.2 Limitations

As discussed in Chapter 4, Section 4.5.2 the task was perhaps over complex. The rapid presentation of multiple trial types may not allow for the analysis of probe positions within distinct trial pairings due to overlapping processes.

Evidence by Smith et al. (2006) suggests that current mood can moderate an attention bias towards negative information. They used both behavioural and ERP methodology to investigate attention bias to negative and positive stimuli in different affective contexts. ERP results showed that when participants were primed with negative information, the P1 amplitude was increased in response to negative stimuli in the testing phase, whereas when participants were primed with positive information, P1 amplitude was increased in response to positive stimuli. They suggested that when participants were primed with the positive information, attention bias to negative stimuli can be eliminated or attenuated. This suggests that participants may only show an attention bias to angry faces (both speedier reaction time and increased P1 amplitude) in negative current mood states. The current mood of the participants at the time of completing the task was not measured and therefore the lack of significant differences in the behavioural data could be explained by the variance of mood states.

There is further evidence to suggest that aggressive individuals have reduced levels of emotional intelligence; for example, a systematic review by García-Sancho, Salguero, and Fernández-Berrocal (2014) found strong evidence to suggest that people with increased levels of aggression have lower emotional intelligence scores. It appears that this relationship is robust across ages, types of aggression, and cultures. Due to the use of emotional stimuli (angry and happy faces) used within this study, allocation of attentional resources to angry faces in participants with increased aggression may be explained by poor emotional intelligence. Individuals with poor emotional intelligence lack the ability to perceive and appraise emotions accurately, understand emotion, or regulate their own emotions (Mayer & Salovey, 1997). Therefore, in future work I would suggest measuring emotional intelligence as well as aggression to aid understanding of attention bias to angry and happy stimuli.

The use of a happy face to measure attention bias for positive emotionality may not be a suitable control for measuring attention bias to angry faces. I was interested in whether physically aggressive individuals show a bias to angry faces or whether they show a more general emotional bias towards angry and happy faces. However, threat-related expressions are much more relevant to the observer compared to happy facial expressions, as they require rapid in-depth processing needed for response formation. In social contexts, if an individual sees an angry face, they will need to attend to the person in order to evaluate the impending aggression and prepare a response. Whereas if an individual encounters a smile in their environment there is no urgent response required. Therefore the response demand-characteristics of angry and happy facial expressions are perhaps not comparable (Brosch et al., 2008).

These fundamental differences in angry and happy faces may provide an explanation as to why angry faces are preferentially attended to, compared to happy faces (shown by both behavioural and ERP evidence in the current study). Individuals are primed to detect possible threat in the environment (e.g. Darwin &

Darwin, 2009; Nesse, 1998) in order to protect oneself from danger. Furthermore, happy faces are consistently used in the literature as a measure of positive emotion, in comparison to either neutral or negative emotion, primarily as there are very few possible effective alternatives, especially when conducting the dot-probe paradigm. However, these are important considerations when interpreting these results and may contribute to the differences in attention bias effects for angry and happy faces.

Furthermore, happy and angry facial expressions represent two distinct emotions. Emotional valence is related to behavioural approach and avoidance inclinations (Chen & Bargh, 1999). When considering motivational tendencies, Carver and Harmon-Jones (2009) suggest that angry stimuli can be met with either approach or avoid motivational tendencies and that different brain areas may be responsible for each of the mechanisms. Anger can be associated with an approach motivational orientation, that is, anger is is experienced when goal behaviour is disrupted meaning that a desired end point can not be reached. Approach tendencies also underlie behavioural responses to anger when individuals aim to remove the violation or disruption to goal directed behaviour. This theory could contribute to findings which show that participants have a heightened vigilance for threatening faces compared to neutral or happy faces (Bradley et al., 1998; Mogg et al., 1997; Santesso et al., 2008). Carver and Harmon-Jones (2009) also suggest that anger and fear are closely linked as the presentation of anger is usually met with fear. Therefore, if an angry stimulus is appraised as threatening and causes fear in the perceiver, this stimulus may be met with an avoid motivational orientation. This theory is consistent with the ‘fight or flight’ repsonse (Cannon, 1929). Therefore, this suggests that individuals will generally avoid information with a possible negative outcome such as negative affect, but will approach a stimulus when a positive outcome or affect is expected (Carver, Avivi, & Laurenceau, 2008). In response to hostile stimuli, such as an angry facial expression, aggressive individuals are more likely to use approach motivational strategies, compared to avoidance strategies.

According to these theories, angry and happy faces may impact the motivational response system in different ways. Therefore, due to the complex nature of emotion, emotions portrayed by facial expressions may have different influences on attentional allocation and subsequent behaviour. In the context of social information processing models where it is presumed that cognitive processes influence behaviour, happy faces may not be a suitable control for emotionality. When exploring whether attention biases are distinct for angry stimuli within aggressive populations, it may be more effective to use another negatively valanced stimuli which may be associated with similar approach or avoid motivational tendencies.

As noted in previous chapters, the results consistently show differences in evoked amplitude in response to congruent and incongruent trials before the presentation of the probe. This effect seemed to be most salient on angry-neutral trials. Theoretically, it is not clear why participants would show differences in the processing of angry and neutral face pairs at 300ms, depending on the upcoming location of the probe at 500ms. This surprising finding will require further investigation in order to provide a valid explanation.

The final consideration is that the study consisted of a non-clinical sample. Although it is important to study increased levels of aggression in a normative sample this may explain why the results revealed no significant interactions with aggression. Conclusions are drawn based on analyses of the whole sample; however conclusions regarding how aggression may influence these biases are drawn from limited evidence and are made with caution. Replication and further research will be crucial in confirming these conclusions.

6.5.3 Future work

In addition to the suggestions made in the preceding discussion, there are a number of further recommendations for future work. In order to establish the specificity of negative attention biases in aggression, a number of different dot-probe tasks could be conducted to explore attentional processes involved with attending to different negatively valenced emotional faces. It would be interesting to explore if participants would still show a quicker response/increased amplitude to angry faces if they were paired with a disgustful or sad face. These two faces are much closer in negative valence and consequently could test if the attention bias effect is unique to angry faces. It would be expected that attentional processes, reflected by ERPs, would be elevated, and reaction times would be quicker, in response to angry targets. Öhman et al. (2001) found evidence for this, threatening angry faces were more quickly and accurately detected than were other negative faces (sad or "scheming"), which suggests that the threat advantage can be attributed to threat rather than to the negative valence or the uniqueness of the target display.

Due to the lack of between-subject effects within this study, a recommendation would be to recruit a clinically aggressive sample. The future aim would be to replicate these findings across a healthy control group, and understand how attentional processes may differ amongst a population with extreme levels of aggression. I believe that using two extreme groups may allow for more robust between-subject conclusions to be drawn.

6.5.4 Contributions

The research contributes to the aggression and attention bias literature in a number of ways. Firstly, to my knowledge it is the first study to investigate selective attention bias to different emotional faces (angry, happy and neutral) in aggression using both behavioural and EEG methodology. It has provided evidence for increased processing of angry faces compared to both neutral and happy faces. There were no differences between evoked amplitude to congruent and incongruent trials on happy-neutral trials. These findings suggest that angry faces have a specific influence on the attentional system which evokes greater processing.

Previous studies on aggression have not been interested in cognitive processing of different types of facial expressions (namely angry versus happy). However, facial expressions convey emotions and therefore being able to effectively interpret different expression is essential for successful social communication (Green & Phillips, 2004). Different emotions may be associated with distinct perceptual and neuro-cognitive processes (Oster, Daily & Goldenthal, 2013). Therefore it is important to understand how aggressive individual perceive different emotions and the role this plays in aggressive behaviour. The findings from this study suggest that when presented with both emotionally positive and negative faces, participants will be quicker to respond and have increased P1 amplitude in response to probes that replace negative faces, suggesting that initial attentional resources are allocated towards such stimuli. Surprisingly, there were no differences in evoked amplitude in response to face pair presentation between angry-neutral, happy-neutral and angry-happy trials. However, this study goes some way to contributing to the complex understanding of neuro-cognitive processes associated with selective attention to angry and happy facial expressions in aggression.

6.5.5 Conclusions

Using behavioural and ERP techniques, this study explored attention bias to happy and angry faces in aggression. The first aim of the study was to replicate findings from Study 3 which showed attenton bias to angry faces compared with neutral faces during a dot-probe task in which they were simultaneously presented. The second aim was to explore attention bias to different emotional faces by including two other trial types; happy-neutral and angry-happy. The behavioural results from Study 3 did not replicate as there was no effect of congruency on angry-neutral trials. On angry-happy trials there was a main effect of trial congruency in which participants were generally quicker to respond to probes on angry trials compared to happy trials. To my knowledge this is the first study to investigate selective attention processes associated with attending to angry-happy stimuli when they are simultaneously presented and therefore, future research will be needed to replicate these results.

The main ERP findings suggest that across both angry-neutral and angry-happy trials, there is a general task effect in which participants have increased amplitude on angry-congruent trials (regardless of the valence of the simultaneously presented distracter stimuli). The ERP results on angry-neutral trials are similar to those found in Study 3, however in Study 3 there was increased amplitude on angry trials in the low physical aggression group only (amplitude was relatively stable in the high physical aggression group), whereas results from this study showed no significant interaction between trial congruency and aggression group, suggesting that both low and high aggression groups show an increased amplitude to angry-congruent trials. This is in keeping with previous literature (Helfritz-Sinville & Stanford, 2015; Thomas et al., 2007) which suggests increased processing of angry stimuli in normative healthy samples. To conclude, using a combination of behavioural and ERP methods, the study has provided initial ERP evidence for a general processing bias for angry faces compared to neutral and happy faces, during a selective attention task. Due to minimal behavioural effects and between-subject differences the conclusions drawn are tentative, however, its suggested that future work is important in understanding how increased P1 amplitude in response to angry trials during a selective attention task may contribute to aggressive behaviour.

7 Study 5 – Hostility-related interpretation bias

7.1 Introduction

The previous four empirical chapters have reported four studies that have investigated *attention* biases to stimuli of different types and different valences across high and low aggression groups. As aggression is also associated with other cognitive biases, such as interpretation bias, it is important to investigate these too. Social information processing theory (Crick & Dodge, 1994) explains how attention and interpretation processes have an effect on other cognitive processes involved with the formation of behavioral responses to the environment (clarification of goals, response access or construction, response decision, and behavioural enactment). In aggression, interpretation bias refers to attributing negative, hostile or angry intentions to the behaviour of individuals in the environment (Nasby et al., 1980). The fifth and final study presented in this thesis investigated the cognitive processes involved with *interpretation* bias in aggression. In this chapter, *hostile interpretation bias* and *attributing hostile intent* are used synonymously.

Attributing hostile intent to peers has been consistently linked to aggressive behaviour in children (Dodge, 1980; Dodge & Coie, 1987; Dodge & Frame, 1982; Dodge & Newman, 1981; Dodge et al., 1990; Dodge & Tomlin, 1987; Fitzgerald & Asher, 1987; Guerra & Slaby, 1989; Quiggle et al., 1992; Sancilio et al., 1989; Steinberg & Dodge, 1983). Findings suggest that aggressive boys aged between 5 and 11 are more likely than non-aggressive boys to attribute hostile rather than accidental behaviour to their peers after an ambiguous provoking event, such as 'getting hit in the back with a ball thrown by a peer' (Dodge & Frame, 1982). This work was influential as it suggested that interpretation biases were evident in children as young as five, and inspired further work into the role of cognitive biases in the development and maintenance of aggressive behaviour. The relationship between attributing hostile intent and aggression has since been demonstrated across multiple adult samples (Dill et al., 1997; Epps & Kendall, 1995; Hall &

Davidson, 1996). These studies suggest that biases in cognitive processing, especially attributing hostile intent, are robust and enduring.

Interpretation bias has been evidenced in forensic, highly aggressive samples (Dodge et al., 1990; Milich & Dodge, 1984; Slaby & Guerra, 1988, and non-forensic samples with high trait aggression (e.g. (Dill et al., 1997; Epps & Kendall, 1995; Hall & Davidson, 1996). Dodge et al. (1990) explored the relationship between interpretation bias and aggression in a sample of juvenile offenders aged between 14 and 19 years. Using a multiple choice format, participants were asked to attribute intent to a protagonist in three different types of video vignettes. Participants with increased levels of reactive aggression, who committed a greater number of violent crimes, made more hostile attributions (stated behaviour of the protagonist was 'to be mean'). This finding is consistent across non-clinical adult populations. Epps and Kendall (1995) found that adults scoring high on self-rated anger gave more negative interpretations to unfamiliar situations which outlined an interpersonal interaction. These results suggest that more aggressive individuals are sensitive to hostile environmental cues; therefore they may disproportionately attribute hostility to the actions of others, even in the presence of dominant non-hostile cues.

There is evidence to suggest that making hostile attributions of intent may be particularly salient in individuals who report high levels of reactive aggression (Crick & Dodge, 1996; Dodge & Coie, 1987; Dodge et al., 1990; Lobbestael et al., 2013). Reactive aggression refers to angry, emotional or affective aggression which is usually expressed in a physical response to provocation (Dodge & Coie, 1987). It is therefore perhaps not surprising that making negative interpretations of instrumental situations has also been associated with physical aggression (Dodge, 1980; Dodge & Somberg, 1987). The first aim of this study was to replicate previous studies and test the association between hostile-related interpretation bias and aggression; however, due to the broad association between interpretation bias and different types of aggression, interpretation bias across anger, hostility, verbal and physical aggression subscales was investigated, with the aim of increasing understanding of cognitive processing of social stimuli that may contribute to aggressive behaviour.

Although it is well established that aggressive individuals have a negative interpretation bias, very little is known about neural processes associated with this bias. Current experimental methods for measuring interpretation bias have relied on participants' subjective reports. These may be influenced by demand characteristics, the mood-congruency hypothesis, or social desirability bias. Therefore, functional neuroimaging methods such as EEG may be useful in determining the underlying neural processes associated with interpreting hostile stimuli. There have been only a small number of studies which have used EEG methodology to examine potential neural correlates of making hostile attributions. However, Moser et al. (2008a) conducted a study in which high and low socially anxious groups completed an ambiguous sentence completion task while EEG was recorded. Participants were required to identify the valence of the resolution word. The ERP results revealed that individuals scoring low on social anxiety were characterized by larger P600 in response to negative sentence resolutions compared to positive, whereas high socially anxious individuals showed similar P600 in response to both types of sentence resolutions. The P600 is similar to the P300 component and is evoked in response to expectancy violations, however the effect appears later (Van Herten et al., 2005). The authors hypothesised that non-anxious individuals have a positive bias whereby social situations are generally interpreted positively, and consequently that unexpected negative resolutions evoke a peak in P600 amplitude. However, anxious samples show no evidence of this positivity bias. These results fit with expectancy models of the P600 component and contribute to the understanding of cognitive processes involved with interpreting the environment in social anxiety.

Gagnon and colleagues have assessed the association between evoked N400 potential and hostile interpretation bias (Gagnon et al., 2016; Gagnon et al., 2017).

The N400 component is associated with semantic processing, that is processing of the meaning of a stimulus in its context (reviewed in Kutas & Federmeier, 2011), and is sensitive to violations of expectancy models (Gagnon et al., 2016). For example, Moreno and Vázquez (2011) found that participants had evoked N400 amplitude to positive and negative sentence stems which were displayed with a nonsense outcome, compared to their emotionally matched expected outcome. Gagnon et al. (2016) investigated the expectations of hostile intent and the N400 component in a healthy sample. Participants were presented with a number of ambiguous hostile or non-hostile scenarios which were disambiguated with the presentation of either a hostile or non-hostile final target word. ERPs in response to the target word of each scenario were recorded. A larger N400 was evoked in response to mis-matching target words (when a non-hostile resolution word was presented for a hostile scenario and vice-versa). Further to this, Gagnon et al. (2017) replicated the previous methods using an aggressive sample. They found that, similar to the healthy sample (Gagnon et al., 2016), aggressive participants showed increased N400 amplitude in response to non-hostile words that resolved the ambiguity of hostile scenarios. They also observed an increased LLP-like component in which there was increased positive amplitude in response to hostile words that resolved the ambiguity of non-hostile scenarios, suggesting that in aggressive individuals the LPP may reflect the difficulty in integrating non-hostile social cues.

Research indicates that the LPP (sometimes referred to as the P600, a late P300 effect, in these studies) and the N400 show differences in hostile attribution bias. The LPP component is evoked in response to both pleasant and unpleasant stimuli compared with neutral (Foti & Hajcak, 2008), and is particularly salient in response to infrequent, surprising or important information (Polich & Criado, 2006). The LPP reflects cognitive processes involved with semantic and thematic expectancy violations, and is particularly sensitive to sentence processing tasks (Van Herten et al., 2005). In parallel, the LPP literature demonstrates that the component is increased in response to emotionally salient stimuli (e.g. Cuthbert et al., 2000; Hajcak & Olvert, 2008; Schupp et al., 2000b) and is particularly enhanced when the stimuli is particularly arousing, such as threat scenes (Schupp et al., 2004a). Therefore the LPP may be an appropriate component for assessing positive and negative (hostile) expectancy outcomes during the recognition task. Due to the limited research exploring the neural correlates of interpretation bias, and the evidence which suggests variation across a number of different components, predictions for the LPP and N400 were made. Both components may be useful in identifying and understanding the cognitive processes that contribute to hostility-related biases.

Tasks used to measure interpretation bias ask participants to attribute thoughts and feelings to unfamiliar situations, therefore participants are making clear and conscious attributions. I chose to use a recognition task as this aims to measure interpretation biases that are present at a more implicit level (Mathews & Mackintosh, 2000). Although the method and format used in this study were consistent with other recognition tasks, presentation of stimuli was modified to ensure EEG compatibility. Due to the novel use of the recognition task with simultaneous EEG recording, the aim was to assess the concurrent validity of these measures when assessing interpretation bias in aggression.

7.2 Aims and rationale

To summarise, the aims of this chapter were twofold. To my knowledge only a small number of studies have used EEG to investigate neuro-cognitive processes involved with hostile related interpretation biases (Gagnon et al., 2016; Gagnon et al., 2017; Godleski et al., 2010; Moser et al., 2008a). The recognition task has not been implemented with simultaneous EEG recording; therefore, the first aim of the study was to assess the validity of this assessment for measuring neural correlates of interpretation bias. To do this behavioural (interpretation bias score) and ERP (evoked amplitude in response to positive and negative statements) results extracted from the recognition task were compared with scores on an explicit measure of interpretation bias (AIHQ). Firstly, it was predicted that behavioural measures of interpretation bias across the implicit and explicit tasks would positively correlate. Additionally, based on findings by Moser et al. (2008a), it was predicted that participants scoring low on an explicit measure of interpretation bias would have a positivity bias, such that they would generally interpret social scenarios positively. Therefore, when responding to negative statements on the implicit recognition task, they would show increased N400/LPP potential. This is also consistent with the LPP and emotion literature which suggests that arousing stimuli of a threatening or hostile nature evoked increased potential (Hajcak, MacNamara & Olvet, 2010). However, participants scoring high on the explicit measure of interpretation bias would not show evidence of this positivity bias and would therefore show similar amplitude in response to positive and negative statements on the implicit measure of interpretation bias.

Crucially, the second aim of the study was to explore whether individuals with increased levels of aggression have a greater interpretation bias using explicit (AIHQ) and implicit (recognition task) measures. Both explicit and implicit behavioural measures were used to investigate the consistency of findings across measures, and included multiple subscales of aggression to explore the specificity of this bias. By investigating differences in ERP patterns between making hostile and non-hostile attributions in low and high aggression groups, the aim was to reveal possible neural correlates of negative interpretations in aggression. It was predicted that the robust association between aggression and interpretation bias would be replicated (e.g. Dill et al., 1997; Dodge & Frame, 1982; Epps & Kendall, 1995; Hall & Davidson, 1996; Lobbestael et al., 2013), such that aggression score would positively correlate with behavioural measures of interpretation bias across both implicit and explicit tasks. It was expected that high aggression participants would attribute hostile intent more frequently, and rate a scenario more negatively, compared to low aggression participants.

Drawing on the small number of previous studies (Gagnon et al., 2016; Gagnon et al., 2017; Moser et al., 2008a) tentative predictions were made regarding the ERP responses to the recognition task within an aggressive sample. The implicit recognition task allows for two types of analyses; effect of valence (positive and negative statements), and effect of similarity rating (similar and dissimilar). Regarding predictions of evoked amplitude in response to differently valenced statements, based on findings by Moser et al. (2008a) it was hypothesized that high aggression participants would show similar amplitude when responding to positive and negative statements, whereas low aggression participants would show increased amplitude in response to negative statements. This is also based on the findings from studies one and three in which, during an attention bias task, high aggression participants showed less differentiation in evoked P300 amplitude in response to angry and neutral stimuli, compared to low aggression participants.

On the recognition task (an implicit measure of interpretation bias), hostile interpretation bias is reflected in increased similarity ratings between an ambiguous scenario and negative statements. Therefore, of particular interest was the complex cognitive processes, reflected in evoked N400/LPP amplitude, when making increased similarity ratings of negative statements. Due to the novelty of using the recognition task with simultaneous EEG recording it was not possible to make firm predictions regarding ERP amplitude in response to making similarity ratings of negative and positive statements; however, in line with the previous predictions, and consistent with the expectancy models of the N400 (Gagnon et al., 2016; Gagnon et al., 2017) and LPP (Moser et al., 2008a), it was predicted that differences in interpretation bias (and the cognitive processes that contribute to this) between aggression groups would result in different ERP patterns when making similarity ratings of positive and negative statements (see Appendix W for example statements). It was proposed that low aggression participants would not have a negative interpretation bias, and therefore N400/LPP amplitude would be increased when making mis-matched responses that were not consistent with their positive expectation outcomes (similar ratings of negative statements and dissimilar ratings of positive statements). However, it was expected that high aggression participants would show evidence of a negative interpretation bias, and therefore it was predicted that N400/LPP amplitude would be evoked when making positive interpretations that were not in line with consistent expectancy models. Therefore they would have increased amplitude when making dissimilar ratings of negative statements and similar ratings of positive statements. Finally, based on the N400 literature, a basic prediction for the main effect of similarity was made; it was suggested that when averaged across both statement types, N400 amplitude would be increased when making dissimilar ratings compared with making similar ratings.

7.2.1 Research questions and hypotheses

Overarching research questions:

- Does the recognition task detect differences in interpretation bias between aggression groups using both behavioural and EEG methods?
- Do high aggression participants have increased negative interpretation bias compared with low aggression participants, and is this reflected in different ERP patterns in response to negative and positive statements between aggression groups?

Research questions and hypotheses:

I. Are the results consistent across implicit and explicit measures of interpretation bias?

Hypothesis 1a: Behavioural interpretation bias scores on the recognition task (implicit) and AIHQ scores across all subscales (explicit) will positively correlate.

Hypothesis 1b: Participants that have a lower score on AIHQ (an explicit measure of interpretation bias) will show increased N400/LPP amplitude in response to negative statements compared to positive, whereas those with a higher score on AIHQ will show relatively undifferentiated N400/LPP amplitude in response to both statements.

II. Do participants with increased levels of aggression show an increased hostility-related interpretation bias across explicit (AIHQ) and implicit (recognition task) measures?

Hypothesis 2a: A greater explicit interpretation bias, reflected by an increased score on the AIHQ, will be positively correlated with aggression.

Hypothesis 2b: When asked to rate the similarity between ambiguous scenarios and positive and negative statements, individuals with an increased aggression score will rate negative statements on the recognition task as more similar in meaning to the ambiguous scenarios compared to the positive statements, reflected in a greater target bias score,

Hypothesis 2c: Target bias (calculated from the similarity ratings in response to target statements) on the recognition task will correlate with aggression; however foil bias (calculated from the similarity ratings in response to foil statements) will not.

III. Are there differences between higher and lower aggression groups in evoked N400/LPP amplitude when responding to positive and negative statements during the recognition task?

Hypothesis 3a: Low aggression individuals will show increased N400/LPP amplitude when responding to negative statements compared to positive statements.

Hypothesis 3b: High aggression individuals will show similar N400/LPP amplitude when responding to both negative and positive statements.

IV. Are there differences between higher and lower aggression groups in evoked N400/LPP when making similar and dissimilar ratings of positive and negative statements during the recognition task?

Hypothesis 4a: Low aggression participants will not have a negative interpretation bias therefore they will show increased N400/P600 amplitude when making similar ratings of negative statements and dissimilar ratings of positive statements.

Hypothesis 4b: High aggression participants will show evidence of a negative interpretation bias, and therefore will have increased N400/P600 amplitude when making dissimilar ratings of negative statements and similar ratings of positive statements.

7.3 Methods

7.3.1 Participants and procedures

Data were collected from 36 male University of East Anglia (UEA) students and staff, and members of the wider community. These participants were recruited as part of a larger research project in which they completed three tasks; a dot-probe word task (results of this are reported in Chapter 3), a dot-probe face task (results of this are reported in Chapter 5) and finally the recognition task reported in this current chapter. Therefore a full description of the sample (see Section 3.3.2) and procedures (see Section 3.3.7) can be found in Chapter 3.

For the recognition task, one participant was ineligible due to their first language not being English (the recognition task requires a relatively high standard of English language comprehension) and was therefore excluded from analysis. A further two participants were excluded due to excessive noise during EEG recording. Therefore for all continuous analyses conducted in this chapter, the final sample consisted of 33 participants (M = 21.77, SD = 4.55). For the first set of analyses, participants were categorised into two groups based on median split of scores achieved on the AIHQ (one participant scored the median resulting in16 participants with a low interpretation bias score and 16 with a high interpretation bias score). For further analyses participants were categorised into high and low aggression groups based on the total aggression score (one participant scored the median resulting in16 participants with low aggression scores and 16 with high aggression scores.

7.3.2 Self-report measures

The current chapter describes an ERP study that was conducted as part of a larger project consisting of a number of studies outlined within this thesis. The overall project aimed to investigate the influence of aggression on both attention bias (chapters 3-6) and interpretation bias. The ERP interpretation paradigm was the last of three computerised tasks that participants completed during the lab session (participants also completed two short dot-probe tasks as a measure of attention bias). Participants completed the following questionnaires; Aggression Questionnaire (AQ; (Buss & Perry, 1992), Ambiguous Intentions Hostility Questionnaire (AIHQ; (Combs et al., 2007), Attentional Control Scale (ACS; (Derryberry & Reed, 2001), Delinquency Questionnaire (DQ; taken from (Tarry & Emler, 2007), and Trait form of the State-Trait Anxiety Inventory (STAI-T; (Spielberger et al., 1983). Information on the AIHQ is presented in this chapter; full information of all questionnaire measures can be found in Chapter 3, Section 3.3.3).

7.3.2.1 ***Ambiguous Intentions Hostility Questionnaire*** (AIHQ; Combs, Penn, Wicher, & Waldheter, 2007) ***(Appendix X).***

This measure is used as an explicit measure of Interpretation Bias. Participants are presented with 15 scenarios, with a sub-set of five scenarios measuring either; intentional, ambiguous or accidental subscales. The participants are asked to respond to five questions relating to each scenario. The first asks them to state the real reason the person behaved in the specific way described. Question *A* is an open question and is rated by the researcher on a scale of 1 to 5 for *hostility* of the perceived intention behind the other person's behaviour in each scenario. Question *B* requires participants to respond on a 6 point Likert scale whether they think the actions described in the scenario were carried out with purpose intent. Questions *C* and *D* ask participants to rate how angry it makes them feel, and to indicate how much they would blame the person for the behaviour on a scale from 1 to 5. Questions B to D are summed to create a *blame* rating. Finally, question 'E' asks participants to write down what they would do in response to the described scenario. The stated behaviour of the participants towards the other person/situation in the scenario was rated by the researcher for *aggression* using a 5 point scale. Ratings for questions A and E for each of the scenarios were rated by the researcher. These items were also coded by a second researcher for a quarter of the sample (18/72; 72 was the total sample who completed the questionnaire as part of the larger study). The intraclass correlation of the sum of the rated items was calculated showing relatively high internal reliability (18 items; $\alpha = 0.79$). A total

AIHQ score was calculated by totalling the score to all questions of each subscale. A higher overall score indicated an increased level of hostile attribution bias.

7.3.3 Recognition Task

Implicit interpretation bias was measured using a recognition task (Mathews & Mackintosh, 2000) (see Appendix W). Initially the participants read twenty ambiguous scenarios designed to evoke hostile attributions. Next, four sentences are shown to the participants in relation to each scenario. Two sentences describe possible (target) items and two sentences describe non-relevant (foil) items. Foil items are used as a control; interpretation bias effects on foil items were not expected. There is one negative (hostile) and one positive interpretation of the scenarios for each of the target and foil items. Participants are asked to indicate how similar in meaning (on a scale of one to four, where one indicates ‘very different in meaning’ and four indicates ‘very similar in meaning’) each of the statements is to the scenario they previously read. When reading the scenario participants are encouraged to imagine themselves in the situation and how they would feel, therefore responses reflect interpretation of the ambiguous scenarios. Rating negative target statements as more similar in meaning to the scenarios compared with the positive target statements reflects a more negative interpretation bias. The task was split into two blocks, such that each block consisted of ten scenarios.

7.3.4 EEG Acquisition

The School’s EEG laboratory protocol (Version 1.1, 24.02.15) was followed throughout to ensure safe and responsible administration of the procedure. The EEG was recorded with a 32-channel active electrode system (Brain Products GmbH) embedded in a nylon cap (10/10 system extended). An additional electrode was placed under the left eye in order to monitor vertical eye movements (lower EOG). The continuous EEG signal was acquired at a 500 Hz sampling rate using FCz as reference. The impedance was kept below 20 kΩ.

7.3.5 Data extraction from the Recognition task

7.3.5.1 Behavioural data

Mean responses for each type of sentence (negative target, negative foil, positive target, and positive foil) were extracted and averaged across both blocks, resulting in an overall mean response to negative statements and positive statements for targets and foils across all trials. A *target bias* was calculated by subtracting the mean response to negative statements from the mean response to positive statements. A *foil bias* was calculated by subtracting the mean response to negative foils from the mean response to positive foils. A minus score reflects a greater interpretation bias of hostility related stimuli (if negative statements are rated as 'more similar in meaning' (higher) than positive statements then this will give a negative bias score). Greater hostility bias reflects a negative interpretation of the scenarios.

During this task EEG was simultaneously recorded, therefore the presentation of the original task was modified slightly. During presentation of the scenarios, each line was displayed until the participant pressed the downward arrow to continue, when the next line of the scenario was then displayed. To standardise reading speed the four response statements were presented between one and three words at a time, in five separate presentations of 500ms each. ERP data were therefore time-locked to presentation of the last word of the sentence. This allows for the measurement of an accurate representation of brain activity during the time taken for the participant to make their similarity rating.

7.3.5.2 EEG data

Offline analyses were conducted using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014), two open source toolboxes running under Matlab 7.12 (R2013a, The Mathworks). High- and low-pass filter half-amplitude cut-offs were set at 0.1 and 40 Hz, respectively. Before averaging, trials contaminated by excessive artifacts were rejected automatically with a step function (Luck, 2005) with a voltage threshold of ± 100 μV in moving windows of

200ms and with a window step of 100ms. Noisy channels were interpolated using the EEGLAB function eeg_interp (spherical interpolation). The data was not re-referenced offline.

The EEG was segmented into epochs of 1000ms (from -200ms prior to, to 800ms after presentation of the final word of each sentence). Data was locked to the last word of each statement (negative/positive). Mean amplitude between 200-300ms, 300-400ms, 400-500ms, 500-600ms, 600-700 and 700-798ms post stimulus onset were extracted for statistical analyses. Data was extracted from a posterior subset of electrode sites including CP1/2, CP5/6, P7/8, P3/P4 and TP9/10. The EEG analyses were conducted for target statements only; interpretation bias effects for foil items were not expected (confirmed by the behavioural analysis), therefore I was interested in evoked amplitude when participants made similarity ratings of positive and negative target statements. For the EEG analyses the 'similarity ratings' were categorised into two conditions; dissimilar (rating one and two) and similar (rating three and four). Therefore, for each electrode, the mean amplitude for the four possible response outcomes were extracted; negative statement and similar rating, negative statement and dissimilar rating, positive statement and similar rating, and positive statement and dissimilar rating. This allowed for the comparison of evoked amplitude in response to differently valenced statements, and investigate whether there is evidence of a processing bias when individuals with increased aggression score make 'similar' ratings of negative statements (hostility-related interpretation bias).

7.3.6 Data analysis plan

The behavioural data was explored using both a correlational and between-subject approach. Pearson correlations were conducted to the relationship between interpretation bias scores on the recognition and AIHQ, and aggression. A repeated measures two (bias type; target and foil) by two (aggression; high and low) ANOVA was also conducted to explore the difference in *target bias* and *foil bias* between low and high aggression samples.

The ERP data was explored using between-subject analyses. The sample was categorised based on a median split of AIHQ scores and aggression (see Section 7.4.2). Firstly, the evoked amplitude in response to negative and positive target statements was explored across both high and low AIHQ and aggression groups. Secondly, it was explored whether amplitude of high and low aggression samples differ depending on their similarity ratings of the positive and negative target statements. Target statement type (positive versus negative trials), response (similar versus dissimilar), electrode (5 levels) and hemisphere (left versus right) were included as within-subject factors. Total aggression score and AIHQ score were added as a between-subject factor. Greenhouse-Geisser (Geisser & Greenhouse, 1958) *F* test is reported throughout for all repeated measures to ensure there are no violations of the sphericity assumption.

A more detailed analysis plan for each hypothesis can be found in Appendix Y.

7.4 Results

7.4.1 Data preparation

7.4.1.1 Missing Items

The BPAQ (physical aggression subscale) had one case of missing data and the AIHQ had five missing items (ambiguous subscale). The missing values were replaced with the mean of the completed items for each appropriate measure (as in (Judah et al., 2014). This simple approach was selected as it is considered to make relatively little difference if missing data represent less than 5% of the dataset (Tabachnick & Fidell, 2013).

7.4.1.2 Distribution of data

All aggression data (BPAQ and subscales) was normally distributed. The AIHQ was assessed for normality; skewness and kurtosis were divided by their corresponding standard error. The calculated statistic was between acceptable limits of ± 2 (Field, 2013), therefore parametric tests were conducted.

Response variables for each of the statement types on the recognition task (negative target, negative foil, positive target, and positive foil) were assessed for normality. Positive and negative targets were normally distributed; however, the foil items were positively skewed and were therefore not normally distributed. Interpretation bias scores for both foils and targets were also assessed for normality. The calculated statistic showed that *target bias* was normally distributed, however *foil bias* was not. Although there were two outliers in the calculated foil bias score these were not adjusted as the data is based upon a numerical key press response and not reaction time. Parametric tests were conducted for analyses on positive and negative target statements and the calculated target bias. Whereas, non-parametric tests were conducted for analyses on foil statements and the calculated foil bias.

7.4.1.3 Reliability of questionnaires

The BPAQ (α = .92), physical aggression subscale from BPAQ (α = .90), anger subscale from BPAQ (α = .81), hostility subscale from BPAQ (α = .88), DQ (α = .81), and STAI-T (α = .94) demonstrated good internal reliability. The AIHQ was internally reliable with a Cronbach's alpha of 0.92. The ambiguous (α = .85), intentional (α = .84), and accidental (α = .81) subscales from the AIHQ also displayed good internal reliability. The verbal aggression subscale from the BPAQ (α = .77) was moderately reliable.

7.4.2 Descriptive Results

7.4.2.1 Aggression Questionnaire (Buss & Perry, 1992)

The sample was categorised based on a median split of the total aggression score of the Aggression Questionnaire (Buss & Perry, 1992) (Median = 75, range = 87). A median split was used as this is not affected by outliers. Any participants scoring the median score were not included in the analysis. The high total aggression group (M = 89.26, SD = 13.45) significantly differed from the low total aggression group (M = 57.13, SD = 10.55; $t(32) = 7.519$, $p < .001$).

For exploratory analyses the participants were also categorised based on a median split of the *physical aggression* subscale (Median = 19.0, Range = 31) and *verbal aggression* subscale (Median = 14, range = 16). The high physical aggression group (M = 28.44, SD = 5.32) significantly differed from the low physical aggression group (M = 13.58, SD = 3.09; $t(32) = 9.667$, $p < .001$). The high verbal aggression group (M = 18.80, SD = 2.62) significantly differed from the low verbal aggression group (M = 11.15, SD = 2.15; $t(28) = 8.343$, $p < .001$).

7.4.2.2 Ambiguous Intentions Hostility Questionnaire

Analyses were conducted to investigate the relationship between explicit and implicit measures of interpretation bias. To do this the participants were categorised into two groups based on median split of scores achieved on the AIHQ (Median = 186, range = 96), where higher scores reflected a higher level of hostility related interpretation bias (high AIHQ: M = 205.44, SD = 14.63; low AIHQ: M = 164.38, SD = 14.68). Both groups consisted of 16 participants (one participant scored the median and therefore could not be categorised).

7.4.2.3 Recognition task

Exploratory descriptive analyses were conducted to explore whether participants responded differently to positive and negative statements. A paired samples t-test revealed significant differences between *negative targets* (M = 2.75, SD = 0.45) and *positive targets* (M = 3.20, SD = 0.29); $t(32) = -5.03$, $p < .001$, and a Wilcoxon rank test showed significant differences between *negative foils* (M = 1.48, SD = 0.35) and *positive foils* (M = 1.70, SD = 0.38); $Z(32) = -3.89$, $p < .001$. Furthermore there was an overall effect of statement type such that targets (both negative and positive) (M = 2.97, SD = 0.27) were rated as more similar in meaning to the ambiguous scenario, compared to foils (M = 1.59, SD = 0.32); $t(32) = 19.24$, $p < .001$. These analyses show that there are significant differences in the ratings of each of the statements and therefore the design of the task is effective in measuring interpretation bias. The ERP analyses focused on targets (negative and positive) only.

7.4.2.4 State-Trait Anxiety Inventory (Spielberger et al., 1983)

Total aggression significantly correlated with anxiety (r = .412, p = .017), however interestingly physical aggression was not correlated with anxiety (r = .098, p = .588). To explore the possible confound of anxiety on the relationship between total aggression score and interpretation bias, the effect of anxiety on interpretation bias was investigated. Pearson's correlations showed a significant relationship between anxiety and interpretation bias for *targets* (r = -.474, p = .005). The correlation between anxiety and an interpretation bias for *foils* approached significance (r = -.326, p = .064). This shows that there may be a valence effect for both targets and foils. However independent samples t-tests revealed no significant differences in interpretation bias for *targets* between high anxiety (M = 0.29, SD = 0.53) and low anxiety participants (M = 0.54, SD = 0.45);

t(30) = -1.39, *p* = .175. There was also no significant difference in interpretation bias for *foils* between high anxiety (*M* = 0.18, *SD* = 0.44) and low anxiety participants (*M* = 0.26, *SD* = 0.19); *t*(30) = -0.600, *p* = .553). Due to the non-significant between-subject effects, anxiety was not included as a covariate in the following analyses.

7.4.3 Results relating to hypotheses

7.4.3.1 Hypothesis one

7.4.3.1.1 *Behavioural*

The total AIHQ score significantly correlated with interpretation bias score for targets on the recognition task (r = -.540, p = .001). Target bias also significantly negatively correlated with all subscales of the AIHQ; accidental scenarios (r = -.398, p = .022), ambiguous scenarios (r = -.435, p = .011), and intentional scenarios (r = -.521, p = .002). This shows that measures of implicit and explicit interpretation bias are consistent. AIHQ or any of the subscales did not correlate with an interpretation bias to foils (p > .089). This shows support for hypothesis 1a and provides evidence of concurrent validity of both measures.

7.4.3.1.2 *ERP*

An ANOVA was conducted to explore the possible interaction between the explicit measure of interpretation bias (AIHQ) and evoked amplitude in response to positive and negative target statements on the recognition task. Analyses were conducted for each epoch.

The ANOVA revealed a significant main effect of statement valence between 200 and 300ms, F(1,30) = 7.60, p = .010, η_p^2 = .202; 300 and 400ms, F(1,30) = 7.76, p = .009, η_p^2 = .206; and 400 and 500ms F(1,30) = 4.17, p = .050, η_p^2 = .122. Positive statements evoked increased positive amplitude compared to negative statements.

There was also a significant interaction between statement valence and AIHQ between 200 and 300ms, F(2,30) = 3.31, p = .050, η_p^2 = .181; and 300 and 400ms, F(2,30) = 3.64, p = .038, η_p^2 = .195. Post-hoc analyses between 200 and 300ms revealed that the main effect of statement valence was significant in the low AIHQ group, F(1,15) = 6.35, p = .024, η_p^2 = .297, but not in the high AIHQ group. Similarly, between 300 and 400ms there was a significant effect of statement valence in the low AIHQ group, F(1,15) = 13.33, p = .002, η_p^2 = .471, but not in the high AIHQ group. There were no significant effects between 500 and 600ms, or 600 and 700ms.

Finally, between 700 and 800ms there was a significant interaction between statement valence, electrode and AIHQ, $F(8,120) = 2.35$, $p = .049$, $\eta_p^2 = .135$. Post-hoc analyses revealed that in the low AIHQ group there was a significant main effect of statement valence, $F(1,15) = 9.46$, $p = .008$, $\eta_p^2 = .387$; and a significant interaction between statement valence and electrode, $F(4,60) = 3.28$, $p = .039$, $\eta_p^2 = .179$. The effect of statement valence was significant at TP10, $F(1,15) = 8.72$, $p = .010$, $\eta_p^2 = .368$; CP6, $F(1,15) = 11.36$, $p = .004$, $\eta_p^2 = .431$; P3, $F(1,15) = 5.30$, $p = .036$, $\eta_p^2 = .261$; and P4, $F(1,15) = 6.80$, $p = .020$, $\eta_p^2 = .312$. It also approached significance at TP9, $F(1,15) = 4.24$, $p = .057$, $\eta_p^2 = .220$; and P8, $F(1,15) = 3.58$, $p = .078$, $\eta_p^2 = .193$. There were no significant findings in the high AIHQ group.

Inspection of the waveform (Figure 36) suggest that participants with low scores on the AIHQ have increased amplitude to positive statements compared to negative. Although the waveform shows that differences in ERP patterns are robust and long lasting, significant effects were found between 200 and 500ms, and 700 and 800ms, suggesting that that interpretation bias effects may reflect a LPP-like component specifically. These results show some support for hypothesis 1b as the high AIHQ show group showed little differentiation between valenced stimuli (Figure 37); however the effect in the low AIHQ group was in the opposite direction to that hypothesised.

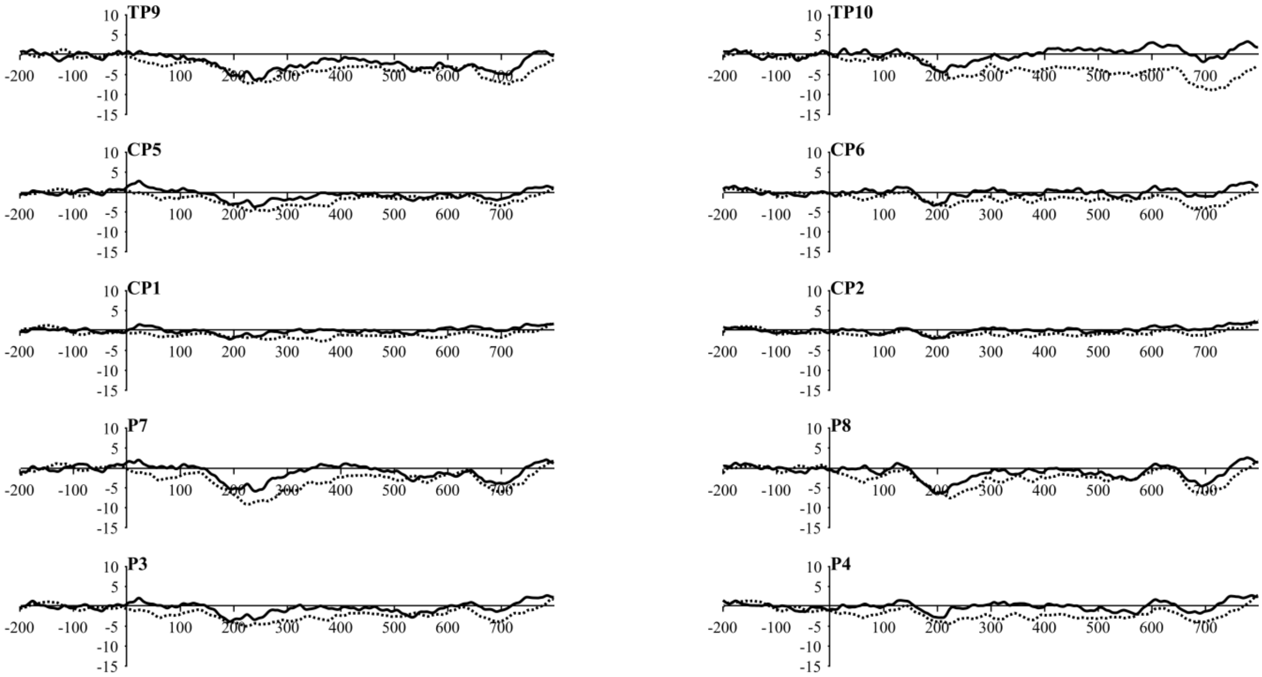

Figure 36: Grand average ERPS for evoked amplitude in response to positive statements (black) compared to negative statements (dotted) in participants scoring low on the AIHQ (n = 16).

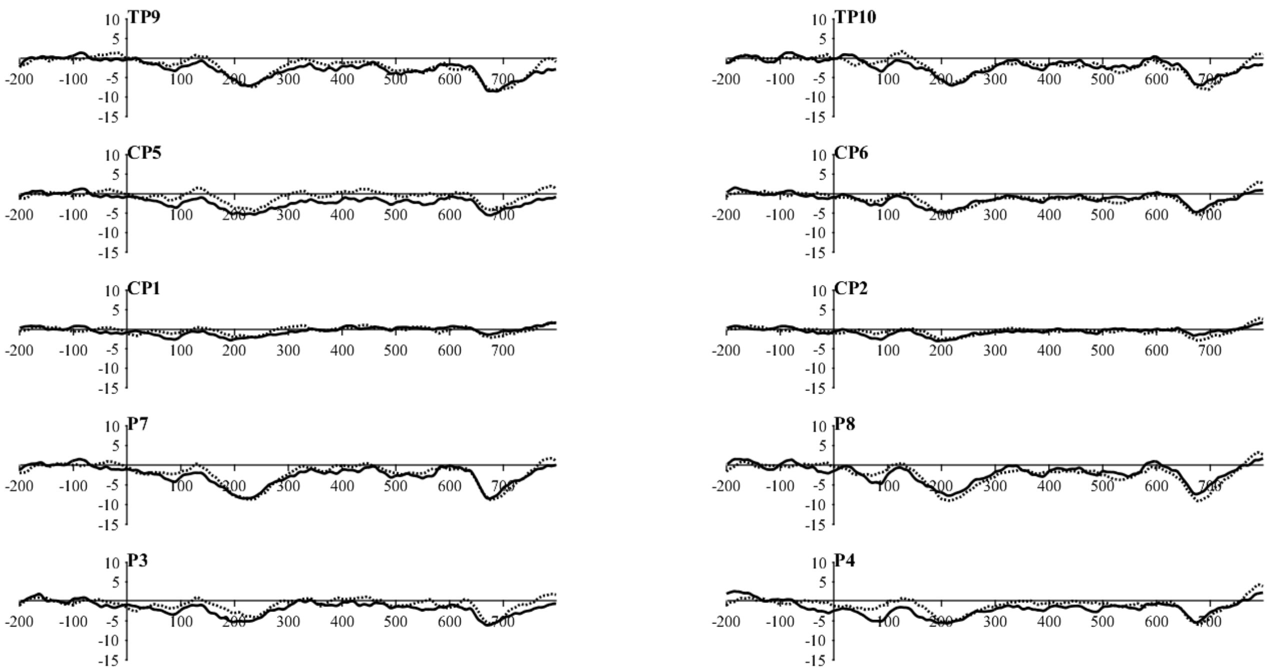

Figure 37: Grand average ERPS for evoked amplitude in response to positive statements (black) compared to negative statements (dotted) in participants scoring high on the AIHQ (n = 16).

7.4.3.2 Hypothesis two

7.4.3.2.1 *AIHQ*

Total AIHQ significantly positively correlates with total aggression score (r = .637, p < .001) and all subscales of aggression; physical aggression (r = .462, p = .007), verbal aggression (r = .702, p < .001), anger (r = .474, p =.005) and hostility (r = .445, p = .009). The AIHQ subscales also frequently correlated with aggression subscales (Table 12). This suggests that individuals with increased levels of aggression also had increased scores on the AIHQ and therefore there is support for hypothesis 2a.

Table 13: Pearson's correlations between Aggression and AIHQ subscales

	Total Aggression	**Physical Aggression**	**Verbal Aggression**	**Anger**	**Hostility**
Ambiguous	.529 (.002)	.492 (.004)	.499 (.003)	.328 (.062)	.339 (.053)
Intentional	.552 (.001)	.399 (.021)	.681 (<.001)	.412 (.017)	.347 (.048)
Accidental	.527 (.002)	.247 (.165)	.600 (<.001)	.475 (.005)	.457 (<.008)

7.4.3.2.2 *Recognition task*

Note: Only the results based on the total aggression score are presented here. Exploratory results were relatively consistent across all aggression subscales (physical aggression, verbal aggression, anger and hostility); therefore it was decided to focus on the effects of the composite total aggression score.

A repeated measures ANOVA was conducted to explore the difference between target and foil bias in the high and low aggression groups. The ANOVA revealed a significant main effect of bias, $F(1,32) = 369.99$, $p < .001$, $\eta_p^2 = .920$; and a significant interaction between bias and aggression, $F(1,32) = 7.65$, $p = .009$, $\eta_p^2 = .193$. Post-hoc tests were conducted to explore whether significant differences in aggression occurred for target or foil bias. Results of an independent samples t-test showed that the mean *target bias* score for high total aggression (M = 0.18, S.D

= 0.50) and low total aggression (M = 0.68, S.D = 0.39) significantly differed; $t(32) = -3.176$, $p = .003$). Mann Whitney U tests revealed that there was no significant differences in *foil bias* between high total aggression (M = 0.19, S.D = 0.43) and low total aggression (M = 0.24, S.D = 0.22); $U(32) = 116.5$, $p = .664$). Further to this, total aggression significantly positively correlated with ratings of negative target statements ($r = .358$; $p = .041$), and negatively correlated with ratings of positive target statements ($r = -.589$; $p < .001$). This suggests that individuals with increased aggression score rate negative statements as more similar in meaning and rate positive statements as more dissimilar in meaning. This shows support for hypothesis 2b.

Crucially, these findings were confirmed by correlation analyses. Pearson's correlation showed that *target bias* and total aggression score significantly negatively correlated ($r = -.640$; $p < .001$) (Figure 38). Spearman's correlation results showed that *foil bias* did not correlate with total aggression ($r = -.091$; $p = .614$). These results show support for hypothesis 2c and suggest that those individuals scoring higher on aggression had a more negative bias for targets (hostility related bias), showing that they rated the negative targets as some similar in meaning to the scenario compared to positive statements. As expected there were no significant differences in foil bias across aggression groups.

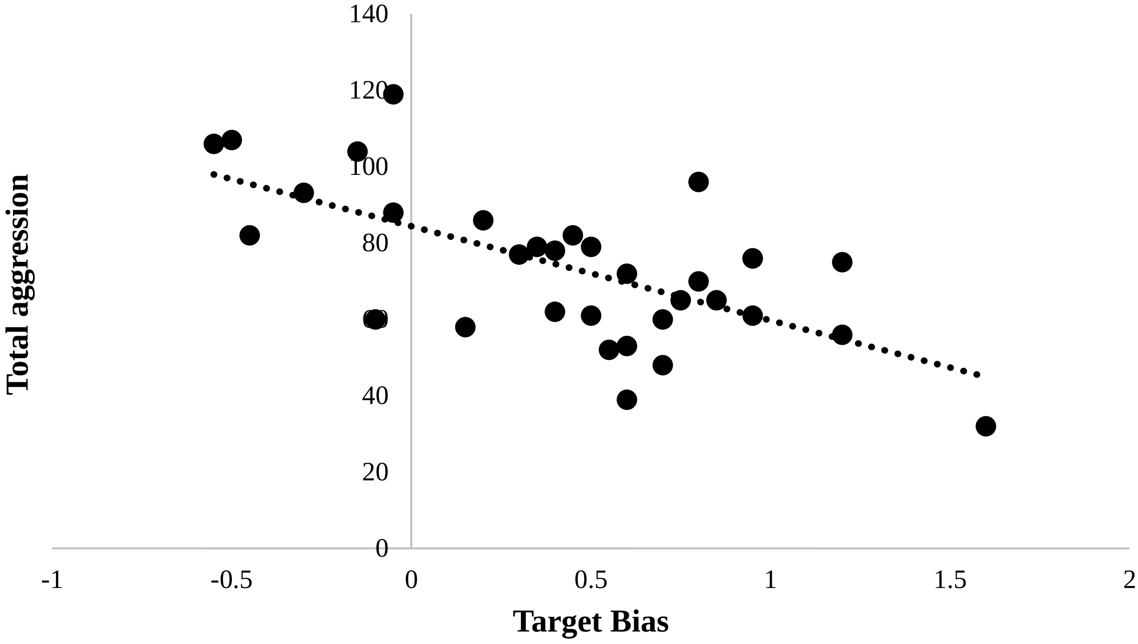

Figure 38: Scatterplot and regression line (r = -.640; p < .001) to show the correlation between target bias and total aggression score (n = 33).

7.4.3.3 Hypothesis three

Only the total aggression results are reported here. This is due to the behavioural result being significant across all subscales of aggression and therefore it was decided to present the analysis of EEG data using the composite score of all aggression items.

7.4.3.3.1 *Effect of statement type*

The mixed model ANOVA showed a significant main effect of statement valence between 300 and 400ms, $F(1,32) = 6.73$, $p = .015$, $\eta_p^2 = .183$. This effect also approached significance between 200 and 300ms, $F(1,32) = 3.97$, $p = .056$, $\eta_p^2 = .117$; and between 700 and 800ms, $F(1,32) = 3.99$, $p = .055$, $\eta_p^2 = .117$. Inspection of the waveform (Figure 39) revealed that there was increased positive amplitude in response to positive target statements compared to negative target statements, across all participants.

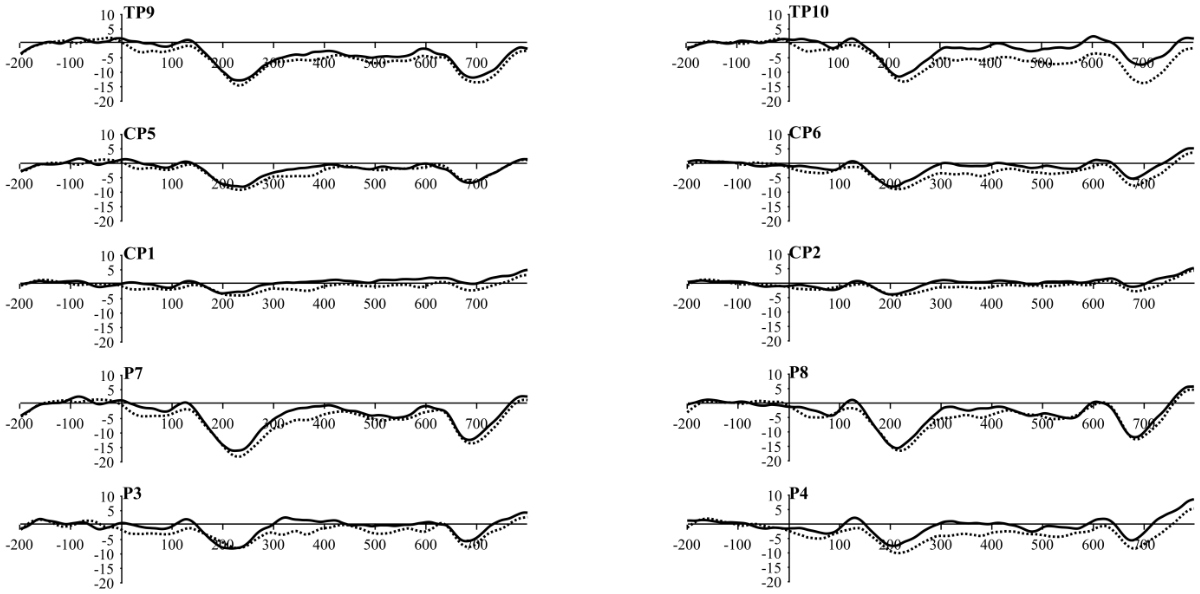

Figure 39: Grand average ERPS for positive statements (black) compared to negative statements (dotted) across the whole sample of participants (n = 33).

7.4.3.3.2 *Interaction between statement type and aggression*

There was a significant interaction between statement valence, electrode, hemisphere and aggression group between 200 and 300ms, $F(4,32) = 4.66$, $p = .008$, $\eta_p^2 = .135$; and between 400 and 500ms, $F(4,32) = 3.29$, $p = .045$, $\eta_p^2 = .099$. This interaction also approached significance between 500 and 600ms, $F(4,32) = 2.62$, $p = .071$, $\eta_p^2 = .080$.

To explore the interaction between statement valence, electrode, hemisphere and aggression group in each epoch, post-hoc analyses were conducted to explore the effect of statement valence on each electrode in the high and low aggression group. Between 200 and 300ms there was a significant effect of statement valence at P4 only, $F(1,16) = 4.75$, $p = .046$, $\eta_p^2 = .241$, in the high aggression group. In the low aggression group the effect of statement valence was significant at P7, $F(1,16) = 4.66$, $p = .048$, $\eta_p^2 = .237$; and approached significance at P8, $F(1,16) = 4.05$, $p = .062$, $\eta_p^2 = .213$; P3, $F(1,16) = 4.05$, $p = .063$, $\eta_p^2 = .212$; and P4, $F(1,16) = 4.18$, $p = .059$, $\eta_p^2 = .218$. Between 400 and 500ms there were no significant effects of statement valence at any electrode site in the high aggression group. In the low aggression group the effect of statement valence approached significance at TP10, $F(1,16) = 3.41$, $p = .085$, $\eta_p^2 = .185$. Between 500 and 600ms the effect of statement valence approached significance at P4, $F(1,16) = 3.56$, $p = .079$, $\eta_p^2 = .192$ in the high aggression group.

It was hypothesised that low aggression participants would show increased LPP amplitude in response to negative statements. However, results suggest that low aggression participants show increased amplitude in response to positive statements, compared to negative. (Figure 40) Therefore, these results show no support for hypothesis 3a.

The evidence in relation to hypothesis 3b is somewhat inconclusive; it was hypothesised that high aggression participants would show relatively stable evoked amplitude in response to both positive and negative statements. However the results show that the effect of statement is significant at P4 between 200 and 300ms in the high aggression group, and approached significance at the same electrode between 500 and 600ms (Figure 41). The effect of statement is perhaps more robust in the low aggression sample as it is significant in a greater number of electrodes, however the effect sizes of significant effects are relatively similar across high and low aggression groups.

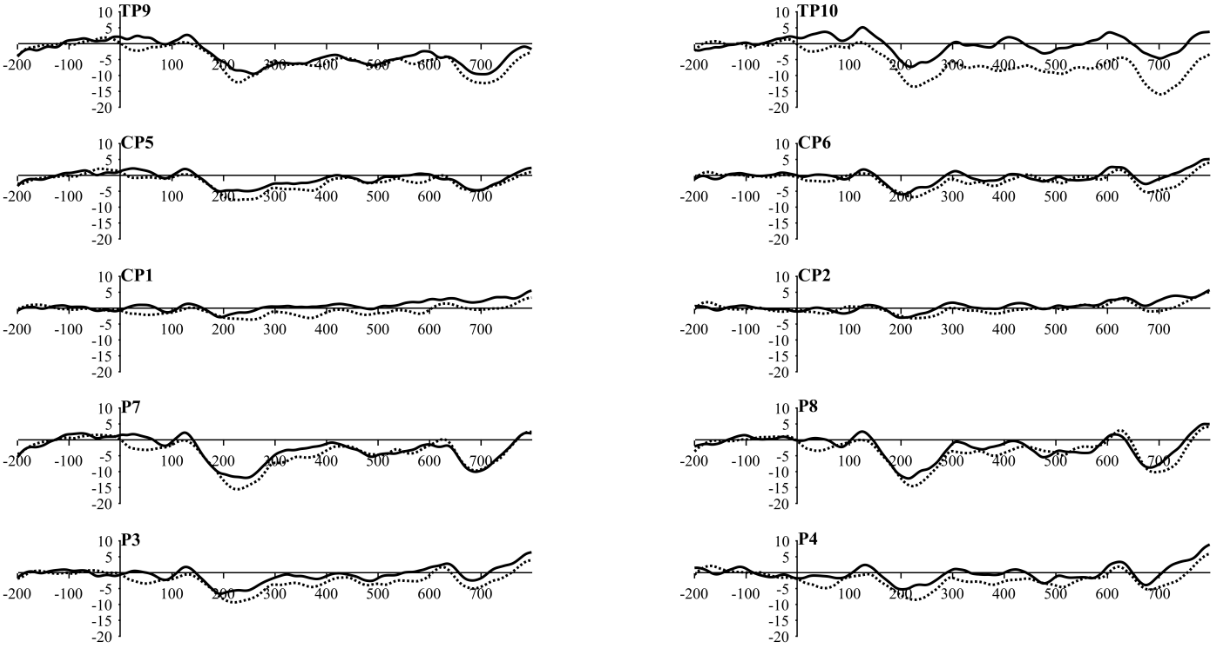

Figure 40: Grand average ERPS for positive statements (black) compared to negative statements (dotted) in low aggression participants (n = 16).

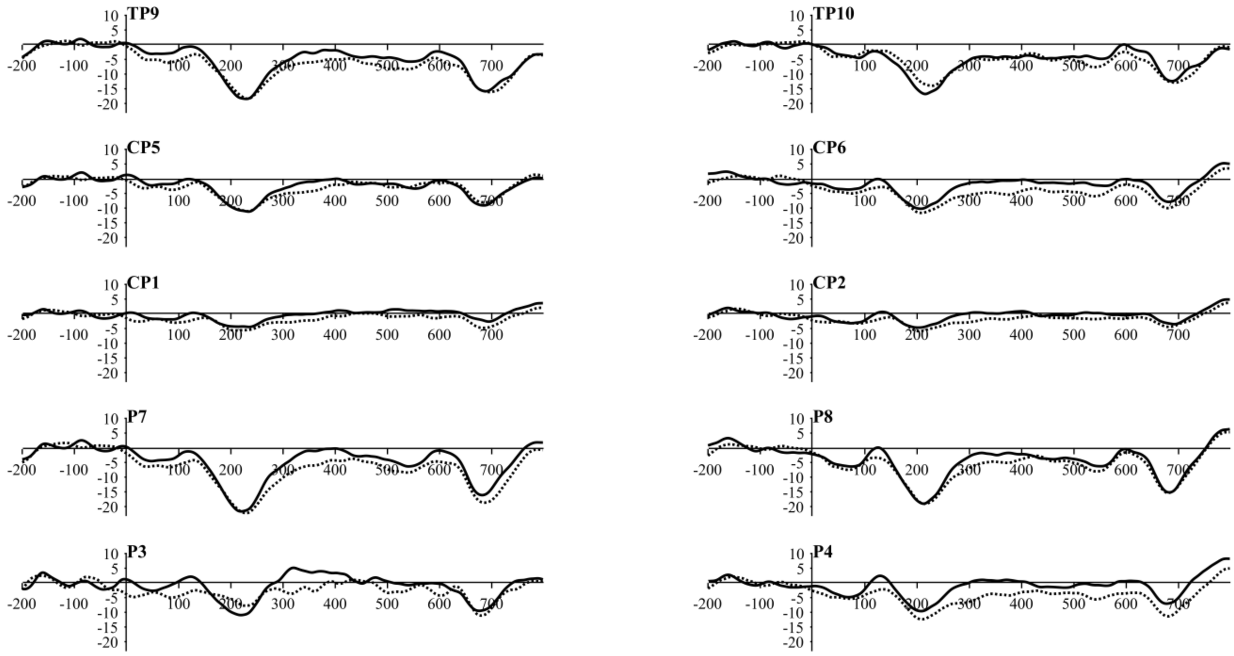

Figure 41: Grand average ERPS for positive statements (black) compared to negative statements (dotted) in high aggression participants (n = 16).

7.4.3.4 Hypothesis four

7.4.3.4.1 *Main effect of similarity ratings*

Initially analyses were conducted to measure the N400; it was predicted that participants would show increased N400 amplitude in response to all dissimilar ratings compared with similar ratings, across both statement types. To do this, mean amplitude between 300ms and 500ms was extracted for each similarity rating (similar/dissimilar), across both statement types (positive and negative). An ANOVA was conducted to explore whether there was any difference in amplitude when participants made similar versus dissimilar ratings across both negative and positive statement types. Similarity rating (2 levels: similar and dissimilar), electrode (6 levels) and hemisphere (2 levels) were added as within subject factors. Total aggression was added as between-subject factors.

Surprisingly, a standard N400 effect was not observed in this paradigm. The results revealed no significant effects which suggests there were no significant diffrences in amplitude when participants made similar and disimilar ratings of statements. These null results could be attributed to the significant interactions found between statement type and similarity ratings (as discussed below), and therefore may overlap with a simulaneously occuring positive component.

7.4.3.4.2 *Interaction between similarity rating (response) and aggression*

Results from the omnibus ANOVA showed a significant interaction between statement valence, response, electrode, and aggression group between 400 and 500ms, $F(4,32) = 3.19$, $p = .044$, $\eta_p^2 = .096$; and also approached significance between 300 and 400ms, $F(4,32) = 2.62$, $p = .078$, $\eta_p^2 = .080$; and 700 and 800ms**,** $F(4,32) = 2.80$, $p = .069$, $\eta_p^2 = .085$. There was a significant interaction between statement valence, response, electrode, hemisphere and aggression group between 200 and 300ms, $F(4,32) = 4.92$, $p = .008$, $\eta_p^2 = .141$; this effect also approached significance between 400 and 500ms, $F(4,32) = 2.52$, $p = .075$, $\eta_p^2 = .078$; and 500 and 600ms, $F(4,32) = 2.40$, $p = .078$, $\eta_p^2 = .074$. Post-hoc analyses were performed to investigate these complex interactions with response type**.** To do this a one-way

ANOVA was conducted to explore the effects of response for negative and positive statements at each electrode site. Separate analyses were conducted for the high and low aggression groups.

Between 200 and 300ms there was no significant effect of response to negative or positive statements in the high aggression group. In the low aggression group there was no significant main effect of response to negative statements at any electrode sites. There was a significant effect of response to positive statements at TP10, $F(1,16) = 6.28$, $p = .024$, $\eta_p^2 = .295$. Inspection of the bar chart (Figure 42) shows that the low aggression group show greater distinction between making similar and dissimilar ratings of positive statements. They show increased negative amplitude when making similar compared to dissimilar ratings.

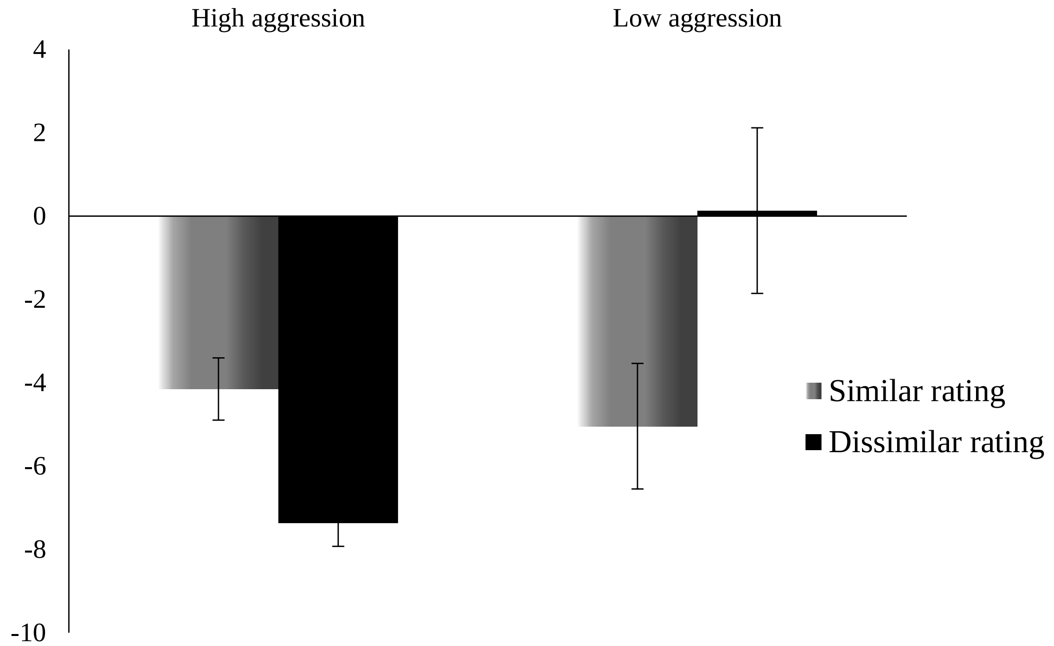

Figure 42: Bar chart to show the evoked mean amplitude at electrode TP10 when making similar and dissimilar ratings in response to positive statements in low (n = 16) and high (n = 16) aggression participants between 200 and 300ms post statement onset (error bars = ± 1 standard error).

Between 300 and 400ms there was no significant main effect of response to negative statements at any electrode sites in the high aggression group. However, there was a close significant effect of response to positive statements at TP10, $F(1,16) = 4.14$, $p = .060$, $\eta_p^2 = .216$; and CP6, $F(1,16) = 4.24$, $p = .057$, $\eta_p^2 = .220$. In the low aggression group there was a significant effect of response to negative statements at CP1, $F(1,16) = 5.83$, $p = .029$, $\eta_p^2 = .281$; and a significant effect of response to positive statements at TP10, $F(1,16) = 6.95$, $p = .019$, $\eta_p^2 = .316$.

Between 400 and 500ms, in the high aggression group there was a significant effect of response to negative statements at P7, $F(1,16) = 6.80$, $p = .020$, $\eta_p^2 = .312$. The main effect of response to positive statements was significant at TP10, $F(1,16) = 4.99$, $p = .041$, $\eta_p^2 = .250$; and approached significance at TP9, $F(1,16) = 4.02$, $p = .063$, $\eta_p^2 = .211$. In the low aggression group, the main effect of response to negative statements approached significance at CP1, $F(1,16) = 3.49$, $p = .081$, $\eta_p^2 = .189$; and the effect of response to positive statements was significant at TP10, $F(1,16) = 7.62$, $p = .015$, $\eta_p^2 = .337$.

Between 500 and 600ms, there was no significant effect of response to negative or positive statements at any electrode sites in the high aggression group. In the low aggression group there were no main effects of response to negative statements. However, the main effect of response to positive statements was significant at TP10, $F(1,16) = 5.91$, $p = .028$, $\eta_p^2 = .283$; and P7, $F(1,16) = 4.53$, $p = .050$, $\eta_p^2 = .232$; and approached significance at CP5, $F(1,16) = 3.79$, $p = .070$, $\eta_p^2 = .202$; CP1, $F(1,16) = 4.41$, $p = .053$, $\eta_p^2 = .227$; and P3, $F(1,16) = 3.47$, $p = .082$, $\eta_p^2 = .188$.

Between 700 and 800ms, in the high aggression group the main effect of response to negative statements approached significance at P7, $F(1,16) = 3.70$, $p = .074$, $\eta_p^2 = .198$. There were no significant main effects of response to positive statements. In the low aggression group there were no main effects of response to negative statements. However, the main effect of response to positive statements was close to significance at CP1, $F(1,16) = 3.31$, $p = .088$, $\eta_p^2 = .182$; and P7 ,$F(1,16) = 4.35$, $p = .054$, $\eta_p^2 = .225$.

These results suggest that across both low and high aggression groups, participants make some differentiations between making similar and dissimilar ratings of both positive and negative statements. However, results suggest that the low aggression group seem to make greater distinctions between similarity ratings, particularly in response to positive statements. This effect is most salient 500 and 600ms following statement presentation and may reflect a P600/LPP type ERP component. Specifically, in line with predictions, the results show that the low aggression group have increased positive amplitude when making dissimilar ratings of positive statements (Figure 43) and increased positive amplitude when making similar rating of negative statements (Figure 45). Therefore hypothesis 4a is supported.

Due to the robust behavioural association between aggression and negative interpretation bias it was predicted that high aggression participants would show increased N400/LPP amplitude when making similar ratings of positive statements. The results and inspection of the waveforms (Figure 44) suggests some evidence for this. However, they also show some evidence of increased amplitude when making similar ratings of negative statements, although this was only significant at electrode P7 (Figure 46). Inspection of the waveform shows that evoked amplitude when making similarity ratings of negative statements is variable across the region of interest. At TP10 the high aggression group show increased amplitude when making dissimilar ratings of negative statements, although this did not reach significance. These results show some support for hypothesis 4b, however due to the mixed evidence, subsequent conclusions are made with caution.

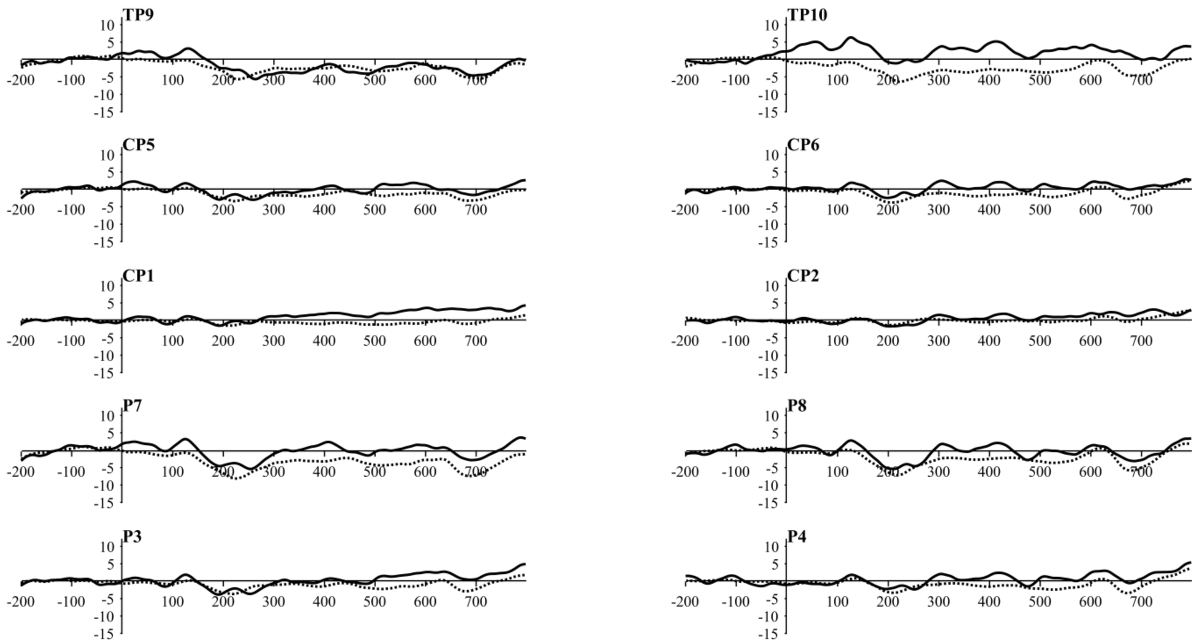

Figure 43: Grand average ERPS for the effect of response to positive statements in low aggression participants. Mean amplitude to dissimilar (black) and similar (dotted) response ratings are compared.

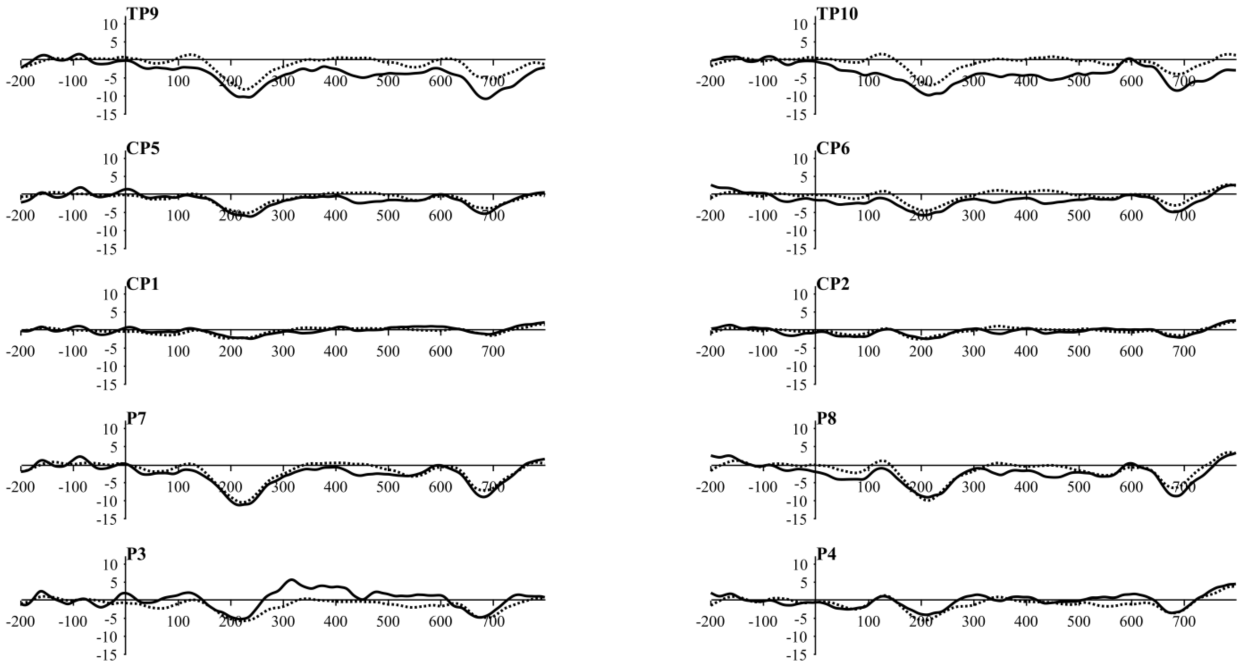

Figure 44: Grand average ERPS for the effect of response to positive statements in high aggression participants. Mean amplitude to dissimilar (black) and similar (dotted) response ratings are compared.

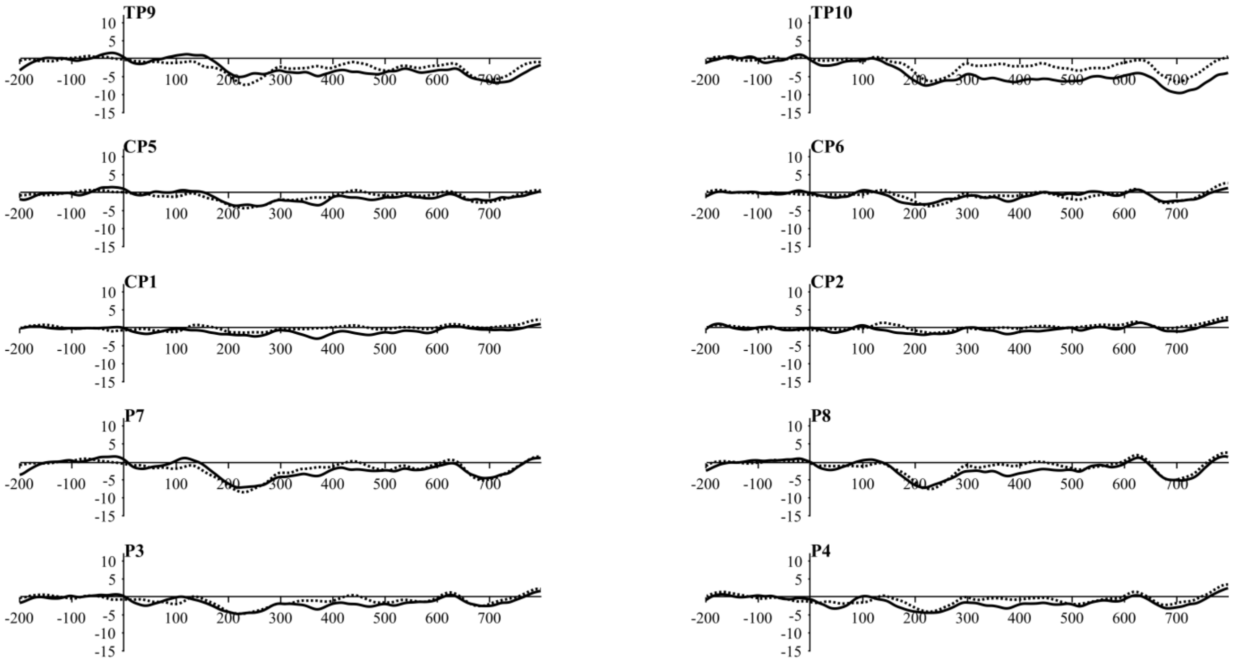

Figure 45: Grand average ERPS for the effect of response to negative statements in low aggression participants. Mean amplitude to dissimilar (black) and similar (dotted) response ratings are compared.

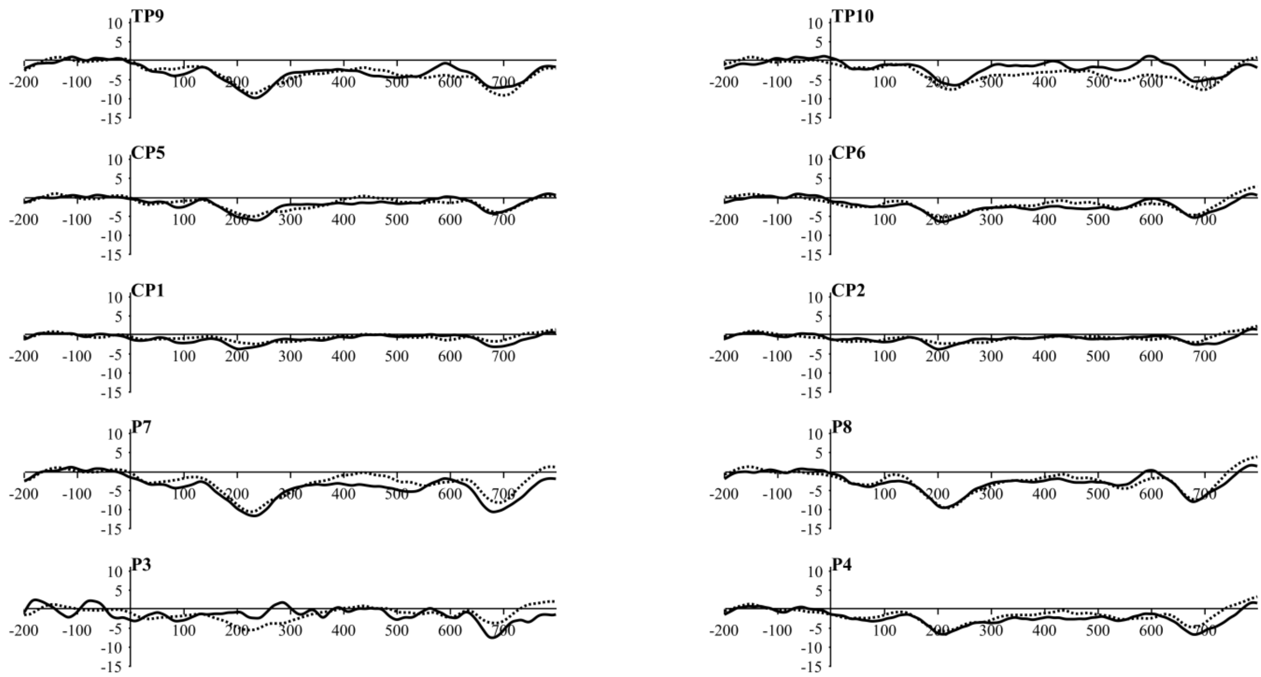

Figure 46: Grand average ERPS for the effect of response to negative statements in high aggression participants. Mean amplitude to dissimilar (black) and similar (dotted) response ratings are compared.

7.5 Discussion

This study investigated whether individuals with increased levels of aggression would show a hostile interpretation bias using two different measures; a frequently used explicit questionnaire measure, and an implicit experimental measure. The first of these measures was the AIHQ in which participants explicitly stated how they would behave in response to a provoking situation, the second task was an implicit recognition task in which participants made similarity ratings of positive and negative statements which related to previously presented ambiguous scenarios. In a unique contribution to the literature, EEG was also simultaneously measured during completion of the experimental measure in order to explore the ERP correlates of interpretation bias in aggression. It is unknown if there is a distinct ERP signature associated with hostility related biases in aggression. Due to the novelty of analysing brain processing during the recognition task (implicit interpretation bias) the concurrent validity of this measure was also of interest.

7.5.1 Main findings and interpretations

7.5.1.1 Hypothesis one

The first hypothesis concerned the comparison of the two measures of interpretation bias. Hypothesis 1a was supported as bias score on the recognition task was associated with scores on all subscales of the AIHQ. This suggests that both measures are sensitive to hostility-related interpretation bias. Hypothesis 1b made predictions regarding the ERP patterns associated with explicit hostile interpretation bias. Specifically, it was hypothesised that participants with high AIHQ scores would show relatively undifferentiated N400/LPP amplitude in response to positive and negative statements during the recognition task, whereas participants with lower AIHQ scores would show increased N400/LPP amplitude in response to negative statements compared to positive. This prediction was partially supported as there were no significant differences in evoked amplitude when participants with high AIHQ scores responded to negative and positive statements. These findings suggest that participants with a negative interpretation bias (as measured by an explicit questionnaire measure), process positive and negative statements similarly. The effect in the low AIHQ group was in the opposite direction to that hypothesised, such that they showed increased amplitude when responding to positive statements, compared to negative statements. Based on limited previous literature (Moser et al., 2008a), it was suggested that participants that show no evidence of a negative interpretation bias would not expect negative outcomes in the environment and therefore these expectancy violations would evoke increased amplitude. Moser et al. (2008a) found that during a sentence completion task, participants with low anxiety had a larger evoked P600 response to negative sentence resolutions compared to positive. It was suggested that non anxious participants have a positivity bias, such that they do not expect negative sentence resolutions, consequently they evoke increased amplitude. However, no support for this explanation was found. It is suggested that due to the lack of interpretation bias in the low AIHQ group, these participants' allocate more resources when processing positive statements.

To my knowledge this is the first study to explore possible neural correlates of hostile interpretation bias. These findings suggest that this bias is characterized by the allocation of similar resources in interpreting both positive and negative statements. This may reflect the efficiency with which participants interpret all ambiguous scenarios as hostile and are therefore less likely to differentiate between statement types. Due to the limited previous evidence, these suggestions are made cautiously and it is recognised that this effect will require replication. However, the current study suggests that participants scoring high and low on an explicit measure of interpretation bias (AIHQ) have different ERP patterns in response to positive and negative statements on the (implicit) recognition task. This suggests that ERPs are sensitive to processes associated with negative interpretations, and that the recognition task is a valid task for measuring interpretation bias using EEG methodology.

7.5.1.2 Hypothesis two

Hypothesis two was that participants with increased levels of aggression would show an increased interpretation bias across both behavioural measures (AIHQ and recognition task). There was evidence to support hypothesis 2a as increased aggression scores were positively correlated with increased AIHQ scores. This demonstrates that on an explicit measure of interpretation bias, participants scoring high on aggression made more hostile attributions of intent compared to low aggression participants. There was also support for hypotheses 2b and 2c; both between-subject and correlational evidence suggests that during the recognition task increased aggression was associated with a negative bias for targets. In comparison to participants with lower levels of aggression, participants with higher levels of aggression rated negative statements as more similar in meaning to the previously presented scenario compared to positive statements. These findings are consistent with previous literature (e.g. Dill et al., 1997; Dodge & Frame, 1982; Dodge et al., 1990; Epps & Kendall, 1995) and provides support for the association between aggression and hostility-related interpretation biases. These findings also suggest that these biases are robust across different methods. I used the AIHQ which is an explicit questionnaire measure, and the recognition task which measures interpretation bias at a relatively implicit level. Therefore, I conclude that aggression-related interpretation biases are evident, and measurable, under conditions of conscious awareness, but also occur automatically with little conscious control. This supports Wilkowski and Robinson's (2010) cognitive model of trait anger and reactive aggression according to which hostile interpretation bias in aggressive individuals is primarily reliant on automatic processes such that they occur spontaneously, efficiently and unconsciously.

7.5.1.3 Hypothesis three

Prior to this study there were only two studies that explored neural correlates of interpretation bias in aggression from which to base the current predictions (Gagnon et al., 2016; Gagnon et al., 2017). Therefore, work by Moser et al. (2008a) and Moser et al. (2012) which explored interpretation bias in anxiety using EEG methodology is also drawn upon. Moser et al. (2008a) found that highly anxious participants showed similar P600 amplitude in response to negative and positive sentence resolutions; however, participants with lower levels of anxiety demonstrated an increased P600 amplitude in response to negative sentence resolutions. It is suggested that hostile interpretation bias in aggression might be characterized by a similar ERP pattern and therefore it was hypothesised that there would be a significant main effect of statement type in the low aggression group but not in the high aggression group.

There was no support for hypothesis 3a, because the main effect in the low aggression group was in the opposite direction to that predicted. In the low aggression group, positive statements evoked increased positive amplitude compared to negative statements. A possible explanation is that individuals with low levels of aggression have a positivity bias in which they avoid allocating attention towards angry words and process positive information in greater detail compared to high aggression participants. These effects contrast with those found by Moser et al. (2008a), and the attention bias literature which suggests increased P300 amplitude to negative stimuli, compared to neutral words in healthy low aggression samples (Helfritz-Sinville & Stanford, 2015; Thomas et al., 2007). However, the recognition task used in the current research is perhaps not comparable to simpler tasks used in previous research. The recognition task is a more complex task used to infer implicit interpretation bias; the task requires a similarity rating in response to the valenced stimuli and therefore ERP effects of interpretation bias may be confounded by decision making processes.

Hypothesis 3b was that high aggression individuals would show similar evoked amplitude in response to positive and negative statements. There was some tentative support for this hypothesis; although the effect of statement was significant at electrode P4 between 200 and 300ms in the high aggression sample, the effect seemed more consistent in the low aggression sample as it was significant at more electrodes at several epochs (at electrode P7 at earlier epochs

(200-400ms) and TP10 at later epochs (400-600ms)). These findings suggest that compared with low aggression participants, high aggression participants differentiate less between positive and negative statements. This finding is consistent with work by Helfritz-Sinville and Stanford (2015) who used a modified oddball task including threat and neutral words to investigate P300 amplitude in attention bias. They reported similar P300 amplitude in response to presentation of both word types in the aggressive sample, whereas control participants exhibited enhanced amplitude to the threat words (social and physical) compared to neutral words. Due to the similarity in evoked P300 amplitude, it is proposed that high aggression participants perceive threatening words in a similar way to the neutral words. Current behavioural and ERP findings from the recognition task suggest that aggressive participants have increased interpretation bias in which they attribute hostile intent more frequently, and interpret scenarios more negatively compared to low aggression participants; however, this interpretation bias is reflected in fewer differences in evoked amplitude. Making negative interpretations of both positive and negative statements (rating positive statements as dissimilar in meaning to the scenario and negative statements as similar) is likely to demand a similar allocation of cognitive resources. This may be explained by desensitization of negative information that results in emotional processing deficits in individuals with high aggression (Helfritz-Sinville & Stanford, 2015).

7.5.1.4 Hypothesis four

To my knowledge there is no previous literature that has used the recognition task with simultaneous EEG recording to measure interpretation bias. Therefore, based on studies using different tasks and the expectancy models of the N400 and LPP, exploratory predictions regarding differences between aggression groups in evoked amplitude when making similar or dissimilar ratings of negative versus positive statements were made. Hypothesis 4a was that low aggression participants would show no evidence of a hostility-related interpretation bias, therefore making negative interpretations would violate their positive expectancy outcomes. Therefore I suggested amplitude would be increased when making similar ratings of negative statements and dissimilar ratings of positive statements. There was evidence to support this hypothesis; however response effects may have been more consistent for the positive statements compared with the negative statements. Findings suggest that amplitude is maximal when low aggression participants make dissimilar ratings of positive statements. The waveform shows that effects begin as early as 200ms after statement presentation and are long lasting; however the effect is maximal between 500 and 600ms and therefore may reflect a LPP-like ERP component. The findings are consistent with the expectancy account of the LPP. The LPP, or the P600 reflects cognitive processing of word tasks and is sensitive to expectancy violations, therefore it is evoked in response to semantic information which does not fit with current cognitive models (Coulson, 1998; Van Herten et al., 2005). It is suggested that low aggression participants will expect positive outcomes in social scenarios; therefore increased amplitude reflects updating of memory when individuals evaluate and interpret an expected negative situation (Coulson, 1998). On the recognition task the unexpected negative outcome could be when rating positive statements as dissimilar to the scenario, or when rating negative statements as similar to the scenario. There was evidence for increased amplitude in response to both of these conditions.

Hypothesis 4b was that high aggression participants would show evidence of a negative interpretation bias and have negative expectancy outcomes; therefore they would have increased N400/LPP amplitude when making positive interpretations, i.e. dissimilar ratings of negative statements and similar ratings of positive statements. There was some evidence to support this hypothesis as the high aggression group showed increased amplitude when making similar ratings of positive statements across a number of electrodes. The findings regarding evoked amplitude when making similar and dissimilar ratings of negative statements were more mixed; for example, at P7 there was a significant effect of response such that high aggression participants had increased amplitude when making similar ratings of negative statements. However, at TP10 there was evidence of increased amplitude when making dissimilar ratings of negative statements, although this did not reach significance. Due to this mixed evidence it is difficult to draw firm conclusions; however, these results suggest that, similar to previous results presented in this thesis, the low aggression group showed greater differences in amplitude when distinguishing between ratings of positive and negative statements.

7.5.2 Limitations and future work

Although this research makes a number of valuable and interesting contributions to this field, I acknowledge that due to the novelty of using the recognition task to measure behavioural hostile interpretation bias and simultaneous neural correlates, the results should be interpreted with some caution. The task had previously not been used in conjunction with EEG methodology and was therefore modified to enable measurement of ERPs time locked to a specific point of interest. The ERP data were time locked to the last word of the sentence, therefore providing as accurate as possible a representation of interpretation time after presentation. However the time in which participants interpreted the statement and made their similarity rating could vary greatly. As EEG has high temporal resolution, specific brain processes relating to the participants response may not be evident at exactly the same time for all participants and could cause variation across grand-average ERPs. Another potential limitation of time-locking data to the last word of the sentence is that participants could have inferred the end of the sentence from the presentation of the first few words and subsequently decided on their interpretation of each statement before the last word was presented. Therefore, the time-locked ERP may not be a true representation of brain processes related to the similarity response. Again the lack of consistency across participants may cause variation in evoked amplitude and potentially may distort ERP patterns. Although this limitation is acknowledged, it is difficult to think of an alternative method for time-locking ERPs to assess interpretation bias using the recognition task.

It is recognised that the sample size was relatively small and therefore some of the analyses may have been underpowered. However, results showed medium to large effect sizes across many of the analyses suggesting that the findings are reasonably robust. Between-subject effects of self-rated aggression were found within a normal healthy sample; however, recruiting forensic samples with increased levels of aggression may provide greater clarity when distinguishing between the ERP patterns of high and low aggression samples.

Having considered the limitations of the current study, I make a number of proposals for future work. To overcome the problems with time-locking the ERP data as outlined, future work could identify participants with high levels of negative interpretation bias using an implicit measure (e.g. recognition task) before subjecting participants to an explicit measure of interpretation bias with simultaneous EEG recording. A simple explicit measure of interpretation bias would require less interpretation time and therefore it would be possible to achieve better temporal consistency across the time-locked data of multiple participants. A suggestion for future work is to conduct a sentence completion task (e.g. Moreno & Vázquez, 2011; Moser et al., 2008a; Moser et al., 2012;) in which sentences are resolved with either a final negative or neutral word. Words could be presented one at a time and ERP data could be time locked to the presentation of the resolution word. This would not necessarily require a response from the participants but a comprehension question could be added to ask participants if the resolution was hostile or not.

An alternative task would involve simple stimulus presentation in which participants are asked to respond yes or no to whether they believe the stimulus is hostile or not. ERP data could be time-locked to the stimulus presentation and participants would be asked to make a quick interpretation and choose one of two responses (yes/no). These tasks remove previous complications or ambiguity of making similarity ratings required in the recognition task.

This study could have important implications for CBM applications. CBM has been used within anxious samples to reduce threat related attentional bias and an increase in positive interpretation bias (e.g. Bowler et al., 2017). The study outlined in this chapter shows that participants that had a high score on AIHQ had different ERP patterns to those that scored low on AIHQ. This may have clinical applications for understanding brain processes involved with making negative interpretations. ERP correlates of interpretation bias could be used to evaluate the effectiveness of CBM intervention. For example, distinguishing between ERP patterns pre-and post-training would enable practitioners to explore whether there are changes in cognitive processes associated with making hostile attributions.

7.5.3 Contributions

Due to the unique use of the recognition task in this study, and the subsequent original results yielded from the ERP analyses, it is acknowledged that the results will require replication before more firm conclusions can be drawn. However, I believe it makes a number of useful contributions to the interpretation bias literature. Firstly it shows evidence of a distinct ERP pattern related to high levels of explicit interpretation bias. To my knowledge this is the first study to reveal that individuals with an explicit hostile interpretation bias have different evoked ERPs in response to positive and negative statements compared to individuals that show no explicit hostile interpretation bias. This suggests that EEG methodology can be used to measure interpretation bias and therefore is an appropriate method for detecting hostility-related biases. The knowledge that explicit interpretation bias, measured using a simple questionnaire, is characterized by a unique ERP pattern, is useful for possible cognitive bias modification work.

Secondly the study provides evidence to support previous literature on aggression and interpretation bias. The association between hostile interpretation bias and aggression has been relatively robust across studies using behavioural measures. This current study replicates and extends this work by providing evidence of this relationship across an explicit and implicit measure. This suggests that biases are detectable at both a conscious and an automatic level of awareness. This provides further support with previous theories that automatic cognitive processes contribute to hostile interpretation bias in aggression (Wilkowski & Robinson, 2010).

Finally the current findings suggest that there are differences in brain processing between high and low aggression individuals when they interpret ambiguous scenarios. Initial results suggest that low aggression individuals showed differentiations in amplitude when responding to differently valenced statements, and when making similar and dissimilar response ratings of such statements; in contrast, individuals with high aggression show more similar ERP patterns in response to both positive and negative statements. Specifically, the low aggression group showed increased evoked positive amplitude when responding to positive statements compared to negative statements. In line with the expectancy account of the LPP (Van Herten et al., 2005), this group also showed increased amplitude when positive expectation outcomes were violated; they rated negative statements as similar and positive statements as dissimilar. It is proposed that a positivity bias in low aggression participants may contribute to the differences in amplitude between aggression groups, but as these findings differ from those of some of the previous literature (Moser et al., 2008a), additional research will be required to further understand these findings.

Overall these findings provide support for previous attention bias results from Chapters one and three and suggest that, compared to high aggression participants, low aggression participants differentiate between stimuli to a greater extent at both attention and interpretation stages of cognitive processing. Due to the exploratory nature of these analyses it is suggested that these findings will require replication; however they provide initial results on which future work can be based. These findings contribute to the understanding of brain processing related to attributing hostile intent to ambiguous scenarios and help to understand why aggressive individuals may respond inappropriately in benign situations.

7.5.4 Conclusions

In summary, firstly this study has validated the recognition task as an appropriate measure of interpretation and shown that explicit interpretation bias may be characterised by a distinct ERP signature. Individuals with an explicit negative interpretation bias showed similar evoked amplitude in response to positive and negative statements, whereas individuals who showed no evidence of an explicit interpretation bias had increased amplitude in response to positive statements compared to negative statements during the recognition task. This may have important implications for future work as it suggests that, in individuals with hostile interpretation bias, brain processing is similar during the interpretation of negative and positive stimuli.

Both between-subject and correlational findings suggested that during the recognition task increased aggression was associated with a negative bias for targets. The findings also suggest that there are differences in the ERP patterns when interpreting positive and negative target statements between aggression groups.

Limitations of the current task are recognised and therefore methods for future work are proposed. It is suggested that an explicit measure of interpretation bias is used during EEG recording, for example sentence completion in which sentences are resolved either negatively or neutrally during presentation of the last word of each sentence. This would allow for more straightforward time-locking of EEG data. Having taken these limitations into consideration, this study has made significant contributions to the literature and could prove instrumental in designing future interpretation and cognitive bias modification studies.

8 General Discussion

8.1 Discussion - part one: Attention bias chapters

8.1.1 Overview of thesis

This thesis comprises a review of the literature and reports of five studies that use complementary behavioural and ERP methods to explore cognitive biases in aggression. Studies one to four investigated attention biases in aggression; the first two studies explored attention bias to different word types, whereas studies three and four explored attention bias to different facial expressions. The final fifth study investigated interpretation bias in aggression. The main aim of the thesis was to improve understanding of the neural correlates associated with selective attention biases to angry words and faces, and negative interpretation bias, in individuals with increased levels of self-rated aggression. Although previous evidence suggests that increased aggression is associated with attention bias to hostile words (e.g. Smith & Waterman, 2003), attention bias to angry faces (e.g. van Honk et al., 2001a), and negative interpretation bias (e.g. Epps & Kendall, 1995), very little is understood about the neural mechanisms which contribute to these behavioural effects. The first section of the discussion will review the four studies relating to attention bias, before drawing comparisons between results and discussing limitations and suggestions for future work. The second section of the discussion will review the final interpretation bias chapter in relation to the findings on attention bias, and present the overall conclusions.

The studies on attention bias have a predominant focus on physical aggression. Although behavioural and ERP effects were explored across multiple subscales of aggression, it was found that physical aggression yielded the most salient between-group differences. I suggest that physical aggression is a measurable explicit behavioural expression of anger. This is also consistent with a study by Smith and Waterman (2005), which found physical aggression to be predictive of hostile attention bias, and Helfritz-Sinville and Stanford (2015), which found distinctive ERP patterns in response to threat words in impulsive and premeditated physically aggressive men.

8.1.2 Overview of Study 1

In the first empirical chapter attention bias to angry and neutral words in aggression was investigated using a dot-probe task and simultaneous EEG recording. Both behavioural data (reaction time to probes), and ERP data (evoked P300 amplitude in response to the words and probes) was analysed. Consistent with previous work (Smith & Waterman, 2003), behavioural results provided both correlational and between-subjects evidence to support the predictions that physically aggressive males would have an increased attention bias to angry words, such that they had faster reaction times on congruent trials compared with incongruent trials. ERP results indicated that, in response to word-pair presentation, the high aggression group had overall increased P300 amplitude across all trials compared to the low aggression group. These findings are in contrast to previous evidence which suggests that aggressive individuals have reduced P300 response to stimuli presented across multiple tasks (Bernat et al., 2007; Gao & Raine, 2009; Gao et al., 2013). The trial congruency effects post-probe presentation suggested that participants generally show increased amplitude in response to incongruent compared with congruent trials. However, this effect did not interact with aggression. This finding is inconsistent with the predictions and with previous evidence that suggests participants with low levels of physical aggression show increased P300 amplitude to aggression-related words compared to neutral words (Helfritz-Sinville & Stanford, 2015; Thomas et al., 2007). Taking into account the previous literature and theoretical accounts of attention bias, it is not clear why participants showed increased P300 amplitude on trials in which the probe replaces the neutral word.

Unexpectedly, congruency effects were most salient pre-probe presentation; low physically aggressive individuals showed enhanced P300 amplitude in response to incongruent trials compared to congruent trials, whereas high aggression individuals showed greater similarity in their evoked amplitude in response to both trial types. Effects of congruency on the dot-probe task before the probe had been presented and the congruency of the trial has been revealed were not predicted. Therefore, these conclusions are made with caution and require replication.

8.1.3 Overview of Study 2

The primary aim of Study 2 was to test whether the main effect found in Study 1 would replicate; that ERP patterns would be different when responding to probes that replace angry and neutral words, and that this effect would interact with aggression. Therefore, the second empirical chapter explored attention bias to angry and neutral words in participants with increased levels of self-rated physical aggression. However, in this study selective attentional processes involved with attending to happy and neutral words, and simultaneously presented angry and happy words was also explored. This allowed for the investigation of whether attention bias effects in aggression were specific to angry words, and to explore the role of the distracter stimuli in hostility-related attention biases.

Unexpectedly, the behavioural results for angry-neutral trials from Study 1 were not replicated; there were no significant differences in reaction times between congruent and incongruent trials across any of the trial types. However, inspection of the means revealed that effects were in the expected direction (results suggested that high aggression participants had attentional facilitation of angry words, whereas low aggression participants avoid angry words). The main ERP findings indicated that high physical aggression participants showed overall increased P300 amplitude in response to all three trial types at word pair onset. There was also a main effect of valence such that *angry-neutral* trials evoke increased amplitude compared to *happy-neutral* and *angry-happy* trials. The effect of congruency following probe presentation was also analysed for each trial type. Unexpectedly, on *angry-neutral* trials, the ERP results post-probe presentation showed no main effect of congruency, therefore the effect from Study 1 was not replicated. With regard to ERP results for happy-neutral trials, there was a main effect of congruency such that participants had increased P300 amplitude on congruent trials (probe replaces happy word) compared to incongruent trials (probe replaces neutral word). This effect seemed to be particularly salient in the high physical aggression group. On angry-happy trials there was an overall main effect such that participants showed significantly increased positive amplitude to probes that replaced happy stimuli compared to probes that replaced angry stimuli. These results suggest that happy words evoke a greater increase in amplitude. The P300 component generally reflects the allocation of neural resources for information processing tasks (Polich, 2007), and reveals different processing patterns of stimuli depending on their task relevance (Coles et al., 1995; Donchin & Coles, 1988; Oliver-Rodríguez et al., 1999; Polich, 2007). Therefore, relatively larger P300 amplitude in response to happy words compared to angry/neutral words may indicate that more cognitive resources are allocated in processing positive stimuli. It is therefore suggested that the main effects may reflect a general positivity bias in which participants preferentially allocate cognitive resources to the processing of positive word stimuli. However, it is unclear why this effect would be most salient in the high aggression group on happy-neutral trials. Therefore, other explanations should be considered; for example, due to the tendency for aggressive individuals to perceive neutral stimuli as hostile (Mellentin et al., 2015), increased amplitude to happy trials could be attributed to the poor emotion regulation and response inhibition (e.g., (Patrick, 2008) which contribute to enhanced recruitment of resources needed to disengage from the simultaneously presented distracter stimuli (neutral word) in order to complete the task efficiently (Koster et al., 2004).

In addition to the predicted findings, evidence of congruency effects pre-probe presentation were also found. These were unexpected and are unexplainable in terms of attentional theory.

8.1.4 Overview of Study 3

The third empirical study was identical in design to Study 1, but this study investigated attention bias to angry faces (compared to neutral) instead of words. I was interested in whether modality of stimuli would influence attention bias effects and subsequent neural processing. Behavioural results provided correlational evidence for the relationship between physical aggression and attention bias to angry faces; however, this was not supported by between-group effects. The ERP results showed a main effect of congruency in the low physical aggression group but not in the high physical aggression group. The low physical aggression group showed increased P1 and P300 amplitude in response to congruent trials compared to incongruent trials, whereas the high aggression group showed relative stable amplitude in response to probes replacing both angry and neutral faces. This finding is consistent with previous research demonstrating an increased positive P300 amplitude to negative words in low aggression (Helfritz-Sinville & Stanford, 2015) and non-aggressive undergraduate (Thomas et al., 2007) samples.

The high physical aggression group showed an attentional bias for angry faces, reflected in their speedier reaction times to probes replacing angry faces; however, the ERP evidence suggests little difference in processing between trial types. This suggests that attention bias in physical aggression is not reflected in distinct differences in ERP patterns. Post-hoc ERP analyses revealed that increased physical aggression levels were related to increased P300 amplitude evoked by neutral trials, while amplitude to negative trials did not correlate with aggression. I suggest that increased amplitude on neutral trials in the high aggression group is a possible explanation for the similarity in amplitude across trial types. Due to the nature of the dot-probe task in which both angry and neutral faces are presented simultaneously, it is suggested that on neutral trials high aggression participants are required to assign greater cognitive resources (reflected in the increased P300 amplitude) to inhibit the response to the angry face distracter in order to effectively complete the task. This also explains why participants with increased physical aggression had an increased attention bias in which they had quicker reaction times to the probe that replaced the angry face, and delayed reaction times to the probe that replaced the neutral face.

As well as these post-probe results, similar patterns occurred pre-probe. Although consistent with similar findings from Study 1 and Study 2, it was again unexpected and inexplicable

8.1.5 Overview of Study 4

The initial aim of Study 4 was to test whether the Study 3 findings would be replicated. Due to a limitation of the previous study, that emotionality and aggression may be confounded in angry faces, and mirroring methods used for Study 2, two other trial types were included, happy-neutral, and angry-happy, with the aim of exploring attention bias to different emotional faces. Contrary to expectations, the results showed no significant interactions with aggression, and therefore firm conclusions cannot be drawn regarding the role that attention bias may play in contributing to aggressive behaviour. However, there were some interesting significant main effects of congruency across both behavioural and ERP data. The behavioural results revealed only one significant difference in reaction time to probes: on angry-happy trials there was a main effect in which participants were generally quicker to respond to probes that replaced angry faces compared to probes that replaced happy faces. There were no significant differences between reaction times to probes on angry-neutral trials, therefore the correlational evidence for the association between physical aggression and attention bias from Study 3 was not replicated. This is somewhat surprising given the previous literature which suggests attention bias to emotional faces over neutral faces (Bradley et al., 1997; Pishyar et al., 2004; van Honk et al., 2001a).

The main ERP findings showed on angry-neutral trials: there was a main task effect in which participants had increased amplitude to congruent trials compared to incongruent trials. This effect is consistent with the effect of trial congruency found in the low aggression group in Study 3. This study also showed a main effect on angry-happy trials in which participants generally had increased amplitude in response to probes replacing angry faces compared to probes replacing happy faces. These results suggest that P1 amplitude is increased in response to probes that replace angry faces. This finding is consistent with work by Santesso et al. (2008) which found increased P1 amplitude to probes replacing angry faces compared to probes replacing neutral faces, and by Smith et al. (2003) which found enhanced P1 amplitude in response to negative affective pictures compared to positive pictures. Results of the current study suggest that people preferentially attend to angry faces, regardless of whether they are simultaneously presented with neutral or happy faces. Crucially there were no significant ERP effects for happy-neutral trials, which suggest that amplitude is increased in response to angry faces and rather than to all emotional faces.

Although the findings for angry-neutral trials from Study 4 are somewhat comparable to Study 3 (the general task effect in which there was increased amplitude for congruent compared with incongruent trials is consistent with that for the low aggression group in Study 3) the interaction between aggression and trial congruency was not replicated. Due to the lack of significant behavioural effects, between-subjects effects, and replication, these findings are interpreted with caution. Nevertheless, these original findings contribute to the understanding of selective attention processes involved with attending to angry and happy faces when they are simultaneously presented.

8.1.6 Comparisons across studies

8.1.6.1 Studies one and three (angry v neutral: words and faces)

The same sample was recruited for studies one and three: the studies were very similar in design except that Study 1 explored attentional bias to angry words, whereas Study 3 explored attentional bias to angry faces. Across both studies the behavioural results showed a correlation between physical aggression and attention bias index. This suggests that across both modalities - angry words and angry faces - individuals with higher physical aggression were quicker to respond to the probes that appeared in place of angry stimuli compared to the probe that appeared in place of neutral stimuli. These results are consistent with previous research which show significant attention bias effects in aggression when using violently themed words (Smith & Waterman, 2003) and angry faces (van Honk et al., 2001a). There are a number of different cognitive processes which could contribute to this attention bias. For example, Wilkowski and Robinson, (2010) and Koster et al. (2004) suggest that facilitated attention (angry words grab attention following presentation) and suboptimal regulatory control, resulting in reduced ability to successfully disengage with angry stimulus once it has been attended to, contribute to faster reaction times to probes replacing angry stimuli. Previous research has predominantly used the Stroop task and therefore these current studies suggest that attention bias effects are robust across selective attention tasks, such as the dot-probe, and across stimulus modalities.

The ERP results indicated a consistent main effect of aggression across studies one and three, such that the high physical aggression group had increased amplitude across all trials in response to stimuli (word and face) presentation, compared to the low aggression group. These findings are in contrast to previous work which suggests that individuals with increased levels of aggression have a reduced P300 in response to presented stimuli (e.g. Barratt et al., 1997; Fanning et al., 2014; Helfritz-Sinville & Stanford, 2015; Surguy & Bond, 2006). However there is some mixed evidence regarding the ERP correlates of attention bias to aggression-related words in aggression. The current findings are consistent with those found by Stewart et al. (2010); they showed that individuals with higher anger-out scores showed increased P300 amplitude in response to the negative words during an emotional Stroop task. Taken together with the behavioural evidence - which showed biased attention towards angry stimuli in the high aggression group - the ERP results suggest that increased amplitude may reflect increased processing of negative stimuli.

Although the behavioural results from Study 3 indicated differences in reaction time between congruent and incongruent trials in the high aggression group, the ERP data suggests relatively stable P300 patterns in response to both angry and neutral faces in the high physical aggression group. A possible explanation for this is that aggressive individuals perceive hostility in ambiguous as well as non-ambiguous hostile expressions (Mellentin et al., 2015), and therefore process both angry and neutral faces similarly. Secondly, relative uniformity in amplitude across stimulus types may be attributed to increased processing on incongruent trials. Neuro-cognitive models of aggression suggest that deficits in regulatory control over incoming perceptual stimuli contribute to visual attention bias (e.g., Wilkowski & Robinson, 2010) in physical aggression, with physically aggressive behaviour being characterized by poor emotion regulation and response inhibition (e.g., Patrick, 2008). Therefore, high aggressive participants may assign greater cognitive resources to inhibit the habitual response to the simultaneously presented angry stimuli (distracter).

Furthermore, Study 3 revealed that low aggression participants, along with high aggression participants, had a slight negative bias for angry faces (quicker reaction times on congruent compared to incongruent trials, although neither between-subject tests reached significance). Consistent with previous literature (Santesso et al., 2008), the ERP results showed that low aggression participants had an increased P1 amplitude on trials where the probe appeared in place of angry faces, compared to neutral. These findings suggest that low aggressive participants have attentional facilitation for angry faces, reflected in quicker reaction times and increased P1 amplitude on angry-congruent trials. Angry faces may command attentional resources and are therefore detected quicker and allocated greater attentional resources, as reflected in increased amplitude

Behavioural results from Study 1 revealed a significant interaction between trial congruency and physical aggression group such that participants with higher physical aggression scores had faster reaction times on congruent trials compared to incongruent trials, whereas participants with lower physical aggression scores had faster reactions time on incongruent trials compared to congruent trials. Although this interaction was not consistent across the ERP results, a main effect revealed that participants showed increased amplitude in response to neutral words compared to angry words. This was opposite to main effect found on the face task, where participants showed increased P300 amplitude to angry trials compared to neutral trials. I suggest that angry words may not facilitate attention in the same way as angry faces and therefore attentional resources are easily directed away from angry words and towards the opposing stimuli (to neutral words).

Surprisingly, across both tasks an interaction between congruency and aggression was found pre-probe presentation. Consistent with the main task effect, congruency effects in the low aggression group showed contrasting results for the word and faces task. These salient pre-probe congruency results were unexpected and therefore do not contribute to answering the research questions. However, further research will be crucial in trying to understand the differences in cognitive processes detected by these ERP patterns.

8.1.6.2 Studies two and four (angry v happy v neutral: words and faces)

Studies two and four were designed to complement studies one and three respectively. Although the same sample was used across studies two and four, this was different from the sample recruited for studies one and three. Studies two and four were identical in design apart from stimulus modality; they explored attention bias to words and faces respectively. Overall, across both studies there were no clear conclusions to be drawn from the behavioural results as there were no between-subject differences for any trial types. The null effects across both studies are surprising given the findings from studies one and three and the theoretical models on which the predictions were based. For example, it is suggested that hostility-related attention bias contributes to aggressive behaviour because anger is a response to perceived provocation and therefore the recipient is motivated to aggressively confront and remove the threat (Smith et al., 1996). Attention bias to hostile stimuli has previously been found to be more salient in aggressive samples compared to non-aggressive samples (e.g. Putman et al., 2004; van Honk et al., 2001a)

Similar to the behavioural results, the ERP results across studies two and four only yielded a small number of between-subject effects. Study 2 (words) showed a main effect of aggression such that in response to all trials the high physical aggression group showed greater evoked amplitude compared to the low aggression group; this effect was most salient on angry-neutral and angry-happy trials. This suggests that high aggression individuals have increased amplitude in response to trials in which angry words are presented. However, in contrast to these findings, Study 4 (faces) showed an opposite effect in which the low physical aggression group generally showed increased amplitude in response to stimuli presentation; this effect was most salient for angry-neutral trials.

The effects of congruency for each trial type (angry-neutral, happy-neutral, and angry-happy) were analysed for both tasks. Although studies two and four yielded very few between-subject effects, there were some interesting general congruency ERP effects. The main findings for Study 2 (words) suggest that, in general, participants have increased amplitude in response to happy words; on *happy-neutral* trials participants had increased P300 amplitude on happy trials compared to neutral trials, and on *angry-happy* trials participants showed increased amplitude to happy compared to angry trials. In contrast to these findings, Study 4 (faces) showed a general task effect in which participants showed increased amplitude in response to angry faces; on *angry-neutral* trails participants showed increased amplitude to angry faces compared to neutral, and on *angry-happy* trials participants showed increased amplitude to angry faces compared to happy. These findings suggest that individuals have increased processing of happy words and angry faces during selective attention tasks.

The overall findings from studies two and four, which show contrasting effects of valence on evoked amplitude in response to angry and happy words and faces, seems to be consistent with studies one and three (in which the general main effect for evoked amplitude on angry-neutral trials was in the opposite direction for words and faces). The literature suggests that healthy participants have increased

P300 amplitude in response to pleasant and unpleasant words compared to neutral words (Sass et al., 2014), and to angry and happy emotional faces, compared to neutral faces (Holmes et al., 2009). Santesso et al. (2008) also showed that during a dot-probe task in which angry-neutral and happy-neutral face pairs were presented, participants had increased P1 amplitude to probes that replaced angry faces, compared to probes that replaced happy faces. Considering these findings the results from the faces dot-probe task are more in line with previous literature (increased amplitude to probes appearing in the prior location of angry faces). It is somewhat surprising that there was a general effect in which participants showed increased P300 amplitude in response to happy words compared to neutral and angry words. However, these findings suggest that word and face stimulus modalities are processed differently and that facilitation and disengagement processes (Cisler & Koster, 2010) may contribute to biases differently depending on the stimuli presented. I suggest that participants are better able to avoid attending to angry words (or disengage faster from such stimuli) and therefore attentional resources (reflected in increased positive amplitude) are allocated towards simultaneously presented happy or neutral words. In comparison, angry faces provide social cues (Argyle, 1994) and therefore, due to the importance for social interaction, are detected quickly (Fox et al., 2000). Angry faces may command attentional resources, consequently participants show increased amplitude on angry-congruent trials regardless of the distracter stimuli (happy or neutral face) (explanations for the differences between stimulus modalities are considered further in Section 8.1.8).

In the absence of behavioural effects for both words and faces tasks, the ERP results should be interpreted with caution. However, the lack of significant differences in reaction times across probe positions in any trial type may explain why there were very few interactions with aggression in the ERP data (another possible explanation is the lack of extreme aggression scores in the high and low aggression groups, see Section 8.1.9). The ERP data reveals differences in processing between trial types in the absence of significant reaction time differences. This suggests that ERP effects are more sensitive to differences in attentional processes. Reaction time measures represent a combination of processes including evaluation, decision-making, and motor processes, whereas EEG detects changes in neural activity evoked by an event of interest directly from the scalp (Luck, 2005).

8.1.7 Addressing the research questions

Overall, studies one and three showed greater between-subjects effects (perhaps attributed to the more extreme aggression scores within this sample, see Section 8.1.9) and therefore contribute to answering the research questions. Studies two and four yielded some interesting results, but, due to the lack of between-subject effects, the conclusions drawn relating to the research questions are somewhat limited.

The main research questions were whether high aggression participants have an increased attention bias to angry stimuli compared with low aggression participants, and whether this bias was reflected in different ERP patterns in response to angry and neutral stimuli (these research questions were consistent across both studies one and three, concerning words and faces respectively). Evidence suggests that high physical aggression was associated with faster reaction time to probes replacing angry stimuli compared to neutral stimuli. ERP patterns differed between aggression groups with high aggression participants showing relatively stable amplitude in response to probes replacing angry and neutral faces, suggesting that attention biases in aggression are reflected in less differentiated ERPs. The low aggression group had greater differences in amplitude between congruent and incongruent trials. However, the main effects of congruency were in opposing directions for word and face modalities. Participants had increased amplitude in response to neutral words compared to angry, and increased amplitude in response to angry faces compared to neutral. These findings contributed to the decision to further explore differences in stimulus modalities in the follow-up studies.

One of the limitations of studies one and three was that findings could in fact reflect a general emotional bias and not specifically an attention bias for angry stimuli. Therefore I was interested in whether aggressive individuals show an attention bias for all emotional stimuli, or whether the effect is unique to angry stimuli. In order to test this two further studies were conducted (two and four, using words and faces respectively) which included a happy-neutral condition. The aim was also to better understand the role of the distracter stimuli in selective attention tasks and therefore included an angry-happy condition. The studies explored whether high aggression participants show differences in evoked P300 amplitude compared to low aggression participants when selectively attending to negative and positive emotionally-valenced stimuli. Across studies two and four, very few significant interactions with aggression were found, therefore limited conclusions were drawn regarding attention bias to either negative or positive stimuli in aggression. The significant interaction found in Study 3 was not replicated and therefore it cannot be firmly concluded that attention bias effects are unique to angry stimuli in aggression. However, the results provide some interesting and valuable insights on attention processes associated with different stimulus pairs across word and face tasks. Generally, and similar to the main effect of congruency in studies one and three, for all trial types participants showed increased amplitude in response to positive word stimuli, but increased amplitude in response to negative face stimuli. Crucially this shows that attention to faces and words may be reflected in different cognitive processing.

Differences in evoked amplitude were evident across a number of different components. Based on previous evidence and the speed of attentional orienting, it was predicted that attentional processes would influence the P1 and P300 component specifically. Evidence suggested that P1 and P300 are different between aggression groups and in response to congruent and incongruent trials across both types of stimuli, but more generally effects were longer lasting and also influenced a number of other ERP components such as the N2 and the P600/LPP. This suggests that attentional biases may affect early stages of spatial attention (e.g. Hillyard & Anllo-Vento, 1998; Woldorff et al., 2002) and more elaborative stages in which attentional resources are allocated and stimuli are categorised (Polich, 2007). The P600 is similar to the P300; however, the increased potential is sustained for a longer latency (Hajcak & Olvet, 2008). It is sensitive to salient and unlikely information (Van Herten et al., 2005). Therefore the findings suggest that, within the general population angry faces and happy words receive greater attentional resources when the stimuli first capture attention (P1) and when they are further appraised in relation to current cognitive models (P300). Particularly salient stimuli may then require further processing, reflected in increased P600 amplitude. There was some evidence to suggest that high aggression participants were less likely to show this ERP pattern and recruited similar resources across all ERP components when attending to angry and neutral faces.

8.1.8 Interpretations

Due to the original finding that individuals show different ERP patterns in response to different trial types, and across modalities, I suggest that the attentional processes contributing to attention bias, especially when investigating selective attention, are extremely complex. Consequently, I propose that during selective attention tasks, the distracter stimulus plays an important role in attention allocation. Therefore, consistent with theories of selective attention (Wilkowski & Robinson, 2010), attention bias is a consequence of both increased stimulus-driven attentional facilitation and suboptimal regulatory control resulting in delayed disengagement. It is a unique combination of both these features when attending to angry words and faces which contribute to the difference in effects across stimulus modalities. However, within the current studies attentional control was measured as a possible moderator of attention bias and found that attentional control only correlated with attention bias to angry words in one of the four studies and therefore did not seem to be a large contributing factor to the current results. To my knowledge this is the first programme of research to compare ERP patterns evoked by words and faces during a selective attention task. There is no solid explanation for the difference in findings between modalities; however, the effect is relatively consistent across both sets of studies. I aim to address some of these differences and provide potential explanations for the differences in findings between the words and faces task.

Both faces and words are recognised based on the features of the stimuli and the configuration of these features. For example, faces are comprised of common features such as eyes, nose and mouth; however, to facilitate emotion recognition, it is necessary to process the face holistically (Ventura, 2014). Subtle differences are essential for successful recognition of a given facial expression (e.g., Maurer et al., 2002). Similarly, recognition of written words is dependent on understanding the letters which make up the words, and the composition of these letters (Ventura, 2014). However, faces provide important social cues and are therefore central to human interaction (Argyle, 1994), whereas words may seem fairly arbitrary. Facial expressions are also generally universally recognised, while recognition of words relies on vocabulary knowledge and language. These distinct differences in stimulus types may contribute to variation in processing and explain current findings which show important differences in how the attentional systems responds when attending to words and faces.

Wang et al. (2012) suggests that faces are processed in a distinctive manner and that face-specific processing is essential for holistic face recognition. This suggests that faces may have a unique influence on the attentional system. Furthermore, Fox et al. (2000) suggest that detection of threat facial expressions are particularly important for social outcomes and therefore may facilitate attention and be more elaborately processed compared to linguistic stimuli. Fox et al. (2000) conducted a number of visual search tasks in which participants had to scan multiple faces and report whether there were any that were different. They report that angry faces were detected amongst a group of neutral faces more efficiently than happy faces amongst neutral faces. They also found that response times to detect a discrepant face were slower when the visual display contained angry faces.

They concluded that angry faces may hold visual attention and act as a distracter to attending to other stimuli. The inconsistency in valence effects between words and images could be due to the increased salience of angry faces. The aggressive words may not hold attention in the same way as the angry faces (or aggression-related words are less distracting than angry faces).

Sternberg, Wiking and Dahl (1998) investigated the role interference of angry faces plays during a task in which participants are presented with words superimposed on different facial expressions. Participants were asked to categorise the words as good or bad, while trying to ignore the face. They found a main effect of word type in which negative words took longer to process compared with positive words. The results also showed that word latencies were longer when angry faces were presented, suggesting that angry faces interfere with other ongoing processes and therefore may inhibit task response. This is consistent with work by Hansen & Hansen (1994) which found that angry faces tend to attract and then hold attention. These findings are important considerations when interpreting the current data, especially given the simultaneous presentation of stimuli in the dot-probe task. These findings suggest that angry faces grab attention, and may interfere with task demands to a greater extent compared to angry words. Further analyses of the current studies showed a general task effect in which participants had quicker reaction times to probes on the word task (Study 1) compared to the face task (Study 3) across both congruent (word task, M = 486.33, SD = 73.67; face task, M = 579.47, SD = 63.11; $t(31)$ = 7.958, $p < .001$, d = 1.358) and incongruent trials (word task, M = 489.05, SD = 76.95; face task, M = 583.34, SD = 64.03; $t(31)$ = 8.658, $p < .001$, d = 1.332). These findings suggest that angry faces interfered with the ability of participants to respond to the probe location efficiently, and therefore overall they had a delayed response compared to words.

When comparing studies one and three in which angry and neutral words and faces were presented respectively, the behavioural data showed a significant interaction between attention bias and aggression for words, but only correlational evidence for faces. Study 1 showed a significant interaction such that high aggression participants attended faster to probes replacing angry words compared to neutral words, whereas low aggression had faster reaction times to probes replacing neutral words compared to angry words. Study 3 showed a significant correlation between attention bias and aggression, although inspection of the means showed that both groups attended more quickly to probes appearing in place of angry faces. Taken together these behavioural results suggest that the high aggression group may have had facilitated attention for both word types, whereas the low aggression group may have had attentional facilitation for angry faces, but be able to avoid attending to angry words. This fits with ERP data which suggested that low aggression participants had increased P300 amplitude when responding to angry-incongruent trials (probe replaced neutral word) on the word task (although this was found across both aggression groups), but increased P1 and P300 amplitude in response to congruent trials (probe replaces angry face) on the faces task.

There is evidence to suggest that speed of processing varies depending on the modality of the stimuli. Schacht and Sommer (2009) compared the effects of emotional words and faces in an ERP within-subject design. Similar ERP patterns were found when processing both words and faces; however, the effects appeared at very different latencies suggesting that stimulus types are processed at different speeds. Specifically, speed of meaning was accessed more directly and faster for facial expressions than for words. This may explain why in the current study early effects of congruency evoked differences in the P1 component for the face task but not the word task. Therefore, when investigating attention bias across different stimuli, it may be appropriate to use a shorter SOA (e.g. 100ms) when investigating attentional processing of faces, and a longer SOA (e.g.750ms) when investigating attentional processing of words.

Due to the faster processing of faces compared to words, it is proposed that these stimuli may be dependent on different processing routes. It is recognised that there are two types of processes involved with visual selection; bottom-up and top-down. Bottom-up processing is an automatic and pre-conscious process which refers to the allocation of attention driven by characteristics of the stimulus. Top down processing is conscious and controlled and refers to the allocation of attention driven by the observer (e.g., Burnham, 2007; Theeuwes, 2010). It is proposed that attention biases are driven by competition for attention (Desimone & Duncan, 1995) and that attended stimuli receive priority over unattended stimuli. Attentional effects are the result of competition between bottom-up and top-down features. When the stimuli is particularly salient, bottom-up processes immediately allocate attention towards such stimuli.

Based on these models of attention I suggest that different types of processing are responsible for the allocation of attention to stimuli on the word and faces dot-probe task. Face stimuli, and particularly angry faces, are subject to quick, automatic bottom-up processing. Therefore, the general task effect found in studies three and four, which shows increased P1/P300 amplitude to angry faces, may indicate consistent engagement and processing of threat stimuli. It is proposed that low physical aggression participants may be able to override these automatic processes, and use top-down resources to disengage with such stimuli and complete the dot-probe task effectively, whereas high aggressive participants become fixated on angry faces. In comparison to angry faces, angry words are perhaps less salient compared to angry faces and therefore do not grab attention in the same way; attention is therefore not stimulus-driven. The slower processing of angry words would allow for conscious allocation of attention to the simultaneously-presented second word. In Study 2, the ERP results shows that P300 (and number of other components such as the P1 on angry-happy trials) were increased in response to happy words compared to neutral or angry words. Therefore it is suggested that, due to the reduced salience of angry words, participants allocated fewer cognitive resources to detecting angry words and were better able to attend to neutral and happy words, resulting in increased processing of such stimuli (reflected in increased amplitude). In Study 2 no significant effects of a behavioural attention bias were found; participants responded fairly similarly across all trial types. This suggests that participants may not have significant differences in reaction times when responding to happy compared to neutral/angry words, however they may have a positive processing bias in which happy words are more elaborately processed and command greater resources.

I suggest that in real life settings individuals are better able to disengage from angry words and therefore attention is allocated to other stimuli within the environment. However, angry faces are subject to quick automatic processing due to the potential threat they may present. Therefore individuals find it more difficult to disengage with such stimuli. I suggest that disengagement processes are crucial when understanding the differences between aggressive and less aggressive individuals due to deficits in regulatory control and response inhibition in individuals with increased aggression. Attentional disengagement also plays an important role in understanding differences in processing biases across modalities; increased salience of angry faces compared to words contributes to greater difficulties in disengagement resulting in differing ERP patterns across tasks. There are very few studies that explore the differential influence of word and face stimuli on attention bias and therefore replication would be recommended to further assess the validity of these findings.

8.1.9 Limitations

Behavioural analyses in studies two and four utilizing a between-subject design based on a median split of physical aggression score failed to provide clear differences in bias indices between groups. The ERP results revealed some interesting effects of trial congruency, however these effects did not significantly interact with physical aggression. These findings may be explained by the lack of statistical power in analyses using between-subjects designs based on dichotomisation of a continuous variable (Gignac & Szodorai, 2016).

The primary limitation of using a median-split design is the lack of power, which reduces chances of finding a significant relationship when there is one (Type II error). Therefore, the significant results in this thesis may be under represented. More robust findings may have been evident if aggression had been used as a continuous variable within a multiple regression analyses. This analysis can include interactions between continuous and categorical predictors. However, with large number of variables, multiple regression can be difficult to interpret, with it usually being necessary to break down the results into sub-sets of variables. Linear regressions are also sensitive to outliers, which can distort the results substantially. It has been argued that using a median-split also increases the chance of a Type I error through false–positive consumer psychology (McClelland, Lynch, Irwin, Spiller, & Fitzsimons, 2015). However, Iacobucci, Posavac, Kardes, Schneider and Popovich, (2015) claim that median splits result in no more Type I errors than a regression on a continuous variable. Although the limitations of the chosen method are recognized, having considered alternative statistical approaches such as multiple regression, I believe that median split analyses were most appropriate for testing the current hypotheses. These allowed for comparison to previous studies that have utilized a between-subject design. This design also allowed for the straightforward interpretation of ERP patterns, for example, it was possible to qualitatively inspect the differences in averaged ERPs between groups scoring high and low on physical aggression. Furthermore, there is evidence to suggest, in the absence of multicollinearity, median splits do not create misleading results

(Iacobucci et al., 2015). Across the studies presented in this thesis, correlations were conducted with a continuous aggression variable to support between-subject analyses and gain a better understanding of the data.

Visual inspection of the means in Studies 1/3 and 2/4 show that, although the overall sample means are comparable, in study 1/3 the high aggression group had a higher mean physical aggression score and the low aggression group had a lower mean physical aggression score, compared to Study 2/4. The extreme scores at both ends of the aggression scale may drive the significant behavioural and between-subject ERP results. Therefore as an alternative to the median split, to explore these further, additional between-subject analyses for Study 2 and four were conducted based on the lower (n = 16) and upper (n = 13) quartile cut-offs. However, there was no significant difference in attention bias scores between aggression groups for Study 2 (words; p > .706) or Study 4 (faces; p > .177). This indicates that the median split was not a key limitation of the current analyses; it is suggested that, although the upper quartile of represents the most extreme scores within the current sample, these scores were not high enough to reveal an aggression-related attention bias. Future work including a forensically aggressive sample and non-aggressive control group would be expected to yield greater between-subject differences and larger effects sizes.

Another explanation for the non-significant between-subjects effects in studies two and four could be that participants completed the dot-probe tasks within a positive environment and context. There is evidence to suggest that the attention bias to negative information is reduced in positive affective contexts. Smith et al. (2006) found that if participants were primed with negative information during an emotional Stroop task, the P1 amplitude was increased in response to negative stimuli, whereas if participants were primed with positive information, P1 amplitude was increased in response to positive stimuli. This is consistent with previous evidence which suggests that attention biases are only observed when trait and state anger levels were increased (Eckhardt & Cohen, 1997). The current study measured only trait aggression and therefore participants reported that they had the ability to behave aggressively, but were not in an aggressive state when they took part in the tasks. Future work could measure both state and trait aggression, or adopt a similar approach to Eckhardt and Cohen (1997) in which participants were provoked to induce increased levels of state aggression, prior to completion of an attention bias measure.

Previous research suggests that there are several processes that contribute to attention bias, including facilitated engagement, difficulty in disengagement and attentional avoidance (e.g. Cisler & Koster, 2010; Koster et al., 2006). The dot-probe task is a paradigm used to capture both facilitative and disengagement biases (Koster et al., 2004). Therefore, the behavioural and ERP data are interpreted drawing upon mechanisms involved with both attention facilitation and poor disengagement with the aim of better understanding how each of these features may contribute to attention bias in aggression. It is suggested that low and high aggression individuals may have attentional facilitation for angry faces, however low aggression individuals are more readily able to disengage from such stimuli. When comparing stimulus modalities I suggest that words may not command attentional resources to the extent that angry faces do, and that low aggression participants can more easily direct attention away from angry words to positive words. However, due to the simultaneous presentation of both words, it is recognized that it is not clear from the data whether differences in amplitude in the low aggression group can be attributed to increased amplitude to neutral words, or decreased amplitude to angry words on Study 1, and increased amplitude to angry faces or decreased amplitude to neutral faces on Study 2. The findings are far from conclusive and interpretations are somewhat speculative. Therefore follow up studies will be essential in trying to separate distinct mechanisms and ERP correlates of attention bias (see Section 8.1.11).

Within the attention bias literature the dot-probe paradigm is a widely used method for measuring attentional allocation to stimuli. However Kappenman, Farrens, Luck, and Proudfit (2015) and Schmukle (2005), suggest that this method is suboptimal for measuring attention bias due to its lack of test-retest reliability. Using a dot-probe task, (Kappenman et al., 2015) studied attentional bias to threat measured by behavioural and ERP methods. They found no attentional bias effect using traditional reaction time measures. In contrast, measuring the N2pc component as a physiological marker for attentional allocation revealed a significant effect of attention bias to threat. However, there was no evidence of a relationship between the attention bias effect and anxiety. They reported that the reaction time measure of threat bias was not internally reliable, whereas the N2pc showed highly significant internal reliability. The research carried out by Kappenman and colleagues suggests a need for more reliable reaction time methods, it also demonstrates the usefulness of ERP analysis and how both methods can be used together to better understand cognitive processing.

A further limitation of the dot-probe tasks used in the studies presented in this thesis is that stimuli were presented vertically (one stimulus appeared above the fixation cross and one appeared below). Eye-tracking evidence suggests that visual attention is inherently directed upwards (e.g. Price et al., 2015; Waechter et al., 2014) therefore there may be a bias for stimuli presented in heightened visual field locations. However, some research suggests that the right hemisphere is dominant in the perception of emotional faces, (e.g. Davidson, 1993), therefore there may be a tendency to allocate attention to the left visual field in emotional dot-probe tasks if the stimuli are presented horizontally (Mogg and Bradley, 1999c). In the current study presentation of stimuli types and locations of the probes were counter balanced to reduce the possible influence of stimuli location.

In Study 2 the positive and negative words from the Brysbaert database (Brysbaert & New, 2009) were matched on length and frequency. Arousal values of the word stimuli were not matched. Research suggests that increased arousal is associated with increased attentional facilitation, and that arousal may have greater influence on attention processes compared to valence (Vogt et al., 2008).

Therefore, in the current studies speedier reaction times and increased processing of positive or negative stimuli may be explained by the arousal value rather than the emotional valence of the stimulus. It is also recognised that the neutral words mainly consisted of nouns, whereas the angry and happy words are made up of nouns, verbs and adjectives, these fundamental differences in word types may influence attentional processes and therefore a direct comparison may not be suitable. It is recommended that future studies using different word types should match stimuli on word type, length, frequency and arousal.

The face stimuli were selected from the Chicago Face database (Ma et al., 2015). This is a standardised database of facial expressions and therefore a number of factors such as head position, exposure, and facial hair are controlled for. However, due to the random selection of stimuli used for Study 1 it is recognised that other factors relating to the facial expressions, such as attractiveness, arousal or dominance, may have influenced the results. For Study 4, 32 out of the possible 35 actors portraying angry, happy and neutral faces were included. Those with the highest angry/happy ratings based on norming data were selected. Due to the limited number of possible stimuli that could be utilised from the Chicago Face database (Ma et al., 2015) other confounding factors were not controlled for. However, as the same actor was used for both facial expressions in each stimulus pairing, any confound of attractiveness should be reduced. Using the norning data I ran some post-hoc tests to explore whether the actors were generally rated as more angry or happy. Although in both Studies 3 and 4 the actors had slighltly increased ratings of anger, compared to happiness, this difference did not reach significance. The Chicago Face database (Ma et al., 2015) only provides norming data for each actor showing a neutral expression, therefore it is unclear how individual features of each of the facial expressions may have influenced the results. Happy expressions may be perceived as more attractive than neutral expressions. For example, Tatarunaite, Playle, Hood, Shaw, and Richmond (2005) found that smiling faces were rated as more attractive compared to non-smiling faces. It has also been found that at later stages of attentional processing (LPP), resources are more likely to be allocated to unattractive faces expressing a negative emotion, compared to more attractive faces showing a positive emotion (Sun, Chan, Fan, Wu, & Lee, 2015). This suggests that attractiveness of facial expressions may influence how the face is processed and subsequent attention bias conclusions. In future research it would be suggested that individual ratings for emotion, attractiveness, and arousal are given by an independent sample before selecting the stimuli for inclusion in the final studies.

Data for studies one and three were collected during the same laboratory session and therefore there could have been order effects. There could be procedural limitations of completing the word dot-probe task before the face dot-probe task. The sequence of task completion may have influenced the results, for example it may have been that participants understood the task better the second time, or were more relaxed or bored during the second task and therefore responded more automatically without thinking about the task too much. Further research simultaneously studying multiple dot-probe tasks should counter-balance these in order to check for these effects. Data for studies two and four were collected during the same testing session; for these follow-up sessions the order in which people completed the word and face task, and the order in which they completed the questionnaire measures and the experimental task, were counterbalanced.

I appreciate that methods of recruitment and instructions provided to participants may have resulted in priming effects or demand characteristics. For example, In studies 1 and 3, posters calling for individuals that ‘lose their temper’ or ‘experience road rage’ were used to recruit participants with higher levels of aggression. Also, due to counterbalancing of questionnaire and experimental tasks, some participants completed the aggression questionnaire before taking part in the dot-probe and recognition tasks. Therefore, participants may have predicted the aims of the study and were therefore more alert to angry stimuli. With the aim of reducing any priming effects, participants were provided with the only minimal information needed to complete the task. Also, out of the final sample, only a small number of participants were recruited using targeted posters.

It is acknowledged that these studies were correlational and quasi-experimental in nature and therefore it is not possible to determine whether attentional bias is a cause or consequence of aggressive behaviour. A review by Van Bockstaele et al. (2014) suggests there is mixed evidence for the causal relationship between attention biases and anxiety. Some studies report that, in line with casual predictions, attention biases precede anxiety, whereas other work indicates that anxious symptoms can occur before vigilance to threat is evident. Van Bockstaele et al. (2014) also found that a change in attention biases is related to a change in anxiety and vice versa, suggesting that there is a reciprocal relationship between the two phenomena. The association between attention biases and aggression are likely to develop in a similar reciprocal manner; therefore, future work exploring the causal nature of hostile-related attention bias would be beneficial in understanding the development of aggression. For example, does attention bias for angry stimuli (identified in studies one and three) predict the likelihood of behaving aggressively? It is proposed that large-scale longitudinal studies measuring biases and aggression at multiple time-points are needed to establish true cause-effect relationships. Furthermore it is very difficult to mimic real life aggression within laboratory settings. Violent behaviour is the result of a complex interaction of trait-like vulnerabilities relating to self control and emotion regulation (Buckholtz, 2015). Therefore, it is a challenge to recreate the dynamic nature of physical aggression within the static nature of laboratory based assessments (Poldrack et al., 2017). In combination with laboratory studies, in depth case studies may be useful when exploring aggression-related attention biases. Observing physically aggressive behaviour within a natural environment may give a more accurate measure of aggression than self-reports. Aggressive participants could then be subjected to further laboratory assessments where attention biases could be studied. Also, since the sample was male-only, predominantly British and young, this study has limited generalizability; further work is required to investigate neural correlates of attention bias in aggression in female, older, and non-British samples.

An important consideration is that these results focus on physical aggression and therefore may not be generalisable to other types of aggression. Physical aggression is arguably the most extreme factor measured by the Aggression Questionnaire (Buss & Perry, 1992). Physical aggression reflects a behavioral response which involves physical contact with another person, whereas the other three factors of this scale (verbal aggression, hostility and anger) relate to feelings or emotion associated with aggression and not necessarily the act itself. Compared to other studies that have measured general aggression or anger, the current studies show that attention biases may influence subtypes of aggression differently; it was found that attention biases may be particularly salient in individuals with increased levels of physical aggression. This could have important implications for interventions as it suggests that attention biases are associated with violent behavior.

8.1.10 Contributions

I believe that these four studies make a considerable contribution to the attention bias and aggression literature. Firstly, results show that methodologically, ERPs are sensitive to differences in attention allocation and therefore an appropriate method for measuring neural processes associated with attention bias in aggression. Specifically, results suggest that the dot-probe assessment of selective attention is compatible with simultaneous EEG recording. In the second and fourth studies there were ERP differences even in the absence of a significant behavioural difference. Previous studies have utilised this method (see Torrence & Troup, 2017 for review), but to my knowledge none of these have explored between-group differences in aggression.

Secondly, the studies show that there are differences in processing of angry and neutral stimuli between aggression groups. There was very limited evidence of the neural correlates of selective attention in aggression, especially using facial stimuli. Results replicate previous work by Helfritz-Sinville and Stanford (2015) and extended this work to explore selective attention using the dot-probe task and face stimuli as well as words. Results suggest that participants scoring low on physical aggression have significantly different evoked amplitude in response to angry and neutral stimuli, whereas participants with increased scores of physical aggression show relatively undifferentiated ERPs. However, this between-group effect was not fully replicated across both follow-up studies.

The third main contribution is the finding that attention bias for words and faces may be driven by different underlying mechanisms. A consistent main task effect was found in which evoked amplitude in response to positively and negatively valenced stimuli was in the opposite direction for faces compared to words. More positively valenced words (happy or neutral) evoked increased positive amplitude compared to negatively valenced (angry) words, whereas negatively valenced (angry) faces evoked increased positive amplitude compared to more positively valenced (happy or neutral) faces.

8.1.11 Future research

Although the studies make a number of contributions to this field, it is recognised that, due to limited previous evidence, some of the hypotheses were exploratory in nature. Therefore, the main aim for future work would be to test the replicability of the key findings. In particular it would be useful to replicate this work in a forensically aggressive sample. The studies outlined in this thesis had original designs and present some unique results. Although I have proposed a number of possible interpretations for these findings, the results are far from conclusive.

A limitation of the dot-probe tasks used in studies one to four, and the subsequent interpretations of the data, is that they did not include a neutral-neutral control condition. I would suggest including this stimulus pairing in future research. It may have also been beneficial to include an angry-angry and happy-happy control condition in studies two and four. A neutral-neutral stimulus pairing would provide a 'baseline' for which to compare the three experimental conditions and evaluate evoked ERP amplitude to angry and happy stimuli. In relation to reaction time data, this may help to better distinguish between the differences in amplitude for each stimuli type and therefore make more informed conclusions regarding facilitation and disengagement processes in aggression (Koster et al., 2004). Differences in reaction time and evoked amplitude were found on angry-neutral trials between probe positions. However, on neutral-neutral trials it would be expected that there would be no difference in reaction time to probes appearing in any position due to the similarity between stimuli. Similarly, no differences in evoked amplitude between trial types would be expected as the probe would be appearing in place of neutral stimuli on all trials. This may also contribute to the understanding of the pre-probe congruency effects found across all studies. It would be possible to compare the ERP effects in response to neutral-neutral, happy-neutral and angry-neutral trial types.

In line with qualitative inspection of the waveform, the ERP results consistently show significant differences in evoked amplitude at TP9 and TP10. This suggests that differences in attentional processes may be most salient at the temporal-parietal region. This is consistent with previous evidence which suggest that the temporo-parietal attentional network situated in the TPJ is a crucial generator of the P300 component (Knight et al., 1989). The TPJ in the right hemisphere has been associated with distinct cognitive processes (Decety & Lamm, 2007), particularly those involved with orienting of attention (Corbetta et al., 2008). However, EEG has poor temporal resolution, and therefore no direct link between electrode site and presumed cortical generation can be made.

A large number of electrodes were included across all studies to overcome limitations of previous literature that presents only a few mid-line electrodes. The aim was to enhance transparency in EEG research by avoiding selecting only a few electrodes and epochs where significant differences were evident. However, due to the large number of statistical analyses that resulted from the current design, suggestions are made for future work. Although it is impossible to identify the location of cognitive mechanisms using EEG analyses alone; driven from the current findings that the TPJ is particularly important when exploring attention biases in aggression, it is suggested that future studies exploring the spatial location of attention bias effects could focus analyses on a smaller region of interest located between the temporal and parietal lobes (for example, a selection of electrodes from; TP9, TP10, TP7, TP8, T7, T8, P7, P8, P3, P4).

The current study found some distinct differences in ERP patterns between modalities. To explore this further, and to test whether the results found across the current studies replicate, I would propose conducting one dot-probe task in which there were randomised blocks of either word or face stimuli. The blocks relating to each modality could then be extracted and analysed separately. This data would reveal if processing patterns and ERP correlates change distinctively between each block. It would be predicted that there would be regular and uniform observed differences between each of the blocks. This would also reduce the chance of order effects, possibly caused by completing two dot-probe tasks in succession.

A number of studies using the dot-probe task with simultaneous EEG recording have analysed evoked amplitude in response to the stimuli (word or face onset) and target (probe onset). A review by Torrence & Troup (2017) shows that both analytical methods have been used when assessing neural correlates of attention bias, and that they yield varied results. Only a few studies (e.g. Santesso et al., 2008) have analysed data time locked to both stimulus and target, as in the studies reported here. In all four of the attention bias studies data was analysed time locked to the onset of the stimulus pairing (words or faces). A pre-stimulus baseline was chosen to avoid the possible confound of introduction of post-probe trial type effects created by pre-arrow change in baseline (Mingtian et al., 2011; Poulsen et al., 2005). It also allowed for the exploration of the neural processes involved with attending to two simultaneously presented stimuli. To better understand the time course of attentional bias in aggression, early ERP components such as P1, and later components such as the P300, were investigated in relation to both stimuli and target onset. I expected to find pre probe presentation (0-500ms) effects of aggression, or valence (in response to the three different stimulus parings used in studies two and four). However, I expected to find congruency effects of each trial type only following the probe presentation at 500ms. Contrary to expectations, the data consistently yielded pre-probe differences in amplitude between congruent and incongruent trials across all four studies. Consistent with previous evidence (Mingtian et al., 2011; Poulsen et al., 2005) this confirmed that data time-locked to the probe onset would not have a valid baseline. It is suggested that, future dot-probe research investigating congruency effects across different trial types using the same methodology should adopt an analytical approach in which the length of the whole trial is statistically analysed based on a pre-stimuli baseline.

To my knowledge, and not surprisingly, no other studies have used the dot-probe paradigm to explore effects of trial congruency pre-probe presentation. It is difficult to explain why participants could show differences in ERP patterns at 300ms post-stimulus presentation based on upcoming probe presentation at 500ms (either in a congruent or incongruent position). Across all studies, combinations of probe type (left or right facing arrow), face type (angry or neutral) and position on screen (top or bottom) were all counterbalanced, with a new random order for each participant. Therefore, any predictions based on probe location should not have been possible. A possible explanation is that differences in amplitude pre-probe presentation reflect long lasting effects based on the probe positioning evident from the previous trial. However due to the length of the trial and the speed of attentional allocation this is somewhat unlikely. In order to rule out this possible explanation I would suggest using a longer epoch, in which there is a larger gap between each stimulus pair. This would ensure that attentional processes associated with attending to the stimuli and then the following probe, would be complete before the next set of stimuli are presented. Participants would be asked to pause with their eyes shut for a few second between each trial.

A further explanation of the pre-probe effects is that high and low aggression groups allocate attention differently to angry and neutral stimuli when they are simultaneously presented prior to the probe presentation. An assumption of the dot-probe task is that attention will be faster to probes replacing angry faces if attention was pre-directed to that region of the visual display. Whereas a more even monitoring of the face-pair display in the normal population is suggested by more equivalent reaction times to the probe when it subsequently replaces either the angry or neutral face with equal probability. Therefore ERP effects may reflect this assumption. Participants scoring high on aggression may have allocated attention to the angry stimuli during the pre-probe stimuli-pair display, and therefore when averaged across fifty percent of the trials (probe later appeared in the position of angry/neutral face) and split by the later trial type, amplitude may be elevated for both later anger-congruent and incongruent trials. Whereas, averaging of ERPs shown by the low aggression group perhaps suggest a more even monitoring of both stimuli presented during simultaneous pair presentation. Therefore, attention was already directed in the region that the 50:50 arrow then appeared. When averaged, these small differences in pre-probe attentional allocation may be reflected in differences in ERP patterns in response to upcoming angry-congruent and angry-incongruent probe positions. However, it is impossible to draw clear conclusions regarding these findings using the current design and therefore future work would be needed to better understand the impact of attentional allocation between stimulus types during the dot-probe task on pre-probe ERP patterns.

The studies presented here have demonstrated that EEG is a useful method for assessing processes associated with attention bias; however, from the current analysis it is not possible to fully distinguish between quickened facilitation and delayed disengagement. Interpretations were made regarding these processes based on the findings; however, using eye tracking software may enable more concrete conclusions to be drawn. Eye tracking software would be useful in tracking initial saccades in response to stimulus presentation (e.g. Armstrong & Olatunji, 2012; Duque & Vazquez, 2015). Used in conjunction with EEG methodology and the dot-probe task this may be useful in understanding the neural correlates associated with attentional facilitation and disengagement and provide a more fine-grained analysis of time course. This method may show how processes contribute to attention bias during selective attention tasks in aggression.

This thesis suggests a number of implications for rehabilitation of aggressive individuals. Studies one and three show that attention bias to angry stimuli is linked to physical aggression and that these biases are reflected in relatively undifferentiated ERPs in response to negative and neutral stimuli. This suggests that modification or reduction in attention bias may have a rehabilitative value in reducing physically aggressive behaviour. This could be particularly useful in reducing youth crime and possibly preventing criminal careers progressing into adulthood.

Cognitive bias modification (CBM) is an experimental paradigm used to change cognitive biases. It can be used to induce a positive or negative cognitive bias. CBM has commonly been used to modify negative biases in anxious and depressed individuals. The premise of CBM is that modifying cognitive biases will produce changes in behaviour, for example reduce anxious and depressed symptoms (Hallion & Ruscio, 2011). There are two types of cognitive bias modification; attentional bias modification (ABM) which addresses hostile/threat-related attention bias, and interpretation bias modification (CBM-I) which targets negative interpretations. In particular, attention training (Amir, Beard, Burns, & Bomyea, 2009) is used to modify attention biases and has been proven to be an effective rehabilitative method for treating anxiety (for example, Amir, Beard, Burns, & Bomyea, 2009; Bar-Haim, 2010; Eldar et al., 2014; Hallion & Ruscio, 2011; Hoppitt et al., 2014; Schmidt, Richey, Buckner, & Timpano, 2009). This relatively new body of research indicates that ABM is successful in changing cognitions and these changes in attention processes are responsible for a change in behaviour. The findings of previous studies suggest that ABM may be an effective method for reducing hostility-related biases in aggression. Furthermore, Bowler et al. (2017) used cognitive bias modification techniques in anxious individuals to investigate whether implementing positive interpretation or attention training also had positive effects on the untrained cognitive domain. They found that attention bias training resulted in a reduced threat-related attention bias and an increase in positive interpretation bias. These results demonstrate the need for further work investigating the cognitive mechanisms which underlie both attention and interpretation processes (this formed part of the rationale for Study 5 (Chapter 7), and findings in relation to attention and interpretation bias are discussed in Section 8.2.2)). To my knowledge ABM has not been used as a rehabilitative method for attention bias in aggression. I suggest that EEG methodology could be used to assess the effectiveness of attention bias training in aggressive individuals.

8.2 Discussion – part two: Attention and interpretation

The thesis explored both attention and interpretation biases in aggression with the aim of better understanding how cognitive biases may contribute to aggressive behaviour. This section of the discussion will give a brief overview of the fifth empirical study before reviewing the results from the attention bias and interpretation bias studies together. Comparisons between the two types of bias will be made and a summary of how these results can contribute to the understanding of cognitive biases in aggression will be presented.

8.2.1 Overview of Study 5

The final chapter of the thesis used novel EEG techniques to investigate interpretation bias in aggression. The study used an explicit questionnaire measure, and implicit recognition task to measure hostile interpretation bias. EEG was recorded during completion of the implicit interpretation bias task. The first aim of the study was to assess the validity of using EEG to investigate interpretation bias during a recognition task. The second aim was to measure variations in interpretation bias in aggression using both behavioural and ERP techniques. The results provide evidence to suggest that ERPs are an appropriate method for assessing interpretation bias. Firstly the explicit and implicit measures of interpretation bias were consistent, showing that participants with increased scores on the AIHQ had a greater interpretation bias score on the recognition task (rated negative targets as more similar in meaning to previously presented scenarios, compared to positive target statements). These findings suggest that both measures are sensitive to hostility-related interpretation bias. Secondly, the evidence indicates that interpretation bias is characterised by a distinct ERP pattern. Results showed that participants showing no evidence of an explicit interpretation bias had increased amplitude in response to positive statements compared to negative statements during the recognition task, whereas participants demonstrating an explicit negative interpretation bias showed similar amplitude in response to positive and negative statements.

Regarding the second aim, the behavioural results provided evidence to support previous work (e.g. Dill et al., 1997; Dodge & Frame, 1982; Dodge et al., 1990; Epps & Kendall, 1995), which suggests a negative interpretation bias (reflected in increased scores on the AIHQ, and increased similarity ratings of negative target statements compared to positive target statements on the recognition task) in participants with increased aggression across both measures. The ERP results showed that low aggression participants had increased P300 amplitude when making similarity ratings of positive statements compared to when they made similarity ratings of negative statements. The high aggression group showed less differentiation in amplitude between statement types. These differences are similar to those found by Moser et al. (2008a) who explored the psycho-physiological correlates of interpretation bias in high and low socially anxious groups. They found that participants scoring low on social anxiety showed significantly different P600 amplitude in response to negative and positive sentence resolutions, whereas, participants scoring high on social anxiety showed similarity between both sentence resolutions. However, there is one main difference between the current findings conducted with high and low aggression groups, and those by Moser et al. (2008a) using low and high social anxiety groups. In the current study, those that reported low on aggression had increased amplitude in response to positive statements compared to negative, whereas, in the study by Moser et al. (2008a) participants scoring low on social anxiety showed increased amplitude to negative resolutions compared to positive resolutions. It is somewhat unexpected that these differences would be found in the participants who self-rate themselves as having only few anxious or aggressive tendencies.

8.2.2 Integration

Attention and interpretation are cognitive processes that not only interact with one another but influence other subsequent processes. The social information processing theory (Crick & Dodge, 1994) explains how attention and interpretation processes have an effect on the formation of behavioral responses to the environment. Therefore attention and interpretation biases should not be studied as distinct processes (White et al., 2011). It is suggested that poor identification of stimuli during the encoding and interpretations stages of processing results in the attribution of hostile intent in social situations (Crick & Dodge, 1994). Therefore, investigating attention and interpretation processes together contributes to the understanding of the larger picture of how cognitive biases potentially influence aggressive behaviour. Across studies presented in this thesis, there seem to be some similar ERP patterns associated with attention and interpretation bias in aggression. The main effect of stimulus valence is consistent across attention bias and interpretation bias results. Study 3 showed that low aggression participants show larger differences in amplitude when responding to a dot-probe task in which angry and neutral faces were presented, whereas high aggression participants show little differentiation between trial types. This effect is relatively consistent with the current results, which show that low aggression participants show significant differences in amplitude in response to differently valenced statements, whereas high aggression participants show relatively undifferentiated amplitudes. This suggests that hostile-related attention bias and interpretation bias in aggression is reflected in similar processing across positive and negative valenced stimuli.

The findings found for the interpretation bias task are particularly relevant to those found for attention bias to words. The effect in the current study is consistent with the main effect found in the attention bias word task (Study 1); participants showed increased amplitude to neutral trials compared to angry words. In the interpretation bias task (Study 5), low aggression participants show increased amplitude in response to positive compared to negative statements. This is in contrast to the study exploring attention bias to angry faces (Study 3) which showed increased amplitude to angry compared to neutral trials in the low aggression group.

These findings suggest that cognitive biases at attention and interpretation stages of processing may be reflected in similar neural patterns in response to a set of stimuli. Behaviourally, participants with increased levels of aggression showed an attention bias to angry words (study 1) and a hostility-related interpretation bias (study 5), and this was reflected in relatively stable amplitude when responding to positive and negative stimuli across both tasks (single words in the dot-probe task and full statements in the recognition task). Whereas, participants with lower aggression scores showed increased amplitude to neutral or positive stimuli (compared to angry) across both tasks. Findings that biases are consistent across processing stages are consistent with the Social Information Processing model (Crick & Dodge, 1994) which suggests that each of the six stages influence one another. The model suggests that biases in attention influence how stimuli is interpreted, and that this interpretation influences subsequent stages such as clarification of goals and response formation.

Taking into account the current findings I suggest that aggressive individuals are vigilant to angry stimuli in the environment, and they subsequently interpret social situations as more hostile. For example, they are more likely to notice an angry face in a crowd, and more likely to interpret an accidental push as aggressive provocation. The ERP results suggest that aggressive individuals require greater levels of attentional resources to disengage from hostile stimuli suggesting difficulties in overriding attention to negative information. These processes contribute to the decision to confront provoking behaviour, resulting in a potentially aggressive response. Consistent with Carver and Harmon-Jones (2009) I suggest that aggressive individuals are also more likely to be oriented towards approach motivations. Anger is associated with approach motivations when goal directed behaviour is disrupted meaning that a desired end point can not be reached. These motivations are subsequently associated with an increased liklihood of an aggressive behavioural response as individuals aim to remove the violation, and change the behaviour of others in order to reach the desired goal.

In line with neurocognitive models of aggression that suggest deficits in regulatory control and emotion regulation in physical aggression (e.g., Patrick, 2008; Wilkowski & Robinson, 2010), I suggest that participants with high aggression process both angry and neutral words similarly. They are likely to expect aggressive stimuli in their environment, and have hostile expectations in response to social situations. This may contribute to relatively stability when presented with angry and neutral stimuli. In contrast, participants scoring low on aggression may have a positivity bias in which they are better able to regulate attentional processes and allocate resources to neutral or happy words. The trait-congruency hypothesis (Blaney, 1986; Bower, 1981; Miranda & Persons, 1988) states that internal traits have a direct impact on the cognitive processes used when attending to the environment. Therefore individuals attend to stimuli that are consistent with their internal traits. Non-aggressive individuals are less likely to attend to angry stimuli in their environment, due to the ease in which they can disengage with angry words/faces and allocate resources elsewhere (supported by reaction time and ERP evidence). As attention bias is a cognitive process which influences subsequent interpretation processes and resulting behavior, they are also less likely to behave aggressively.

9 Concluding comments

This thesis aimed to investigate the neural correlates of attention and interpretation bias within individuals with increased levels of aggression. The review of the attention and interpretation bias literature revealed that there is a fairly robust behavioural association between hostile-related biases and aggressive behaviour. However, very little is known about the processes that contribute to these biases. Therefore, across five studies which recruited two undergraduate volunteer samples, both behavioural and ERP measures were used to explore between-group differences in cognitive biases. Attention bias was tested using the dot-probe paradigm and interpretation bias was tested using the recognition task; EEG was recorded during task completion. Across the four attention bias studies the aim was to explore if attentional processes involved with attending to stimuli during the dot-probe varied between aggression groups, between modalities, and between emotion of the presented stimuli. The final study tested the validity of the recognition task as a suitable measure for assessing neural correlates of interpretation bias, and investigated differences in aggression-related interpretation biases between-groups. Both behavioural data (reaction time/response data) and ERP data (evoked amplitude) was analysed to inform my conclusions.

The behavioural data from two of the four attention bias studies supported previous findings and was in line with predictions; individuals with increased levels of aggression had reduced reaction times to probes replacing angry stimuli compared to probes replacing neutral stimuli. The data from the fifth study also showed behavioural evidence of hostility-related interpretation biases in aggression.

Overall, the first set of attention bias findings, in which only angry and neutral stimuli were presented, indicated that low aggression participants differentiated between congruent and incongruent trials during the dot-probe task, whereas the high aggression participants showed much greater similarity in amplitude across trial types. It is suggested that uniformity in amplitude may be attributed to increase attentional resources recruited on incongruent trials to disengage with the simultaneously presented angry stimuli. The main effect of congruency was in opposite directions for different stimulus modalities. On the word task, participants showed increased amplitude on incongruent trials compared to congruent trials, whereas on the face task, participants exhibited increased amplitude to congruent trials compared with incongruent trials. This was consistent with the general task effect found in the replication attention bias studies (studies 2 and 4), in which angry, neutral and happy stimuli were presented. These follow-up results showed that on the word task participants had increased amplitude to positive stimuli (happy words), whereas on the face task individuals had increased amplitude to negative stimuli (angry faces) regardless of the simultaneously presented distracter stimuli. In contrast to predictions and to the findings of Study 3, this did not interact with aggression. I suggest that increased amplitude in response to happy words may reflect a positive bias in which individuals are able to avoid angry words (in line with reaction time data), whereas it may be harder to avoid attending to angry faces as they command attentional resources. Therefore, faces are detected quicker and are allocated greater attentional resources than words, as reflected in increased amplitude.

Echoing attention bias results from Study 3, results from Study 5 suggest that low aggression individuals show differentiations in amplitude when responding to differently valenced statements, and when making similar and dissimilar response ratings of such statements, whereas individuals with high aggression show more similar ERP patterns in response to both positive and negative statements. Taking together the results across all five studies, suggest that, compared to high aggression individuals, low aggression individuals differentiate between stimuli to a greater extent at both attention and interpretation stages of cognitive processing.

Although the interaction with aggression did not replicate across all studies, I believe that using an original design that uses both behavioural and ERP methods, this research advances the understanding of cognitive processes that contribute to attention and interpretation biases in aggression. Previous evidence suggests that modifying cognitions is an appropriate treatment method for changing behaviours. Therefore, it is suggested that understanding the cognitive processes that contribute to aggressive behaviour may be essential in designing rehabilitation programmes for aggressive offenders.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213.

Amir, N., Beard, C., Burns, M., & Bomyea, J. (2009). Attention modification program in individuals with generalized anxiety disorder. *Journal of Abnormal Psychology, 118*(1), 28.

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Psychology, 53*(1), 27.

Archer, J. (2004). Sex differences in aggression in real-world settings: a meta-analytic review. *Review of General Psychology, 8*(4), 291-322.

Argyle, M. (1994). *The psychology of interpersonal behaviour*. London: Penguin.

Armstrong, T., & Olatunji, B. O. (2012). Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis. *Clinical Psychology Review*, *32*(8), 704-723.

Aydin, O., & Markova, I. (1979). Attribution tendencies of popular and unpopular children. *British Journal of Social and Clinical Psychology, 18*(3), 291-298.

Bantin, T., Stevens, S., Gerlach, A. L., & Hermann, C. (2016). What does the facial dot-probe task tell us about attentional processes in social anxiety? A systematic review. *Journal of Behavior Therapy and Experimental Psychiatry, 50*, 40-51.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin, 133*(1), 1-24.

Bar-Haim, Y. (2010). Research review: attention bias modification (ABM): a novel treatment for anxiety disorders. *Journal of Child Psychology and Psychiatry, 51*(8), 859-870.

Barnes, L. L., Harp, D., & Jung, W. S. (2002). Reliability generalization of scores on the Spielberger state-trait anxiety inventory. *Educational and Psychological Measurement, 62*(4), 603-618.

Baron, R. A., & Richardson, D. R. (2004). *Human aggression*. New York: Springer.

Barratt, E. S., Stanford, M. S., Kent, T. A., & Alan, F. (1997). Neuropsychological and cognitive psychophysiological substrates of impulsive aggression. *Biological Psychiatry, 41*(10), 1045-1061.

Bartholow, B. D., Fabiani, M., Gratton, G., & Bettencourt, B. A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science, 12*(3), 197-204.

Bem, S. L. (1985). *Androgyny and gender schema theory: A conceptual and empirical integration.* Paper presented at the Nebraska symposium on motivation, 32, 179-226.

Berggren, N., Koster, E. H., & Derakshan, N. (2012). The effect of cognitive load in emotional attention and trait anxiety: An eye movement study. *Journal of Cognitive Psychology, 24*(1), 79-91.

Berkowitz, L. (1993). *Aggression: Its causes, consequences, and control*. New York: McGraw-Hill.

Bernat, E. M., Hall, J. R., Steffen, B. V., & Patrick, C. J. (2007). Violent offending predicts P300 amplitude. *International Journal of Psychophysiology, 66*(2), 161-167.

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.

Bishop, S. J. (2008). Neural mechanisms underlying selective attention to threat. *Annals of the New York Academy of Sciences, 1129*(1), 141-152.

Bjorkqvist, K., Osterman, K., & Lagerspetz, K. M. (1994). Sex differences in covert aggression among adults. *Aggressive Behavior, 20*(1), 27-33.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin, 99*(2), 229-246.

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*(2), 129-148.

Bowler, J., Hoppitt, L., Illingworth, J., Dalgleish, T., Ononaiye, M., Perez-Olivas, G., & Mackintosh, B. (2017). Asymmetrical transfer effects of cognitive bias modification: Modifying attention to threat influences interpretation of emotional ambiguity, but not vice versa. *Journal of Behavior Therapy and Experimental Psychiatry, 54*, 239-246.

Bradley, B. P., Mogg, K., Falla, S. J., & Hamilton, L. R. (1998). Attentional bias for threatening facial expressions in anxiety: Manipulation of stimulus duration. *Cognition & Emotion, 12*(6), 737-753.

Bradley, B. P., Mogg, K., & Lee, S. C. (1997). Attentional biases for negative information in induced and naturally occurring dysphoria. *Behaviour Research and Therapy, 35*, 911-927.

Bradley, B. P., Mogg, K., Millar, N., Bonham-Carter, C., Fergusson, E., Jenkins, J., & Parr, M. (1997). Attentional biases for emotional faces. *Cognition & Emotion, 11*(1), 25-42.

Bradley, B. P., Mogg, K., White, J., Groom, C., & Bono, J. (1999). Attentional bias for emotional faces in generalized anxiety disorder. *British Journal of Clinical Psychology, 38*(3), 267-278.

Brosch, T., Sander, D., Pourtois, G., & Scherer, K. R. (2008). Beyond fear: Rapid spatial orienting toward positive emotional stimuli. *Psychological Science, 19*(4), 362-370.

Brooks-Crozier, J. (2011). The nature and nurture of violence: early intervention services for the families of MAOA-low children as a means to reduce violent crime and the costs of violent crime. *Conneticut Law Review, 44*, 531-572.

Brugman, S., Lobbestael, J., Arntz, A., Cima, M., Schuhmann, T., Dambacher, F., & Sack, A. T. (2015). Identifying cognitive predictors of reactive and proactive aggression. *Aggressive Behavior*, *41*(1), 51-64.

Brugman, S., Lobbestael, J., von Borries, A. K. L., Bulten, B. E. H., Cima, M., Schuhmann, T., . . . Arntz, A. (2016). Cognitive predictors of violent incidents in forensic psychiatric inpatients. *Psychiatry Research, 237*, 229-237.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977-990.

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

Burnham, B. R. (2007). Displaywide visual features associated with a search display's appearance can mediate attentional capture. *Psychonomic Bulletin & Review*, *14*(3), 392-422.

Bushman, B. J., & Anderson, C. A. (2001). Is it time to pull the plug on hostile versus instrumental aggression dichotomy? *Psychological Review*, *108*(1), 273-279.

Buss, A. H., & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology, 21*, 343-349.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63*(3), 452-459.

Buss, D. M., & Shackelford, T. K. (1997). Human aggression in evolutionary psychological perspective. *Clinical Psychology Review, 17*(6), 605-619.

Cairns, R. B., Cairns, B. D., Neckerman, H. J., Ferguson, L. L., & Gariepy, J.-L. (1989). Growth and aggression: I. Childhood to early adolescence. *Developmental Psychology, 25*(2), 320-330.

Cannon, W. B. (1929). *Bodily changes in pain, hunger, fear, and rage*. Boston: Branford.

Carretié, L., Mercado, F., Tapia, M., & Hinojosa, J. A. (2001). Emotion, attention, and the ‘negativity bias’, studied through event-related potentials. *International Journal of Psychophysiology*, *41*(1), 75-85.

Carver, C. S., Avivi, Y. E., & Laurenceau, J. P. (2008). Approach, avoidance, and emotional experiences. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation* (pp. 385–397). New York: Psychology Press.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: evidence and implications. *Psychological Bulletin*, *135*(2), 183-204.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology, 67*(2), 319-333.

Chan, S.-C., Raine, A., & Lee, T. M. (2010). Attentional bias towards negative affect stimuli and reactive aggression in male batterers. *Psychiatry Research, 176*(2), 246-249.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. Personality and Social Psychology Bulletin, *25*, 215–224.

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.

Ciucci, E., Kimonis, E., Frick, P. J., Righi, S., Baroncelli, A., Tambasco, G., & Facci, C. (2018). Attentional orienting to emotional faces moderates the association between callous-unemotional traits and peer-nominated aggression in young adolescent school children. *Journal of Abnormal Child Psychology, 46*(5), 1011-1019.

Clarke, P. J., MacLeod, C., & Guastella, A. J. (2013). Assessing the role of spatial engagement and disengagement of attention in anxiety-linked attentional bias: a critique of current paradigms and suggestions for future research directions. *Anxiety, Stress & Coping*, *26*(1), 1-19.

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.

Cohen, D. J., Eckhardt, C. I., & Schagat, K. D. (1998). Attention allocation and habituation to anger-related stimuli during a visual search task. *Aggressive Behavior, 24*(6), 399-409.

Coles, M. G. H., Smid, H. G. O. M., Scheffers, M. K., & Otten, L. J. (1995). Mental chronometry and the study of human information processing. In M. D. Rugg & M. G. H. Coles (Eds.), *Oxford psychology series, No. 25. Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 86-131). New York: Oxford University Press.

Combs, D. R., Penn, D. L., Wicher, M., & Waldheter, E. (2007). The Ambiguous Intentions Hostility Questionnaire (AIHQ): a new measure for evaluating hostile social-cognitive biases in paranoia. *Cognitive Neuropsychiatry, 12*(2), 128-143.

Connor, K. M., Davidson, J. R., Churchill, L. E., Sherwood, A., Weisler, R. H., & Foa, E. (2000). Psychometric properties of the Social Phobia Inventory (SPIN): New self-rating scale. *The British Journal of Psychiatry*, *176*(4), 379-386.

Cooper, R. M., & Langton, S. R. (2006). Attentional bias to angry faces using the dot-probe task? It depends when you look for it. *Behaviour Research and Therapy*, *44*(9), 1321-1329.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron, 58*(3), 306-324.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201-215.

Coulson, S. (1998). ERPs and domain specificity: Beating a straw horse. *Language and Cognitive Processes, 13*(6), 653-672.

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), 74-101.

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

Crick, N. R., & Ladd, G. W. (1993). Children's perceptions of their peer experiences: Attributions, loneliness, social anxiety, and social avoidance. *Developmental Psychology, 29*(2), 244-254.

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biological Psychology*, *52*(2), 95-111.

Darwin, C., & Darwin, F. (2009). *The Expression of the Emotions in Man and Animals*. London: Cambridge University Press.

Davidson, R. J. (1993). Cerebral asymmetry and emotion: Conceptual and methodological conundrums. *Cognition & Emotion*, *7*(1), 115-138.

Davis, M., & Whalen, P. J. (2001). The amygdala: vigilance and emotion. *Molecular Psychiatry, 6*(1), 13-24.

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

Decety, J., & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *The Neuroscientist, 13*(6), 580-593.

Dehghani, M., Sharpe, L., & Nicholas, M. K. (2003). Selective attention to pain-related information in chronic musculoskeletal pain patients. *Pain, 105*(1), 37-46.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9-21.

Derryberry, D., & Reed, M. A. (2001). A multidisciplinary perspective on attentional control. In C. L. Folk & B. S. Gibson (Eds.), *Advances in psychology, 133. Attraction, distraction and action: Multiple perspectives on attentional capture* (pp. 325-347). New York: Elsevier Science.

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology, 111*(2), 225-236.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193-222.

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.

Dimberg, U., & Öhman, A. (1996). Behold the wrath: Psychophysiological responses to facial stimuli. *Motivation and Emotion, 20*(2), 149-182.

Dodge, K. A. (1980). Social Cognition and Children's Aggressive Behavior. *Child Development, 51*(1), 162-170.

Dodge, K. A. (1985). Attributional bias in aggressive children. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy* (pp. 73-110). San Diego: Academic Press.

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), 1146-1158.

Dodge, K. A., & Frame, C. L. (1982). Social Cognitive Biases and Deficits in Aggressive Boys. *Child Development, 53*(3), 620-635.

Dodge, K. A., & Newman, J. P. (1981). Biased decision-making processes in aggressive boys. *Journal of Abnormal Psychology, 90*(4), 375-379.

Dodge, K. A., Pettit, G. S., McClaskey, C. L., Brown, M. M., & Gottman, J. M. (1986). Social competence in children. *Monographs of the Society for Research in Child Development*, i-85.

Dodge, K. A., Price, J. M., Bachorowski, J.-A., & Newman, J. P. (1990). Hostile attributional biases in severely aggressive adolescents. *Journal of Abnormal Psychology, 99*(4), 385-392.

Dodge, K. A., & Somberg, D. R. (1987). Hostile attributional biases among aggressive boys are exacerbated under conditions of threats to the self. *Child Development*, 58(1), 213-224.

Dodge, K. A., & Tomlin, A. M. (1987). Utilization of self-schemas as a mechanism of interpretational bias in aggressive children. *Social Cognition, 5*(3), 280-300.

Dollard, J., Miller, N. E., Doob, L. W., Mowrer, O. H., & Sears, R. R. (1939). *Frustration and aggression*. New Haven: Yale University Press.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating. *Behavioral and Brain Sciences, 11*(3), 357-427.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412-431.

Duque, A., & Vázquez, C. (2015). Double attention bias for positive and negative emotional faces in clinical depression: Evidence from an eye-tracking study. *Journal of Behavior Therapy and Experimental Psychiatry*, *46*, 107-114.

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology, 14*(5), 456-467.

Eckhardt, C. I., & Cohen, D. J. (1997). Attention to anger-relevant and irrelevant stimuli following naturalistic insult. *Personality and Individual Differences, 23*(4), 619-629.

Ekman, P., & Friesen, W. V. (2003). *Unmasking the face: A guide to recognizing emotions from facial clues*: Cambridge: Malor Books.

Eldar, S., Apter, A., Lotan, D., Edgar, K. P., Naim, R., Fox, N. A., . . . Bar-Haim, Y. (2014). Attention bias modification treatment for pediatric anxiety disorders: a randomized controlled trial. *American Journal of Psychiatry*, 169(2), 213-230.

Eldar, S., & Bar-Haim, Y. (2010). Neural plasticity in response to attention training in anxiety. *Psychological Medicine, 40*(04), 667-677.

Eldar, S., Yankelevitch, R., Lamy, D., & Bar-Haim, Y. (2010). Enhanced neural reactivity and selective attention to threat in anxiety. *Biological Psychology, 85*(2), 252-257.

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

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion, 7*(2), 336-353.

Fanning, J. R., Berman, M. E., & Long, J. M. (2014). P3 and provoked aggressive behavior. *Social Neuroscience, 9*(2), 118-129.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160.

Faunce, G. J., Mapledoram, P. K., & Job, R. S. (2004). Type A behaviour pattern and attentional bias in relation to anger/hostility, achievement, and failure. *Personality and Individual Differences, 36*(8), 1975-1988.

Feldman, J. A., & Ballard, D. H. (1982). Connectionist models and their properties. *Cognitive Science, 6*(3), 205-254.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage.

Fitzgerald, P., & Asher, S. (1987). *Aggressive-rejected children's attributional biases about liked and disliked peers.* Paper presented at the annual meeting of the American Psychological Association, New York.

Ford, B. Q., Tamir, M., Gagnon, S. A., Taylor, H. A., & Brunyé, T. T. (2012). The angry spotlight: Trait anger and selective visual attention to rewards. *European Journal of Personality, 26*(2), 90-98.

Foti, D., & Hajcak, G. (2008). Deconstructing reappraisal: Descriptions preceding arousing pictures modulate the subsequent neural response. *Journal of Cognitive Neuroscience, 20*(6), 977-988.

Fox, E. (1993). Attentional bias in anxiety: selective or not? *Behaviour Research and Therapy, 31*(5), 487-493.

Fox, E., Derakshan, N., & Shoker, L. (2008). Trait anxiety modulates the electrophysiological indices of rapid spatial orienting towards angry faces. *Neuroreport, 19*(3), 259-263.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition & Emotion, 14*(1), 61-92.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130*(4), 681-700.

Fox, E., Russo, R., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition & Emotion, 16*(3), 355-379.

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.

Gao, Y., & Raine, A. (2009). P3 event-related potential impairments in antisocial and psychopathic individuals: A meta-analysis. *Biological Psychology, 82*(3), 199-210.

Gao, Y., Raine, A., Venables, P. H., & Mednick, S. A. (2013). The association between P3 amplitude at age 11 and criminal offending at age 23. *Journal of Clinical Child & Adolescent Psychology, 42*(1), 120-130.

García-Sancho, E., Salguero, J. M., & Fernández-Berrocal, P. (2014). Relationship between emotional intelligence and aggression: A systematic review. *Aggression and Violent Behavior*, *19*(5), 584-591.

Geen, R. G. (2001). *Human aggression*. Buckingham: Open University Press.

Geisser, S., & Greenhouse, S. W. (1958). An extension of box's results on the use of the $ F $ distribution in multivariate analysis. *The Annals of Mathematical Statistics*, *29*(3), 885-891.

Gerstle, J. E., Mathias, C. W., & Stanford, M. S. (1998). Auditory P300 and self-reported impulsive aggression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 22*(4), 575-583.

Giancola, P. R., & Parrott, D. J. (2008). Further evidence for the validity of the Taylor aggression paradigm. *Aggressive Behavior, 34*(2), 214-229.

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74-78.

Gilbert, P. (1998). The evolved basis and adaptive functions of cognitive distortions. *British Journal of Medical Psychology, 71*(4), 447-463.

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.

Goetz, T. E., & Dweck, C. S. (1980). Learned helplessness in social situations. *Journal of Personality and Social Psychology, 39*(2), 246-255.

Gotlib, I. H., & McCann, C. D. (1984). Construct accessibility and depression: An examination of cognitive and affective factors. *Journal of Personality and Social Psychology, 47*(2), 427-439.

Green, M. J., & Phillips, M. L. (2004). Social threat perception and the evolution of paranoia. *Neuroscience & Biobehavioral Reviews, 28*(3), 333-342.

Green, M. J., Williams, L. M., & Davidson, D. (2003). Visual scanpaths to threat-related faces in deluded schizophrenia. *Psychiatry Research, 119*(3), 271-285.

Guerra, N. G., & Slaby, R. G. (1989). Evaluative factors in social problem solving by aggressive boys. *Journal of Abnormal Child Psychology, 17*(3), 277-289.

Hadwin, J., Frost, S., French, C. C., & Richards, A. (1997). Cognitive processing and trait anxiety in typically developing children: Evidence for an interpretation bias. *Journal of Abnormal Psychology, 106*(3), 486-490.

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental Neuropsychology, 35*(2), 129-155.

Hajcak, G., & Olvet, D. M. (2008). The persistence of attention to emotion: brain potentials during and after picture presentation. *Emotion, 8*(2), 250-255.

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. A., Leibenluft, E., . . . Pine, D. S. (2010). Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. *Biological Psychiatry, 68*(11), 982-990.

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

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin, 137*(6), 940-958.

Hansen, C. H., & Hansen, R. D. (1994). Automatic emotion: Attention and facial efference. In P.M. Niedenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention (pp. 217-243)*. San Diego: Academic Press.

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology, 54*(6), 917-924.

Harmon-Jones, E., Barratt, E. S., & Wigg, C. (1997). Impulsiveness, aggression, reading, and the P300 of the event-related potential. *Personality and Individual Differences, 22*(4), 439-445.

Harris, J. A. (1997). A further evaluation of the Aggression Questionnaire: Issues of validity and reliability. *Behaviour Research and Therapy, 35*(11), 1047-1053.

Hazaleus, S. L., & Deffenbacher, J. L. (1986). Relaxation and cognitive treatments of anger. *Journal of Consulting and Clinical Psychology, 54*(2), 222-226.

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

Helfritz-Sinville, L. E., & Stanford, M. S. (2015). Looking for Trouble? Processing of Physical and Social Threat Words in Impulsive and Premeditated Aggression. *The Psychological Record, 65*(2), 301-314.

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences, 95*(3), 781-787.

Hillyard, S. A., & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual Review of Psychology, 34*(1), 33-61.

Hillyard, S. A., Mangun, G. R., Woldorff, M. G., & Luck, S. J. (1995). Neural systems mediating selective attention. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 665-681). Cambridge: The MIT Press.

Holmes, A., Bradley, B. P., Nielsen, M. K., & Mogg, K. (2009). Attentional selectivity for emotional faces: Evidence from human electrophysiology. *Psychophysiology, 46*(1), 62-68.

Holmes, A., Mogg, K., de Fockert, J., Nielsen, M. K., & Bradley, B. P. (2014). Electrophysiological evidence for greater attention to threat when cognitive control resources are depleted. *Cognitive, Affective, & Behavioral Neuroscience, 14*(2), 827-835.

Hoppitt, L., Illingworth, J. L., MacLeod, C., Hampshire, A., Dunn, B. D., & Mackintosh, B. (2014). Modifying social anxiety related to a real-life stressor using online Cognitive Bias Modification for interpretation. *Behaviour Research and Therapy, 52*, 45-52.

Houston, R. J., Stanford, M. S., Villemarette Pittman, N. R., Conklin, S. M., & Helfritz, L. E. (2004). Neurobiological correlates and clinical implications of aggressive subtypes. *Journal of Forensic Neuropsychology, 3*(4), 67-87.

Hu, B., Rao, J., Li, X., Cao, T., Li, J., Majoe, D., & Gutknecht, J. (2017). Emotion Regulating Attentional Control Abnormalities In Major Depressive Disorder: An Event-Related Potential Study. *Scientific Reports*, *7*(1), 13530.

Huesmann, L. R. (1998). The role of social information processing and cognitive schema in the acquisition and maintenance of habitual aggressive behavior. New York: Academic Press.

Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2015). The median split: Robust, refined, and revived. *Journal of Consumer Psychology, 25*(4), 690-704.

Iwaki, S., Sutani, K., Kou, H., & Tonoike, M. (2007). *Modeling multiple neuromagnetic activities during the visual infrequent target detection processing.* Paper presented at the International Congress Series.

Jinghua, L., & Yong, Z. (2014). Attentional Bias in Individuals with Different Level of Implicit/Explicit Aggression: Behavioral and ERP Evidence. *Journal of Psychological Science, 1,* 007.

Judah, M. R., Grant, D. M., Mills, A. C., & Lechner, W. V. (2014). Factor structure and validation of the attentional control scale. *Cognition & Emotion, 28*(3), 433-451.

Kappenman, E. S., Farrens, J. L., Luck, S. J., & Proudfit, G. H. (2015). Behavioral and ERP measures of attentional bias to threat in the dot-probe task: Poor reliability and lack of correlation with anxiety. *Frontiers in Psychology, 5,* 1-9.

Kirita, T., & Endo, M. (1995). Happy face advantage in recognizing facial expressions. *Acta Psychologica, 89*(2), 149-163.

Knight, G. P., Fabes, R. A., & Higgins, D. A. (1996). Concerns about drawing causal inferences from meta-analyses: An example in the study of gender differences in aggression. *Psychological Bulletin, 119*(3), 410-421.

Knight, R. T., Scabini, D., Woods, D. L., & Clayworth, C. C. (1989). Contributions of temporal-parietal junction to the human auditory P3. *Brain Research, 502*(1), 109-116.

Koster, E. H., Crombez, G., Verschuere, B., & De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: Differentiating vigilance and difficulty to disengage. *Behaviour Research and Therapy, 42*(10), 1183-1192.

Koster, E. H., Crombez, G., Verschuere, B., Van Damme, S., & Wiersema, J. R. (2006). Components of attentional bias to threat in high trait anxiety: Facilitated engagement, impaired disengagement, and attentional avoidance. *Behaviour Research and Therapy, 44*(12), 1757-1771.

Kovacs, M., & Beck, A. T. (1978). Maladaptive cognitive structures in depression. *American Journal of Psychiatry, 135*(5), 525-533.

Krall, S., Rottschy, C., Oberwelland, E., Bzdok, D., Fox, P., Eickhoff, S., . . . Konrad, K. (2015). The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. *Brain Structure and Function, 220*(2), 587-604.

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.

Ladd, G. W., & Mize, J. (1983). A cognitive–social learning model of social-skill training. *Psychological Review, 90*(2), 127-157.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133*(3), 339-354.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23*(1), 155-184.

LeDoux, J. (2003). Emotion circuits in the brain. *Annual Review of Neuroscience, 23*, 155-184.

LeDoux, J. E. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon & Schuster.

Leppänen, J. M., Moulson, M. C., Vogel-Farley, V. K., & Nelson, C. A. (2007). An ERP study of emotional face processing in the adult and infant brain. *Child Development, 78*(1), 232-245.

Liossi, C., White, P., & Schoth, D. E. (2011). Time-course of attentional bias for threat-related cues in patients with chronic daily headache–tension type: Evidence for the role of anger. *European Journal of Pain, 15*(1), 92-98.

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.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8,* 1-14.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, London: MIT Press.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*(4), 1122-1135.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology, 95*(1), 15-20.

MacLeod, C., & Rutherford, E. M. (1992). Anxiety and the selective processing of emotional information: Mediating roles of awareness, trait and state variables, and personal relevance of stimulus materials. *Behaviour Research and Therapy, 30*(5), 479-491.

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: Assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology, 111*(1), 107-123.

Maoz, K., Adler, A. B., Bliese, P. D., Sipos, M. L., Quartana, P. J., & Bar-Haim, Y. (2017). Attention and interpretation processes and trait anger experience, expression, and control. *Cognition and Emotion, 31*(7), 1453-1464.

Martin, M., Williams, R. M., & Clark, D. M. (1991). Does anxiety lead to selective processing of threat-related information? *Behaviour Research and Therapy, 29*(2), 147-160.

Mathews, A., & Mackintosh, B. (1998). A cognitive model of selective processing in anxiety. *Cognitive Therapy and Research, 22*(6), 539-560.

Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology, 109*(4), 602-615.

Mathews, A., & MacLeod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy, 23*(5), 563-569.

Mathews, A., & MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. *Annual Review of Psychology, 45*(1), 25-50.

Mathias, C. W., & Stanford, M. S. (1999). P300 under standard and surprise conditions in self-reported impulsive aggression. *Progress in Neuro-psychopharmacology and Biological Psychiatry, 23*(6), 1037-1051.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255-260.

Mayer, J. D., & Salovey, P. (1997).What is emotional intelligence? In Salovey, P., & Sluyter, D. (Eds.) *Emotional development and emotional intelligence: Implications for educators* (pp.3-31). New York: Basic Books.

McClelland, G. H., Lynch Jr, J. G., Irwin, J. R., Spiller, S. A., & Fitzsimons, G. J. (2015). Median splits, Type II errors, and false–positive consumer psychology: Don't fight the power. *Journal of Consumer Psychology*, *25*(4), 679-689.

McKenna, C. J. (1986). Theories of individual search behaviour. *Bulletin of Economic Research, 38*(3), 189-207.

Meesters, C., Muris, P., & van Rooijen, B. (2007). Relations of neuroticism and attentional control with symptoms of anxiety and aggression in non-clinical children. *Journal of Psychopathology and Behavioral Assessment, 29*(3), 149-158.

Mellentin, A. I., Dervisevic, A., Stenager, E., Pilegaard, M., & Kirk, U. (2015). Seeing enemies? A systematic review of anger bias in the perception of facial expressions among anger-prone and aggressive populations. *Aggression and Violent Behavior, 25*, 373-383.

Micco, J. A., Henin, A., & Hirshfeld-Becker, D. R. (2014). Efficacy of interpretation bias modification in depressed adolescents and young adults. *Cognitive Therapy and Research, 38*(2), 89-102.

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

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*(1), 167-202.

Mingtian, Z., Xiongzhao, Z., Jinyao, Y., Shuqiao, Y., & Atchley, R. A. (2011). Do the early attentional components of ERPs reflect attentional bias in depression? It depends on the stimulus presentation time. *Clinical Neurophysiology*, *122*(7), 1371-1381.

Miranda, J., & Persons, J. B. (1988). Dysfunctional attitudes are mood-state dependent. *Journal of Abnormal Psychology, 97*(1), 76-79.

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy, 36*(9), 809-848.

Mogg, K., & Bradley, B. P. (1999a). Orienting of attention to threatening facial expressions presented under conditions of restricted awareness. *Cognition & Emotion*, *13*(6), 713-740.

Mogg, K., & Bradley, B. P. (1999b). Selective attention and anxiety: A cognitive–motivational perspective. In T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 145-170). New York: John Wiley & Sons Ltd.

Mogg, K., & Bradley, B. P. (1999c). Some methodological issues in assessing attentional biases for threatening faces in anxiety: A replication study using a modified version of the probe detection task. *Behaviour Research and Therapy, 37*(6), 595-604.

Mogg, K., & Bradley, B. P. (2005). Attentional bias in generalized anxiety disorder versus depressive disorder. *Cognitive Therapy and Research, 29*(1), 29-45.

Mogg, K., Bradley, B. P., De Bono, J., & Painter, M. (1997). Time course of attentional bias for threat information in non-clinical anxiety. *Behaviour Research and Therapy, 35*(4), 297-303.

Mogg, K., Bradley, B. P., Millar, N., & White, J. (1995a). A follow-up study of cognitive bias in generalized anxiety disorder. *Behaviour Research and Therapy, 33*(8), 927-935.

Mogg, K., Bradley, B. P., & Williams, R. (1995b). Attentional bias in anxiety and depression: The role of awareness. *British Journal of Clinical Psychology, 34*(1), 17-36.

Mogg, K., Bradley, B. P., Williams, R., & Mathews, A. (1993). Subliminal processing of emotional information in anxiety and depression. *Journal of Abnormal Psychology, 102*(2), 304-311.

Mogg, K., Mathews, A., & Eysenck, M. (1992). Attentional bias to threat in clinical anxiety states. *Cognition & Emotion, 6*(2), 149-159.

Mogg, K., Mathews, A., & Weinman, J. (1989). Selective processing of threat cues in anxiety states: A replication. *Behaviour Research and Therapy, 27*(4), 317-323.

Mogg, K., Millar, N., & Bradley, B. P. (2000). Biases in eye movements to threatening facial expressions in generalized anxiety disorder and depressive disorder. *Journal of Abnormal Psychology, 109*(4), 695-704.

Mogg, K., Philippot, P., & Bradley, B. P. (2004). Selective attention to angry faces in clinical social phobia. *Journal of Abnormal Psychology, 113*(1), 160-165.

Mogoaşe, C., David, D., & Koster, E. H. (2014). Clinical efficacy of attentional bias modification procedures: An updated meta-analysis. *Journal of Clinical Psychology, 70*(12), 1133-1157.

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

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

Moser, J. S., Huppert, J. D., Duval, E., & Simons, R. F. (2008b). Face processing biases in social anxiety: an electrophysiological study. *Biological Psychology, 78*(1), 93-103.

Moser, J. S., Huppert, J. D., Foa, E. B., & Simons, R. F. (2012). Interpretation of ambiguous social scenarios in social phobia and depression: evidence from event-related brain potentials. *Biological Psychology, 89*(2), 387-397.

Mueller, E., Hofmann, S., Santesso, D., Meuret, A., Bitran, S., & Pizzagalli, D. A. (2009). Electrophysiological evidence of attentional biases in social anxiety disorder. *Psychological Medicine, 39*(07), 1141-1152.

Mueller, E. M., Hofmann, S. G., Santesso, D. L., Meuret, A. E., Bitran, S., & Pizzagalli, D. A. (2009). Electrophysiological evidence of attentional biases in social anxiety disorder. *Psychological Medicine, 39*(07), 1141-1152.

Muris, P., & Field, A. P. (2008). Distorted cognition and pathological anxiety in children and adolescents. *Cognition and Emotion, 22*(3), 395-421.

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), 459-468.

Nesse, R. (1998). Emotional disorders in evolutionary perspective. *British Journal of Medical Psychology, 71*(4), 397-415.

Nesse, R. M. (1994). Fear and fitness: An evolutionary analysis of anxiety disorders. *Ethology and Sociobiology, 15*(5), 247-261.

O'Toole, L., & Dennis, T. A. (2012). Attention training and the threat bias: An ERP study. *Brain and Cognition, 78*(1), 63-73.

Öhman, A. (2005). The role of the amygdala in human fear: automatic detection of threat. *Psychoneuroendocrinology, 30*(10), 953-958.

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: a threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80*(3), 381-396.

Oliver-Rodríguez, J. C., Guan, Z., & Johnston, V. S. (1999). Gender differences in late positive components evoked by human faces. *Psychophysiology, 36*(02), 176-185.

Oster, H., Daily, L., & Goldenthal, P. (2013). Processing facial affect. In: Young, A. W. & Ellis, H. D. (Eds). *Handbook of research on face processing*. Amsterdam: Elsevier.

Parrott, D. J., & Zeichner, A. (2002). Effects of alcohol and trait anger on physical aggression in men. *Journal of Studies on Alcohol, 63*(2), 196-204.

Pashler, H. E., & Sutherland, S. (1998). *The psychology of attention* (Vol. 15). Cambridge, MA: MIT press.

Patrick, C. J. (2008). Psychophysiological correlates of aggression and violence: an integrative review. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363*(1503), 2543-2555.

Piaget, J., & Cook, M. (1953). *The Origin of Intelligence in the Child*. New York: Routledge & Kegan Paul Limited.

Pineles, S. L., & Mineka, S. (2005). Attentional biases to internal and external sources of potential threat in social anxiety. *Journal of Abnormal Psychology, 114*(2), 314-318.

Pishyar, R., Harris, L. M., & Menzies, R. G. (2004). Attentional bias for words and faces in social anxiety. *Anxiety, Stress & Coping, 17*(1), 23-36.

Placanica, J. L., Faunce, G. J., & Soames Job, R. (2002). The effect of fasting on attentional biases for food and body shape/weight words in high and low eating disorder inventory scorers. *International Journal of Eating Disorders, 32*(1), 79-90.

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.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128-2148.

Polich, J., & Criado, J. R. (2006). Neuropsychology and neuropharmacology of P3a and P3b. *International Journal of Psychophysiology, 60*(2), 172-185.

Poulsen, C., Luu, P., Davey, C., & Tucker, D. M. (2005). Dynamics of task sets: evidence from dense-array event-related potentials. *Cognitive Brain Research*, *24*(1), 133-154.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3-25.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*(1), 25-42.

Price, R. B., Kuckertz, J. M., Siegle, G. J., Ladouceur, C. D., Silk, J. S., Ryan, N. D., ... & Amir, N. (2015). Empirical recommendations for improving the stability of the dot-probe task in clinical research. *Psychological Assessment*, *27*(2), 365-376.

Pritchard, W. S. (1981). Psychophysiology of P300. *Psychological Bulletin, 89*(3), 506-540.

Putman, P., Hermans, E., & van Honk, J. (2004). Emotional stroop performance for masked angry faces: it's BAS, not BIS. *Emotion, 4*(3), 305-311.

Qualtrics, Version [October, 2015]. Copyright © [2017] Qualtrics. Qualtrics and all other Qualtrics product or service names are registered trademarks or trademarks of Qualtrics, Provo, UT, USA. http://www.qualtrics.com

Quay, H. C., & Peterson, D. R. (1987). *Manual for the Revised Behavior Problem Checklist.* Miami: Coral Gables.

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.

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,* 32, 159–171.

Richards, A., & French, C. C. (1990). Central versus peripheral prsentation of stimuli in an emotional stroop task. *Anxiety Research, 3*(1), 41-49.

Richards, A., & Millwood, B. (1989). Colour-identification of differentially valenced words in anxiety. *Cognition and Emotion, 3*(2), 171-176.

Roy, A. K., Vasa, R. A., Bruck, M., Mogg, K., Bradley, B. P., Sweeney, M., . . . Team, C. (2008). Attention bias toward threat in pediatric anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(10), 1189-1196.

Rubin, K. H., & Krasnor, L. R. (1986). Social-cognitive and social behavioral perspectives on problem solving. In *Cognitive perspectives on children's social and behavioral development. The Minnesota Symposia on Child Psychology* (pp. 1-68). Hillsdale: Erlbaum.

Rule, B. G., Taylor, B. R., & Dobbs, A. R. (1987). Priming effects of heat on aggressive thoughts. *Social Cognition, 5*(2), 131-143.

Salemink, E., & van den Hout, M. (2010). Validation of the “recognition task” used in the training of interpretation biases. *Journal of Behavior Therapy and Experimental Psychiatry, 41*(2), 140-144.

Salum, G. A., Mogg, K., Bradley, B. P., Gadelha, A., Pan, P., Tamanaha, A. C., ... & do Rosario, M. C. (2013). Threat bias in attention orienting: evidence of specificity in a large community-based study. *Psychological Medicine, 43*(4), 733-745.

Sancilio, M. F., Plumert, J. M., & Hartup, W. W. (1989). Friendship and aggressiveness as determinants of conflict outcomes in middle childhood. *Developmental Psychology, 25*(5), 812-819.

Santesso, D. L., Meuret, A. E., Hofmann, S. G., Mueller, E. M., Ratner, K. G., Roesch, E. B., & Pizzagalli, D. A. (2008). Electrophysiological correlates of spatial orienting towards angry faces: a source localization study. *Neuropsychologia, 46*(5), 1338-1348.

Sass, S. M., Evans, T. C., Xiong, K., Mirghassemi, F., & Tran, H. (2017). Attention training to pleasant stimuli in anxiety. *Biological psychology*, *122*, 80-92.

Sass, S. M., Heller, W., Fisher, J. E., Silton, R. L., Stewart, J. L., Crocker, L. D., . . . Miller, G. A. (2014). Electrophysiological evidence of the time course of attentional bias in non-patients reporting symptoms of depression with and without co-occurring anxiety. *Frontiers in Psychology, 5*, 1-11.

Sass, S. M., Heller, W., Stewart, J. L., Silton, R. L., Edgar, J. C., Fisher, J. E., & Miller, G. A. (2010). Time course of attentional bias in anxiety: emotion and gender specificity. *Psychophysiology*, *47*(2), 247-259.

Schacht, A., & Sommer, W. (2009). Emotions in word and face processing: early and late cortical responses. *Brain and Cognition, 69*(3), 538-550.

Schmukle, S. C. (2005). Unreliability of the dot probe task. *European Journal of Personality, 19*(7), 595-605.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*: Psychology Software Incorporated.

Schmidt, N. B., Richey, J. A., Buckner, J. D., & Timpano, K. R. (2009). Attention training for generalized social anxiety disorder. *Journal of Abnormal Psychology, 118*(1), 5-14.

Schulkin, J. (2003). *Rethinking homeostasis: Allostatic regulation in physiology and pathophysiology*. Cambridge, MA: MIT Press.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology*, 37(2), 257-261.

Schupp, H., Cuthbert, B., Bradley, M., Hillman, C., Hamm, A., & Lang, P. (2004a). Brain processes in emotional perception: Motivated attention. *Cognition and Emotion*, *18*(5), 593-611.

Schupp, H. T., Öhman, A., Junghöfer, M., Weike, A. I., Stockburger, J., & Hamm, A. O. (2004b). The facilitated processing of threatening faces: an ERP analysis. *Emotion*, *4*(2), 189-200.

See, J., MacLeod, C., & Bridle, R. (2009). The reduction of anxiety vulnerability through the modification of attentional bias: a real-world study using a home-based cognitive bias modification procedure. *Journal of Abnormal Psychology, 118*(1), 65-75.

Serin, R. C. (1991). Psychopathy and violence in criminals. *Journal of Interpersonal Violence*, *6*(4), 423-431.

Slaby, R. G., & Guerra, N. G. (1988). Cognitive mediators of aggression in adolescent offenders: I. Assessment. *Developmental Psychology*, *24*(4), 580-588.

Smith, N. K., Cacioppo, J. T., Larsen, J. T., & Chartrand, T. L. (2003). May I have your attention, please: Electrocortical responses to positive and negative stimuli. *Neuropsychologia*, *41*(2), 171-183.

Smith, N. K., Larsen, J. T., Chartrand, T. L., Cacioppo, J. T., Katafiasz, H. A., & Moran, K. E. (2006). Being bad isn't always good: Affective context moderates the attention bias toward negative information. *Journal of Personality and Social Psychology*, *90*(2), 210-220.

Smith, C. A., McHugo, G. J., & Kappas, A. (1996). Epilogue: Overarching themes and enduring contributions of the Lanzetta research program. *Motivation and Emotion*, *20*(3), 237-253.

Smith, P., & Waterman, M. (2003). Processing bias for aggression words in forensic and nonforensic samples. *Cognition & Emotion, 17*(5), 681-701.

Smith, P., & Waterman, M. (2005). Sex differences in processing aggression words using the emotional Stroop task. *Aggressive Behavior, 31*(3), 271-282.

Spielberger, C. D. (1991). *State-Trait Anger Expression Inventory: Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Spielberger, C. D. (1999). *STAXI-2: State-Trait Anger Expression Inventory-2 : Professional Manual*. Lutz, FL: Psychological Assessment Resources.

Spielberger, C. D., & Gorsuch, R. L. (1970). *STAI Manual for the State-trait Anxiety Inventory ("Self-evaluation Questionnaire")*. Consulting Psychologists Press.

Spielberger, C. D., Jacobs, G., Russell, S., & Crane, R. S. (1983). Assessment of anger: The state-trait anger scale. *Advances in Personality Assessment, 2*, 159-187.

Steinberg, M. S., & Dodge, K. A. (1983). Attributional Bias in Aggressive Adolescent Boys and Girls. *Journal of Social and Clinical Psychology, 1*(4), 312-321.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, *57*(4), 421-457.

Stenberg, G., Wiking, S., & Dahl, M. (1998). Judging words at face value: Interference in a word processing task reveals automatic processing of affective facial expressions. *Cognition & Emotion*, *12*(6), 755-782.

Stewart, J. L., Silton, R. L., Sass, S. M., Fisher, J. E., Edgar, J. C., Heller, W., & Miller, G. A. (2010). Attentional bias to negative emotion as a function of approach and withdrawal anger styles: an ERP investigation. *International Journal of Psychophysiology, 76*(1), 9-18.

Stormark, K. M., Nordby, H., & Hugdahl, K. (1995). Attentional shifts to emotionally charged cues: Behavioural and ERP data. *Cognition & Emotion*, *9*(5), 507-523.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*(3), 220-247.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643-662.

Sun, D., Chan, C. C., Fan, J., Wu, Y., & Lee, T. (2015). Are happy faces attractive? The roles of early vs. late processing. *Frontiers in Psychology*, *6*, 1812e.

Surguy, S. M., & Bond, A. J. (2006). P300 to emotionally relevant stimuli as an indicator of aggression levels. *Aggressive Behavior*, *32*(3), 253-260.

Sutton, T. M., & Altarriba, J. (2011). The automatic activation and perception of emotion in word processing: Evidence from a modified dot probe paradigm. *Journal of Cognitive Psychology*, *23*(6), 736-747.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics*. Boston: Pearson Education.

Tarry, H., & Emler, N. (2007). Attitudes, values and moral reasoning as predictors of delinquency. *British Journal of Developmental Psychology*, *25*(2), 169-183.

Tatarunaite, E., Playle, R., Hood, K., Shaw, W., & Richmond, S. (2005). Facial attractiveness: a longitudinal study. *American Journal of Orthodontics and Dentofacial Orthopedics*, *127*(6), 676-682.

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

Telzer, E. H., Mogg, K., Bradley, B. P., Mai, X., Ernst, M., Pine, D. S., & Monk, C. S. (2008). Relationship between trait anxiety, prefrontal cortex, and attention bias to angry faces in children and adolescents. *Biological Psychology, 79*(2), 216-222.

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77-99.

Thomas, S. J., Johnstone, S. J., & Gonsalvez, C. J. (2007). Event-related potentials during an emotional Stroop task. *International Journal of Psychophysiology, 63*(3), 221-231.

Todorov, A., & Bargh, J. A. (2002). Automatic sources of aggression. *Aggression and Violent Behavior, 7*(1), 53-68.

Torrence, R. D., & Troup, L. J. (2018). Event-related potentials of attentional bias toward faces in the dot-probe task: A systematic review. *Psychophysiology, 55*(6), 1-20.

Van Bockstaele, B., Verschuere, B., Tibboel, H., De Houwer, J., Crombez, G., & Koster, E. H. (2014). A review of current evidence for the causal impact of attentional bias on fear and anxiety. *Psychological Bulletin, 140*(3), 682-721.

Van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22*(2), 241-255.

van Honk, J., Tuiten, A., de Haan, E., vann de Hout, M., & Stam, H. (2001). Attentional biases for angry faces: Relationships to trait anger and anxiety. *Cognition & Emotion, 15*(3), 279-297.

van Honk, J., Tuiten, A., van den Hout, M., Putman, P., de Haan, E., & Stam, H. (2001). Selective attention to unmasked and masked threatening words: Relationships to trait anger and anxiety. *Personality and Individual Differences, 30*(4), 711-720.

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

Verleger, R., Heide, W., Butt, C., & Kömpf, D. (1994). Reduction of P3 b in patients with temporo-parietal lesions. *Cognitive Brain Research, 2*(2), 103-116.

Verwoerd, J., de Jong, P. J., & Wessel, I. (2008). Low attentional control and the development of intrusive memories following a laboratory stressor. *Journal of Psychopathology and Behavioral Assessment, 30*(4), 291-297.

Vogt, J., De Houwer, J., Crombez, G., Koster, E.H.W., & Van Damme, S. (2008). Examining the attentional bias to emotional stimuli: A crucial role for arousal level. *Emotion, 8,* 880-885

Wadsworth, B. J. (1996). *Piaget's theory of cognitive and affective development: Foundations of constructivism*. White Plains: Longman Publishing.

Waechter, S., Nelson, A. L., Wright, C., Hyatt, A., & Oakman, J. (2014). Measuring attentional bias to threat: Reliability of dot probe and eye movement indices. *Cognitive Therapy and Research*, *38*(3), 313-333.

Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science*, *23*(2), 169-177.

Waters, A. M., Craske, M. G., Bergman, R. L., & Treanor, M. (2008a). Threat interpretation bias as a vulnerability factor in childhood anxiety disorders. *Behaviour Research and Therapy, 46*(1), 39-47.

Waters, A. M., Henry, J., Mogg, K., Bradley, B. P., & Pine, D. S. (2010). Attentional bias towards angry faces in childhood anxiety disorders. *Journal of Behavior Therapy and Experimental Psychiatry, 41*(2), 158-164.

Waters, A. M., Mogg, K., Bradley, B. P., & Pine, D. S. (2008b). Attentional bias for emotional faces in children with generalized anxiety disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(4), 435-442.

Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Colour naming of phobia-related words. *British Journal of Psychology*, *77*(1), 97-108.

Weems, C. F., Costa, N. M., Watts, S. E., Taylor, L. K., & Cannon, M. F. (2007). Cognitive Errors, Anxiety Sensitivity, and Anxiety Control

Beliefs: Their Unique and Specific Associations With Childhood Anxiety Symptoms. *Behavior Modification, 31*(2), 174-201.

Weiner, B., & Graham, S. (1984). An attributional approach to emotional development. *Emotions, Cognition, and Behavior*, 167-191.

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record, 58*(3), 475-482.

White, L. K., Suway, J. G., Pine, D. S., Bar-Haim, Y., & Fox, N. A. (2011). Cascading effects: The influence of attention bias to threat on the interpretation of ambiguous information. *Behaviour Research and Therapy, 49*(4), 244-251.

Wilkowski, B. M., & Robinson, M. D. (2008a). The cognitive basis of trait anger and reactive aggression: An integrative analysis. *Personality and Social Psychology Review, 12*(1), 3-21.

Wilkowski, B. M., & Robinson, M. D. (2008b). Guarding against hostile thoughts: Trait anger and the recruitment of cognitive control. *Emotion, 8*(4), 578-583.

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.

Woldorff, M. G., Liotti, M., Seabolt, M., Busse, L., Lancaster, J. L., & Fox, P. T. (2002). The temporal dynamics of the effects in occipital cortex of visual-spatial selective attention. *Cognitive Brain Research, 15*(1), 1-15.

Yamaguchi, S., & Knight, R. T. (1992). Effects of temporal-parietal lesions on the somatosensory P3 to lower limb stimulation. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 84*(2), 139-148.

Appendix A – Study Poster

University of East Anglia

PARTICIPATE IN RESEARCH!

Do you tend to lose your temper?

You could be eligible to take part in a research study investigating **emotion and cognition**.

We are looking for healthy male volunteers aged between 18 and 35.

The study is taking place on campus from 1st January 2017. Participants will receive **£20** as a thank you, and as a volunteer you could be helping to advance research.

If you would like to express interest and find out more about the study, please email

r.crago@uea.ac.uk

and quote reference A300

Appendix B – Demographic information questions

Age: _________________

Date of Birth: ____

What is your marital status?

Single, never married____ ☐

Married/Civil partnership_______________________________________ ☐

Widowed____ ☐

Divorced____ ☐

Separated____ ☐

A member of an unmarried couple________________________________ ☐

What is your ethnic group?

White
English/Welsh/Scottish/Northern Irish/British___________________ ☐

Irish____ ☐

Gypsy or Irish traveller_____________________________________ ☐

Any other White background, please describe____________________ ☐

Mixed/multiple ethnic groups

White and Black African_________________________________ ☐

White and Black Caribbean_______________________________ ☐

White and Asian_______________________________________ ☐

Any other Mixed/Multiple ethnic background, please

describe____ ☐

Asian/Asian British

Indian____ ☐

Pakistani____ ☐

Bangladeshi____ ☐

Chinese____ ☐

Any other Asian background, please ☐

describe____

Black/African/Caribbean/Black British

African____ ☐

Caribbean____ ☐

Any other Black/African/Caribbean background, please describe___ ☐

____

Other ethnic group

Arab____ ☐

Any other ethnic group, please describe______________________ ☐

____

What is the highest level of education you have completed? (If currently enrolled in education, please mark the last level you completed).

No schooling completed_________________________________ ☐

High School (GCSEs or equivalent)________________________ ☐

Sixth-form (A-Levels or equivalent)_____________________________ ☐

Some University credit, but less than one year______________________ ☐

One or more years of University, no degree________________________ ☐

☐

Bachelor's degree (for example: BA, BSc, LLB)______________________

Master's degree (for example: MA, MSc, MChem)____________________

Doctoral degree (for example: PhD, LLD, EngD)_____________________

Other, please specify: ____

How would you describe your current employment status?

Employed full time_______________________________________

Employed part time____

Unemployed/looking for work___________________________________

Unable to work____

Student____

Homemaker____

Retired____

Appendix C – Aggression Questionnaire

Instructions:

Using the 5 point scale shown below, indicate how uncharacteristic or characteristic each of the following statements is in describing you.

1 = extremely uncharacteristic of me
2 = somewhat uncharacteristic of me
3 = neither uncharacteristic nor characteristic of me
4 = somewhat characteristic of me
5 = extremely characteristic of me

1. Some of my friends think I am a hothead
2. If I have to resort to violence to protect my rights, I will.
3. When people are especially nice to me, I wonder what they want.
4. I tell my friends openly when I disagree with them.
5. I have become so mad that I have broken things.
6. I can't help getting into arguments when people disagree with me.
7. I wonder why sometimes I feel so bitter about things.
8. Once in a while, I can't control the urge to strike another person.
9. I am an even-tempered person.
10. I am suspicious of overly friendly strangers.
11. I have threatened people I know.
12. I flare up quickly but get over it quickly.
13. Given enough provocation, I may hit another person.
14. When people annoy me, I may tell them what I think of them.
15. I am sometimes eaten up with jealousy.
16. I can think of no good reason for ever hitting a person.
17. At times I feel I have gotten a raw deal out of life.
18. I have trouble controlling my temper.
19. When frustrated, I let my irritation show.
20. I sometimes feel that people are laughing at me behind my back.
21. I often find myself disagreeing with people.
22. If somebody hits me, I hit back.
23. I sometimes feel like a powder keg ready to explode.
24. Other people always seem to get the breaks.
25. There are people who pushed me so far that we came to blows.

26. I know that “friends” talk about me behind my back.
27. My friends say that I’m somewhat argumentative.
28. Sometimes I fly off the handle for no good reason.
29. I get into fights a little more than the average person.

Scoring:

The two questions with the asterisk are reverse scored.

The Aggression scale consists of 4 factors, Physical Aggression (PA), Verbal Aggression (VA), Anger (A) and Hostility (H). The total score for Aggression is the sum of the factor scores.

Physical Aggression = 2, 5, 8, 11, 13, 16, 22, 25, 29
Verbal Aggression = 4, 6, 14, 21, 27
Anger = 1, 9, 12, 18, 19, 23, 28
Hostility = 3, 7, 10, 15, 17, 20, 24, 26

Appendix D – Attentional Control Scale

A number of statements which people have used to describe themselves are given below. Please read each statement and then circle the appropriate number to the right of the statement to indicate how much you think it applies to you. Please do not spend long answering each question.

		Almost never	Sometimes	Often	Always
1.	It's very hard for me to concentrate on a difficult task when there are noises around.	1	2	3	4
2.	When I need to concentrate and solve a problem, I have trouble focussing my attention.	1	2	3	4
3.	When I am working hard on something, I still get distracted by events around me.	1	2	3	4
4.	My concentration is good even if there is music in the room around me.	1	2	3	4
5.	When concentrating, I can focus my attention so that I become unaware of what's going on in the room around me.	1	2	3	4
6.	When I am reading or studying, I am easily distracted if there are people talking in the same room	1	2	3	4
7.	When trying to focus my attention on something, I have difficulty blocking out distracting thoughts.	1	2	3	4
8.	I have a hard time concentrating when I'm excited about something.	1	2	3	4
9.	When concentrating I ignore feelings of hunger or thirst.	1	2	3	4
10.	I can quickly switch from one task to another.	1	2	3	4
11.	It takes me a while to get really involved in a new task.	1	2	3	4
12.	It is difficult for me to coordinate my attention between the listening and writing required when taking notes during lectures.	1	2	3	4
13.	I can become interested in a new topic very quickly when I need to.	1	2	3	4
14.	It is easy for me to read or write while I'm also talking on the phone.	1	2	3	4
15.	I have trouble carrying on two conversations at once.	1	2	3	4

16.	I have a hard time coming up with new ideas quickly	1	2	3	4
17.	After being interrupted or distracted, I can easily shift my attention back to what I was doing before.	1	2	3	4
18.	When a distracting thought comes to mind, it is easy for me to shift my attention away from it.	1	2	3	4
19.	It is easy for me to alternate between two different tasks.	1	2	3	4
20.	It is hard for me to break from one way of thinking about something and look at it from another point of view.	1	2	3	4

Appendix E – Delinquency Questionnaire

Please read each statement and then circle the appropriate number to the right of the statement to indicate how many times you have behaved in this way in the last 12 months

		Never	**Once or twice**	**A few times**	**Several times**
1	Thrown stones at cars, trains, buses or other vehicles	0	1	2	3
2	Purposely destroyed, damaged or defaced people's private property or belongings	0	1	2	3
3	Smashed, slashed or damaged things in public places, e.g. in streets, cinemas, pubs, clubs, trains, buses, etc.	0	1	2	3
4	Sold illegal drugs to other people	0	1	2	3
5	Purposely annoyed, insulted or taunted strangers in the street	0	1	2	3
6	Thrown things, such as stones, at other people	0	1	2	3
7	Struggled or fought to get away from a police officer	0	1	2	3
8	Written on walls in public places with spray paint	0	1	2	3
9	Drunk alcohol whilst not at home and not in a pub, e.g. in a park	0	1	2	3
10	Trespassed in places you were not supposed to go, e.g. railway lines, good yards, private gardens, empty houses, factories etc.	0	1	2	3
11	Broken the windows of empty houses	0	1	2	3
12	Stolen school/University property worth more than about £0.00	0	1	2	3
13	Driven a car on the roads without a licence	0	1	2	3
14	Stolen money from slot machines, juke boxes, public telephones, etc.	0	1	2	3
15	Deliberately littered the street or pavement by smashing bottles, tipping over dustbins, etc.	0	1	2	3
16	Stolen property from a deserted house or flat	0	1	2	3
17	Purposely annoyed, insulted or taunted one of your tutors/lecturers	0	1	2	3

18	Found property belonging to other people and failed to return it	0	1	2	3
19	Been involved in a group fight	0	1	2	3
20	Got money by lying	0	1	2	3
21	Purposely annoyed, insulted or defied a police officer	0	1	2	3
22	Set fire on purpose to something not belonging to you	0	1	2	3
23	Threatened someone with a weapon	0	1	2	3
24	Refused to tell a police officer or other official what you knew about a crime	0	1	2	3

Appendix F – State-Trait Anxiety Inventory

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of each statement to indicate how you <u>*generally*</u> feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe how you generally feel.

		almost never	sometimes	often	almost always
1	I feel pleasant	1	2	3	4
2	I feel nervous and restless	1	2	3	4
3	I feel satisfied with myself	1	2	3	4
4	I wish I could be as happy as others seem to be	1	2	3	4
5	I feel like a failure	1	2	3	4
6	I feel rested	1	2	3	4
7	I am 'calm, cool and collected'	1	2	3	4
8	I feel that difficulties are piling up so that I cannot overcome them	1	2	3	4
9	I worry too much over something that really doesn’t matter	1	2	3	4
10	I am happy	1	2	3	4
11	I have disturbing thoughts	1	2	3	4
12	I lack self-confidence	1	2	3	4
13	I feel secure	1	2	3	4
14	I make decisions easily	1	2	3	4
15	I feel inadequate	1	2	3	4
16	I am content	1	2	3	4
17	Some unimportant thought runs through my mind and bothers me	1	2	3	4
18	I take disappointments so keenly that I can’t put them out of my mind	1	2	3	4
19	I am a steady person	1	2	3	4
20	I get in a state of tension or turmoil as I think over my recent concerns and interests	1	2	3	4

Appendix G – Eligibility Questionnaire

Please tick as appropriate

	True	False
I am male	☐	☐
I am aged between 18 and 35	☐	☐
I speak English as my first language	☐	☐
I am right-handed	☐	☐
I have normal or corrected-to-normal vision (glasses or contact lenses)	☐	☐
I am able to read and understand text displayed on a computer screen	☐	☐
I am able to use a computer keyboard comfortably for 30 minutes at a time	☐	☐
I have not been diagnosed by the GP with a neurological or psychological condition such as anxiety or depression within the last 12 months	☐	☐
I am not currently be receiving psychological treatment such as cognitive behavioural therapy, and have not done so within the past three months	☐	☐
I am not currently taking psychiatric medication (e.g. Zoloft, Xanax etc), and have not done so within the past three months.	☐	☐
I am not currently taking anabolic steroids or testosterone supplements	☐	☐

If your answer to any of these questions is 'FASLE' please inform the experimenter now.

Appendix H – Unique ID Code

Before you begin, so that data can be collected anonymously, please create your own personal identification code. To do this, we suggest combining the last four digits of your telephone number with your first initial

E.g. Jane Smith, 07777 123456 = 3456J

You will need to remember this ID code as you will be asked to provide it in a few moments

Please write your identification code below:

____

Appendix I – Study 1/3 Information sheet

'Anger and cognition'

Participant Information Sheet

Thank you for your interest in this study. Before you decide whether to take part, please read the following information carefully (this sheet is for you to keep). You may ask me any questions if you would like more information.

What is this research looking at?
The purpose of the study is to investigate how people process different types of emotional information, and how this relates to their experience of anger.

Do I have to take part?
It is up to you to decide to join the study. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any point during the experimental session, without giving a reason.

What will happen if I agree to take part?
Part 1 of the study is in two stages, you will have already completed the initial stage which involved answering a short ten minute online questionnaire. For the lab session, you will be asked to fill out four short questionnaires and take part in a laboratory based experiment. This experiment is a simple cognitive task that involves pressing a key in response to a simple stimulus on a screen. Further, you will be asked to read a few scenarios and answer questions in relation to each scenario. A researcher will provide you with all the information you need for completing the tasks. You may additionally be required to put on a channel cap with electrodes in order to record your brain activity in relation to the tasks using electroencephalography (EEG). EEG is a safe and non-invasive technique that measures the electrical activity of the brain using electrodes placed on the scalp. You will have already been informed if EEG recordings will be carried out during this session.

It is estimated that the session will take approximately 1 hour without EEG and an hour and a half with EEG recording. This is part 1 of a two part study, therefore at the end of this experimental session you will be asked whether you would be interested in being invited back for part 2 of the study. It is estimated that part 2 will take a similar amount of time to part 1, however, participants are under no obligation to consent to participating in part 2 of the study.

.

EEG – head measurement and gel use (only relevant to those having EEG recordings)

EEG involves measuring your head to choose an appropriate cap. We will then place the cap on your head and attach 32 electrodes to it. One or two electrodes will also be placed on your face to record your eye movements. We will tell you at each point what we are doing.

To record accurately, we need to put a water-based gel into your hair under each electrode using blunt syringes. This gel is easy to wash out after the experiment. We have facilities and private space for you to wash your hair. We will give you as much time as you like to wash your hair at the laboratory. During set-up, we will also carefully part your hair beneath the electrodes – this may involve making contact with your scalp, but should never hurt. We will ask you to provide feedback on any part of the procedure and will stop immediately if you feel uncomfortable at any point.

<u>EEG – Movement and Blinking</u> (only relevant to those having EEG recordings)

The EEG recording can be disrupted if you move or blink excessively. So, you will be invited to find a comfortable position in your chair to limit movement as much as possible and to minimise eye-blinks and face movements. Your experimenter will give you very clear instructions about when it is OK to move and blink and when it is best to keep as still as possible, but ask for clarification if anything is not clear. We will give you breaks and water will be available whenever you need it, but please ask for additional breaks as needed.

<u>EEG – Brain measurement</u> (only relevant to those having EEG recordings)

EEG only allows to record neural activity naturally occurring in your brain. It does not stimulate any part of your brain, nor allow to “read your mind”. It will not be used to diagnose any condition.

If you wish to receive more information about EEG before you decide to take part, please feel free to ask us.

Are there any problems with taking part?
Some questionnaires ask personal questions which can, in some cases, cause discomfort. If you do not wish to answer these questions, you have the right to omit them without giving reason.

The placement of the EEG cap (only relevant to those having EEG recordings) is not painful, although there may be minor discomfort. Some people find that their skin may be slightly reddened after the electrodes are removed. This reddening will disappear within a few hours. If you experience any irritation or inconvenience during the study, you can choose to stop at any time.

Will it help me if I take part?
No, but it will benefit the programme of research and contribute to our understanding. Also, we hope you might find the experience interesting.

How will you store the information that I give you?
All information which you provide during the study will be stored in accordance with the 1998 Data Protection Act and kept strictly confidential. The chief investigator will be the custodian of the anonymous research data. Any identifiable data will be stored separately in a password protected file and will be securely disposed of as soon as it is no longer necessary, and within 5 years. Electronic data will be stored on a password protected computer and paper information will be stored in an academic's filing cabinet in a locked office. The data will be stored anonymously and will not be linked to any participant. All data will only be accessible to members of the research team and academic staff reviewing the project.

How will the data be used?
The data will be analysed and reported in an academic journal or conference. Only group data will be presented and participants will never be identified.

What happens if I agree to take part, but change my mind later?
If, at any point, you no longer wish to take part in the study, you have the right to withdraw from the study without giving any reason. Your data will be destroyed and will not be included in the final report. If you wish to withdraw, please inform the researcher before the end of the experimental session.

If you give permission to be contacted via email about participating in part 2 of the study you can withdraw your data by contacting the researcher via email any time between now and the date in which you are contacted about part 2 data collection (approximately 6 months). After this you will be unable to withdraw your data.

How do I know that this research is safe for me to take part in?
All research in the University is looked at by an independent group of people, called a Research Ethics Committee, to protect your safety, rights, wellbeing and dignity. This research was approved by the Psychology Research Ethics Committee at the University of East Anglia on 2nd September 2015.

You are under no obligation to agree to take part in this research.
If you do agree you can withdraw at any time without giving a reason.

Please note that any declared recent events which would put you or any other member of the community in danger of harm will be reported to the relevant authorities.

Do also contact us if you have any worries or concerns about this research.

Contact details:
Rebecca Crago: r.crago@uea.ac.uk
Jennifer Bowler: j.bowler@uea.ac.uk
Dr Gavin Nobes: g.nobes@uea.ac.uk

Dr Laura Biggart: l.biggart@uea.ac.uk
Dr Louis Renoult: l.renoult@uea.ac.uk

School of Psychology Ethics Committee:
ethics.psychology@uea.ac.uk; Phone 01603 597146

Head of School Professor Kenny Coventry:
k.coventry@uea.ac.uk; Phone 01603 597145

Appendix J – study 1/3 Consent form

University of East Anglia

'Anger and cognition'

Name of Researcher: Gavin Nobes, Laura Biggart, Louis Renoult, Jennifer Bowler, Rebecca Crago

Please **INITIAL** all boxes

I have read and understand the information sheet 'Anger and cognition' and have had the opportunity to ask questions and have had these answered satisfactorily. ☐

I agree to have my EEG recorded, which involves to have electrodes placed on my head and face, and a water-based gel placed into my hair. I understand that EEG only records neural activity naturally occurring in my brain. It does not send or emit current. ☐

I understand that my participation is voluntary and I am free to withdraw from the session at any time, for any reason and without prejudice. I understand that due to the anonymous nature of the data, unless I give permission to take part in the second stage of testing, I cannot withdraw my data once I have left the session. ☐

I understand breaks will be provided and that I can request additional breaks if needed. ☐

I know that no personal information (such as my name) will be shared outside of the research team or published in the final report(s) from this research. ☐

I am happy for the researchers to contact me by email in approximately six months about the second part of the study. ☐

I agree to take part in the above study. ☐

Participant I.D code (last four digits of telephone number and first initial).................................

Participant's signature…………………………………………….........................Date………..

Participant EEG testing number…………………………. (Researchers use only)

Researcher Contact details:
Rebecca Crago: r.crago@uea.ac.uk; *Jennifer Bowler:* j.bowler@uea.ac.uk; *Dr Gavin Nobes:* g.nobes@uea.ac.uk; *Dr Laura Biggart:* l.biggart@uea.ac.uk; *Dr Louis Renoult*: l.renoult@uea.ac.uk

Do also contact us if you have any worries or concerns about this research.

School of Psychology Ethics Committee:
ethics.psychology@uea.ac.uk; Phone 01603 597146

Head of School Professor Kenny Coventry:
k.coventry@uea.ac.uk; Phone 01603 597145

Appendix K – Study 1/3 Debrief

'Anger and cognition'

Thank you for participating in this study. Your time and efforts are much appreciated.

Theories of aggression and antisocial behaviour highlight the role of increased attention to aggressive cues in favour of non-aggressive cues and the tendency to make hostile attributions for others' ambiguous behaviour. We are interested in understanding how differences in preference for hostile and benign cues as well as tendency to perceive hostile meanings are implicated in aggressive behaviour.

The main aim of this study is to provide greater insight into whether: a) a susceptibility to identify faces as hostile predicts the frequency of aggressive behaviour, b) hostile attribution bias (over attribute hostile intentions to peers even when a hostile attribution is not warranted by circumstances) predicts aggressive behaviour and c) increased bias will be associated with increased P300 (event related potential) on the electroencephalography recordings (not all participants had EEG recordings).

If you have any questions regarding this study please feel free to ask or contact the researcher or supervisor of this study now, or at a later date. If you wish to withdraw your data please let the researcher know immediately. If you consent to being contacted about participating in part 2 of the study you will have further opportunity to withdraw by emailing the researcher and specifying that you no longer wish to be included. At the end of the testing session for part 2 you will have no more opportunities to withdraw. Testing for part 2 takes place in approximately 6 months.

If you would like to receive a report of the main findings of the study (or a summary of the findings) when it is completed please contact the researcher, however individual feedback on your results cannot be given.

<u>Sources of support</u>
Sometimes people taking part in research projects are interested in finding out more information about dealing with emotional difficulties, either for themselves or their friends. Below are some sources of support if you are interested.

General sources of support

1. Seeking help or information for emotional difficulties
The first step in accessing help is to discuss the problem with your GP. They will be able to advise you on access to local resources and refer you on if appropriate.

2. Useful web sites
The British Association of Behavioural and Cognitive Psychotherapies (http://www.babcp.org.uk). This site offers a 'user's area' with information on mental health difficulties and a facility to help you find an accredited cognitive behavioural therapist.

The Changing Minds website (http://www.rcpsych.ac.uk/campaigns/cminds/). This site is produced by the Royal College of Psychiatrists and provides information and advice about mental health issues. The website contains on-line leaflets about several topics including anxiety, depression, anorexia and bulimia.

Mind website (http://www.mind.org.uk/) is supported by a leading mental health charity in England and Wales and also provides high-quality information and advice about mental health issues.

Sources of support for UEA members
At UEA there are a number of options. Information about them is available through the UEA website (see below) or through Student Services. You can get in touch with the mental health coordinator, Beckie Davies, directly, or someone who knows you can make initial contact on your behalf, either by calling in to reception at the Dean of Students' Office (Upper Street, opposite Waterstones Bookshop), by telephone (01603 593032) or by email: beckie.davies@uea.ac.uk .The service is usually available Monday-Friday, 9am - 5pm.

On the UEA Portal page, select the Help and Advice Tab.
Under the Health and Well-being heading you will find many useful links including:

‘Medical Services Unit’ and a route for contacting a GP for advice.

‘Mental Health Coordinator’ where you will find information about advice and support and links to useful leaflets about mental health issues.

‘Counselling Services’ where you will find information about counselling and links for ‘crisis information’ which includes the Samaritans and the student led Nightline 01603 503504.

Do also contact us if you have any worries or concerns about this research.
Researchers:

Rebecca Crago: r.crago@uea.ac.uk
Jennifer Bowler: j.bowler@uea.ac.uk
Dr Gavin Nobes: g.nobes@uea.ac.uk
Dr Laura Biggart: l.biggart@uea.ac.uk
Dr Louis Renoult: l.renoult@uea.ac.uk

School of Psychology Ethics Committee:

ethics.psychology@uea.ac.uk; Phone 01603 597146

Head of School Professor Kenny Coventry:
k.coventry@uea.ac.uk; Phone 01603 597145
Thank you again for your participation.

Appendix L – Study 1/3 Normality of data

	Skewness			Kurtosis		
	Statistic	Standard Error	Calculated score	Statistic	Standard Error	Calculated score
Total Aggression	.340	.409	.831	.334	.798	.419
Physical Aggression	.581	.409	1.421	-.547	.798	-.685
Verbal Aggression	.129	.409	.315	-.377	.798	-.472
Anger	.120	.409	.293	-.803	.798	-1.006
Hostility	.144	.409	.352	-.150	.798	-.188
STAI-T	.630	.409	1.540	.529	.798	.663
ACS	.199	.409	.487	-.477	.798	-.598
Delinquency	1.455	.409	3.557	1.622	.798	2.033
DPTW - congruent	1.881	.409	4.599	4.637	.798	5.811
DPTW - incongruent	2.004	.409	4.900	5.371	.798	6.731
DPTW - bias	-.553	.409	-1.352	-.053	.798	-.066
DPTI - congruent	.398	.409	.973	-.667	.798	-.836
DPTI - incongruent	.171	.409	.418	-.648	.798	-.812
DPTI - bias	-.003	.409	-.007	-.225	.798	-.282

Appendix M – Study 1 – Total aggression results

100-200ms
Main effect of congruency; $F(1,30) = 10.119$, $p = .003$, $\eta_p^2 = .254$)
Congruency and aggression interaction; $F(1,30) = 4.467$, $p = .043$, $\eta_p^2 = .130$)
Hemisphere and aggression interaction; $F(1,30) = 4.773$, $p = .037$, $\eta_p^2 = .137$)
Congruency, electrode, and aggression interaction; $F(4,120) = 2.376$, $p = .074$, $\eta_p^2 = .073$)

200-300ms
Hemisphere and aggression interaction; $F(1,30) = 8.629$, $p = .006$, $\eta_p^2 = .223$)
Congruency, electrode, and aggression interaction; $F(4,120) = 4.253$, $p = .005$, $\eta_p^2 = .124$)
Congruency, electrode, hemisphere, and aggression interaction; $F(4,120) = 3.635$, $p = .032$, $\eta_p^2 = .108$)

300-400ms
Hemisphere and aggression interaction; $F(1,30) = 10.651$, $p = .003$, $\eta_p^2 = .266$)
Congruency, electrode, and aggression interaction; $F(4,120) = 2.454$, $p = .059$, $\eta_p^2 = .076$)
Congruency, electrode, hemisphere, and aggression interaction; $F(4,120) = 2.686$, $p = .068$, $\eta_p^2 = .082$)

400-500ms
Hemisphere and aggression interaction; $F(1,30) = 4.974$, $p = .033$, $\eta_p^2 = .142$)
Congruency, electrode, and aggression interaction; $F(4,120) = 3.232$, $p = .021$, $\eta_p^2 = .097$)

500-600ms
Hemisphere and aggression interaction; $F(1,30) = 5.822$, $p = .022$, $\eta_p^2 = .163$)

600-700ms
Hemisphere and aggression interaction; $F(1,30) = 5.203$, $p = .030$, $\eta_p^2 = .148$)

700-800ms
Hemisphere and aggression interaction; $F(1,30) = 4.974$, $p = .033$, $\eta_p^2 = .142$)

800-900ms
None

900-1000ms
Congruency and hemisphere interaction; $F(1,30) = 4.139$, $p = .051$, $\eta_p^2 = .121$)

Appendix N – Study 1 – Correlations between physical aggression and amplitude

	Electrode	PA - Congruent	PA - Incongruent
100-200ms	*TP9*	.279	.915
	TP10	.369*	.771
	CP5	.231	.692
	CP6	.456**	.628
	CP1	.276	.548
	CP2	.286	.297
	P7	.218	.066
	P8	.354*	.180
	P3	.257	.162
	P4	.360*	.302
200-300ms	*TP9*	.248	.147
	TP10	.391*	.264
	CP5	.221	.357*
	CP6	.491**	.217
	CP1	.337	.495**
	CP2	.286	.583***
	P7	.180	.171
	P8	.317	.326
	P3	.248	.353*
	P4	.344	.437*
300-400ms	*TP9*	.142	.145
	TP10	.248	.172
	CP5	.089	.295
	CP6	.385*	.263
	CP1	.293	.411*
	CP2	.252	.555**
	P7	.132	.170
	P8	.255	.363*
	P3	.221	.335
	P4	.366*	.489**
400-500ms	*TP9*	.109	.155
	TP10	.204	.082
	CP5	-.041	.222
	CP6	310	.204
	CP1	.139	.312
	CP2	.097	.488**
	P7	.126	.164
	P8	.205	.352*
	P3	.103	.329
	P4	.190	.470**

500-600ms	*TP9*	.127	.191
	TP10	.273	.124
	CP5	.012	.207
	CP6	.303	.186
	CP1	.261	.295
	CP2	.194	.386*
	P7	.163	.190
	P8	.232	.285
	P3	.157	.323
	P4	.175	.405*
600-700ms	*TP9*	.337	.216
	TP10	.356*	.188
	CP5	.075	.136
	CP6	.386*	-.001
	CP1	.433*	.172
	CP2	.398*	.249
	P7	.288	.142
	P8	.352*	.261
	P3	.283	.264
	P4	.298	.307
700-800ms	*TP9*	.207	.134
	TP10	.317	.232
	CP5	-.191	.030
	CP6	.375*	-.045
	CP1	.152	.137
	CP2	.228	.078
	P7	.044	-.009
	P8	.291	.210
	P3	.015	.059
	P4	.139	.235
800-900ms	*TP9*	.310	.197
	TP10	.373*	.269
	CP5	-.021	.086
	CP6	.418*	-.011
	CP1	.244	.150
	CP2	.253	.082
	P7	.262	.101
	P8	.289	.256
	P3	.180	.128
	P4	.211	.224

Appendix O – Study 2 Information sheet

University of East Anglia

'Anger and cognition'

Thank you for your interest in this study. Before you decide whether to take part, please read the following information carefully (this sheet is for you to keep). You may ask me any questions if you would like more information.

What is this research looking at?
The purpose of the study is to investigate how people process different types of emotional information, and how this relates to their experience of anger.

Do I have to take part?
It is up to you to decide to join the study. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any point during the experimental session, without giving a reason.

What will happen if I agree to take part?
During the lab session, you will be asked to complete four short questionnaires online and take part in a laboratory based experiment. Before completing the experimental tasks, you will asked to put on a channel cap with electrodes in order to record your brain activity in relation to the tasks using electroencephalography (EEG). EEG is a safe and non-invasive technique that measures the electrical activity of the brain using electrodes placed on the scalp. This experiment is a simple cognitive task that involves pressing a key in response to a simple stimulus on a screen. A researcher will provide you with all the information you need for completing the tasks. Before and after completion of the cognitive tasks you will be asked to sit quietly for three minutes while EEG is recorded. During this time you will be required to close your eyes, try and clear your mind, but not fall asleep. It is estimated that the whole session will take approximately an hour and a half.

<u>**EEG** – head measurement and gel use</u>

EEG involves measuring your head to choose an appropriate cap. We will then place the cap on your head and attach 32 electrodes to it. One or two electrodes will also be placed on your face to record your eye movements. We will tell you at each point what we are doing.

To record accurately, we need to put a water-based gel into your hair under each electrode using blunt syringes. This gel is easy to wash out after the experiment. We have facilities and private space for you to wash your hair. We will give you as much time as you like to wash your hair at the laboratory. During set-up, we will also carefully part your hair beneath the electrodes – this may involve making contact with your scalp, but should never hurt. We will ask you to provide feedback on any part of the procedure and will stop immediately if you feel uncomfortable at any point.

EEG – Movement and Blinking

The EEG recording can be disrupted if you move or blink excessively. So, you will be invited to find a comfortable position in your chair to limit movement as much as possible and to minimise eye-blinks and face movements. Your experimenter will give you very clear instructions about when it is OK to move and blink and when it is best to keep as still as possible, but ask for clarification if anything is not clear. We will give you breaks and water will be available whenever you need it, but please ask for additional breaks as needed.

EEG – Brain measurement

EEG only allows to record neural activity naturally occurring in your brain. It does not stimulate any part of your brain, nor allow to "read your mind". It will not be used to diagnose any condition.

If you wish to receive more information about EEG before you decide to take part, please feel free to ask us.

Are there any problems with taking part?
Some questionnaires ask personal questions which can, in some cases, cause mild distress. If you do not wish to answer these questions, you have the right to omit them without giving reason.

The placement of the EEG cap is not painful, although there may be minor discomfort. Some people find that their skin may be slightly reddened after the electrodes are removed. This reddening will disappear within a few hours. If you experience any irritation or inconvenience during the study, you can choose to stop at any time.

Will it help me if I take part?
No, but it will benefit the programme of research and contribute to our understanding. You will however receive SONA credits or payment as a thank you and appreciation for your time. Also, we hope you might find the experience interesting.

How will you store the information that I give you?
All information which you provide during the study will be stored in accordance with the 1998 Data Protection Act and kept strictly confidential. The chief investigator will be the custodian of the anonymous research data. Any identifiable data will be stored separately in a password protected file and will be securely disposed of as soon as it is no longer necessary, and within 5 years. Electronic data will be stored on a password protected computer and paper information will be stored in an academic's filing cabinet in a locked office. The data will be stored

anonymously and will not be linked to any participant. All data will only be accessible to members of the research team and academic staff reviewing the project.

How will the data be used?
The data will be analysed and reported in a PhD thesis, as well as academic journals or conferences. Only group data will be presented and participants will never be identified.

What happens if I agree to take part, but change my mind later?
If, at any point, you no longer wish to take part in the study, you have the right to withdraw from the study without giving any reason. Your data will be destroyed and will not be included in the final report. If you wish to withdraw, please inform the researcher before the end of the experimental session.

How do I know that this research is safe for me to take part in?
All research in the University is looked at by an independent group of people, called a Research Ethics Committee, to protect your safety, rights, wellbeing and dignity. This research was approved by the Psychology Research Ethics Committee at the University of East Anglia on 7th October 2016.

You are under no obligation to agree to take part in this research.

If you do agree you can **withdraw at any time without giving a reason.**

Please note that any declared recent events which would put you or any other member of the community in danger of harm will be reported to the relevant authorities.

Do also contact us if you have any worries or concerns about this research.

Contact details:
Rebecca Crago: r.crago@uea.ac.uk
Jennifer Bowler: j.bowler@uea.ac.uk
Dr Gavin Nobes: g.nobes@uea.ac.uk
Dr Laura Biggart: l.biggart@uea.ac.uk
Dr Louis Renoult: l.renoult@uea.ac.uk

School of Psychology Ethics Committee:
ethics.psychology@uea.ac.uk; Phone 01603 597146

Head of School Professor Kenny Coventry:
k.coventry@uea.ac.uk; Phone 01603 597145

Appendix P – Study 2 Consent form

Name of Researchers: Rebecca Crago, Jennifer Bowler, Gavin Nobes, Laura Biggart, Louis Renoult

Please **INITIAL** all boxes

I have read and understand the information sheet 'Anger and cognition' and have had the opportunity to ask questions and have had these answered satisfactorily. ☐

I agree to have my EEG recorded, which involves to have electrodes placed on my head and face, and a water-based gel placed into my hair. I understand that EEG only records neural activity naturally occurring in my brain. It does not send or emit current. ☐

I understand that my participation is voluntary and I am free to withdraw from the session at any time, for any reason and without prejudice. I understand that due to the anonymous nature of the data, I cannot withdraw my data once I have left the session. ☐

I understand breaks will be provided and that I can request additional breaks if needed. ☐

I know that no personal information (such as my name) will be shared outside of the research team or published in the final report(s) from this research. ☐

I agree to take part in the above study. ☐

Participant's signature……………………………………..........................Date…………………

Participant EEG testing number…………………………. (Researchers use only)

Researcher Contact details:

Rebecca Crago: r.crago@uea.ac.uk; *Jennifer Bowler:* j.bowler@uea.ac.uk; *Dr Gavin Nobes:* g.nobes@uea.ac.uk; *Dr Laura Biggart:* l.biggart@uea.ac.uk; *Dr Louis Renoult*: l.renoult@uea.ac.uk

Do also contact us if you have any worries or concerns about this research.

School of Psychology Ethics Committee:
ethics.psychology@uea.ac.uk; Phone 01603 597146

Head of School Professor Kenny Coventry:
k.coventry@uea.ac.uk; Phone 01603 597145

Appendix Q – Study 2 Debrief

University of East Anglia

'Anger and cognition'

Thank you for participating in this study. Your time and efforts are much appreciated.

Theories of aggression and antisocial behaviour highlight the role of increased attention to aggressive cues in favour of non-aggressive cues. We are interested in understanding how differences in preference for hostile and benign cues are implicated in aggressive behaviour.

The main aim of this study is to provide greater insight into whether: a) a susceptibility to identify faces as hostile predicts the frequency of aggressive behaviour, b) increased attention bias will be associated with increased P300 (event related potential) on the electroencephalography recordings. You were asked to sit quietly before and after the tasks in order to measure resting state ERPs and whether these differentiate from ERPs recorded during task completion.

If you have any questions regarding this study please feel free to ask or contact the researcher or supervisor of this study now, or at a later date. If you wish to withdraw your data please let the researcher know immediately.

If you would like to receive a report of the main findings of the study (or a summary of the findings) when it is completed please contact the researcher, however individual feedback on your results cannot be given.

<u>Sources of support</u>
Sometimes people taking part in research projects are interested in finding out more information about dealing with emotional difficulties, either for themselves or their friends. Below are some sources of support if you are interested.

General sources of support

1. Seeking help or information for emotional difficulties
The first step in accessing help is to discuss the problem with your GP. They will be able to advise you on access to local resources and refer you on if appropriate.

2. Useful web sites
The British Association of Behavioural and Cognitive Psychotherapies (http://www.babcp.org.uk). This site offers a 'user's area' with information on mental health difficulties and a facility to help you find an accredited cognitive behavioural therapist.

The Changing Minds website (http://www.rcpsych.ac.uk/campaigns/cminds/). This site is produced by the Royal College of Psychiatrists and provides information and advice about mental health issues. The website contains on-line leaflets about several topics including anxiety, depression, anorexia and bulimia.
Mind website (http://www.mind.org.uk/) is supported by a leading mental health charity in England and Wales and also provides high-quality information and advice about mental health issues.

Sources of support for UEA members
At UEA there are a number of options. Information about them is available through the UEA website (see below) or through Student Services. You can get in touch with the mental health coordinator, Beckie Davies, directly, or someone who knows you can make initial contact on your behalf, either by calling in to reception at the Dean of Students' Office (Upper Street, opposite Waterstones Bookshop), by telephone (01603 593032) or by email: beckie.davies@uea.ac.uk .The service is usually available Monday-Friday, 9am - 5pm.

On the UEA Portal page, select the Help and Advice Tab.
Under the Health and Well-being heading you will find many useful links including:
'Medical Services Unit' and a route for contacting a GP for advice.
'Mental Health Coordinator' where you will find information about advice and support and links to useful leaflets about mental health issues.
'Counselling Services' where you will find information about counselling and links for 'crisis information' which includes the Samaritans and the student led Nightline 01603 503504.

Do also contact us if you have any worries or concerns about this research.
Researchers:

Rebecca Crago: r.crago@uea.ac.uk
Jennifer Bowler: j.bowler@uea.ac.uk
Dr Gavin Nobes: g.nobes@uea.ac.uk
Dr Laura Biggart: l.biggart@uea.ac.uk
Dr Louis Renoult: l.renoult@uea.ac.uk

School of Psychology Ethics Committee:
ethics.psychology@uea.ac.uk; Phone 01603 597146

Head of School Professor Kenny Coventry:
k.coventry@uea.ac.uk; Phone 01603 597145

Thank you again for your participation.

Appendix R – Study 2/4 Normality of data

	Skewness			Kurtosis		
	Statistic	Standard Error	Calculated score	Statistic	Standard Error	Calculated score
Total Aggression	.233	.333	0.700	-.305	.656	-0.465
Physical Aggression	.693	.333	2.081	.572	.656	0.872
Verbal Aggression	.086	.333	0.258	-.226	.656	-0.345
Anger	0.69	.333	2.072	-.466	.656	-0.710
Hostility	.069	.333	0.207	-.496	.656	-0.756
STAI-T	.035	.333	0.105	-.443	.656	-0.675
ACS	.340	.333	1.021	.774	.656	1.180
Delinquency	1.759	.333	5.282	3.452	.656	5.262
DPTW - AN congruent	1.298	.333	3.898	1.573	.656	2.398
DPTW - AN incongruent	0.848	.333	2.547	.007	.656	0.011
DPTW - AH congruent	1.137	.333	3.414	.801	.656	1.221
DPTW - AH incongruent	.724	.333	2.174	-.035	.656	-0.053
DPTW - HN congruent	1.202	.333	3.610	.692	.656	1.055
DPTW - HN incongruent	1.108	.333	3.327	.793	.656	1.209
DPTW - AN bias	.806	.333	2.420	.973	.656	1.483
DPTW - AH bias	-.188	.333	-0.565	.259	.656	0.395
DPTW - HN bias	.044	.333	0.132	.118	.656	0.180
DPTI - AN congruent	1.364	.333	4.096	1.315	.656	2.005
DPTI - AN incongruent	1.294	.333	3.886	1.170	.656	1.784
DPTI - AH congruent	1.376	.333	4.132	1.597	.656	2.434
DPTI - AH incongruent	1.529	.333	4.592	2.054	.656	3.131
DPTI - HN congruent	1.412	.333	4.240	1.572	.656	2.396
DPTI - HN incongruent	1.503	.333	4.514	2.086	.656	3.180
DPTI - AN bias	.533	.333	1.601	.830	.656	1.265
DPTI - AH bias	.478	.333	1.435	-.168	.656	-0.256
DPTI - HN bias	.254	.333	0.763	.511	.656	0.779

Appendix S – Study 3/4 stimuli – example of face pairs

Study 3 – example of angry-neutral face pairs

Study 4 – example of angry-neutral, happy-neutral and angry-happy face pairs

Appendix T – Study 3 - Total aggression results

100-200ms
None

200-300ms
Hemisphere and aggression interaction; $F(1,30) = 5.669$, $p = .024$, $\eta_p^2 = .159$)

300-400ms
Main effect of congruency; $F(1,30) = 4.226$, $p = .049$, $\eta_p^2 = .123$)
Congruency and aggression interaction; $F(1,30) = 4.348$, $p = .046$, $\eta_p^2 = .127$)
Congruency and electrode interaction; $F(4,120) = 3.082$, $p = .057$, $\eta_p^2 = .093$)

400-500ms
Congruency and electrode interaction; $F(4,120) = 3.047$, $p = .055$, $\eta_p^2 = .092$)

500-600ms
Congruency and aggression interaction; $F(1,30) = 3.629$, $p = .066$, $\eta_p^2 = .108$)
Congruency and electrode interaction; $F(4,120) = 3.599$, $p = .040$, $\eta_p^2 = .107$)

600-700ms
Main effect of congruency; $F(1,30) = 3.757$, $p = .062$, $\eta_p^2 = .111$)

700-800ms
Hemisphere and aggression interaction; $F(1,30) = 6.009$, $p = .062$, $\eta_p^2 = .112$)

800-900ms
Congruency and aggression interaction; $F(1,30) = 3.891$, $p = .058$, $\eta_p^2 = .115$)
Hemisphere and aggression interaction; $F(1,30) = 4.205$, $p = .049$, $\eta_p^2 = .123$)

900-1000ms
Main effect of congruency; $F(1,30) = 3.780$, $p = .061$, $\eta_p^2 = .112$)
Hemisphere and aggression interaction; $F(1,30) = 5.110$, $p = .031$, $\eta_p^2 = .146$)
Congruency, hemisphere and aggression interaction; $F(1,30) = 4.306$, $p = .047$, $\eta_p^2 = .126$)

Appendix U – T-test results to show differences in mean amplitude between high and low aggression groups across congruent and incongruent trials at multiple epochs and electrode sites

			High physical aggression group (N = 15)		Low physical aggression group (N = 15)		Overall participant mean (N = 32)		Significance (p)
	Time band	Electrode	Mean	SD	Mean	SD	Mean	SD	
Congruent trials	2-300ms	TP9	4.35	4.32	2.04	4.48	3.45	4.44	0.156
	3-400ms	TP9	4.70	3.34	3.51	4.00	4.29	3.61	0.372
	4-500ms	TP9	6.01	3.72	4.81	4.39	5.43	3.91	0.415
	5-600ms	TP9	4.63	3.55	3.84	3.89	4.33	3.58	0.559
	6-700ms	TP9	5.63	6.09	4.86	4.03	5.23	4.69	0.684
	7-800ms	TP9	3.93	5.61	3.40	4.16	3.58	4.77	0.772
	8-900ms	TP9	1.23	5.91	0.92	4.61	1.06	5.07	0.872
Congruent trials	2-300ms	TP10	4.47	4.80	2.26	5.12	3.60	4.95	0.226
	3-400ms	TP10	4.81	3.10	4.02	5.03	4.62	4.04	0.601
	4-500ms	TP10	5.88	3.71	5.39	4.42	5.60	3.89	0.739
	5-600ms	TP10	4.36	3.25	4.43	4.62	4.42	3.78	0.961
	6-700ms	TP10	5.00	4.92	4.79	4.86	4.76	4.69	0.907
	7-800ms	TP10	4.12	4.93	3.69	4.00	3.76	4.32	0.792
	8-900ms	TP10	0.93	4.65	1.16	3.92	1.08	4.1	0.886
Congruent trials	2-300ms	CP1	2.52	1.73	2.20	1.83	2.49	1.77	0.625

	3-400ms	CP1	3.00	1.14	2.84	1.56	3.03	1.45	0.752
	4-500ms	CP1	2.84	1.65	3.10	1.91	2.98	1.76	0.692
	5-600ms	CP1	1.57	1.49	2.08	1.89	1.85	1.70	0.408
	6-700ms	CP1	2.04	1.88	1.56	2.59	1.81	2.21	0.563
	7-800ms	CP1	1.81	2.04	1.48	2.99	1.69	2.45	0.728
	8-900ms	CP1	4.86	2.99	3.97	3.86	4.57	3.40	0.487
Congruent trials	2-300ms	CP2	2.50	1.79	1.54	2.28	2.17	2.13	0.198
	3-400ms	CP2	2.99	1.16	2.34	2.01	2.81	1.67	0.273
	4-500ms	CP2	2.77	1.75	2.51	1.89	2.68	1.74	0.695
	5-600ms	CP2	1.57	1.68	1.60	1.80	1.64	1.68	0.968
	6-700ms	CP2	1.72	2.11	0.89	2.48	1.36	2.24	0.333
	7-800ms	CP2	2.12	1.81	1.45	2.75	1.84	2.26	0.438
	8-900ms	CP2	5.59	3.19	4.52	3.28	5.18	3.27	0.376
Congruent trials	2-300ms	CP5	3.78	2.99	2.46	3.01	3.35	3.05	0.230
	3-400ms	CP5	4.02	2.09	3.40	2.52	3.89	2.34	0.461
	4-500ms	CP5	4.00	2.29	3.93	2.90	4.02	2.49	0.943
	5-600ms	CP5	2.36	2.18	2.91	2.53	2.75	2.31	0.525
	6-700ms	CP5	3.05	3.54	3.08	3.56	3.12	3.39	0.981
	7-800ms	CP5	1.98	3.46	2.16	4.03	2.16	3.60	0.896
	8-900ms	CP5	2.90	4.24	2.48	4.35	2.95	4.22	0.788
Congruent trials	2-300ms	CP6	3.80	2.81	2.37	3.26	3.18	2.99	0.200
	3-400ms	CP6	4.19	1.82	3.47	2.93	3.96	2.40	0.417
	4-500ms	CP6	4.43	1.96	3.43	2.61	3.96	2.25	0.236

	5-600ms	CP6	3.18	2.39	2.60	2.48	2.96	2.35	0.513
	6-700ms	CP6	3.28	2.67	2.48	3.40	2.88	2.94	0.482
	7-800ms	CP6	3.50	2.44	2.48	3.37	3.03	2.85	0.354
	8-900ms	CP6	4.13	2.72	3.28	4.26	3.86	3.48	0.519
Congruent trials	2-300ms	P3	5.00	3.26	3.95	3.15	4.71	3.23	0.367
	3-400ms	P3	5.43	2.28	5.00	2.83	5.45	2.69	0.636
	4-500ms	P3	5.08	2.80	5.14	2.71	5.18	2.72	0.945
	5-600ms	P3	2.90	2.38	3.34	1.98	3.23	2.20	0.580
	6-700ms	P3	3.32	3.81	2.61	3.36	3.08	3.47	0.594
	7-800ms	P3	2.51	4.45	2.13	3.53	2.46	3.89	0.798
	8-900ms	P3	5.87	5.55	4.80	4.14	5.66	4.86	0.554
Congruent trials	2-300ms	P4	5.07	2.88	3.61	3.41	4.55	3.18	0.208
	3-400ms	P4	5.71	2.21	4.53	2.91	5.40	2.75	0.214
	4-500ms	P4	5.34	3.02	4.61	2.76	5.07	2.86	0.488
	5-600ms	P4	3.08	2.47	2.91	2.50	3.07	2.43	0.847
	6-700ms	P4	3.23	3.48	2.14	3.19	2.73	3.23	0.377
	7-800ms	P4	3.66	3.06	2.65	2.73	3.22	2.83	0.347
	8-900ms	P4	6.73	4.63	5.85	3.87	6.46	4.14	0.577
Congruent trials	2-300ms	P7	6.65	4.95	4.41	4.04	5.75	4.52	0.180
	3-400ms	P7	6.36	3.10	5.36	3.78	6.08	3.49	0.424
	4-500ms	P7	6.48	3.37	5.66	3.80	6.17	3.50	0.529
	5-600ms	P7	4.03	2.88	3.76	3.40	4.01	3.04	0.809
	6-700ms	P7	4.76	5.05	4.01	4.44	4.42	4.54	0.669

	7-800ms	P7	3.02	5.04	2.70	4.59	2.82	4.71	0.857
	8-900ms	P7	2.81	5.75	2.53	4.77	2.78	5.09	0.884
Congruent trials	2-300ms	P8	7.37	5.79	4.83	4.68	6.30	5.24	0.193
	3-400ms	P8	7.62	3.78	5.88	4.32	7.01	4.11	0.240
	4-500ms	P8	7.81	4.24	6.07	3.90	7.01	4.01	0.244
	5-600ms	P8	4.95	3.51	4.21	3.63	4.67	3.43	0.572
	6-700ms	P8	5.40	4.98	3.78	3.55	4.50	4.20	0.312
	7-800ms	P8	5.00	4.86	3.37	2.94	4.10	3.98	0.279
	8-900ms	P8	3.83	5.20	2.66	3.58	3.46	4.41	0.478
Incongruent trails	2-300ms	TP9	4.64	4.22	0.47	4.89	2.76	4.92	0.017*
	3-400ms	TP9	4.08	4.26	1.34	4.28	2.94	4.39	0.084
	4-500ms	TP9	5.16	5.12	3.49	4.89	4.32	4.85	0.362
	5-600ms	TP9	3.44	4.79	2.20	3.62	2.82	4.11	0.425
	6-700ms	TP9	3.71	7.17	2.83	4.18	3.27	5.60	0.684
	7-800ms	TP9	2.11	7.59	1.27	4.92	1.59	2.53	0.722
	8-900ms	TP9	-0.40	7.24	-1.14	4.65	-0.88	5.81	0.739
Incongruent trails	2-300ms	TP10	4.59	4.75	0.55	5.46	2.68	5.44	0.036*
	3-400ms	TP10	4.10	4.63	1.99	4.82	3.22	4.73	0.225
	4-500ms	TP10	5.10	5.15	3.81	4.57	4.36	4.80	0.469
	5-600ms	TP10	3.69	4.84	2.38	3.52	2.96	4.14	0.398
	6-700ms	TP10	4.40	6.28	2.77	3.62	3.41	4.98	0.392

	7-800ms	TP10	3.71	6.75	2.04	3.90	2.53	5.56	0.415
	8-900ms	TP10	1.28	6.21	-0.87	4.05	0.02	5.18	0.273
Incongruent trails	2-300ms	CP1	2.81	1.75	1.63	2.08	2.28	1.93	0.098
	3-400ms	CP1	3.05	1.57	1.90	2.42	2.59	2.07	0.126
	4-500ms	CP1	2.88	2.16	2.53	2.81	2.68	2.42	0.699
	5-600ms	CP1	2.05	2.08	1.20	2.63	1.56	2.33	0.324
	6-700ms	CP1	1.70	2.18	1.07	3.02	1.33	2.53	0.518
	7-800ms	CP1	1.82	2.92	1.24	3.08	1.45	2.89	0.600
	8-900ms	CP1	4.78	3.70	3.47	2.76	4.14	3.28	0.280
Incongruent trails	2-300ms	CP2	3.11	1.58	1.35	2.33	2.13	2.16	0.020*
	3-400ms	CP2	3.45	1.19	1.50	2.43	2.44	2.11	0.007*
	4-500ms	CP2	2.94	2.06	1.66	2.47	2.19	2.36	0.127
	5-600ms	CP2	2.15	1.72	0.41	2.38	1.13	2.32	0.026*
	6-700ms	CP2	1.90	1.44	0.16	2.60	0.81	2.42	0.034*
	7-800ms	CP2	2.80	1.65	0.88	2.82	1.54	2.68	0.033*
	8-900ms	CP2	6.14	2.79	3.73	2.89	4.69	3.11	0.028*
Incongruent trails	2-300ms	CP5	4.29	2.29	1.55	2.99	3.10	2.91	0.007*
	3-400ms	CP5	4.28	1.66	2.18	2.96	3.42	2.63	0.020*
	4-500ms	CP5	4.38	2.36	2.92	3.48	3.67	2.99	0.182
	5-600ms	CP5	3.25	2.13	1.65	2.94	2.44	2.64	0.092
	6-700ms	CP5	3.25	3.34	2.11	3.34	2.68	3.22	0.358
	7-800ms	CP5	2.66	3.67	1.80	4.04	2.13	3.75	0.548
	8-900ms	CP5	3.46	4.02	2.10	3.50	2.76	3.65	0.331

Incongruent trails	2-300ms	CP6	3.75	3.28	1.71	3.76	2.68	3.58	0.117
	3-400ms	CP6	4.09	2.93	2.43	3.30	3.25	3.23	0.149
	4-500ms	CP6	3.78	4.30	2.86	3.40	3.17	3.95	0.513
	5-600ms	CP6	2.42	3.85	1.61	2.81	1.83	3.48	0.508
	6-700ms	CP6	2.34	4.99	1.69	2.90	1.80	4.03	0.663
	7-800ms	CP6	3.28	5.52	2.06	3.64	2.33	4.73	0.481
	8-900ms	CP6	3.89	5.70	2.96	3.45	3.19	4.66	0.594
Incongruent trails	2-300ms	P3	5.65	2.75	3.07	3.24	4.45	3.14	0.023*
	3-400ms	P3	5.56	2.18	3.88	3.29	4.85	3.02	0.103
	4-500ms	P3	5.14	3.41	4.38	3.83	4.71	3.66	0.561
	5-600ms	P3	2.99	3.14	2.52	3.32	2.66	3.25	0.687
	6-700ms	P3	2.90	4.02	2.36	4.25	2.51	4.00	0.725
	7-800ms	P3	2.70	5.19	2.04	4.29	2.18	4.71	0.707
	8-900ms	P3	5.98	5.48	4.52	3.77	5.17	4.55	0.403
Incongruent trails	2-300ms	P4	5.80	2.76	2.91	3.91	4.39	3.60	0.024*
	3-400ms	P4	5.90	1.81	3.59	3.80	4.82	3.25	0.037*
	4-500ms	P4	5.57	3.60	3.87	3.53	4.63	3.79	0.194
	5-600ms	P4	3.73	2.62	1.65	2.93	2.53	3.29	0.046*
	6-700ms	P4	4.09	3.04	0.82	3.22	2.24	3.59	0.008*
	7-800ms	P4	5.16	3.81	1.37	3.27	3.04	4.05	0.007*
	8-900ms	P4	8.27	4.56	4.37	3.78	6.22	4.45	0.017*
Incongruent trails	2-300ms	P7	7.24	4.75	3.47	4.67	5.51	4.92	0.034*

	3-400ms	P7	6.62	3.89	4.37	4.26	5.73	4.25	0.134
	4-500ms	P7	7.19	4.86	5.43	4.85	6.35	4.80	0.321
	5-600ms	P7	4.93	4.03	3.34	3.47	4.15	3.74	0.249
	6-700ms	P7	5.16	5.80	3.81	4.54	4.42	5.01	0.484
	7-800ms	P7	4.15	6.43	2.71	4.92	3.21	5.71	0.497
	8-900ms	P7	4.03	6.24	2.36	4.21	3.05	5.21	0.398
Incongruent trails	2-300ms	P8	8.36	5.45	3.32	5.85	5.94	5.97	0.019*
	3-400ms	P8	8.12	3.54	4.26	4.90	6.36	4.59	0.017*
	4-500ms	P8	8.16	4.93	5.02	4.54	6.54	4.90	0.076
	5-600ms	P8	5.57	3.73	2.52	3.74	3.98	3.93	0.030*
	6-700ms	P8	6.09	4.95	2.28	3.53	3.91	4.62	0.022*
	7-800ms	P8	6.00	5.46	1.98	3.83	3.64	5.17	0.028*
	8-900ms	P8	5.11	5.10	1.45	3.86	3.20	4.74	0.036*

Appendix V – Study 3 – Correlations between physical aggression and amplitude

	Electrode	PA - Congruent	PA - Incongruent
200-300ms	*TP9*	.237	.356*
	TP10	.292	.393*
	CP5	.171	.404*
	CP6	.315	.310
	CP1	.129	.287
	CP2	.246	.472**
	P7	.241	.342
	P8	.302	.439*
	P3	.181	.380*
	P4	.255	.420*
300-400ms	*TP9*	.113	.234
	TP10	.123	.227
	CP5	.119	.363*
	CP6	.244	.237
	CP1	.199	.240
	CP2	.288	.491**
	P7	.123	.205
	P8	.245	.395*
	P3	.169	.259
	P4	.270	.358*
400-500ms	*TP9*	.184	.168
	TP10	.163	.216
	CP5	.135	.306
	CP6	.331	.180
	CP1	.204	.120
	CP2	.263	.352*
	P7	.236	.221
	P8	.323	.398*
	P3	.219	.186
	P4	.264	.318
500-600ms	*TP9*	-.006	.050
	TP10	-.045	.111
	CP5	-.160	.271
	CP6	.180	.135
	CP1	-.024	.161
	CP2	.119	.392*
	P7	.008	.137
	P8	.095	.368*
	P3	.015	.100
	P4	.120	.374*
600-700ms	*TP9*	-.085	.001

	TP10	-.046	.098
	CP5	-.133	.064
	CP6	.108	.057
	CP1	.037	.050
	CP2	.132	.309
	P7	-.035	.015
	P8	.068	.270
	P3	.000	-.027
	P4	.058	.381*
700-800ms	*TP9*	-.097	.026
	TP10	.021	.121
	CP5	-.185	.002
	CP6	.180	.108
	CP1	-.038	.020
	CP2	.063	.282
	P7	-.085	.025
	P8	.142	.325
	P3	-.082	-.036
	P4	.071	.385*
800-900ms	*TP9*	-.114	.065
	TP10	-.039	.207
	CP5	-.098	.069
	CP6	.134	.121
	CP1	.015	.098
	CP2	.043	.305
	P7	-.070	.095
	P8	.100	.367*
	P3	-.019	.051
	P4	.014	.386*

Appendix W – Recognition task

The evening class
You have just started going to an evening class. The instructor asks a question and no one in the group volunteers an answer, so he looks directly at you. You answer the question, and then other people in the class speak up and disagree with your answer.

Have you been going to the evening class for a long time?

Classmates offer opinions that differ from yours
You give a good answer
Classmates are being very argumentative
When answering you make a mistake

The supermarket
You are shopping for groceries at the supermarket. As you walk down the cereal aisle you see a man walking towards you. As you get closer to him, he stops in the middle of the aisle and blocks your way.

Are you in the cereal aisle?

As you approach the man stops for a moment
The cereal you want is on sale
The man refuses to let you pass
Your favourite cereal is sold out

The race
You are running in a race with a few of your friends. The winner of the race gets a small prize. As you near the finish line, you are very close to another runner. The two of you are in the lead. As you turn a corner, you trip on the other runner's foot and fall. The other runner is the winner.

Did you win?

You trip over by accident
The race weather is nice
The runner trips you up on purpose
The race weather is disappointing

The car park
You drive to the mall on a Saturday afternoon. You are trying to find a good parking spot near the entrance. You see an open spot and drive up to take it but as you approach the spot, someone else drives through the spot behind it and takes it from the other direction.

Are you at the shops?

The other car does not notice you
You usually enjoy grocery shopping
Another car steals your space
The supermarket is very crowded

The frisbee
You are walking in the park on a sunny summer afternoon. Up ahead you see a group of teenagers throwing a Frisbee. There is no way to avoid walking through their game. You walk hurriedly between them. As you continue to walk away from them, you feel something hit the back of your head. You were hit by the frisbee.

Is it winter?

The teenagers accidentally hit you
The park is full of flowers
Teenagers throw the frisbee at your head
There are a lot of insects

The party
You are at a crowded party on a Friday night. The music is loud and a lot of people are dancing. You get really thirsty and head to the kitchen for a drink. As you are walking back out to the party, you get pushed and you spill your drink all over your shirt

Are people dancing?

You get bumped by a dancer
You enjoy dancing at the party
Someone pushes you on purpose
You don't enjoy the party

The classroom
You arrive at class a few minutes early and take a seat at a desk. You arrange your books and pens on the desk and wait for the teacher to arrive. Other students start arriving and taking seats near you. One student quickly walks past your desk and hits it with his arm causing your books, papers, and pens to all fall on the floor.

Were you early for class?

Another student accidentally bumps your desk
You really enjoy this class
Another student rudely hits your desk
You really do no enjoy this class

The coffee shop

You are in a coffee shop studying and in walks an old friend that you have not spoken with for months. You begin a really important conversation with your friend catching up on what is going on in both of your lives. At a nearby table, a group of people start talking loudly and laughing. You are having a very hard time hearing your friend.

Are you at the library?

The nearby table do not realise how loud they are
You study a lot in the cafe
The nearby table is being obnoxious
You cannot concentrate on work at the cafe

The airport
You are flying to a different state to visit a friend. Your first flight is delayed, but you think that you will have enough time to make your connection. When your first flight lands, you realise that the gate for your next flight is across the airport. You sprint to the gate and, just as you approach it, the airline representative closes the door and does not let you board.

Was your first flight on time?

It's unlucky that you just missed your flight
Your flight reachesits destination
Airline staff purposefully preventyou from boarding
All your flights are cancelled

The new job
You have just started working at a new job. When you were hired, you and your new boss discussed how much money you would be earning. You check your mail and see that you have received your first pay cheque. When you open the envelope, you see that the number is lower than you had anticipated.

Have you been working here long?

There is a mistake in the pay cheque
You love your new job
Your boss is lying about your pay
You hate your new job

The chemist
You are waiting in the check-out queue at the chemist. It is a busy shop and there are a few people ahead of you. You are starting to worry that you will be late for class. As you wait, a lady joins the person in front of you

Are you buying anything?

A lady in front keeps her friend company
You save money using your points card
A lady joins her friend to push in front of you
You cannot find an item you needed

Motorway
You are driving down the motorway on your way to the beach. Near your exit you realise you are in the wrong lane. You immediately signal and try to move to the left lane. As you try to move a car prevents you from taking your exit

Are you going to the mountains?

The other car does not see you
You get to the beach in time for sunset
The other car makes you miss the exit
It is too hot in the car

Busy airport
You are walking through a busy airport terminal on your way to the departure gate. As you switch your bag to your other hand you drop your boarding pass. You bend down to pick it up. Just as you're grabbing it, someone steps on your fingers

Are you on your way to the departure gate?

Someone accidentally steps on your fingers
You are upgraded to first class
Someone carelessly stomps on your fingers
You are late and miss your flight

Shopping bags
You are returning home with many heavy shopping bags. As you walk up to your building, you realise you don't have your key. You see another tenant and think they see you. You say Wait for me! and try to walk faster but the door closes

Are you entering the shop?

The other tenant does not hear you
You make a delicious meal with your shopping
The other tenant shuts the door on you
You forget to buy some items

Paying the bill
You go for pizza with a large group of classmates. When the bill comes it cannot be separated. Someone calculates how much each person should put in. You are paying by credit card, so everyone hands you cash and leaves while you pay. You realise you're short by several pounds.

Did you eat burgers and fries?

The person who does the sums miscalculates
You allhave a great time eating pizza
Your classmates purposefully rip you off
You later get food poisoning

Doctor’s office
You have had a terrible headache for several days and worry there may be something wrong. You make a doctor's appointment. The doctor is busy, but able to see you briefly. You tell him about your headache and he simply says to take an aspirin and leaves.

Did you have a toothache?

You are okay so the doctor is not concerned
After an aspirin your headache gets much better
The doctor is disrespectful and does not listen to you
Your headache gets a lot worse

The restaurant
You go out for a meal with friends. You look at the menu and see your favourite dish! However, you read that it has nuts sprinkled on it and you are allergic to nuts. You explain this to the waiter who promises to tell the kitchen. However, the meal arrives covered in nuts

Are you out for dinner with your family?

The waiter makes an inadvertent mistake
The dinner is absolutely delicious
The waiter does not take you seriously
The dinner tastes really bad

The gym
You are at the gym working out. The cardio machine you want to use in being used by someone else. You do something else but keep an eye on the machine. After about 15 minutes the girl using it gets a phone call. She stops the machine, but stays on it whilst she has a 20 minute chat

Do you want to use a cardio machine?

The girl does not realise you arewaiting
You feel great after your workout
The girl using the machine is selfish
You injure yourself on the machine

The neighbour
You live in an apartment block with an assigned space outside. You and your neighbour have adjacent spaces. One day you get home and try to park, but see that your neighbour is parked halfway in your spot so you cannot fit in the space

Do you live in a house?

Another car must be halfway in her space
When you get home there's a parcel you wanted
Your neighbour parked without concern for you
You notice you have a flat tyre

The beach
You are lying at the beach relaxing with your eyes closed. Nearby, a group of people are playing catch, but do not seem to be throwing the ball near you. You are feeling really relaxed and about to fall asleep when, "Whack!" the ball lands on your stomach

Are you at the pool?

They accidentally throw it too far and it hits you
You have a really fun day
The people playingcatch disregard you
You get really bad sunburn

Appendix X – AIHQ

PLEASE READ EACH OF THE SITUATIONS LISTED ON THE NEXT FEW PAGES AND IMAGINE THE SITUATION HAPPENING TO YOU. FOR EACH SITUATION, WRITE DOWN A BRIEF REASON FOR IT. THEN, RATE WHETHER YOU THINK THE PERSON ACTED THAT WAY TOWARD YOU ON PURPOSE. YOU WILL THEN BE ASKED TO RATE HOW ANGRY THAT SITUATION MAKES YOU FEEL AND HOW MUCH YOU BLAME THE OTHER PERSON. FINALLY, PLEASE WRITE DOWN WHAT YOU WOULD DO ABOUT THAT SITUATION. A RESPONSE OF "I DON'T KNOW" IS NOT ACCEPTABLE. YOU NEED TO DESCRIBE SOME TYPE OF BEHAVIORAL RESPONSE.

Please turn over to begin the questionnaire

1. Someone jumps in front of you on the grocery line and says, "I'm in a rush."

A. What do you think was the real reason why someone jumped in line in front of you?

____

B. Did that person jump in front of you on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame that person for jumping in front of you on line?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

2. A friend of yours slips on the ice, knocking you onto the ground.

A. What do you think was the real reason why your friend knocked you to the ground?

____

B. Do you think your friend knocked you onto the ground on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame your friend for knocking you onto the ground?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

3. You've been at a new job for three weeks. One day, you see one of your new co-workers on the street. You start to walk up to this person and start to say hello, but she/he passes by you without saying hello.

A. What do you think was the real reason why your coworker passed by you without saying hello?

____

B. Do you think your co-worker did this to you on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame the co-worker for passing by you?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

4. While walking outside during the rain, a car swerves to avoid hitting a cat, and drives into a puddle, splashing water onto you.

A. What do you think was the real reason why the car splashed water onto you?

____

B. Do you think the driver of the car splashed water onto you on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame the person in the car for splashing water onto you?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

5. You have an appointment with an important person. When you arrive at your appointment, the secretary informs you that the person is not in; they took the day off.

A. What do you think was the real reason why the person didn't keep your appointment?

____

B. Do you think the person did this to you on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame the person for not keeping your appointment?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

6. You are on a bus sitting in an aisle seat. A person gets on the bus at the next stop, begins walking as the bus moves, and steps on your foot.

A. What do you think was the real reason why the person stepped on your foot?

____

B. Do you think the person did this to you on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame the person for stepping on your foot?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

7. Your neighbours are playing loud music. You knock on the door and ask them to turn it down. Fifteen minutes later, the music is loud again.

A. What do you think was the real reason why your neighbours made the music loud again?

____

B. Do you think your neighbours raised the music on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame them for raising the music again?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

8. You walk past a bunch of teenagers at a mall and your hear them start to laugh.

A. What do think was the real reason why the teenagers started to laugh after you walked past them?

____

B. Do you think the teenagers did this to you on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame the teenagers for laughing as you walked past them?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

9. While driving, the person in the car behind you honks their horn and then cuts you off.

A. What do you think was the real reason why the person cut you off while driving?

____

B. Do you think the person cut you off on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame the driver of the car for cutting you off on the road?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

10. You are supposed to meet a new friend for lunch at a restaurant but she/he never shows up.

A. What do you think was the real reason why your new friend didn't show up at the restaurant?

____

B. Do you think your new friend did this to you on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame your new friend for not showing up at the restaurant?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

11. You've been looking for a parking spot for awhile, when you see one up ahead. You put your signal on, proceed toward the spot, but someone passes your car and takes the parking space.

A. What do you think was the real reason why the person in the other car took your parking space?

____

B. Do you think the person in the other car took your parking space on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame the person in the other car for taking your parking space?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

12. You're dancing at a club and someone bumps into you from behind.

A. What do you think was the real reason why the person in the club bumped into you from behind?

____

B. Do you think the person bumped into you on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame the person for bumping into you at the club?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

13. You call a friend and leave a message on their answering machine, asking them to call you back. One week passes and they have not called you back.

A. What do you think was the real reason why your friend didn't call you back?

____

B. Do you think your friend didn't call you back on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame your friend for not calling you back?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

14. You're at a bar watching a football game and having a drink. Suddenly, the home team scores, people begin to cheer, and someone hits your arm, spilling the drink onto your clothes.

A. What do you think was the real reason why the other person hit your arm?

____

B. Did the other person hit your arm on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	Maybe Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame the other person for hitting your arm?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

15. A day before meeting someone for a date, she/he calls to cancel. This is the third straight time they've done that.

A. What do you think was the real reason why the other person cancelled the date with you?

____

B. Did the other person cancel the date on purpose?

1	2	3	4	5	6
Definitely No	Probably No	Maybe No	May Yes	Probably Yes	Definitely Yes

C. How angry would this make you feel?

1	2	3	4	5
Not at all Angry				Very Angry

D. How much would you blame the other person for cancelling the date?

1	2	3	4	5
Not at All				Very Much

E. What would you do about it?

____

Appendix Y – Study 5 Detailed analysis plan

Hypothesis one

a) A Pearson correlation was conducted to explore the relationship between interpretation bias for target and foil statements on the recognition task and AIHQ score across all subscales.
b) An ANOVA was conducted to explore the possible interaction between the explicit measure of interpretation bias (AIHQ) and evoked amplitude in response to the positive and negative statements on the recognition task, for each epoch. Target statement (2 levels: positive and negative), electrode (5 levels) and hemisphere (2 levels) were added as within subject factors. AIHQ (2 levels: high and low based on a median split of scores) was added as a between-subject factor.

Hypothesis two

a) A Pearson correlation was conducted to explore the relationship between interpretation bias score on the AIHQ and aggression score.
b) A repeated measures ANOVA was conducted to explore the difference in *target bias* and *foil bias* between low and high aggression samples when categorised based on a median split of total aggression score. The ANOVA consisted of bias type (2 levels; target and foil) as a within subject factors. Aggression score was added as a between-subject factor. Post-hoc paired samples t-tests were used to explore significant interactions.
c) A Pearson correlation was conducted to explore the relationship between interpretation bias for target and foil statements on the recognition task and aggression score.

Hypothesis three and four

A repeated measures mixed model analysis of variance (ANOVA) was performed on ERP measures for all selected epochs. The ANOVA was used to address research questions three and four. Firstly, the evoked amplitude in response to negative and positive target statements was explored in both aggression groups. Secondly, we investigated whether amplitude of high and low aggression samples differ depending on their similarity ratings of the positive and negative target statements. Target statement type (positive versus negative trials), response (similar versus dissimilar), electrode (5 levels) and hemisphere (left versus right) were included as within-subject factors. Total aggression score was added as a between-subject factor. Greenhouse-Geisser (Geisser & Greenhouse, 1958) *F* test is reported throughout for all repeated measures to ensure there are no violations of the sphericity assumption.