A Behavioural and Electrophysiological Exploration of the Effect of Vocal Emotion on Visual Spatial Attention in Adolescents with Traits of Attention-Deficit Hyperactivity Disorder

by

Rohanna Clare Sells

A thesis submitted in partial fulfilment for the requirements for the degree of Doctor of Philosophy at the University of Central Lancashire

January 2025

RESEARCH STUDENT DECLARATION FORM

Type of Award __PhD____

School ___School of Psychology and Humanities______________________

*Sections marked * delete as appropriate*

1. Concurrent registration for two or more academic awards

I declare that while registered as a candidate for the research degree, I have not been a registered candidate or enrolled student for another award of the University or other academic or professional institution

____

____

2. Material submitted for another award

I declare that no material contained in the thesis has been used in any other submission for an academic award and is solely my own work

3. Collaboration

Where a candidate's research programme is part of a collaborative project, the thesis must indicate in addition clearly the candidate's individual contribution and the extent of the collaboration. Please state below:

____

4. Use of a Proof-reader

No proof-reading service was used in the compilation of this thesis.

Signature of Candidate _ RCsells ____

Print name: ___**ROHANNA SELLS**_________________

Abstract

Accurate recognition of emotions from vocal expressions is associated with emotion regulation and social competence in children and adolescents (Lemerise & Arsenio, 2000; Crick & Dodge, 1994). Emotion dysregulation is highly prevalent in individuals with ADHD (Shaw et al., 2014), who also display vocal emotion recognition (VER) difficulties (Sells et al., 2023). However, it is debated whether VER difficulties in ADHD reflect generic attention deficits, in favour of cognitive-behavioral theories (Barkley, 1997, 2012) or specific perceptual processing atypicalities in favour of motivational-dysregulation (Sonuga-Barke, 2012) and socio-cognitive models (Lemerise & Arsenio, 2000). This thesis aimed to contribute to this theoretical debate within the literature. Chapter One introduces the vocal emotion processing literature and the theoretical debates in ADHD. Chapter Two presents a systematic review of the VER literature and provides meta-analytic evidence for the presence of generic VER deficits in ADHD. However, a gap in the literature was identified to more robustly examine the potential presence of specific perceptual processing atypicalities in ADHD using preattentive paradigms and neuroimaging measures. Therefore, the present thesis aimed to disentangle a) the preattentive processing of vocal emotion from b) the capture of attention by vocal emotion in ADHD using a novel emotional spatial cueing (ESC) paradigm. Chapter Three presents a study which successfully validated a set of vocal emotional stimuli in adults and adolescents. Chapter Four presents four experiments which validated the ESC paradigm using reaction times, and event-related potential (ERP) measures in adults. This study isolated the Cue-P2 component as a neural marker of preattentive vocal emotion processing in adults. Finally, Chapter Five provides evidence for a positive association between the neural marker of preattentive vocal emotion processing and traits of ADHD (inattention and hyperactivity) in a community sample of 60 adolescents. This is the first study to provide evidence for a preattentive hypersensitivity to vocal emotion in ADHD, in favour of motivational-dysfunction theories (Sonuga-Barke, 2012).

Table of Contents

Acknowledgements

First and foremost, I would like to thank Dr Georgia Chronaki for being my main supervisor, not just throughout my PhD studies, but for the past ten years, from my undergraduate studies at the University of Manchester to my postgraduate studies at the University of Central Lancashire. Without her continued encouragement and support, and thorough feedback, this PhD would not have been possible. I have learnt so much from her, and I cannot thank her enough.

I also want to acknowledge Professor Simon Liversedge, my second supervisor, for his invaluable and wise advice throughout the PhD process. His guidance has undoubtedly enabled me to excel further than I ever thought possible and has motivated me to 'keep going' through challenging times of the PhD.

Many thanks to the School of Psychology for providing funding for this PhD and for the Developmental Cognitive Neuroscience Lab, which has been a great space to conduct my EEG studies. I also extend my thanks to all the adults, adolescents, families and schools who participated in this research.

I feel incredibly fortunate to have had the support of some wonderful friends and colleagues. Laura Rettie has been so supportive and kind since day one of starting my role as a GTA and PhD student when I moved to Preston five years ago. I thank her for sharing both the peaks and troughs of a PhD, along with the rollercoaster of emotions that comes with it. Thanks also to Dr Cristina Foderella for regular coffee catch-ups, and to Petar Atanasov and Dr Zoe Hughes in office DB124 for moments of encouragement and offloading. Special thanks to Leanne Shearsmith, my best friend for fifteen years, for always being there when I've needed it most, and for organising the best trips away, and to the rest of the girls from home – Amy, Kobs, Anna, Rosie and Rachel - thank you for being absolute rocks and helping me to take my mind of the PhD.

Of course, I am eternally grateful to my family, who have always supported me in everything I have chosen to do. Thank you, Mum, Duncan, Dad, and Jane.

Lastly, I want to thank Kieren Banlin. The past write-up year has been especially challenging, but he has been a source of daily love and laughter. Thank you for van trips to Scotland, paddleboarding days on the lakes, and for bringing Honey into our lives—she has been the emotional support dog I never would have thought I needed these last six months.

List of Figures

.

List of Tables

Glossary of Terms, Abbreviations, List of Symbols

ADHD	Attention Deficit Hyperactivity Disorder
ANOVA	Analysis of Variance
APA	American Psychological or Psychiatric Association
CD	Conduct Disorder
CNV	Contingent Negative Variation
dB	Decibel
ED	Emotion Dysregulation
EDI	Escape Delay Incentive Task
EF	Executive Function
EEG	Electroencephalogram
ERP	Event-Related Potential
ESC	Emotional Spatial Cueing Paradigm
F0	Fundamental Frequency
fMRI	Functional Magnetic Imaging
GHQ	General Health Questionnaire
ITI	Inter Trial Interval
ODD	Oppositional Defiant Disorder
SDQ	Strengths and Difficulties Questionnaire
SOA	Stimulus Onset Asyncrony
STC	Superior Temporal Cortex
STG	Superior Temporal Gyrus
STS	Superior Temporal Sulcus
VSA	Visual Spatial Attention

1 INTRODUCTION

1.1 Vocal Emotion Processing and Socio-Emotional Functioning

Emotions, including happiness, anger, fear, sadness, surprise and disgust, are constantly changing in ourselves, and others (Ekman, 1993; Scherer, 1995). They are expressed as a response to events in our environment and can be influenced by intrinsic factors, such as temperament and arousal, and extrinsic factors, such as social norms and context (Johnstone & Scherer, 2000; Fox & Calkins, 2003). Emotions are expressed outwardly via a number of non-verbal cues, such as the face or voice. The tone of voice (i.e., prosody) is particularly good at communicating emotions over long distances, and for conveying subtleties in the speaker's emotion which may otherwise not be expressed via other emotion cues (i.e., facial expressions; Johnstone & Scherer, 2000). Specifically, prosody refers to changes in the intonation and rhythm evident in everyday speech, rather than the more primitive short form of vocal emotion (i.e., laughs, screams, sighs), named affective bursts (Scherer, 1995). Emotion types portrayed via prosody are characterised by distinct acoustic profiles in speech, such as changes in fundamental frequency (i.e., pitch; F0), amplitude (i.e., intensity; dB), duration and unfolding (e.g., the pattern in which fundamental frequency and amplitude change; Scherer, 1995). For example, anger is characterised by an increase in mean fundamental frequency and mean amplitude, whereas sadness is characterised by a decrease in mean amplitude and fundamental frequency (Scherer, 1995).

Perceiving and correctly interpreting emotions (i.e., anger, happiness) from non-verbal cues, such as prosody, are theorised to be the first steps to competent socio-emotional functioning, as outlined in the Social Information Processing Model (See Figure 1.1; Lemerise & Arsenio, 2000; Crick & Dodge, 1994). Socio-emotional functioning involves correctly interpreting another's emotional state, and regulating one's own emotional state (i.e., emotion regulation), to respond appropriately in social situations, and build meaningful relationships with others (Lemerise and Arsenio, 2000; Crick & Dodge, 1994). Crick and Dodges (1994) Social Information Processing model originally outlined how a) encoding cues from one's own expressions, and b) correctly interpreting the emotion can lead to successful behavioural responses in social interactions. Lemerise and Arsenio (2000) revised the model by adding that the interpretation of others emotional cues also contribute to social responses. For example, correctly inferring sadness in a peer's vocal expression may correspond to appropriate responses of warmth and comfort as opposed to incorrectly inferring anger, which may lead to more hostile responses. Lemerise and Arsenio (2000) additionally integrated into the model other emotional processes, such as the intensity with which an emotion is felt and emotion regulation abilities. For example, a child who is good at regulating their emotions may notice a friend appearing angry in a social situation, and rather than assuming the anger is aimed at them and reacting immediately with hostility, they may approach the friend calmly and ask them if they want to talk about it. This ability to consider a situation from multiple perspectives, and regulate one's own emotional response, can help facilitate more supportive responses (Lemerise & Arsenio, 2000). In turn, supportive responses can lead to positive peer evaluation and can help build relationships among peers (Neves et al., 2021). The rest of this section reviews the research which explores associations between vocal emotion recognition, emotion regulation and socio-emotional functioning, in support of the Social Information Processing Model (Lemerise & Arsenio; Crick & Dodge, 1994).

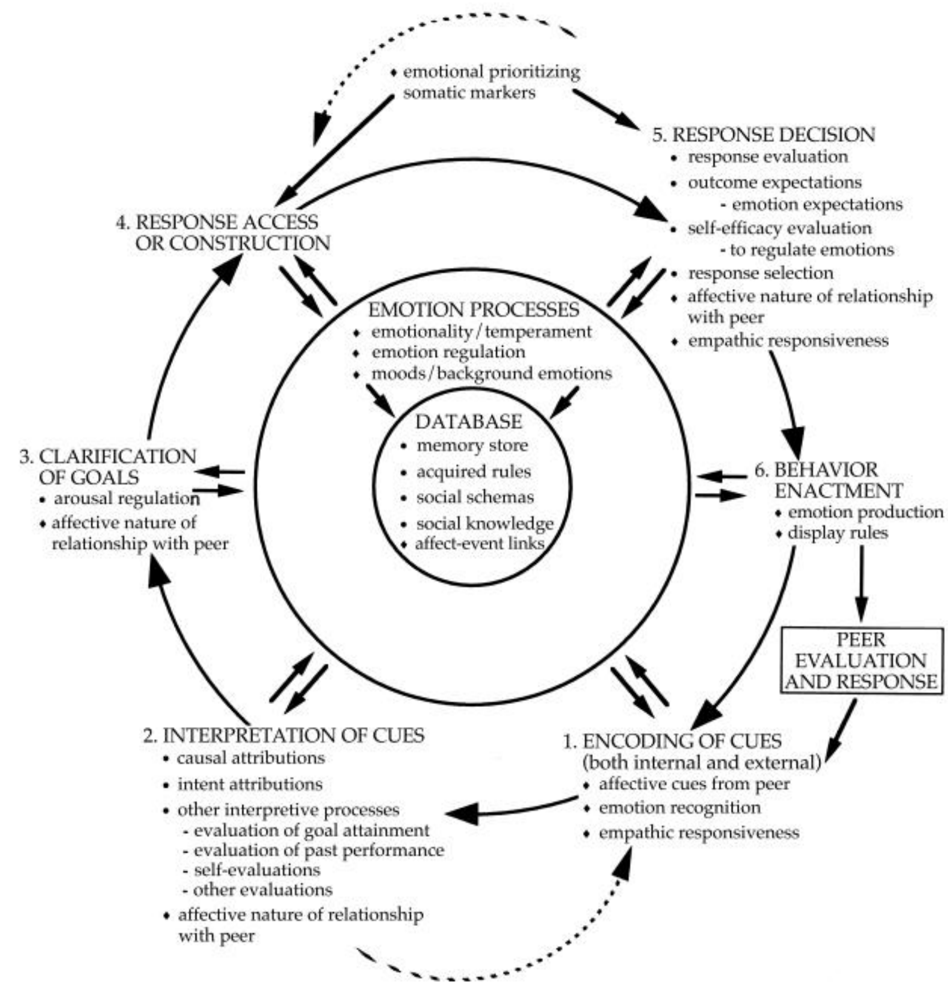

Figure 1.1

The Revised Social Information Processing Model (Lemerise & Arsenio, 2000). The model outlines the integrated emotion and cognitive processes towards behaving with competence in social situations. Bullet pointed items were originally formulated by Crick and Dodge (1994) in the Social Information Processing model. Reprinted from “An integrated model of emotion processes and cognition in social information processing” E.A Lemerise & W.F. Arsenio, Child Development, 71, pages 107-118. Copyright 2000 with permission from John Wiley and Sons. Copyright notice included in the original article: From “A review and reformulation of social-information-processing mechanisms of children’s social adjustment,” N. R. Crick & K. A.Dodge, Psychological Bulletin, 115, p. 74. Copyright 1994 by the American Psychological Association. Adapted with permission by Lemerise & Arsenio (2000).

There has been a wealth of research which supports the association between emotion recognition, emotion regulation and social competence in children (Philippot & Feldman, 1990; Della Longa et al., 2022) and adults (Suslow et al., 2023). Emotion recognition tasks are employed in research which explores associations with social competence. In an emotion recognition task, participants are presented with facial and/or vocal expressions, and instructed to label each expression, typically with one of six basic emotions (i.e., happiness, anger, sadness, fear, disgust, surprise), or neutrality. However, most of the previous research which has explored links between emotion processing and social competence in recognition tasks have primarily only included facial expressions (Philippot & Feldman, 1990; Della Longa et al.,2022; Suslow et al., 2023; Schultz et al., 2001). For example, Philippot and Feldman (1990) presented children with videotaped scenarios and instructed children to only label the emotional expression (i.e., happiness, sadness and fear) displayed in the actor's facial expression, rather than the vocal expression. The authors found that facial emotion recognition accuracy was positively associated with social competence, assessed via a questionnaire completed by parents (Philippot & Feldman, 1990). Similarly, a recent emotion recognition study, which included six facial expressions (happiness, surprise, anger, disgust, sadness and fear) reported significant negative associations between accuracy and self-reported interpersonal problems, such as social dominance and intrusiveness, in adults (Suslow et al., 2023). There has been significantly less research conducted to support these associations within the vocal modality.

The predominance of research on facial emotion processing in the socio-emotional literature is somewhat surprising, given that facial and vocal cues are not always perceived simultaneously, and vocal cues can be recognised independently from facial cues (Scherer, 1995).

Moreover, despite evidence of similarities between the perception of facial and vocal cues, there are notable differences in their nature and transmission (Young et al., 2020). For example, vocal cues are the only form of emotional communication which do not need to be visible to be recognised (Young et al., 2020) and can be conveyed over long distances (or via the telephone) allowing for more efficient signalling of danger, represented by angry and fearful emotions (Johnstone & Scherer, 2000). This also allows for vocal emotions to be more effective at conveying emotions in group settings, when individuals are less able to focus on just one face in a group. Moreover, vocal emotions are transmitted dynamically in a stimulus which changes in intensity, pitch and loudness over time, whereas facial expressions are perceived as static stimuli (Schirmer & Adolphs, 2017). This dynamic nature of vocal expressions adds a layer of complexity when perceiving the emotion and allows more subtleties in emotion changes to be perceived. These substantial differences in both the nature and perception of vocal emotions warrants the exploration of vocal emotion recognition and socio-emotional functioning, independently from facial emotion recognition.

Existing research incorporating vocal, as well as facial expressions, in emotion recognition tasks have reported similar associations with indices of socio-emotional functioning within the two modalities in children (Leppanen & Hietanen, 2001; Nowicki & Duke., 1992; Nowicki & Mitchell, 1998) and adults (Carton et al., 1999). For example, Leppanen & Hietanen (2001) indicated that both vocal and facial emotion recognition accuracy scores in school-aged children (aged 7 to 10 years) were positively associated with social competence, as reported by teachers. Moreover, researchers employing alternative indicators of socio-emotional functioning, such as peer popularity measured through class voting systems, have similarly revealed positive associations between both facial and vocal emotion recognition abilities and social competence in school-aged children (Nowicki & Mitchell, 1998; Nowicki & Duke, 1992). In adult populations, a poorer ability to recognise emotions from both facial and vocal cues correlated with lower relationship wellbeing and greater depressive symptoms; dimensions commonly associated with socio-emotional functioning in adulthood (Carton et al., 1999). Overall, these findings from studies incorporating vocal and facial emotional expressions collectively align with research in the facial modality only, and support theories which suggest vocal and facial emotion recognition is necessary for social competence (Lemerise & Arsenio, 2000; Crick & Dodge, 1994).

There have been few studies which have reported more specific associations between vocal emotion recognition accuracy and certain social behaviours, including adverse behaviours, such as social avoidance (McClure & Nowicki, 2001) and prosocial behaviours, such as cooperation, sharing and helping (Neves et al., 2021) in children. For example, in a study which recruited 62 children aged eight to ten years and asked them to complete an emotion recognition task including both vocal and facial expressions (happy, sad, angry and fearful) and self-report measures of socio-emotional functioning, McClure and Nowicki (2001) identified a negative association between social avoidance and distress, and the ability to recognise vocal emotions. Interestingly, the study did not report similar results for facial emotions, indicating that recognising emotions from voices may be particularly important for children's confidence to engage in social situations (McClure & Nowicki, 2001). Positive associations between social engagement and vocal emotion recognition are supported in a study by Neves et al (2021) who asked six- to eight-year-olds to complete an emotion recognition task, with vocal prosody, affective bursts (i.e., laughs, screams) and facial expressions and asked their parents to complete a questionnaire measuring the child's socio-emotional adjustment. Neves et al (2021) reported parent-rated pro-social behaviours (e.g., cooperation, sharing, helping) were specifically associated with vocal emotion recognition accuracy in six-to-eight-year-olds, and not with affective bursts or facial expressions. Taken together, findings that vocal emotion recognition is specifically associated with increased social engagement (i.e., increased pro-social behaviours and reduced social avoidance) in children may stem from the prominence of vocal communication in group social settings, common in the playground or classroom, and the nature of vocal emotions conveying more subtle changes of emotion (Young et al., 2020; Johnson & Scherer, 2000). For example, recognising subtleties in voices may allow children to better decipher implicit expectations or instructions, such as helping or sharing with others (Neves et al., 2021) and may lead to increased confidence in engaging socially with peers (i.e., less social avoidance; McClure & Nowicki, 2001).

The link between vocal emotion recognition and emotion regulation abilities was also supported in the study by Neves et al (2021), who reported higher vocal emotion recognition accuracy was positively associated with greater abilities for young children to self-regulate, as measured by parent-ratings. This supports social information processing models which suggest recognition and regulation abilities are highly linked (Lemerise & Arsenio, 2000). However, as there has not yet been any research exploring direct links between vocal emotion recognition and emotion regulation, it is not known if emotion recognition abilities contribute to the development of better regulation abilities, or vice versa.

Finally, research has shown links between processing specifically negative vocal emotions, poorer social adjustment and the development of externalising behaviour problems. For example, Maxim and Nowicki (2003) reported that errors to recognise specifically negative vocal emotions (angry and sad voices) in children (aged 4-6 years) were associated with lower social competence, defined as the quality of the relationship children had with adults. Furthermore, the ability to accurately recognise fearful voices in five-year-old boys was positively associated with teacher-rated social competence and the ability to positively interact with peers (Maxim & Nowicki, 1998). Additionally, Davis and colleagues (2020) revealed adolescents (aged 11 to 18 years) who scored higher in socio-emotional abilities were less likely to incorrectly label vocal expressions as angry. The link between a bias in perceiving negative emotions and lower socio-emotional functioning is supported by a study which reported poor regulation capabilities in adolescents (11-14 years) were associated with the self-reported experience of more negative emotions (i.e., anger, frustration, anxiety) towards schoolwork (Gumora and Arsenio, 2002). The limited evidence suggesting an association between atypicalities in processing specifically negative emotions and poorer socio-emotional functioning in childhood and adolescence could be due to the more adverse social consequences of incorrectly perceiving negative emotions, such as more hostile externalising reactions towards others (Lemerise & Arsenio, 2000). The link between vocal emotion recognition abilities and externalising behaviour problems is supported in a study which found that higher hyperactivity and conduct problem scores in children aged 3–6 years were negatively associated with the ability to accurately recognise angry, happy and sad vocal expressions (Chronaki, Garner et al., 2015). Additionally, children aged 8 who recognised fewer vocal expressions were reported by teachers to display more externalising behavioural difficulties at age 10 (Nowicki, Bliwise & Joinson., 2019).

In summary, it is clear from the literature that vocal emotion recognition abilities are associated with socio-emotional functioning, in support of Social Information Processing Theory (Lemerise & Arsenio, 2000; Crick & Dodge, 1994). These associations are evident in children as young as four years (Maxim & Nowicki, 2003), and persist in childhood (Neves et al., 2021), and adulthood (Carton et al., 1999). Specifically, vocal emotion recognition has been shown to be positively associated with pro-social behaviour and emotion regulation (Neves et al., 2021) and negatively associated with social avoidance (McClure & Nowicki, 2001) and externalising behavioural difficulties (Chronaki, Gardner., 2015; Nowicki, Bliwise & Johnson., 2019) in children. This indicates the importance of children developing effective vocal emotion recognition abilities to build and maintain relationships with peers from childhood. The research exploring associations between vocal emotion processing and socio-emotional functioning in adolescents is less clear, however there is some limited evidence which suggests the experiencing more negative vocal emotions may be specifically linked with poorer emotion regulation abilities (Gumuro and Arsenio, 2002). Taken together, findings relating to vocal emotion processing and specific dimensions of socio-emotional functioning reinforces the importance of independently exploring vocal emotion processing, separately from facial emotion processing in children and adolescents.

1.2 The Development of Vocal Emotion Processing

It is argued that the perception of vocal emotions is influenced by both inherent and environmental factors in development (Banziger et al., 2015; Laukka et al., 2016). The innate nature of vocal emotion perception is evidenced by studies which demonstrate that basic emotions (happy, anger, fearful, sadness, disgust, and surprise) can be recognised from the voice universally, across cultures in adults (Sauter et al., 2010), as well as in children and adolescents (Chronaki et al., 2018). However, it is now widely accepted that in addition to inherent factors, environmental factors (e.g., experience, culture, and social influence) play a substantial role in the development of vocal emotion processing from infancy to adolescence (Mesquita & Frijda, 1992; Chronaki et al., 2018; Elfenbein & Ambady, 2002). For example, Chronaki et al (2018) demonstrated that native English-speaking children and adolescents aged between 8 and 13 years were significantly more accurate to recognise emotions in their own native language compared to other foreign languages, supporting the role of socio-cultural factors influencing vocal emotion recognition. In support, studies employing vocal emotion recognition tasks have demonstrated that the ability to recognise vocal emotions does not reach maturity until mid to late adolescence (Chronaki, Hadwin et al., 2015; Grosbras et al., 2018), a trajectory which is longer than that found for other non-verbal cues such as vocal bursts (Sauter et al., 2013) or facial expressions which reach maturity in childhood (Chronaki, Hadwin et al., 2015). This section aims to review the research in the development of vocal emotion processing from infancy to adolescence and explore inherent and social-cultural factors which contribute to its development.

When fully developed in adulthood, accuracy to recognise vocal emotions is high, with rates ranging from 55% to 93%. For example, Banse & Scherer (1996) found that adults could accurately recognise 14 different emotions from meaningless pseudo-utterances (e.g., 'Hat sundig pron you venzy') at above chance levels (55% accuracy). When limited to basic vocal emotions (happiness, anger, sadness, fear and disgust), adult recognition accuracy increased to between 66% (Scherer et al., 2001) and 93% (Chronaki et al., 2018). Studies employing other types of prosodic linguistic stimuli in vocal emotion recognition tasks, such as meaningless interjections ('ah'; Maurage et al., 2007) or semantically neutral sentences ('I felt this feeling before; it was just a few days ago'; Nelson & Russell., 2011) similarly reported high accuracy rates (70%-90%) for emotions (anger, happiness, fear, sadness) in adults. Interestingly, vocal emotion recognition studies consistently demonstrate that adults are more accurate at recognising vocal emotions conveying anger across languages, compared to positive emotions (i.e., happiness; Pell et al., 2009; Chronaki et al., 2018; Liu & Pell, 2012; Zupan, 2015). It is possible that vocal anger is easier to identify due to more distinct changes in acoustic properties, such as pitch or intensity (Gobl & Chasaide, 2003; Zupan, 2015). Moreover, the higher accuracy rates reported for negative vocal emotions may be linked to evolutionary pressures favouring the recognition of emotions which signal threat or danger when the speaker is not in visual sight, such as over long distances (Liu & Pell, 2012; Vuilleumier, 2005). However, there have also been reports that positive vocal emotions are also identified more efficiently than other emotions. For example, Burra et al (2019) reported faster responses to happy compared to both negative and neutral voices in a task which presented emotional and neutral vocal cues to participants and asked them to identify if an emotion was present in the signal. The faster perception of positive emotions has also been linked to adaptive behaviours, such as approaching positive stimuli within the environment, including potential mates and food sources (Anderson, 2016; Schultz, 2004; Wang et al., 2022). Higher recognition of certain emotion types, such as positive vocal emotions, has been related with individual differences. For example, adolescents who self-reported higher levels of loneliness were better at recognising friendliness from the voice and poorer at recognising fearful voices (Morningstar et al., 2020).

There is debate in the literature as to whether infants and young children can accurately perceive and recognise emotions in voices (Grossman, 2010). In line with evolutionary accounts (Liu & Pell, 2012; Vuilleumier, 2005), it is thought inferring vocal emotions, specifically emotions which signal threatening or dangerous situations, is adaptive for young infants who are learning to navigate the world. For example, if an infant is reaching for a dangerous object, such as pan of boiling water, a parent's voice conveying negative affect (such as anger or fear), will deter the infant from doing so more efficiently than a negative facial expression (Mumme et al., 1996). This early sensitivity to vocal emotions is supported by Mumme et al (1996) who reported that 12-month-old infants changed their behaviour towards a novel toy based on their mother's vocal expression. When the mothers produced a negative vocal emotion, infants were less likely to engage with the toy and expressed more negative emotion themselves, suggesting that infants change their behaviour based on vocal emotion perception early in development, which can have adaptive benefits (Mumme et al., 1996). In further support, Flom and Bahrick (2007) presented vocal emotional expressions, consisting of female infant-directed speech, to 3-, 5- and 7-month-old infants, and demonstrated that the 5- and 7-month-old infants, but not the 3-month-old infants displayed longer visual fixations and looking time towards the location of the voice. This indicates that infants as young as 5 months can discriminate changes in vocal emotion (Flom & Bahrick, 2007; Grossman, 2010). However, it is possible these reactions may be due to sensitivities to the changes in the acoustic properties which characterise vocal emotions (i.e., frequency, intensity, unfolding), rather than infants processing the emotion itself. In dispute of this, Mastropieri and Turkewitz (1999) demonstrated that newborn infants (aged 12-72 hours) showed an increase in eye opening responses to only vocal emotions expressed in their native language, and not non-native language. This implies cultural influences contribute to vocal emotion perception from birth. Overall, evidence from infant studies suggest that from birth, infants rapidly develop the ability to implicitly perceive basic changes in emotional valence within voices for adaptive benefits, as measured by eye opening response and behaviours. However, the ability to explicitly recognise emotion types from voices, as measured by vocal emotion recognition tasks, develops more gradually during childhood (Chronaki et al., 2015; Zupan, 2015).

Research employing recognition tasks indicate that although pre-school children can label emotions from voices at above chance accuracy, the ability to recognise vocal emotions at adult-levels develops at a much slower pace than the recognition of emotion from other cues, such as vocal bursts, facial expressions (Zupan, 2015; Chronaki et al., 2015; Nelson & Russell, 2011) and semantics (Aguert et al., 2013; Morton & Trehub, 2001). For example, Nelson and Russell (2011) demonstrated that pre-school children (aged 3-5 years) could accurately infer emotions (happiness, sadness, anger and fear) from facial and body posture stimuli with over 65% accuracy, while recognition rates for vocal emotions were only between 35-50%; significantly lower than those observed in adults (80%). Similarly, other research showed that children aged between 4 and 12 years were consistently more accurate when recognising facial compared to vocal emotions (Chronaki et al., 2015; Zupan, 2015). However, consistent with adult patterns, children displayed higher accuracy rates when recognising negative vocal emotions (particularly angry voices), compared to positive vocal emotions and neutral voices (Nelson & Russell, 2011; Chronaki et al., 2015; Zupan, 2015). This favours evolutionary accounts which argue that developing the ability to recognise vocal emotions which signal danger (i.e., anger) has adaptive benefits (Liu & Pell, 2012; Vuilleumier, 2005). Further research suggests children under the age of 10 are more likely to rely on semantic or contextual information to understand emotions in others, rather than prosody. For example, in an auditory Stroop task which presented children aged 4-10 years with vocal cues, in which prosody and semantics were incongruent in terms of emotion, the children were more likely to infer the emotion from the semantic cues as opposed to prosody (Aguert et al., 2013; Morton & Trehub, 2001). This differs from adults who more rapidly identified emotion from prosody compared to semantics in an auditory Stroop task (Filippi et al., 2017). In addition, 6-11-year-old children are more accurate to recognise vocal emotional cues when the prosody and semantic content are congruent (e.g., as in the sentence 'it's my birthday today' with a happy tone of voice) compared to when these cues are incongruent (e.g. as in the sentence 'it's my birthday today' with a sad tone of voice (Cartwright., 2024). Taken together, research employing Stroop-like paradigms suggests that relying solely on prosody versus semantic context (Aguert et al., 2013; Morton & Trehub, 2001) or facial expressions (Chronaki et al., 2015) develops with age.

Additional findings from studies employing vocal emotion recognition tasks have demonstrated that vocal emotion processing develops gradually throughout childhood and continues to develop after 11 years (Chronaki et al., 2015; McClanahan, 1996, Mitchell, 1995;

Zupan et al., 2015). Specifically, Chronaki et al (2015) reported that vocal emotion recognition was significantly higher for all emotion types (angry, happy and sad) in older children (6-11years) compared to pre-school children (aged 4-5 years). This is consistent with other studies which reported that recognition abilities for angry, happy and sad voices were higher in older children (9-10 year olds) compared to younger children (4-5 years olds; McClanahan, 1996; Mitchell, 1995). Additionally, recognition for vocal emotions were higher in 11–12-year-olds compared to 8–10-year-olds (Zupan, 2015). Importantly, studies have further demonstrated that 11-year-olds are still less accurate to recognise vocal emotions, specifically sad and neutral vocal emotions, compared to adults (Chronaki et al., 2015; Zupan, 2015). Interestingly, the slower developmental rate for the recognition of sad vocal emotions was further highlighted by findings which reported sad vocal emotions spoken with low intensity were recognised with the least accuracy compared to other positive and negative vocal emotions throughout childhood (Chronaki et al., 2015). This finding aligns with accounts which argue that sadness is a more ambiguously expressed emotion (Stifter & Fox, 1986) and evidenced to be influenced more by context and social norms (Gray & Heatherington, 2003). This suggests that vocal emotion recognition, at least for more subtle or ambiguous vocal emotions, continues to develop into adolescence, and may be influenced to a greater extent by experience and learning during this developmental period.

It is now widely accepted that vocal emotion recognition accuracy gradually continues to improve until mid to late adolescence (aged 15-17 years; Filippa et al., 2022; Grosbras et al., 2018; Morningstar, Ly et al., 2018). For example, a study which compared vocal emotion recognition accuracy in 13 to 15-year-olds relative to adults, showed that adolescents' performance was less accurate than adults' only when hearing vocal stimuli produced by other adolescents, but not by adults (Morningstar, Ly, et al., 2018). This could be because adolescent expressions of emotions are less distinct in terms of changes in pitch compared to adult expressions of emotions, resulting in more subtle changes in emotion, which are more difficult to identify (Morningstar, Ly, et al., 2018). In further support of the gradual improvement of vocal emotion recognition in adolescence, Grosbras et al (2018) instructed participants aged 5-17 years to recognise emotions from interjections sounds ('ah') and reported that although vocal emotion recognition improved quickly in childhood, it continued to slowly improve throughout early adolescence and didn't reach adult-like levels until 15 years. Interestingly, Grosbras et al (2018) additionally reported that happiness was recognised with the most accuracy in children and adolescents, which differed from previous reports consistently suggesting negative vocal emotions are recognised with the highest accuracy in adults and children (Pell et al., 2009; Chronaki et al., 2018; Liu & Pell, 2012; Zupan, 2015; Chronaki et al., 2015). However, a higher recognition of happy voices in adolescence may be due to a greater sensitivity to more rewarding stimuli during this period (Galvan, 2013).

Fillipa et al (2022) similarly revealed that the developmental trajectory for vocal emotions improved in adolescence up until 17 years, with vocal anger being recognised with the highest accuracy across development. Within this study a non-forced emotion recognition task was employed, in which children and adolescents (aged 6 to 17 years) judged the emotion and intensity of the vocal expression (i.e., pseudo-utterances) on continuous scales, depicting happiness, fear, sadness, anger, surprise, and neutrality. Using this method, participants could perceive and label the expression as multiple emotion types of varying intensity. For example, participants could label a vocal expression as 75% anger and 25% fear. Using this method, Fillipa et al (2022) reported that older children and adolescents (up to around 14 years) improved in their ability to infer multiple emotions in one vocal stimulus. This novel finding suggests that vocal emotion perception becomes more complex during adolescence and may be linked to other maturations of socio-emotional functioning, such as mentalising (Morningstar et al., 2022), which involves the ability to understand that others can experience multiple, and possibly contradictory emotions, simultaneously (Filippa et al., 2022). This ability is particularly useful in adolescence; a period when individuals are more sensitive to social experiences (Blakemore & Mills, 2014) and tend to spend more time with peers, whom they share more emotional experiences, and build more fulfilling interpersonal relationships with (Trentacosta & Fine, 2010). The link between the development of vocal emotion recognition and socio-emotional functioning in adolescence is supported by evidence from a neuroimaging study which measured changes in the neural activation to emotional voices in adolescents who took part in a vocal emotion recognition task one year apart (Morningstar et al., 2022). The results revealed that significant changes in activation of brain regions (the inferior frontal gyrus, dorsomedial prefrontal cortex and temporo-parietal junction), involved in emotion interpretation and mentalising abilities, were associated with an increase in vocal emotion recognition accuracy across the two time points.

In summary, the developmental literature converges to suggest that although vocal emotion recognition improves significantly in late childhood, the ability to recognise subtleties (i.e., ambiguous, low intensity emotions) and more complexities (i.e., recognising multiple emotions) in vocal expressions gradually improves during mid to late adolescence (Chronaki et al., 2015; Filippa et al., 2022; Grosbras et al., 2018; Morningstar, Nelson, et al., 2018). This suggests that adolescence is an important period for the developmental of vocal emotion processing, which aligns with the development of other forms of social cognition (i.e., mentalising), and may be crucially influenced by changes in the social environment during this time (Blakemore & Mills, 2014). The extended developmental trajectory of vocal emotion recognition in childhood and adolescence support theories which suggests learning and experience influence vocal emotion processing, in addition to inherent factors. This aligns with findings from cross-cultural studies in children and adolescence (Chronaki et al., 2018), which highlight the influence of socio-cultural factors on vocal emotion recognition.

1.3 Vocal Emotion Processing and Attention-Deficit/Hyperactivity Disorder

As evidenced in Section 1.1, the ability to recognise vocal emotions, throughout childhood and adolescence is positively associated with emotion regulation, behavioural adjustment and healthy social-emotional functioning (Neves et al., 2021). Similarly, difficulties recognising vocal emotions are associated with emotion dysregulation (ED) and psychopathology in children and adolescents (Uekermann et al., 2010). ED is defined as impairments in effectively controlling one's own emotions (D'Agostino et al., 2017; Thompson, 1994) and is often conveyed through temper outbursts, reactive anger, and irritability (Bunford et al., 2015; Drechsler et al., 2020; Shaw et al., 2014). ED has been implicated in a wide range of psychopathological conditions, such as anxiety, depression and conduct problems in children and adolescents (Compas et al., 2017; McLaughlin et al., 2011; Moehler et al., 2022). Notably, evidence suggests ED is present in 40-50% of children and 70% of adults diagnosed with Attention-Deficit/Hyperactivity Disorder (ADHD; Faraone et al., 2019; Shaw et al., 2014). This rate is higher compared to the presence of ED in children and adolescents admitted to mental health facilities (26-30%; Moehler et al., 2022) and significantly higher compared to healthy adult controls (Barkley & Fischer, 2010). Given the high prevalence of ED in ADHD, this section first outlines different perspectives of ED in ADHD, and the importance of investigating its underlying causes. Next, the underlying causes and theoretical debates to ED in ADHD are discussed, drawing on the evidence from the vocal emotion processing literature.

Categorical and dimensional perspectives have been proposed to conceptualise ADHD and the presence of ED within it. The categorical perspective defines ADHD as a neuro-developmental disorder outlined by criteria, such as those within the Diagnostic and Statistical

Manual of Mental Disorders (5th edition; DSM-IV, APA, 2013). These criteria specify the high frequency of core symptoms (inattention and hyperactivity/impulsivity) and emphasise that these symptoms must negatively impact both academic and social functioning, in home and school settings, for an individual to reach the threshold for a diagnosis (APA, 2013). When categorically defining ADHD by these criteria, the disorder is prevalent in 5-8% of children and adolescents (Polanczyk et al., 2014; Thomas et al., 2015), and 2.8% of adults (Faraone & Biederman, 2005; Fayyad et al., 2017). Although ED is not explicitly stated within the diagnostic criteria, the presence of ED in ADHD is argued to be clinically significant due to its association with more severe ADHD symptomatology and an increased likelihood of comorbidities with anxiety and depression, and poorer academic and socio-emotional outcomes (Barkley & Fischer, 2010; Bunford et al., 2018; Faraone et al., 2019; McKay et al., 2023). For example, the presence of ED in ADHD has been found to exacerbate social problems, such as difficulties with friendship and social rejection, in adolescents (McKay et al., 2023; Bunford et al., 2018). Alternatively, the dimensional approach defines ADHD not as a disorder, but as a form of neuro-divergence, in which core symptoms (i.e., inattention and hyperactivity/impulsivity) are presented at an extreme end of a continuous scale (Sonuga-Barke et al., 2023; Astle et al., 2022). In this perspective, inattention and hyperactivity/impulsivity, and ED are all separate but highly correlated traits within the population.

Although historically ADHD has been defined as a categorical entity, recent perspectives are evolving to favour the dimensional approach to ADHD (Astle et al., 2022; Cuthbert & Insel, 2013; Sonuga-Barke et al., 2023). Research in support of the dimensional approach provides evidence that symptoms of inattention and hyperactivity/impulsivity are present at varying levels in individuals with, and without, a diagnosis of ADHD, and within other neurodevelopmental disorders, such as learning-related difficulties and autism spectrum disorders (Astle et al., 2022; Sokolova et al., 2017). Furthermore, studies find sub-threshold levels of ADHD are prevalent in up to 23% of children and adolescents (Balázs & Keresztény, 2014), who display impairments, and neurobiological correlates (i.e., cortical thinning; Shaw et al., 2011) similar to those seen in individuals with a clinical diagnosis of ADHD (McLennan, 2016; Sonuga-Barke et al., 2023). Similarly, behavioural interventions have been proven to be effective at reducing behaviour problems in children with ADHD and without a diagnosis of ADHD, who display similar (albeit milder) difficulties (McLennan, 2016). Crucially, in further support of a dimensional approach, longitudinal studies which recruited community samples, in which the full range of ADHD symptom severity was represented from no symptoms to high symptoms, have reported strong associations between inattention, hyperactivity/ impulsivity, ED and social impairments, from pre-school to adolescence (Brocki et al., 2019; Sjowall et al., 2017; Taylor et al., 1996).

Interestingly, further evidence from community samples suggests conduct and emotional problems (i.e., depressive symptoms) are positively associated with inattention and hyperactivity/impulsivity symptoms, particularly in adolescence (Park & Chang, 2022; Bird et al., 1993). For example, a longitudinal study reported that inattentive symptoms in children predicted later depressive symptoms in adolescents (Park & Chang, 2022). Moreover, longitudinal studies have reported that hyperactivity symptoms are a significant predictor of conduct problems in later childhood (Gustafsson et al., 2018) and adolescence (Taylor et al., 1996). This is in accordance with categorical studies which report that adolescents with a diagnosis of ADHD also meet the criteria for Conduct Disorder (CD) or Oppositional Defiant

Disorder in 25 to 75% of cases (Barkley, 1998) and have an increased risk of developing depression (Fischer et al., 2002). Moreover, a recent review suggested that children with a diagnosis of ADHD present with at least one comorbidity, such as CD or depression, in 44% of cases (Steinberg & Drabick, 2015). Importantly, it is suggested that the combination of ED and ADHD symptomatology is a risk factor for developing these comorbidities (Steinberg & Drabick, 2015; Seymour et al., 2014), further highlighting the need to understand the cause of ED in ADHD.

Given the high prevalence of ED in ADHD (Shaw et al., 2014), and its association with adverse academic and socio-emotional outcomes (Barkley & Fischer, 2010; Bunford et al., 2015; Faraone et al., 2019; Stringaris et al., 2012), as well as the increased risk of comorbidities in children and adolescents (Steinberg & Drabick, 2015; Barkley, 1998; Fischer et al., 2002), it is important to consider the underlying causes of ED in ADHD. Most theories converge to suggest that difficulties in recognising and/or allocating attention to emotional expressions, such as the face or voice, contribute to the development of ED in ADHD (Shaw et al., 2014; Uekermann et al., 2010). For example, there has been evidence from studies employing emotion recognition tasks, which consistently suggest individuals with ADHD present with difficulties to recognise emotions from faces (Bora & Pantelis, 2016; Borhani & Nekati, 2018; Collin et al., 2013; Cooper et al., 2020). Moreover, in a meta-analysis presented in Chapter Four, the present thesis provides robust evidence for the presence of vocal emotion recognition difficulties in children, adolescents and adults with ADHD (Sells et al., 2023). However, theoretical perspectives diverge in their explanations of the neurodevelopmental pathways which underlie facial and vocal emotion recognition difficulties in ADHD. Cognitive-behavioural theories suggest difficulties in recognising emotional expressions are secondary symptoms of ADHD, caused from atypicalities in higher-level cognitive processing pathways necessary to allocate attention towards emotional expressions (Barkley, 1997; Nigg et al., 2005). Alternatively, Perceptual Processing theories, including Socio-Cognitive (Crick & Dodge, 1994; Lemerise & Arsenio, 2000) and Motivational Dysfunction (Sonuga-Barke, 2005, 2012) models argue that emotion recognition difficulties are primary symptoms in individuals with ADHD. These models argue that emotion recognition difficulties are caused by specific atypicalities in the lower-level perceptual processing pathways during the encoding of emotional or motivational stimuli, which interact with and contribute to variations in the allocation of attention towards such stimuli (Sonuga-Barke, 2005).

The next section aims to review each theoretical perspective (Cognitive-Behavioural and Perceptual Processing) for the presence of ED in ADHD, with a focus on discussing evidence from the vocal emotion processing literature. As research is evolving towards a more dimensional approach to ADHD pathophysiology (Astle et al., 2022; Cuthbert & Insel, 2013; Sonuga-Barke et al., 2023), evidence from research which conceptualises ADHD as both a categorical and dimensional variable is discussed. First, research which supports the presence of atypicalities in the higher-level cognitive pathways during vocal emotion processing is discussed, in support of cognitive-behavioural theories (Barkley et al., 1997, 2015). This section highlights the limitations with the view that ED in ADHD is only derived from atypicalities in the higher-level cognitive pathways. Next (on page 29), as an alternative, perceptual processing models are discussed, alongside the behavioural and neuroimaging evidence which supports these theories. It is highlighted here that there is not sufficient evidence from the behavioural vocal emotion recognition literature to support these theories. On page 35, the Event-Related Potential (ERP) technique is introduced and the ERP literature, which supports cognitive-behavioural and perceptual processing theories of ADHD, is discussed in detail. This leads to outlining the gaps in the ERP evidence from the vocal emotion processing literature and the theoretical questions the present thesis aims to address to test hypotheses deriving from the existing theoretical models.

Theoretical Perspectives to Vocal Emotion Processing in ADHD

The Cognitive-Behavioural Model. The first model which attempted to explain the presence of ED in ADHD was Barkley's (1997) unifying cognitive-behavioural model. Barkley (1997) suggested that deficiencies in behavioural inhibition (i.e., the ability to inhibit responses in favour of a more goal-oriented response) was the essential impairment in children who display inattentive and hyperactive/impulsive traits. The model stated that inhibition is a prerequisite for other higher-level executive functions (EF), such as problem-solving, goal-directed thinking and self-regulation (e.g., of emotions; Barkley, 1997). Therefore, Barkley (1997) argued that ED is a secondary impairment of ADHD, resulting primarily from atypicalities in higher-level executive functions, initially stemming from deficiencies in inhibition. Specifically, the cognitive-behavioural model states that emotional and social difficulties in ADHD result from processes such as: a) lacking the cognitive control to inhibit emotional responses or anticipate emotional reactions to future events, resulting in greater reactivity in social situations, b) a decreased ability to pause and think about how emotional reactions may impact others in social situations, resulting in inappropriate social responses and, c) a reduced ability to induce an emotional or motivational state in ones-self in order to achieve a certain goal (Barkley, 1997). To support Barkley's (1997) theory, individuals with injuries to the prefrontal cortex, a brain region responsible for inhibition and higher-level cognitive abilities, exhibited ED impairments similar to those observed in individuals with high levels of hyperactivity/impulsivity and inattention (Rolls et al., 1994; Stuss et al., 1992). These findings are consistent with research suggesting that ADHD is a dysfunction of the fronto-striatal structures of the brain, associated with cognitive processing (Brennan & Arnsten, 2008; Bush et al., 2005).

In support for the cognitive-behavioural theory of ED in ADHD (Barkley, 1997), studies employing vocal emotion recognition tasks (as outlined in Section 1.1) have reported a generic pattern of errors (i.e., lower performance overall) in adults diagnosed with ADHD compared to healthy controls (Bisch et al., 2016; Zuberer et al., 2020). The finding that adults with ADHD made generic errors during vocal emotion recognition tasks suggests that errors occurred randomly, implying difficulties to sustain attention throughout the task (Bisch et al., 2016; Zuberer et al., 2020). This contrasts with emotion-specific errors, which would indicate difficulties with perceiving certain emotions. To investigate the pattern of vocal emotion recognition more comprehensively in ADHD, a recent meta-analysis (presented in Chapter Two of the present thesis) synthesised results from 21 vocal emotion recognition studies in children, adolescents, and adults, and similarly reported the presence of generic vocal emotion recognition difficulties in ADHD (Sells et al., 2023). Consequently, results from behavioural vocal emotion recognition tasks support the view that ED in ADHD may be a secondary problem, underpinned by core attention deficits (Barkley, 1997, 2015). This is supported by research indicating that facial emotion recognition errors are associated with attention and inhibition deficits in children with ADHD (Sinzig et al., 2008).

Although evidence from vocal emotion recognition tasks support the presence of generic vocal emotion recognition difficulties in ADHD (Sells et al., 2023), there are limitations to interpreting these difficulties as solely arising from attention deficits. Firstly, the pattern of errors reported from facial emotion recognition tasks differs compared to that reported within the vocal emotion recognition literature. Facial emotion recognition studies consistently show individuals with ADHD present with specific deficits to recognise negative facial emotions, such as anger and fear (Bora & Pantelis, 2016; Borhani & Nejati, 2018; Collin et al., 2013; Cooper et al., 2020). As a consistent pattern of error towards a specific emotion type cannot be due to inattention alone, findings from the facial emotion recognition literature challenge the view that ED in ADHD is only underpinned by attention/inhibition difficulties. Moreover, within the vocal emotion recognition literature, a similar emotion-specific pattern of error was not found in our recent meta-analysis due to insufficient and inconsistent emotion-specific data reported within individual studies (Sells et al., 2023). For example, many vocal emotion recognition studies only included generic accuracy scores or only consistently included happy and angry vocal emotion types, making it difficult to reliably synthesise patterns of emotion-specific recognition atypicalities in ADHD (Sells et al., 2023).

In further support of the argument that the neural underpinnings of ED in ADHD are not limited to atypicalities in higher-level executive functions, a longitudinal study which recruited a community sample of children found that deficits in executive function (e.g., inhibition, selective attention and working memory) at age 5 predicted inattention symptoms at age 6, but not hyperactive-impulsive symptoms (Brocki et al., 2010). Therefore, while cognitive-behavioural models might well explain ED as a secondary symptom in individuals with high levels of inattention, challenges arise when accounting for ED in individuals with high levels of hyperactivity/impulsivity, who do not always present with cognitive difficulties. This is reinforced by a study which found specific ED symptoms relating to the processing of negative emotions, such as anger reactivity, were associated with higher levels of hyperactive/impulsive traits, and not inattentive traits in community samples (Sjowall et al., 2017). Specific difficulties relating to the processing of negative expressions in individuals with hyperactive/impulsive traits gives rise to the question as to whether ED in ADHD may not be exclusively linked to deficiencies in the attentional pathways of the brain but additionally explained by atypicalities in other pathways of the brain, such as those that are implicated in socio-emotion processing, considered to be essential for the correct interpretation of emotions.

Perceptual Processing Models. Alternative models of explaining ED in ADHD, such as Social-Cognitive (Crick & Dodge, 1994; Lemerise & Arsenio, 2000) and Motivational Dysfunction (Sonuga-Barke, 2005) models, state that emotional and social difficulties in ADHD arise from atypicalities in the perceptual processing of emotional and motivational stimuli. Social-Cognitive theories state that inappropriate reactions in social situations may stem from difficulties encoding sensory cues (e.g., vocal emotions), as outlined in the first step of the Social Information-Processing Model (See Figure 1.1, section 1.1; Lemerise & Arsenio, 2000). For example, if an individual misinterprets a sad vocal emotional expression as anger, this may lead to inappropriate behaviours in social situations, such as aggressive or hostile behaviours (i.e., temper tantrums, anger reactivity; Bunford et al., 2015). Moreover, socio-cognitive theories additionally state that hyperactive/impulsive behaviours, such as being too noisy, talking too much or interrupting conversations, might stem from initially misinterpreting emotional cues in social situations (Lavigne-Cervan et al., 2022). For example, if a child misinterprets a teacher's frustrated tone of voice as enthusiasm, the child may understand this as a cue of encouragement to share more thoughts or ideas. As a result, the child may inappropriately talk too much and potentially disrupt the class, rather than modulating their behaviour appropriately in response to frustration.

Motivational dysfunction models of ADHD similarly suggest ED in ADHD may result from atypicalities within motivational or affective pathways of the brain when processing vocal emotion (Sonuga-Barke, 2005; Chronaki, Benikos et al., 2015). However, motivational theories specifically argue that atypicalities in regions of the brain involved in emotion and reward processing (e.g., amygdala, ventral striatum, thalamus, nucleus acumbens), independently contribute to the allocation of attention in individuals with ADHD (Sonuga-Barke, 2005). For example, the delay-aversion theory states that individuals with ADHD present with a hypersensitivity to negative affective states associated with waiting for a delayed reward which contributes to inattentive and hyperactive/impulsive behaviours (Sonuga-Barke, 2005). More specifically, the delay-aversion theory states that to avoid the negative affect associated with delay, individuals with ADHD are more likely to seek out immediate and smaller rewards within their environment, rather than wait for the delayed outcome (Sonuga-Barke, 2005). For example, an adolescent with ADHD may have difficulty completing schoolwork due to perceiving negative emotions, such as frustration or boredom, associated with waiting for a delayed outcome (e.g., meeting up with their friends). Thus, an adolescent with ADHD may be more likely to allocate attention instead to more immediate salient positive stimuli, such as messages on their phone (i.e., a smaller immediate reward), to avoid negative emotions associate with waiting, and this is argued to give rise to symptoms of inattention (e.g., being distracted). Similarly, a child in a restaurant waiting for a delayed reward (an ice cream) will experience delay as aversive which may give rise to hyperactivity (e.g. fidgeting in their seat). Behavioural research which supports the delay-aversion theory has consistently demonstrated that children with ADHD have difficulty waiting for larger delayed rewards compared to smaller immediate rewards (Marco et al., 2009; Schweitzer & Sulzer-Azaroff, 1995; Tripp & Alsop, 2001).

In further support of the delay-aversion hypothesis, neuro-imaging research has reported the presence of atypicalities in the amygdala, a region primarily thought to be involved in the perceptual processing of negative emotions, in children and adolescents with ADHD when processing the negative affect associated with delay (Van Dessel et al., 2020). For example, Van Dessel et al (2020) employed an Escape Delay Incentive (EDI) task to explore the neural processing of avoidable compared to unavoidable delay in 10-18-year-olds with ADHD and age-matched controls. Within the EDI task, participants were instructed to respond as quickly as possible to a target, which was previously signaled by one of three symbols indicating different outcomes. The outcomes consisted of 1) an ‘avoidable delay’ condition, in which a delay between the next trial could be escaped by responding faster to the target; 2) a ‘certain delay’ condition, in which a delay was imposed regardless of speed of response and 3) a ‘no delay’ condition in which there was not a delay, regardless of speed of response (Van Dessel et al., 2020). In ‘certain delay’ conditions, findings from structural MRI scans demonstrated that adolescents with ADHD presented with reductions in the volume of the amygdala, and other subcortical structures responsible for processing the salience of negative emotional stimuli (Van Dessel et al., 2020). This provides support for the argument that ADHD is associated with atypicalities in the regions of the brain which are involved in processing the motivational significance of cues, such as the negative affect associated with delay (Sonuga-Barke, 1994).

Importantly, both social-cognitive theories and motivational dysfunction models do not dismiss the role of cognitive mechanisms which may also contribute to the perception and/or allocation of attention towards emotional or motivationally significant cues (Lemerise & Arsenio, 2000; Nigg et al., 2004). However, unlike cognitive-behavioural theories (Barkley, 1997, 2015), it is acknowledged that atypicalities in both lower-level perceptual and higher-level cognitive pathways may be involved in difficulties recognising and orienting attention to emotional expressions in ADHD (Crick & Dodge, 1994; Lemerise & Arsenio, 2000; Nigg et al., 2004). However, notably, perceptual processing theories suggest that atypicalities in social perception processes independently contribute to the symptoms of ADHD (Sonuga-Barke, 2005; Nigg et al., 2004)

There is not sufficient behavioral evidence from the vocal emotion recognition literature for emotion-specific atypicalities in ADHD (see meta-analysis by Sells et al. 2023), in support of perceptual processing theories. However, neuro-imaging studies, such as those employing functional Magnetic Resonance Imaging (fMRI; Zuberer et al., 2022; Kochel et al., 2015) and Event-Related Potential (ERP; Chronaki et al., 2015) techniques, have provided evidence to support this theory. For example, Zuberer et al (2022) employed a vocal emotion recognition task to explore differences in the brain regions activated in adults with ADHD and controls. Findings showed that adults with ADHD presented with reduced activation in the posterior superior temporal gyrus (STG), and the thalamus, regions both thought to be involved in the early perceptual stages of vocal emotion processing (Schirmer & Kotz, 2006; Zuberer et al., 2022). Additionally, Kochel et al (2015) employed a near-infrared spectroscopy study to investigate the patterns of brain activation during a vocal emotion recognition task in 10-year-old boys with and without ADHD. Similarly to Zuberer et al (2022), findings demonstrated a reduced activation of the right STG during the processing of vocal anger. This supports the argument that there are atypicalities in the brain regions associated with the perceptual processing of vocal emotions, specifically anger in children with ADHD (Sonuga-Barke, 2005).

It is, however, difficult to conclude from fMRI research that atypicalities in the perceptual processing pathways independently contribute to emotion processing difficulties. For instance, fMRI research takes snapshots of the brain every few seconds (Matthews & Jezzard., 2004). As neuro-cognitive processes occur within the space of a second, including the early perceptual processing and later higher-order cognitive evaluation of a stimulus, it is possible that reduced activation of the amygdala, thalamus and STG are attributed to either or both perceptual and cognitive mechanisms (Pessoa, 2008; Zuberer et al., 2020). This argument is in accordance with interactive theories of amygdala functioning, such as the multiple-waves model, which suggest the amygdala's role is to integrate inputs from both lower-level perceptual and higher-level attentional signals (Pessoa, 2008; Schaefer & Gray, 2007). Therefore, it is difficult to disentangle from fMRI research as to whether reduced amygdala function in ADHD reflects dysfunctions in executive functions in support of cognitive-behavioural theories, or atypicalities in the perceptual processing of vocal emotions in support of socio-cognitive (Lemerise & Arsenio, 2000) and motivational dysfunction (Sonuga-Barke, 2005) models.

To test the hypothesis that atypicalities in perceptual pathways independently contribute to vocal emotion processing difficulties in ADHD, it is necessary to explore the time course of vocal emotion processing. Preattentive processing refers to the perceptual processes which occur prior to the attentional capture of irrelevant auditory stimuli (Jääskeläinen et al., 2004; Chronaki & Marsh, 2024), and thus are independent from attentional mechanisms. Electrophysiological techniques allow us to elucidate between preattentive perceptual and cognitive processes, as they measure changes in brain activity millisecond by millisecond in response to a stimulus or event (Luck, 2014). The next section aims to give a more detailed overview of electroencephalography (EEG), and how Event-Related Potentials (ERPs) can be extracted to isolate atypicalities relating to attentional and preattentive vocal emotion processing in ADHD.

Event-Related Potential (ERP) Evidence of Attention and Preattentive Vocal Emotion Processing Difficulties in ADHD.

Electroencephalography (EEG) uses electrodes placed on the scalp to measure electrical activity within the brain (Luck, 2014). The continuous EEG waveform represents hundreds of overlapping neuro-cognitive processes occurring in quick procession (Luck, 2014). However, certain signals within the EEG waveform reflect sensory, perceptual, or cognitive processes triggered by a particular stimulus or event (Luck, 2014). When signals after a specific stimulus or event are averaged over several trials and participants, an Event-Related Potential (ERP) is extracted. An ERP component is a temporal window of neural activity occurring in specific scalp regions, which represent perceptual or cognitive processes related to the event (Luck, 2014). Therefore, compared to fMRI techniques which extracts brain activity every few seconds (Matthews & Jezzard., 2004), the ERP technique has higher temporal resolution and can elucidate between early perceptual (possibly preattentive) and attentional processes, such as those which occur during vocal emotion processing.

Typically, early components occurring between 50 to 250ms are thought to reflect more preattentive processes (e.g., the P2 component; Schirmer & Kotz, 2006), and components occurring later between 300-600ms are thought to reflect later higher-level attentional processes (e.g., the P3 component; Polich, 2007). However, this is not always the case, as demonstrated by early components (at 100ms) to visual (i.e., the P1; Luck et al., 2000; Mangun et al., 1998; Mangun & Hillyard, 1991) or auditory (i.e., the N1; Woldorff et al., 1993; Alho et al., 1994) stimuli, which are enhanced by attentional mechanisms when stimuli are explicitly attended too compared to ignored (See page 37). In this respect, the most effective way to measure the preattentive processing of events (i.e., vocal emotions) using ERP techniques is to include tasks in which vocal emotions are ignored, rather than attended to. Therefore, to explore the preattentive processing of vocal emotions prior to the influence of attentional demands, tasks should measure the processing of task-irrelevant vocal emotions, where the participant is not instructed to attend to the emotion in the voice (Mauchand & Zhang, 2022). For example, participants might attend to another aspect of the vocal stimuli (i.e., the speaker's gender) or neutral targets when a task-irrelevant vocal cue is ignored.

In terms of testing perceptual processing theories of ADHD, as of yet, no research has isolated the ERP components reflecting the preattentive processing of vocal emotions in ADHD when they are ignored, rather than attended to. However, there has been a rich ERP literature which has identified atypicalities in the ERP components reflecting attentional processing in ADHD. Moreover, models of vocal emotion processing have identified ERP components thought to reflect early preattentive, and later cognitive stages of vocal emotion processing in healthy adults (Schirmer & Kotz, 2006). The following sections give an overview of the ERP components reflecting attentional, and vocal emotion, processing in healthy adults, adolescents and individuals with ADHD.

ERP Components of Attentional Processing. ERP components of attentional processing can include early components, reflecting a rapid allocation of attention (i.e., the P1) or later components, reflecting more controlled executive functioning (i.e. the P3).

The Visual P1 Component. In healthy adults, the P1 component emerges between 90-190ms in the parietal occipital electrodes in response to visual stimuli and is thought to be generated in the visual cortex (Hillyard et al., 1998; Vogel & Luck, 2000; Luck, 2014). The P1 has not been found to change significantly with age. For example, a similar time window (80-180ms) and topography for the P1 has been evidenced in children, 6-11 years (Perchet et al., 2007). However, there have been some reports that the P1 decreases in amplitude and latency from childhood (age 10) across adolescence until age 18 (Mahajan & McArthur, 2012). The P1 is primarily modulated by perceptual and physical features of a visual stimulus but has additionally been found to be enhanced by top-down goal-related attention (Hillyard et al., 1998; Vogel & Luck, 2000; Luck, 2005). For example, Hillyard and Colleagues (1998) demonstrated that when adult participants were asked to attend to a visual stimulus in the left hemifield, the P1 was larger compared to when participants were asked to ignore the stimulus (Hillyard et al., 1998). Moreover, the P1 has been found to be enhanced when a visual stimulus appears in an already attended location in healthy adults (Luck et al., 2000; Mangun et al., 1998; Mangun & Hillyard, 1991). This attentional effect on the P1 has been demonstrated in studies which employed spatial cueing paradigms in healthy adults (Mangun & Hillyard, 1991; Störmer et al., 2009) and children (Perchet et al., 2007). For example, in a spatial cueing paradigm (See Figure 1.2) when a visual cue (e.g., a rectangle on the left or right hemifield) correctly indicated the location of a visual target, the P1 amplitude is larger in trials correctly cued compared to trials incorrectly cued, or un-cued (Perchet et al., 2007). Chapter Four (page 129) details variations of spatial cueing paradigms which find similar attentional effects on the P1 in adults. In this respect, the P1 is thought to reflect top-down attentional mechanisms which amplify early sensory signals to a visual stimulus appearing in the attended spatial location (Bruin et al 1998, Harter et al 1982, Luck et al 1990). Thus, the P1 can be measured as an indicator of the rapid allocation of visual spatial attention, after the onset of a visual target in spatial cueing paradigms (Mangun & Hillyard, 1991).

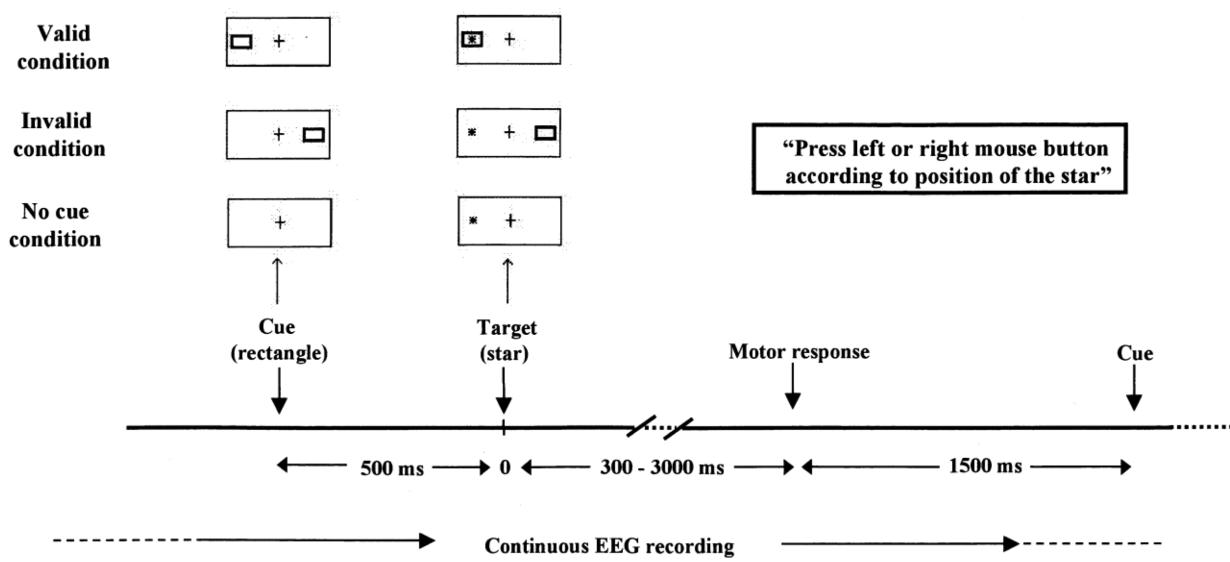

Figure 1.2

Graphical representation of a spatial cueing paradigm (the Posner paradigm) Designed and Implemented by Perchet et al (2007). The paradigm explores how spatial information from a cue affects visual spatial attention, via the P1 component. Reprinted from Biological Psychiatry, Vol 50 /Issue 1, Perchet and colleagues, "Attention shifts and anticipatory mechanisms in hyperactive children: an ERP study using the Posner paradigm", Pages 44-57, Copyright (2001), with permission from Elsevier.

The P3 Component. The P3 is a centro-parietal component observed between 300-600ms, associated with attentional processing to a task-relevant stimulus (Polich, 2007). In adolescence, the Target-P3 has been reported to have a similar midline centro-parietal topography, occurring with a latency slightly later (~25ms) than their adult counterparts within the same study (Broyd et al., 2012). The P3 represents evaluation, categorisation, and attentional resource allocation, observed to a task-relevant stimulus in target detection paradigms (Polich, 2007). For example, in spatial cueing paradigms, the P3 has been found to be larger in incongruent compared to congruent trials in 6-11-year-old children (Perchet et al., 2007). This is thought to reflect the recruitment of higher-level executive functioning processes, necessary to engage spatial attention towards the task-relevant stimulus, in line with task demands (Perchet et al., 2007).

The CNV Component. The Contingent Negative Variation (CNV) component is a slow negative brain potential which occurs in the frontocentral sites in the 100ms interval before a particular target event, which has been cued via the presentation of a prior cueing stimulus (Walter et al., 1964; Wild-Wall & Falkenstein, 2010). In adolescents, the CNV has been evidenced to emerge at a latency 50ms later than adults (Broyd et al., 2012). The CNV is thought to reflect the anticipation of a target stimulus, and the preparation of a motor response to that stimulus (Walter et al., 1964; Brunia et al., 2011). In addition, the CNV is thought to reflect motivation to respond to a particular stimulus (Irwin et al., 1966). For example, in healthy adults, the CNV amplitude has been found to increase when a target stimulus is associated with monetary incentives (Novak & Foti, 2015).

ERP Atypicalities in Attentional Processing in ADHD. ERP components reflecting attentional processing have consistently been reported to be reduced in individuals with ADHD compared to healthy controls. For example, the P1 component as indicator of visual spatial attention, has been found to be attenuated in children with ADHD, compared to controls when employing a spatial cueing paradigm (See Figure 1.2; Perchet et al., 2007). Similarly, attenuated P1 amplitudes have been reported in 7-10-year-old children (Nazari et al., 2010) and adults (Papp et al., 2020) with

ADHD in other studies employing paradigms which include a visual target preceded by cueing stimuli. This suggests children and adults with ADHD do not effectively use information from cueing stimuli to orientate their attention to a target (Perchet et al., 2007). Similarly, the P3b and CNV components have consistently been reported to be attenuated in ADHD in studies involving target stimulus evaluation or categorisation, such as go/no-go tasks (Barry et al., 2003; Rodriguez & Baylis., 2007; Szuromi et al., 2011; Johnstone, Barry & Clarke, 2013). Overall, an attenuated P3 amplitude to a task-relevant target stimulus is thought to reflect difficulties in attention allocation in children (Barry et al., 2003) and adolescents (Chronaki et al., 2017) with ADHD. An attenuated CNV in children and adolescents with ADHD is thought to reflect difficulties with anticipatory attention, necessary to prepare for a target stimulus, or response (Johnstone, Barry & Clarke, 2013). However, interestingly, the CNV has been reported to be larger in ADHD in conditions of avoidable delay and motivational significance in EDI tasks (Chronaki et al., 2019; Chronaki et al., 2017), similar to those employed by Van Dessel et al (2019; See page 31). This provides partial evidence in support of motivational theories which suggest atypicalities in the processing of motivationally significant stimuli interact with the allocation of attention in ADHD (Sonuga-Barke, 2005). However, robust evidence for atypicalities in the preattentive processing of motivationally or emotionally significant stimuli in ADHD is lacking. This leads onto a discussion of the ERP components involved in vocal emotion processing, both prior (i.e., preattentive) and after the attentional capture of the vocal emotion, in healthy individuals.

ERP Components of Vocal Emotion Processing. In healthy adults and children, vocal emotional expressions have been found to modulate ERPs in both early and later processing stages (Schirmer & Kotz, 2006; Chronaki et al., 2012). Schirmer and Kotz (2006) present a model of vocal emotion processing in healthy adults (See Figure 1.3), in which the emotional significance of a voice is processed preattentively in the early perceptual regions (i.e., the STS) via amygdala activation prior to capturing attention. The N1 and P2 components are thought to reflect these early perceptual processes. Subsequently, vocal emotions are processed via higher-level cognitive regions (i.e., the inferior frontal gyrus, and orbitofrontal cortex), during the evaluation and interpretation of vocal cues (Grandjean, 2020; Schirmer & Kotz, 2006), thought to be reflected by later P3 and N400 components. These early and late ERP components are presented in more detail below, followed by discussion on atypicalities in these components in individuals with ADHD.

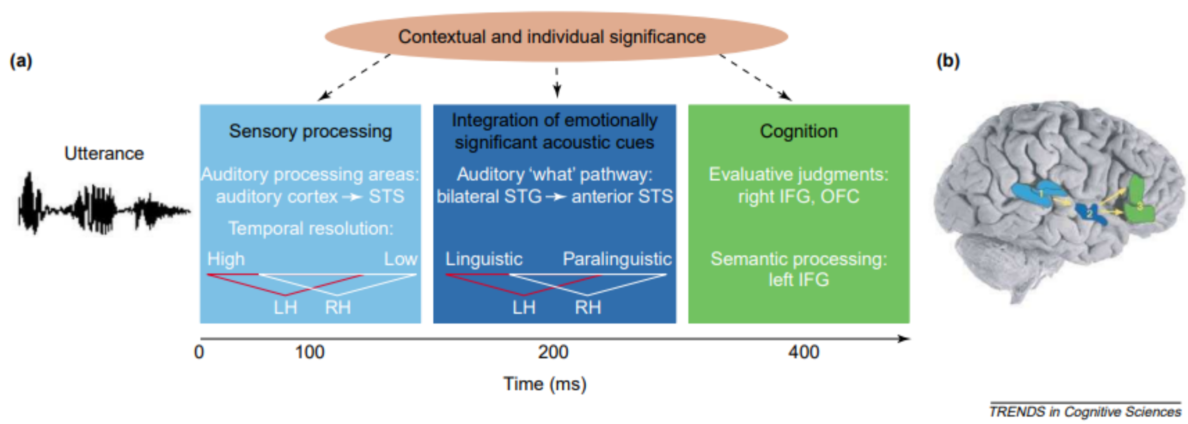

Figure 1.3

The Model Of Vocal Emotion Processing Proposed by Schirmer and Kotz (2006). a) shows a representation of the three stages of processing (sensory processing, integration of emotionally significant acoustic cues, and cognition) (b) shows the areas of the brain implicated in each stage of processing including the auditory cortex in light blue, the STS in dark blue and the IFG and orbitofrontal cortex in green. Reprinted from Trends in Cognitive Sciences, Vol 10 / Issue 1, Schirmer & Kotz, “Beyond the right hemisphere: brain mechanisms mediating vocal emotion processing, Pages 24-30., Copyright (2006), with permission from Elsevier.

The N1 Component. The first sensory stage of vocal emotion processing (See Figure 1.3; Schirmer & Kotz, 2006) is reflected by the N1 component occurring between 90-110ms in fronto-central scalp regions in healthy adults (Alho et al., 1994; Näätänen & Picton, 1987), and parietal-occipital regions between 80 and 150ms in children (Pang & Taylor, 2000; Chronaki et al., 2012). The N1 first emerges in the frontal central sites from around 9 years, increasing in size throughout adolescence until stabalising in adulthood (Mahajan & McArthur, 2012). It is thought the N1 reflects the processing of low-level acoustic features of stimuli within the bilateral auditory cortex. Moreover, the early N1 to auditory stimuli has been evidence to be enhanced by early selective attentional processes relating to task demands (Woldorff et al., 1993; Alho et al., 1994). When low-level acoustic properties of voices are standardized, the N1 to voices is not modulated by emotions in adults (Jessen & Kotz, 2011; Gädeke et al., 2013) or children (Chronaki et al., 2012), suggesting that the preattentive processing of emotional significance in the voice occurs after 100ms in typical adults and children, in support of the model by Schirmer & Kotz (2006).

The P2 Component. The second stage of the model is reflected by the P2 component, observed between 120-260ms within the frontocentral electrode sites in adults and adolescents (Mahajan & McArthur, 2012); a shift from appearing in the parietal sites in children (Oades et al., 1997; Ponton et al., 2000). The P2 is thought to reflect the preattentive integration of emotionally significant acoustic cues, via the activation of the amygdala feeding into the Superior Temporal Cortex (STC; including the Bilateral STG and Anterior Superior Temporal Sulcus; Schirmer & Kotz., 2007). In support, the P2 has been found to be heightened for emotional (i.e., angry and happy) compared to neutral voices in vocal emotion recognition tasks (Gädeke et al., 2013; Iredale et al., 2013; Jessen & Kotz, 2011). Importantly, the P2 has also been found to be modulated by emotion when vocal cues are presented as irrelevant to task goals in adults (Deng et al., 2022; Gädeke et al., 2013; Kao & Zhang, 2023; Lin et al., 2022). For example, Gadeke et al (2013) instructed adults to detect deviant stimuli depending on the syllables in pseudowords, and not the prosody, and reported enhanced frontocentral P2 amplitudes to irrelevant happy and threatening voices compared to neutral voices. Similarly, enhanced P2 amplitudes have also been found for happy compared to neutral stimuli in both explicit (emotion identification) and implicit (gender identification) tasks (Lin et al., 2022). These findings support the idea that the P2 component reflects the preattentive processing of vocal emotions within the second stage of Schirmer & Kotz (2006) model, independent from attentional mechanisms. However, to date, no ERP studies have isolated the P2 as a marker of preattentive vocal emotion processing in adolescents in similar preattentive tasks with irrelevant vocal emotions.

The P3 and N400 Components. Later ERP components such as the auditory P3 observed in parietal sites between 260-380ms in children (Chronaki et al., 2015) and the N400 observed in frontocentral and parietal scalp regions between 300-550ms in adults (Kutas & Hillyard, 1980), are thought to reflect the cognitive processes, such as attentional engagement, and the evaluation of emotional voices and integration with context (Schirmer & Kotz, 2006; Bostanov and Kotchoubey, 2004; Paulmann and Kotz, 2008). For example, the effect of emotional context on the N400 has been demonstrated in a study in healthy adults by Schirmer et al (2005) who presented sentences such as 'she had her exam' in either happy or sad voices, which were followed by either a congruous or incongruous emotional word (i.e., success or failure). Participants were asked whether the word matched the emotional prosody. Schirmer and colleagues (2005) reported an attenuated N400 in congruent compared to incongruent conditions, suggesting the N400 is modulated by the context of the vocal emotion. Moreover, the N400 has been found to be attenuated to negative (e.g. angry) vocal emotions in 6-11-year-old children (Chronaki et al., 2012) and adults (De Pascalis et al., 2009; Gootjes et al., 2011; Kanske & Kotz, 2007; Stewart et al., 2010). This finding has been interpreted as reflecting more controlled cognitive processing when facilitating attention towards negative stimuli (Kanske & Kotz, 2007; Stewart et al., 2010).

ERP Atypicalities in Vocal Emotion Processing in ADHD. Research in the ERPs implicated in emotion processing in ADHD is more limited compared to ERP research implicated in attention processing in ADHD. The first study to examine the electrophysiological correlates of vocal emotion in ADHD was that by Chronaki et al (2015) who employed an explicit vocal emotion recognition task in children aged 6-11 years old with and without ADHD and reported an enhanced N1 amplitude to vocal emotions (anger and happiness) in ADHD. It was hypothesised that this pattern reflects an initial, and possibly preattentive, hyper orientation to vocal emotions. A heightened N1 amplitude to vocal anger reflecting an initial hyper-orientation to threat is based on research which suggests the N1 reflects a rapid "early selection" mechanism of auditory attention (Woldorff et al. 1993). However, since participants attended to vocal emotions in an explicit vocal emotion recognition task within the study by Chronaki et al (2015), it is not yet understood whether a heightened N1 to vocal anger in ADHD may reflect a preattentive sensitivity to vocal anger, prior to the influence of attention, in support of perceptual processing theories. Additionally, the study by Chronaki et al (2015) reported an attenuated P3 amplitude in response to vocal anger in children with ADHD compared to controls. In support, a study employing a facial emotion recognition task reported reduced P3 amplitudes to anger and fearful expressions in 8-17-year-old adolescents with ADHD (Williams et al., 2008). These findings on the P3 suggests individuals with ADHD present with less controlled cognitive processing of vocal emotions in later stages of processing in support of the presence of generic attentional deficiencies during vocal emotion processing in ADHD (Barkley, 1997, 2015).

No ERP research to date has explored the preattentive processing of vocal emotions in ADHD, independently from early selective attention mechanisms. This is important because exploring the preattentive processing of vocal emotions can help us understand more conclusively if there are atypicalities during the perceptual stage of vocal emotion processing in ADHD, prior to atypicalities at the later cognitive stages, and independently from the attentional demands of the task. Moreover, measuring preattentive, rather than the explicit processing of vocal emotions is necessary, as in day-to-day life, individuals often automatically perceive and allocate attention towards emotions from voices without any instruction to do so.

Vocal Emotion Processing and ADHD: Summary and Gaps in Literature

In summary, ED in ADHD is argued to be associated with difficulties recognising and/or allocating attention towards emotional stimuli, such as vocal and facial expressions of emotion (Shaw et al., 2014). There is robust evidence that individuals with ADHD present with difficulties recognising vocal emotions (Sells et al., 2023). However, it is debated whether these difficulties are caused primarily from general executive function deficiencies affecting attention during explicit VER tasks (Barkley, 1997, 2015), and/or atypicalities within the perceptual processing pathways necessary to encode emotions (Crick & Dodge, 1997; Lemerise & Arsenio, 2000; Sonuga-Barke, 2005). Based on the review of the literature so far, there is robust behavioural and ERP evidence in support of the presence of attentional deficits during vocal emotion processing in ADHD. For example, behavioural studies have shown individuals with ADHD present with generic vocal emotion recognition difficulties (Sells et al., 2023). Moreover, ERP studies support the presence of executive function deficits (e.g. disrupted attentional processing) in ADHD with non-emotional stimuli, as reflected by reduced P1, P3 and CNV amplitudes. In support of perceptual processing theories, ERP research has brought new evidence that very early stages of emotion processing (e.g., N1) may be implicated in ADHD (Chronaki et al., 2015). This ERP finding raised the hypothesis that the N100 component may be a marker for automatic hyper-orientation to vocal threat stimuli in ADHD (Chronaki et al., 2015). This is consistent with ERP studies showing that motivationally significant stimuli (e.g., rewards, delay) can enhance ERP components (i.e., CNV, P3) reflecting higher-level attentional processes, such as response preparation in adolescents with ADHD (Chronaki et al., 2017; 2019). However, it is not known whether a) a preattentive hyper-sensitivity to vocal emotions (i.e., anger or happiness) is present in ADHD and b) whether the preattentive processing of vocal emotions independently influences the allocation of attention in ADHD.

The present thesis aimed to address these gaps in the literature. Importantly, because explicit vocal emotion recognition tasks do not allow us to explore the preattentive processes, the present thesis aimed to design and optimise a novel emotion processing task, as a more reliable measure to help disentangle preattentive from attentive processes. Moreover, in line with the view that ADHD can be conceptualised as a dimensional condition (Section 1.3), the present thesis aimed to recruit a community sample of adolescents to explore associations between traits of inattention and hyperactivity and the pre-attentive processing of vocal emotions. The next section presents the novel paradigm, designed and employed in the present thesis, and discusses its theoretical basis.

1.4 Proposed Research

Considering the gaps in the literature highlighted in the previous section, a novel paradigm has been developed to directly address the following theoretical questions:

a) Do adolescents with high levels of inattentive and hyperactive traits present with a preattentive hypersensitivity to vocal emotions?

b) Does the preattentive processing of vocal emotions affect the allocation of attention in adolescents with high levels of inattentive and hyperactive traits?

The emotional spatial cueing paradigm (thereafter ESC) is a novel task designed and optimised to be used with behavioural and ERP techniques to specifically measure ERP components (e.g., P2) to a vocal cue stimulus, as a measure of the preattentive processing of vocal emotions, and the ERP components (e.g., P1) to a visual target stimulus, as a measure of the effect of vocal emotions on the rapid capture of visual spatial attention (thereafter VSA). Reaction times (thereafter RTs) and the P3 component to the target can also be measured to further explore the effects of vocal emotions on the allocation of attention within the ESC paradigm

Figure 1.4 presents a graphical representation of one trial of the ESC paradigm. The basic procedure of one trial is as follows: Participants are presented with 1) a fixation cross which stays on the computer screen for the remainder of the trial, 2) an irrelevant vocal cue stimulus (e.g., an angry voice) for 700ms in either the left or right ear 3) a visual target (i.e., a yellow star) for 100ms in the left or right visual field. Participants respond to indicate whether the visual target appeared on the left or right side of the screen via a button press. This response signifies the end of the trial, after which there is an 1000ms interstimulus interval before the next trial begins. Trials are either ‘congruent’ when the vocal cue and visual target are presented in the same spatial location (e.g., the left side of space) or ‘incongruent’ when the vocal cue and visual target are presented in opposite spatial locations (e.g., the vocal cue is presented in the right ear and the visual target in the left visual field). Participants are instructed to fixate their eyes on the cross, ignore the voices, and respond as fast as they can. The following sections provides a summary of each phase within one trial of the paradigm, with reference to what each phase is expected to do psychologically in healthy adults and adolescents and outlines the predicted atypicalities in each phase in adolescents with ADHD traits.

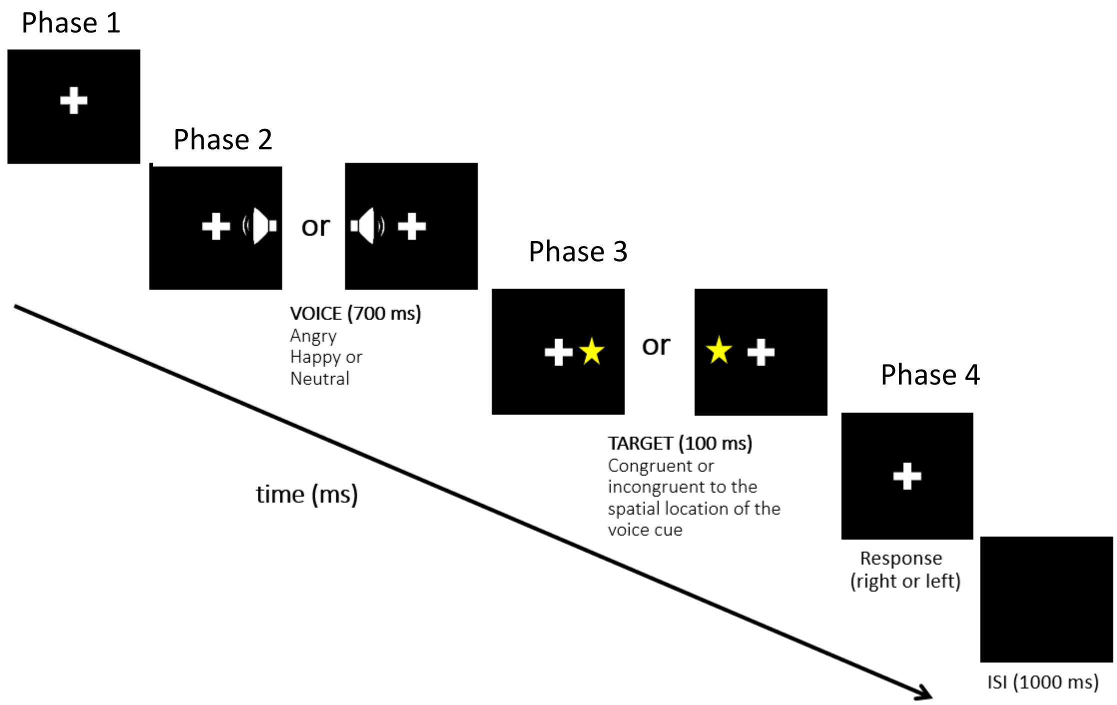

Figure 1.4

A Graphical Representation of One Trial from the Novel Emotional Spatial Cueing Paradigm. The diagram shows four phases: 1) Fixation cross, 2) Vocal Cue, 3) Visual Target and 4) Button Press response.

The Theoretical Basis of Each Phase of The Emotional Spatial Cueing Paradigm (ESC)

First phase: Fixation Cross. The presentation of the fixation cross signifies to participants that a new trial has started and ensures participants eye movements are limited.

Second phase: Vocal Cue. The second phase of the ESC paradigm presents an irrelevant vocal cue stimulus in either the left or right ear, which participants are asked to ignore. Previous theory and literature suggest that even when participants are asked to ignore a vocal emotion, the emotional salience of the irrelevant voice is still be processed preattentively within the corresponding contralateral (i.e., left or right; Schonwiesner et al., 2007; Woldorff et al., 1999) auditory cortex (STG; Schirmer and Kotz, 2006). To support this, the P2 component, thought to reflect the processing of emotional salience (Schirmer & Kotz, 2006), has been found to be enhanced to task-irrelevant vocal emotions (angry and happy) compared to neutral voices in adults (Deng et al., 2022; Gädeke et al., 2013; Kao & Zhang, 2023; Lin et al., 2022). Therefore, within the present ESC paradigm, it was expected that first, the P2 would be enhanced to angry and happy voices compared to neutral voices in healthy adults and adolescents, reflecting the preattentive processing of vocals emotions. Second, if individuals with ADHD present with a hypersensitivity to vocal emotions, as suggested in previous research with children (Chronaki et al., 2015), it was predicted that the difference in the P2 amplitude to vocal emotions (angry and happy) compared to neutral voices would be significantly larger in adolescents with higher levels of ADHD traits.

Third phase: Visual Target. During the next phase of the ESC paradigm, a visual target (i.e., a yellow star) is presented in either the left or right visual field, 750ms after the onset of the spatially presented irrelevant vocal cueing stimuli. When a spatially presented visual stimulus is presented to participants, perceptual features, such as the colour and size of the stimulus are initially processed within the contralateral visual cortex to the hemifield the stimulus is presented in (Hillyard, Vogel & Luck., 1998). In ERP studies, this early visual processing in the visual cortex, produces a P1 component in the parietal occipital sites (Hillyard, Vogel, & Luck, 1998; Vogel & Luck, 2000). Previous research and theory suggest that spatially presented auditory cues capture attention cross-modally due to acting as an 'early warning system', and therefore similarly activate the contralateral visual cortex, even when participants are asked to ignore the auditory cues (Brosch et al., 2009; Spence & Driver, 1997; Dalton & Lavie, 2004; Murphy et al., 2017). Moreover, the P1 to a visual target, has been evidenced to be larger if VSA is already in the spatial location of the target (Luck et al., 2000; Mangun et al., 1998; Mangun & Hillyard, 1991). Therefore, within the present ESC paradigm, if irrelevant vocal cues capture VSA prior to the onset of the visual target, the P1 component is expected to be larger to the visual target in congruent compared to incongruent trials in healthy adults and adolescents. A larger Target-P1 in congruent trials would reflect the summation of activity in the contralateral visual cortex towards the congruently spatially located vocal cue and visual target, as reported in previous spatial cueing paradigms with irrelevant auditory cues (McDonald et al., 2013; Feng et al., 2014). In adolescents with higher levels of ADHD traits, it is expected that the difference in the P1 component between congruent and incongruent trials might be smaller, at least to targets preceded by neutral voices, which would reflect generic attention atypicalities, consistent with that previously reported in a spatial cueing paradigm with neutral visual stimuli (Perchet et al., 2007).

However, within this third phase of the present ESC paradigm, there is an added effect of emotion from the previous vocal cue which may manipulate the capture of VSA, and therefore the P1 response to the visual target. Theory suggests irrelevant threatening stimuli (i.e., anger) more rapidly captures VSA than neutral stimuli to enable a quick response to aid survival (Hansen & Hansen, 1988; Ohman & Mineka, 2001). The rapid capture of VSA by vocal anger, as reflected by the P1 to a visual target, has been demonstrated previously in adults in a study employing dot-probe paradigm which presented angry and neutral vocal cues simultaneously in the left and right ear prior to a visual target presented in either left or right hemifield (Brosch et al., 2009). Brosch et al (2009) reported an enhanced P1 amplitude to visual targets preceded by a spatially congruent angry vocal cue compared to a spatially congruent neutral vocal cue. Therefore, within the present ESC paradigm, in healthy adults and adolescents, it was expected that there would be a larger P1 response to visual targets preceded by spatially congruent angry voices compared to spatially congruent neutral voices. Additionally, if individuals with ADHD present with a rapid and automatic hyper-orientation to vocal anger, as hypothesised by Chronaki et al (2015), it was predicted that the difference in the P1 amplitude to visual targets preceded by spatially congruent angry vocal cues compared neutral voices would be enhanced in adolescents with higher compared to lower levels of inattentive and hyperactive traits.

The P3 component to the visual target in the third phase of the ESC paradigm can additionally be measured as reflecting later attentional processes within the ESC paradigm (See page 38; Polich, 2007). For example, a larger P3 to incongruent compared to congruent trials in spatial cueing paradigms is thought to reflect more effortful processing to engage on the visual target (Perchet et al., 2007). Similarly, a larger Target-P3 component to trials preceded by emotional voices compared to neutral voices thought is also thought to reflect more effortful processing to engage to the visual target (Wang et al., 2022). Therefore, it was expected that if vocal emotions involuntarily captured attention in the ESC paradigm, this may have a spillover effect later attentional processes, capture by the P3 component. For instance, if vocal emotions captured attention this may result in the greater use of cognitive resources to engage on the target downstream, and thus this would be reflected in a larger P3 component to incongruent compared to congruent trials, and a larger P3 component to trials preceded by emotional compared to neutral voices. If adolescents with ADHD displayed generic attention deficiencies, it was expected that the P3 would be smaller, at least to trials preceded by neutral voices, consistent with previous research on the P3 component in ADHD (Barry et al., 2003). However, if emotions within the voice interact with later attention allocation in ADHD as previously reported (Chronaki et al., 2019; 2017), the difference in the P3 amplitude between trials preceded by neutral compared to emotional voices may be larger in adolescents with higher compared to lower traits of ADHD.

Fourth Phase: Button Press Response. Finally, within the ESC paradigm, participants are instructed to make a button-press response to indicate the location of the visual target as fast as possible. Previous research employing spatial cueing paradigms with irrelevant auditory cues have demonstrated that adults respond faster to congruent compared to incongruent trials, thought to reflect the capture of VSA by spatially congruent auditory cues on a behavioural level (Spence & Driver,1997; Mazza et al., 2007; Spence & McDonald, 2004). Moreover, studies employing dot-probe tasks have demonstrated that vocal anger has a facilitating effect on RTs due to more rapidly capturing VSA (Brosch et al., 2009). There are, however, additional, attentional processes which may influence the speed with which participants respond to the visual target, such as EF processes when categorising the visual stimulus, and sustained attention to the task. Therefore, reaction times are not as accurate at measuring the capture of VSA as the P1 component, and within the ESC, RTs may also reflect the effects of vocal emotions on later attentional mechanisms (as discussed in relation to the P3 component above). For example, there has been some research demonstrating that irrelevant vocal anger has a distracting effect on later attentional mechanisms (Kattner & Ellermeier, 2018; Sander et al., 2005), which result in slower RTs (Chronaki & Marsh, 2024). Therefore, within the present ESC paradigm, it was first expected that response times would be faster to congruent compared to incongruent trials in healthy adults and adolescents, consistent with previous spatial cueing paradigms (Spence & Driver, 1997). However, it was possible that the effect of emotion within the vocal cues could have either a facilitating or distracting effect on RTs. Our predictions relating to atypicalities in RTs within the ESC paradigm in adolescents with higher traits of ADHD therefore depended on what was first found in healthy adults and adolescents. However, if adolescents with higher levels of ADHD displayed generic attention deficiencies, it was expected that overall, they would display slower RTs to the visual target.

Overview of Theses Chapters.

Overall, to address the aims of the thesis presented on page 48, Chapter Two presents our meta-analysis which systematically reviewed the current literature on explicit vocal emotion recognition in ADHD. This chapter was vital to inform the gaps in the vocal emotion processing literature in ADHD. Next, Chapter Three presents an experimental study which validated the vocal stimuli employed in the ESC paradigm. Chapter Four presents a study containing four experiments which optimised and validated the ESC paradigm on a behavioural and neural level in adults. Results from Chapter Four relating to the effects of emotion and congruency on RTs and ERPs in the ESC paradigm in healthy adults, informed our predictions and analyses presented in Chapter Five. Finally, Chapter Five presents the experimental work which investigated a) the preattentive processing of vocal emotions in adolescents with ADHD (via the P2 component) and b) effects of vocal emotion on VSA in adolescents with traits of ADHD (via RTs and the P1 components) using the ESC paradigm with behavioural and electrophysiological methods.

2 VOCAL EMOTION RECOGNITION IN ADHD: A META-ANALYSIS

2.1 Chapter Overview

As presented in Chapter One (section 1.3), there is debate in the literature as to whether emotion dysregulation (ED) in Attention-Deficit Hyperactivity Disorder (ADHD) is reflective of deviant attention mechanisms, in support of cognitive-behavioural theories (Barkley, 1997; Nigg et al., 2005) or atypical perceptual processing, in support of social-cognitive and motivational theories (Shaw et al., 2014; Sonuga-Barke, 2005). To add to this debate, existing research has employed explicit facial and vocal emotion recognition tasks to explore the presence and nature of emotion processing difficulties, which are linked to ED (Uekermann et al., 2010). Previous reviews have reliably quantified the presence and nature of facial, but not vocal, emotion recognition atypicalities in ADHD, suggesting that FER atypicalities are specific to negative emotions, such as anger and fear (Bora & Pantelis, 2016). However, less is known about the presence and nature of vocal emotion recognition (VER) difficulties in ADHD. This chapter presents the first study, which aimed to systematically review the behavioural literature up to year 2022 on the presence of VER atypicalities in ADHD and explore whether these are generic or emotion-specific in nature. Although we take a dimensional approach to ADHD within the later experimental chapters of this thesis, in line with recent perspectives (Sonuga-Barke, 2023), for the purpose of the present meta-analysis, ADHD was defined as a categorical variable, in accordance with the Diagnostic Statistics Manual (DSM-5; APA, 2013) or similar diagnostic criteria, such as the Diagnostic Interview Schedule for Children (DISC-IV; (Shaffer et al., 2000). ADHD was defined as a categorical variable in this meta-analysis because the majority of previous studies reviewed defined ADHD as a categorical variable, however we also included studies community samples.

Please note, this Chapter is based on the publication titled 'Vocal Emotion Recognition in ADHD: A Meta-Analysis' (Sells, Liversedge & Chronaki., 2023), published in the Journal of Cognition and Emotion© copyright # [2023]. This work is the author's own apart from the contribution from her supervisors. A substantial portion of this published paper is presented verbatim within this Chapter. Permission to include this piece of work in the present thesis has been granted by the publishers: Informa UK Limited, trading as Taylor & Taylor & Francis Group, http://www.tandfonline.com.

2.2 Introduction

When high levels of inattention and hyperactivity impair functioning at home, school or work or in social situations, individuals can be diagnosed with ADHD (APA, 2013). Although not included in the diagnostic criteria, emotion dysregulation (ED) is often present within those diagnosed with ADHD and is proposed as a clinical feature of the disorder (Shaw et al., 2014). As outlined in Chapter One, there are differing theoretical perspectives on the causes and underpinning cognitive processes of ED in ADHD (Shaw et al., 2014). Cognitive-behavioural models propose a generic deficit in the allocation of attention towards emotional expression, such as the face and voice (Barkley, 1997; Nigg et al., 2005). Alternatively, socio-cognitive (Crick & Dodge, 1994), and motivational (Sonuga-Barke, 2005) theories of ADHD state ED may arise from specific deficits in interpreting emotions, such as those from the face and voice, at a perceptual level (Borhani & Nejati, 2018).

Previous research has compared emotion recognition accuracy, from the face and voice, in participants with ADHD and controls using explicit emotion recognition tasks. If participants with ADHD perform significantly worse on emotion recognition tasks, compared to TD participants, it can be concluded that individuals with ADHD present with emotion recognition deficits. In addition, it is important to explore the nature of emotion recognition deficits by analysing the pattern of errors participants display during emotion recognition tasks. A generic deficit is defined as a reduced ability to recognise all types of emotions in a random pattern, thought to arise from a general deviancy to allocate attention to emotional expressions, in support of cognitive-behavioural theories (Barkley, 1997). An emotion-specific deficit is defined as a consistent pattern of error towards one or several types of emotions. For example, an anger-specific deficit would arise if participants with ADHD consistently failed to recognise angry emotional expressions more than other emotion types (i.e., happy) included in the task. Emotion-specific deficits may be indicative of an atypicality at the perceptual processing stage of emotion recognition, in support of socio-cognitive (Crick & Dodge, 1994) and motivational accounts (Sonuga-Barke, 2005), as a consistent pattern of error towards a specific emotion cannot be due to inattention alone.

To date, reviews indicate robust support for the presence of FER deficits in ADHD and suggest an emotion-specific deficit for negative facial expression, such as anger and fear, in ADHD (Borhani & Nejati, 2018; Collin et al., 2013; Cooper et al., 2020). However, thus far no reviews have reliably synthesised an effect for the presence of VER deficits in ADHD or explored the nature of VER deficits to the same extent as in FER research. It is important to explore the nature of VER accuracy in ADHD to the same rigour as FER accuracy because understanding vocal expressions requires interpreting how the intensity, pitch and loudness of a voice change over time, whereas facial stimuli do not have to be dynamic to be interpreted correctly (Schirmer & Adolphs, 2017). Thus, facial and vocal emotions activate different areas of the brain (Schirmer & Adolphs, 2017), and need to be examined separately.

There are mixed findings within the behavioural literature relating to the nature of VER accuracy in ADHD. Studies employing VER tasks, such as the Diagnostic Analysis of Nonverbal Accuracy (DANVA; Nowicki & Duke, 1994) reported significantly more generic errors across happy, sad, angry, fearful and neutral vocal expressions in children with ADHD compared to typically developing children (Cadesky et al., 2000; Chronaki, Garner, et al., 2015; Corbett & Glidden, 2000). Similar findings have also been shown in adults (Bisch et al., 2016; Zuberer et al., 2020). In contrast, several VER studies have reported that children and adults with ADHD have a specific deficit to recognise negative vocal emotions, such as anger (Kis et al., 2017; Plećević et al., 2021), in line with neuro-imaging research suggesting children with ADHD may present with a hypervigilance to anger (Chronaki, Benikos, et al., 2015). Finally, there is some evidence that the nature of VER accuracy changes depending on the subtype of ADHD or comorbid conduct disorder. For example, ADHD-inattentive subtype is associated with more VER errors overall, whilst hyperactive/impulsive symptoms are associated with greater difficulties to recognise negative (e.g., sad) vocal emotions (Miller et al., 2011). Moreover, a bias to recognise vocal anger in individuals with ADHD has been associated with comorbid conduct disorder (Cadesky et al., 2000). Findings from a number of studies are limited (Egan et al., 1998; Hall et al., 1999; Manassis et al., 2007; Rapport et al., 2002; Waddington et al., 2018) as they did not report accuracy scores for specific emotions, so it is inconclusive whether the VER deficits reported in these articles are reflective of a generic or an emotion-specific deficit in ADHD.

The present meta-analysis aims to provide robust evidence for the presence of VER deficits in ADHD by gathering and synthesising effects from all up-to-date literature, both published and unpublished, which measures the difference in accuracy scores between participants with ADHD and TD controls when completing an explicit VER task. Our primary prediction is that VER deficits will be present in ADHD, in accordance with previous reviews demonstrating the presence of FER deficits in ADHD (Bora & Pantelis, 2016; Borhani & Nejati, 2018; Collin et al., 2013; Cooper et al., 2020). Secondly, we aimed to explore the nature of VER deficits in ADHD by quantifying the moderating effect of emotion type (e.g., happiness and anger) and valence (i.e., positive and negative) on VER accuracy in ADHD. There are two possible outcomes to this exploratory aim. If emotion does not have a moderating effect on accuracy, this would suggest individuals with ADHD make random errors when recognising vocal emotions, due to general attention deviancies. If, however, emotion type has a moderating effect on accuracy, this would suggest individuals with ADHD present with a consistent pattern of error towards a specific emotion type (e.g., anger), as found in FER research (Bora & Pantelis, 2016). This would reflect a distinct emotion processing atypicality in ADHD.

2.3 Method

Search Strategy

The databases utilised to conduct searches for relevant studies published until 14th April 2022 included PsychINFO and PsychArticles, Pubmed, ProQuest Central. To address publication bias, ProQuest Dissertations and Theses Global databases were searched and authors were contacted for unpublished work. The Boolean expression used for the literature search was:

(ADHD OR "attention-deficit hyperactivity disorder" OR "attention-deficit/ hyperactivity disorder" OR "externalising behaviour*" OR "externalizing behavior*" OR "externalising problems" OR "externalizing problems" OR hyperact* OR inattent* OR impuls* OR 'attention problem*' OR 'hyperkinetic disorder) AND (emotion* OR affect* OR ang* OR happ* OR joy OR sad* OR surprise* OR fear* OR disgust* OR fright*)) AND (prosod* OR voice* OR speech OR vocal* OR auditory OR paralanguage OR nonverbal) AND (recogni* OR percept* OR decod* OR identify* OR process* OR naming OR detect* OR match* OR interpret* OR understand*).

The asterisk on some terms allowed for different possible endings e.g., processing, identifying. We formulated our search expression based on terms included in the titles of a list of pre-collected studies exploring associations between vocal emotion recognition and ADHD. Most searches were limited to title and/or abstracts for most terms, apart from words similar to 'vocal', which were searched throughout the whole text. This was because some studies which measured emotion recognition in different modalities do not specify the vocal modality in the title or abstract. To identify additional studies, we searched the reference lists of all included studies, and we searched (via google scholar) for any papers we had missed which had cited one key and most cited paper (Cadesky et al., 2000). See Appendix A (page 252) for the full search strategy.

Inclusion and Exclusion Criteria

For a study to be included in the meta-analysis, it was required to meet a number of criteria. First, the study had to include a task which measured VER accuracy. We included tasks which presented vocally emotive stimuli to the participant and instructed the participant to explicitly identify the emotion in the voice. Tasks were labelled as explicit if the experimenter instructed the participant to recognise and label the emotion, and thus, the participant was aware they were recognising emotions. Participants could respond by pressing a corresponding button on a computer, verbally naming the emotion, or pointing to the corresponding emotion word. Generally, VER tasks (i.e., the DANVA) are valid to employ with both adults (Miller et al., 2011) and children (Chronaki, Garner, et al., 2015), so studies which employed various ages of participants, from childhood to adulthood, were included in this meta-analysis. The literature search yielded some studies with tasks instructing participants to match the vocal expression with a facial expression, or to detect whether the semantic content of the vocal expression matched the prosody (e.g., the MNTAP; (Greenbaum et al., 2009), as well as dichotic listening tasks (Manassis et al., 2000). Although face matching tasks and dichotic listening tasks included elements of recognising emotions from voices, these studies were excluded from the meta-analysis because they were not a direct measure of VER. One study was excluded because participants were selected due to having difficulties recognising emotions as reported by parents or a professional, therefore this measure was biased (Loytomaki et al., 2020). Second, the emotions presented in the task needed to include at least two basic emotions, portraying happiness, anger, sadness, fear, disgust, or surprise. Third, all studies had to include a measure of ADHD, either defined as a categorical or a continuous variable. We included studies which employed only community samples and studies which compared clinical cases of ADHD and controls. These latter studies needed to include a sample of clinically diagnosed individuals with ADHD who met the criteria of the Diagnostic and Statistical Manual (DSM) or Diagnostic Interview Schedule for Children (DISC). Studies which included ADHD as a comorbid diagnosis alongside one other disorder, apart from Conduct Disorder (CD) and Oppositional Defiant Disorder (ODD), were excluded. Studies in which ADHD participants were on medication for ADHD symptoms were not excluded due to the ambiguity of medication status in some studies. However, we coded for variations in medication status as an exploratory variable. Fourth, studies which defined ADHD as a categorical variable had to include a TD group of individuals with no diagnosis of ADHD or other psychiatric disorder. Finally, studies that defined ADHD as a continuous variable and recruited only community samples had to include a measure of ADHD-related behaviours (i.e., inattention, and/or hyperactivity/impulsivity) on a dimensional scale by including a questionnaire measure consistent with the diagnostic criteria of ADHD.

The literature search produced 1,626 published and unpublished research reports. After removing duplicates (n=275), titles and abstracts were assessed for inclusion. 1,296 were removed on the basis they did not include an emotion recognition task, or a measure of ADHD or ADHD-related behaviour. Two articles were removed as the full-text version could not be retrieved. Finally, based on full-text examination of 53 potentially relevant articles, 26 articles were excluded because they did not measure VER in ADHD and TD participants. Three articles were excluded because they included duplicate data. Finally, we contacted 17 authors regarding missing emotion-specific accuracy scores. Six of these authors provided us with the missing emotion-specific data and five responded but could not access the data due to it being deleted in line with ethics guidelines. A total of six authors did not respond despite being contacted multiple times. Two authors could not be contacted (e.g., due to no forwarding email). Due to the lack of relevant statistics provided to calculate effect sizes for either a generic or emotion-specific VER score, three papers were excluded (Bisch et al., 2016; Egan et al., 1998; Manassis et al., 2000). We were left with a final number of 21 studies which met our inclusion criteria. See Figure 2.1 for a flowchart of the literature search, which followed PRISMA guidelines for reporting systematic reviews and meta-analyses (Page et al., 2021).

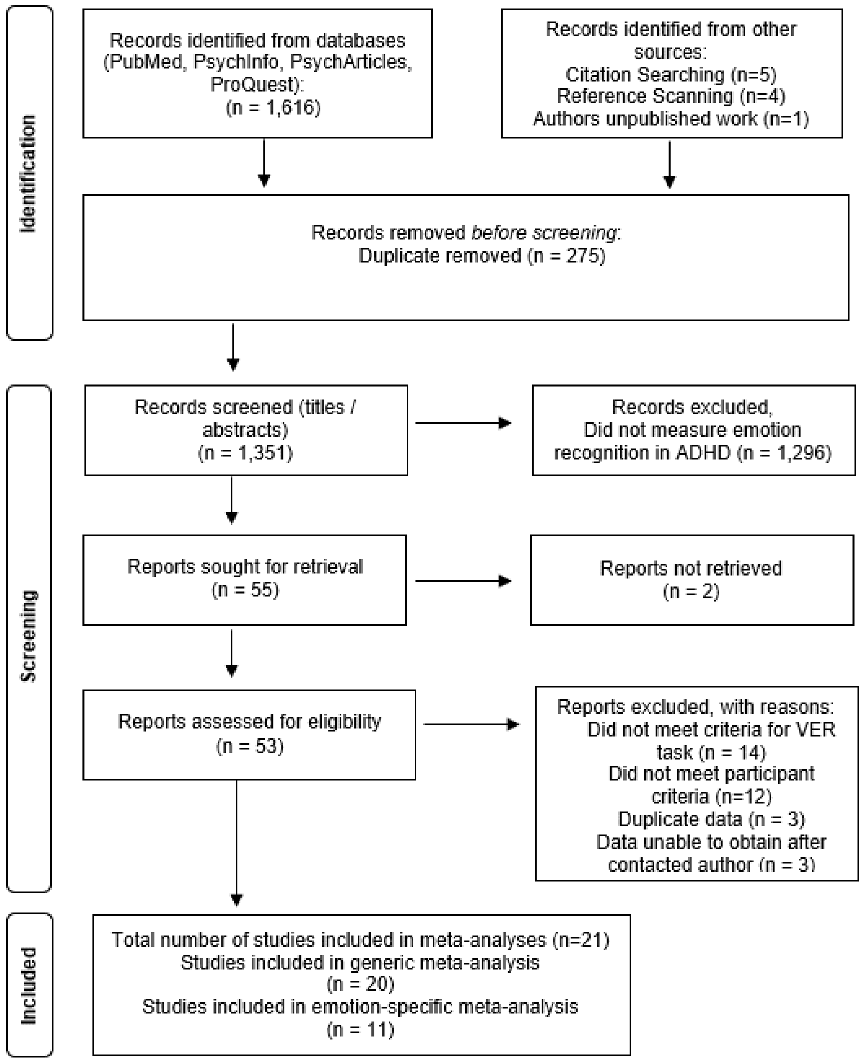

Figure 2.1

PRISMA Flowchart of Literature Search.

Study Coding

All studies were double coded by two independent raters. Any coding disagreements were resolved by discussion. To assess intercoder reliability, intraclass coefficient and kappa coefficient were calculated for continuous and categorical characteristics, respectively (Orwin & Vevea, 2009). The mean reliability coefficient was 0.95, and the mean percentage agreement was 95% on coded study characteristics.

Theoretical Moderator Variables. When articles reported a VER accuracy score for all included emotions, this was coded as an overall VER score. When articles reported several VER accuracy scores for separate emotions, these were coded as emotion-specific VER scores. Twenty studies provided relevant statistics for an overall VER score, and 11 for emotion-specific VER scores. Emotion type and valence of emotion-specific VER scores were coded to assess the effect of emotion as a moderator variable. All 11 studies included scores for happy and angry vocal emotions, eight for sad, six for fearful and six for neutral expressions. Only one study provided scores for disgust and erotic vocal emotions (Zuberer et al., 2020), so these emotion types were excluded from the moderator analyses. Surprise was not coded as it was not included in any papers. Regarding valence, anger, sadness, and fear were classified as negative emotions. Happiness was classified as the only positive emotion.

Exploratory Moderator Variables. Participant characteristics coded for included the sample size, mean age, age range, and the percentage of female participants in each sample. Within the included 21 studies, there was a total of 763 participants with ADHD (32% female) and 888 control participants (42% female). Ages ranged from 3 to 65. Samples with mean ages from 3 to 11 were coded as children, and 12 to 18 as adolescents. This was based on the definition of the start of adolescence being the onset of puberty, which on average starts around 12 years (Blakemore, 2008). Those aged 18 and upwards were classified as adults, as the ability to recognise vocal emotions reaches adult levels by 18 (Chronaki, Hadwin, et al., 2015; Grosbras et al., 2018). Based on this criterion, there were eight adult, nine child, and four adolescent samples included.

Within each study, it was noted whether ADHD was defined as a continuous or categorical variable. Within the 19 studies which defined ADHD as a categorical variable, participants were generally recruited due to being previously diagnosed. The diagnosis was confirmed with a mix of interview and questionnaire measures. The ADHD diagnosis was based on the DSM criteria (APA, 2013), apart from one study (Chronaki, Benikos, et al., 2015), which used the DISC-IV (Shaffer et al., 2000). Only two papers defined ADHD as a continuous variable. To measure ADHD traits, both continuous studies used the Strengths and Difficulties Questionnaire (Goodman, 1997), which maps onto the DSM criteria of ADHD (see Table 2.2 for all continuous measures of ADHD employed by each included study).

For the purpose of this meta-analysis, we defined ADHD as a categorical variable, and therefore we coded the subtypes of ADHD to fit this definition. The categorical subtypes of ADHD included a) ADHD-predominantly inattentive, b) ADHD-predominantly hyperactive/impulsive, and c) ADHD-combined subtype. If the outcome variable included a mix of different subtypes, we coded it as 'ADHD-Combined type' but noted the number of ADHD participants with each subtype included in the sample, if reported. If ADHD was defined as a continuous variable in the included paper, then we coded the subscale used to gain a score of ADHD symptoms (i.e., inattentive-subscale, or hyperactivity/impulsivity-subscale from the SDQ) under the respective subtype category outlined above. For example, if the inattentive subscale was used to quantify a score for ADHD symptoms, the outcome variable was coded as ‘ADHD-predominantly inattentive.’ If both inattentive and hyperactive/impulsive subscales were combined to give one overall score for ADHD symptoms, this was coded as ‘ADHD-combined subtype. Eighteen samples included a participant group of ‘ADHD-combined subtype’. Two studies included samples of ‘ADHD-predominantly hyperactive/impulsive’, and three studies included samples of ‘ADHD-predominantly inattentive’. We also coded the presence of comorbid Conduct Disorder and medication status within the ADHD participants. Three studies included ADHD participants with CD or ODD (Cadesky et al., 2000; Chronaki, Benikos, et al., 2015; Noordermeer et al., 2020). In 11 studies, participants with ADHD were on medication but withdrew from their medication between 4- and 48-hours prior to the study. In one study, all ADHD participants were non-medicated (Plećević et al., 2021), and in another, all ADHD participants were on medication (Norvilitis et al., 2000). In three studies, the medication status of ADHD participants was mixed, in that some participants were medicated, and some were not. In three studies, medication status was not reported.

Publication status and year of study were coded. Six studies were unpublished, and 15 were published between the years of 1999 and 2021. Design variables coded included the name of the VER task, the type of response participants made, the linguistic properties of the stimuli, and the characteristics (i.e., age and gender) of the actor who voiced the stimuli. See Table 2.1 and Table 2.2 for full details of participant, and design characteristics associated with all included studies.

Table 2.1

Participant Characteristics of Studies Included in the Meta-Analysis

Study	Age Group	ADHD Outcome Variable / Type	ADHD (N)	Age, mean years (SD)	Female (%)	IQ, mean	Medication Status	Presence of CD or ODD	TD (N)	Age, mean years (SD)	Female (%)	IQ, mean (SD)
Chronaki, Garner, et al. (2015)	Children	Hyperactive/impulsive (Continuous)	16	5.2 (0.9)	25	N/A	N/A	Mix	41	4.2 (0.7)	49	N/A
Chronaki, Benikos, et al. (2015)	Children	Combined (Categorical)	25	8.9 (1.5)	4	99.0 (10.8)	Medicated but withdrew (24hrs)	3 samples: with CD, excluded CD, mix	25	9.0 (1.5)	8	102.6 (10.2)
Sells et al. (Unpublished)	Adults	Inattentive, Hyperactive/Impulsive (Continuous)	0	N/A	N/A	N/A	N/A	N/A	26	26.7 (4.0)	65	N/A
	Adolescents		0	N/A	N/A	N/A	N/A	N/A	23	14.3 (1.4)	65	N/A
Kis et al. (2017)	Adults	Combined (Categorical)	28	33.8 (8.9)	36	109.4 (9.30)	N/A	CD excluded	29	36.5 (11.4)	52	110.2 (8.9)
Kochel et al. (2015)	Children	Combined (Categorical)	14	10.3 (17.4)	0	99.3 (11.9)	Medicated but withdrew (24 hrs)	N/A	14	10.2 (11.3)	0	107.3 (8.9)
Cadesky et al. (2000)	Children	Combined (Categorical)	86	9.0 (1.4)	17	97.3 (12.5)	Medicated but withdrew (24 hrs)	2 samples: with CD, excluded CD	27	9.3 (1.4)	40	108.4 (17.1)

Grabemann et al. (2013)	Adults	Combined (Categorical)	20	30.3 (9.4)	0	112.7 (11.2)	Medicated but withdrew (24 hrs)	N/A	20	27.9 (6.0)	0	112.7 (10.2)
Corbett and Glidden (2000)	Children	Combined (Categorical)	37	10.1 (1.78)	30	100 (11.1)	Medicated but withdrew (24 hrs)	N/A	37	9.5 (1.9)	49	113 (14.5)
Waddington et al. (2018)	Children and Adolescents	Combined (Categorical)	111	12.6 (1.7)	49	99.5 (14.1)	Mix	N/A	220	13.1 (2.3)	50	105.2 (12.4)
Noordermeer et al. (2020)	Adolescents	Combined (Categorical)	82	16.3 (3.0)	33	96.9 (11.0)	Medicated but withdrew (48 hrs)	2 samples: mix, with ODD	82	16.1 (3.3)	33	98.3 (7.3)
Hall et al. (1999)	Children	Combined (Categorical)	15	N/A	N/A	N/A	2 Samples: On Medication, Medicated but withdrew (24 hrs)	N/A	15	N/A	N/A	N/A
Plećević et al. (2021)	Children	Combined (Categorical)	31	9.25 (1.9)	N/A	N/A	Non-Medicated	N/A	29	9.25 (1.9)	N/A	N/A
Miller et al. (2011)	Adults	2 samples: Combined, (Categorical) Inattentive	17	33.4 (11.45) 36.4 (10.7)	35	114.2 (8.3) 116.3 (14.4)	Medicated but withdrew (24 hrs)	N/A	18	31.9 (7.7)	28	121.9 (11.7)
Rapport et al. (2002)	Adults	Combined (Categorical)	28	36.3 (10.9)	42	106.7 (8.8)	N/A	N/A	28	33.4 (11.5)	46	105.2 (8.4)
Norvilitis et al. (2000)	Children	Combined (Categorical)	44	10.0 (2.1)	30	N/A	On Medication	N/A	36	10.6 (3.1)	45	N/A

Hanford (2000) (Unpublished)	Adults	2 samples(Categorical): Combined, Inattentive	17	33.4 (11.5) 36.4 (10.7)	35	114.2 (8.3) 116.3 (14.3)	Mix	N/A	18	31.9 (7.7)	28	121.9 (11.7)
Techentin (2009) (Unpublished)	Adolescents	Combined (Categorical)	39	12.4 (1.0)	0	N/A	Medicated but withdrew (24 hrs)	N/A	39	12.4 (1.04)	0	N/A
Abraham (2004) (Unpublished)	Adults	Combined (Categorical)	31	20.7 (SD not reported)	51	N/A	Mix	N/A	31	23.8 (SD not reported)	77	N/A
Friedman (2000) (Unpublished)	Adults	Combined (Categorical)	32	36.9 (11.2)	42	102.6 (14.6)	N/A	N/A	41	34.6 (11.6)	55	97.9 (15.1)
Zuberer et al. (2020)	Adults	Combined (Categorical)	44	30 (7.0)	23	N/A	Medicated but withdrew (24hrs)	N/A	43	28.2 (6.5)	42	N/A
Aldea (2013) (Unpublished)	Children	Combined (Categorical)	13	8.5 (SD not reported)	23	100.5 (15.5)	Medicated but withdrew (24 hrs)	N/A	16	8.4 (SD not reported)	56	118.3 (11.3)

Table 2.2

Design Characteristics of Studies Included in the Meta-Analysis

Study	How study diagnosed ADHD	Definition of ADHD	Informants of Categorical Diagnosis	Continuous Measure of ADHD	Informants of Continuous Measure	VER Task / Battery	Task Response	Linguistic Stimuli / Voice	Emotion Types Included	Total no. Trials
Chronaki, Garner, et al. (2015)	NA (Continuous Study)	DSM-IV	Professional	SDQ (Goodman, 1997)	Parent	DANVA 2	Verbal Naming	Semantically neutral phrase (English Adult Female)	Generic, Ang, Sad, Hap, Neu	70
Chronaki, Benikos, et al. (2015)	Research and Clinical Diagnosis	DISC-IV	Professional	SDQ (Goodman, 1997)	Parent and Teacher	Maurage et al. (2007)	Button Press	Interjection sound ‘ahh’ (Adult Female)	Generic, Ang, Sad, Neu	360
Sells et al. (Unpublished)	NA (Continuous Study)	DSM-IV	NA	SDQ (Goodman, 1997)	Self	Maurage et al (2007)	Button Press	Interjection sound ‘ahh’ (Adult)	Generic, Ang, Hap, Neu	156
Kis et al. (2017)	Clinical Diagnosis	DSM-IV-TR	Professional	ADHS (Rosler et al., 2004), WURS-K (Retz-Junginger et al, 2003)	Self	Tübinger Affect Battery (TAB; (Breitenstein et al., 1996)	Verbal Naming	Semantically neutral phrase (German Adult Female)	Generic, Ang, Sad, Fear, Hap, Neu	20
Kochel et al. (2015)	Clinical Diagnosis	DSM-IV-TR	Professional and Parent	FBB-ADHS (Döpfner, Görtz-Dorten, Lehmku, 2008), CBCL (Conners, 2008)	Parent/ Teacher	TAB	Verbal Naming	Semantically neutral phrase (German Adult Female)	Generic, Ang, Sad, Fear, Hap, Neu	12

Cadesky et al. (2000)	Research Diagnosis	DSM-IV	Parent and Teacher	Parent and Teacher Interview (Schacher et al., 1995),	Parent / Teacher	DANVA 2	Verbal Naming	Semantically Neutral (English Adult & Child)	Generic	40
Grabemann et al. (2013)	Clinical Diagnosis	DSM-IV	Professional	ADHS (Rosler et al., 2004), WURS-K (Retz-Junginger et al, 2003)	Self	TAB	Verbal Naming	Semantically Neutral Phrase (German Adult Female)	Generic	15
Corbett and Glidden (2000)	Clinical Diagnosis and Parent Account	DSM-IV	Professional and Parent	CPRS (Conners, 1998)	Parent	The Prosody Test (Tucker et al., 1977)	Point	Semantically Neutral Phrase (English)	Generic	16
Waddington et al. (2018)	Clinical and Research Diagnosis and Parent Account	DSM-IV	Professional and Parent	SCQ (De Giacomo, 2015), CPRS (Conners, 1998)	Self and Parent	Amsterdam Neuropsychological Tasks (ANT; (De Sonneville, 1999),	Verbal Naming	Semantically Neutral Phrase (Dutch Adult)	Generic, Ang, Hap, Sad, Fear	48
Noordermeer et al. (2020)	Clinical and Research Diagnosis	DSM-IV	Professional, Parent, Teacher	CTRS-R:L, CAARS-S:L (Conners, 1998, 1999)	Self and Teacher	ANT	Verbal Naming	Semantically Neutral Phrase (Dutch Adult)	Ang, Hap, Sad, Fear	48
Hall et al. (1999)	Clinical Diagnosis	DSM-IV	Professional	HESB, WPBIC	Parent and Teacher	DANVA	Verbal Naming	Semantically Neutral (English Child Female)	Generic	16
Plećević et al. (2021)	Clinical Diagnosis	DSM-IV-TR	Professional and Parent	SNAP-IV (Swanson et al., 2001)	Parent and Teacher	GEES (Jovičić et al., 2004)	Verbal Naming	Semantically Neutral Phrase (Serbian Adult)	Generic, Ang, Hap, Sad, Fear	18
Miller et al. (2011)	Clinical and Research Diagnosis	DSM-IV	Professional	DSM-IV Subscales (Murphy & Barkley, 1995)	Self and Other (Spouse,	DANVA 2	Verbal Naming	Semantically Neutral Phrase (English Adult)	Generic, Ang, Hap, Sad, Fear	24

					Friend or Parent)					
Rapport et al. (2002)	Clinical Diagnosis	DSM-IV	Professional	ADHD Symptom Checklists (Barkley & Murphy, 1998)	Self	DANVA 2	Verbal Naming	Semantically Neutral Phrase (English Adult)	Generic	24
Norvilitis et al. (2000)	Clinical and Research Diagnosis	DSM-IV	Professional and Parent	CPRS (Conners, 1998, 1999)	Parent	Rothenberg (1970)	Verbal Naming	Semantically Neutral Phrase	Generic	N/A
Hanford (2000) (Unpublished)	Clinical and Research Diagnosis	DSM-IV	Professional	Adult ADHD DSM-IV Rating Scale (Murphy & Barkley, 1995)	Self	DANVA 2	Verbal Naming	Semantically Neutral Phrase (Adult)	Generic	24
Techentin (2009) (Unpublished)	Clinical Diagnosis	DSM-IV	Professional	CBLC – Attention Scale (Conners, 2008)	Parent	Binaural Word Task	Verbal Naming	Pseudo Word (Adult Female)	Generic	18
Abraham (2004) (Unpublished)	Standardised Self-Report Questionnaire, and Clinical Diagnosis	DSM-IV	Self and Professional	CAARS (Conners, 1999)	Self	DANVA 2	Verbal Naming	Semantically Neutral Phrase (English Adult)	Generic, Ang, Hap, Sad, Fear	24
Friedman (2000) (Unpublished)	Clinical and Research Diagnosis	DSM-IV	Professional, Patients, Parents, Spouse	ADHD Symptom Checklist (Barkley & Murphy, 1998)	Self	DANVA 2	Verbal Naming	Semantically Neutral Phrase (English Adult)	Generic	24

Zuberer et al. (2020)	Clinical Diagnosis	DSM-IV	Professional	ADHS (Rosler et al., 2004), WURS-K (Retz-Junginger et al, 2003)	Self	Bisch et al. (2016)	Button Press	Meaningful Word (German Adult)	Generic, Ang, Hap, Sad, Dis, Ero	60
Aldea (2013) (Unpublished)	Clinical Diagnosis and Parent Account	DSM-IV-TR	Professional and Parent	CBCL (Conners, 2008)	Parent	DANVA 2	Verbal Naming	Semantically Neutral Phrase (English Adult & Child)	Generic	24

2.4 Analysis Strategy

Since we defined ADHD as a categorical variable, effect sizes were computed as standardised mean differences (SMD; hedges g) for VER accuracy scores between individuals with ADHD and controls (ADHD group minus TD group; Hedges & Olkin, 1985). All effects from correlational studies were converted from Pearson's r to hedges g. In total there were 93 effect sizes from 21 included studies. The variance for each effect size was also computed. As we can assume that the effect sizes gathered differ due to heterogeneity in samples recruited and study methods, the data were analysed using random-effects models (Borenstein et al., 2010). In typical random-effects models, there is the assumption that all effect sizes are independent from one another (Borenstein et al., 2010). However, in the present meta-analysis, this assumption was violated because we gathered more than one relevant effect size per included study. For example, some studies reported multiple effect sizes for each emotion type (e.g., anger, happiness, neutral), and ADHD participant group. This resulted in clusters of correlated effects for each study. To handle multiple dependent effect sizes within our meta-analyses, the method of robust variance estimation (RVE) was used (Tanner-Smith et al., 2016). There was also evidence of some hierarchical effects within our included studies (e.g., multiple studies from the same lab group). Note though, that most of the dependent effects were correlated, and therefore variance weights for correlated effects were employed (Tanner-Smith et al., 2016).

Two meta-analyses using RVE with variance weights for correlated effects in random-effects models were conducted separately. The first meta-analyses computed the average synthesised effect size for the presence of VER deficits. Twenty studies were included in this analysis as these studies provided the relevant statistics for overall VER accuracy scores. The second analysis explored the moderating effect of the theoretical moderator variables - emotion type and valence - on VER deficits in ADHD. Eleven studies were included in the moderator analyses as these studies provided relevant statistics for each specific emotion type included in the study. All analyses were conducted in the R environment (version 4.1.0). The Metafor package (Viechtbauer, 2010) was used to calculate and convert effect sizes into hedges g. The Robumeta package was used to estimate mean effect sizes using the RVE method (Tanner-Smith et al., 2016). The Grid package was used to create forest plots.

All exploratory moderator variables included in this meta-analysis contained distributions from fewer than 10 studies reporting relevant effect sizes for each variable condition. Therefore, following Kjærvik and Bushman (2021) and Sterne et al. (2011), we judged there was insufficient data to undertake formal statistical exploratory moderator analyses in line with best practice for conducting comprehensive meta-analyses. For transparency of data, we report the numerical trends (SMD and Standard Error) of all exploratory categorical variables in Table 2.3. However, these should be interpreted with caution due to paucity of data.

Table 2.3

Numerical Trends (SMD and Standard Error) of Vocal Emotion Recognition Deficits for All Categorical Exploratory Moderator Variables

Variable	Condition	SMD	Standard Error	Outcomes (k)	No. Studies
Theoretical Moderators					
Emotion Type	**Angry**	**-0.45****	**0.14**	**18**	**11**
	Happy	**-0.31*****	**0.07**	**16**	**11**
	Sad	**-0.43****	**0.13**	**10**	**8**
	Fearful	-0.30	0.20	8	6
	Neutral	-0.44	0.06	9	6
Emotion Valence	**Positive**	**-0.31*****	**0.07**	**16**	**11**
	Negative	**-0.47*****	**0.12**	**34**	**11**
Participant Characteristics					
Age Group	**Children**	**-0.75*****	**0.21**	**14**	**9**
	Adolescents	-0.195	0.18	11	4
	Adults	**-0.47*****	**0.07**	**11**	**9**
Outcome Variable	Continuous	-0.56	0.03	3	2
	Categorical	**-0.53*****	**0.19**	**33**	**19**
ADHD- Subtype	ADHD-Predominantly Inattentive	-0.65	0.11	8	3
	ADHD- Predominantly Hyperactive/Impulsive	-0.54	0.05	7	2
	ADHD- Combined	**-0.51*****	**0.11**	**33**	**20**
ADHD Comorbid CD	ADHD with comorbid CD	-0.43	0.58	5	2
	ADHD without CD	-0.20	0.13	10	4

ADHD Medication Status	**Withdrew for study**	**-0.39****	**0.13**	**23**	**11**
	Not reported or Mixed Status	**0.46*****	**0.08**	**10**	**8**
	Non-medicated	-2.25	0.11	1	1
	On medication	-0.52	0.34	2	2
Design Characteristics					
VER Task	**DANVA**	**-0.45*****	**0.11**	**16**	**9**
	Maurage et al	-0.61	0.13	3	2
	TAB	-0.65	0.09	3	3
	Prosody Test	-0.78	0.06	1	1
	ANT	-0.17	0.14	9	2
	GEES	-2.25	0.11	1	1
	Rothenberg	-0.79	0.01	1	1
	Binaural Word	0.26	0.05	1	1
	Bisch et al	-0.38	0.05	1	1
Response Type	**Verbal Naming**	**-0.52*****	**0.13**	**31**	**17**
	Button Press	-0.63	0.14	5	3
	Point	-0.78	0.06	1	1
Linguistic Stimuli Type	**Semantic Neutral Phrase**	**-0.58*****	**0.12**	**31**	**17**
	Interjection Sound	-0.61	0.13	3	2
	Nonsense Word	0.26	0.05	1	1
	Meaningful Word	-0.38	0.05	1	1
Voice Type	**Adult Female**	**-0.48****	**0.17**	**6**	**6**
	Adult Mix Gender	**-0.54*****	**0.17**	**23**	**12**
	Child	-0.39	0.30	5	3
Language of Stimuli	**English**	**-0.54*****	**0.1**	**18**	**11**
	German	-0.54	0.11	4	4
	Dutch	-0.17	0.14	9	2
	No Language	-0.33	0.31	4	3

	Serbian	-2.25	0.11	1	1
Source Characteristics					
Publication Status	**Published**	**-0.58*** **	**0.12**	**29**	**16**
	Unpublished	-0.37	0.22	7	5

Note. We can only reliably interpret those in bold, as the df was more than 4 when undertaking the analysis using RVE. All other results should be interpreted with caution due to paucity of data.

2.5 Results

When quantifying VER accuracy scores from 20 studies, there was a medium statistically significant effect size (SMD = -0.56, $\mathbf{CI_{95\%}}$[-0.78, -0.34]). Figure 2.2 presents a forest plot showing the weight of each effect size within the analyses, taking into consideration sample size and clusters of correlated effects. Similarly, across all emotion-specific VER accuracy scores from 11 studies, there was a small effect size (SMD = -0.39, $\mathbf{CI_{95\%}}$[-0.59, -0.18]). The slightly smaller effect size in the emotion-specific compared to the overall analyses is likely due to fewer effect sizes and studies included in the emotion-specific analysis. The I squared statistic indicated that 72% of the heterogeneity in the effects was due to between-study differences, suggesting that 28% of the variance is due to sampling error. These results suggest that despite heterogeneity in study methods and participant characteristics, there is strong evidence for VER deficits in ADHD, in accordance with our primary prediction. Moderator analyses showed that neither emotion type nor valence significantly influenced the magnitude of the effect size for emotion-specific accuracy scores (see Table 2.4). Therefore, there is no consistent pattern within the literature to indicate that individuals with ADHD have greater difficulties to recognise negative emotions, such as anger, when completing VER tasks. This contrasts with our prediction and does not provide support for an emotion-specific VER deficit in ADHD. This indicates the presence of a generic VER deficit in ADHD during explicit recognition tasks, supporting cognitive-behavioural theories (Barkley, 1997; Nigg et al., 2005). Alternatively, it is possible there is insufficient power within the literature to provide evidence that emotion moderates VER deficits in ADHD.

Sensitivity analyses were conducted to explore the robustness and validity of the effect for the presence of VER deficits in ADHD. Sensitivity analyses aimed to address what happens to our results, if aspects of the analyses change (Iyengar & Greenhouse, 2009). First, we calculated the estimated mean effect and associated statistics, without any adjustment for potential biases due to correlated effects for the generic VER analyses. This is known as the naïve SMD (Copas & Shi, 2000). To calculate the naïve SMD, a typical random effects model was employed using the Metafor package. Within these sensitivity analyses, we also calculated the prediction intervals. Prediction intervals capture the range of true effects which could be reported in an individual study, and thus suggest how consistent VER deficits are within the literature (Riley, Higgins & Deeks, 2011). The naïve SMD (Copas & Shi, 2000), along with prediction intervals (Riley et al., 2011), and tests for publication bias (Duval & Tweedie, 2000; Egger et al., 1997) are shown in Table 2.5. Potential outliers were identified from funnel plots (see Figures 2.3 and 2.4). Thus, we re-ran all analyses[1] with two extreme outliers (Plećević et al., 2021; Techentin, 2009) excluded. Additionally, to rule out the possible influence of medication of VER deficits in ADHD, we re-ran the analyses excluding all 7 studies in which the medication status of ADHD participants was not reported, mixed or 'on medication'. Excluding all 7 studies did not change the results. We did not conduct sensitivity analyses on the emotion-specific analyses because there were no significant effects. Overall, sensitivity analyses showed publication bias was not present within the current meta-analyses, nor did outliers or medication status influence the overall effect size for VER scores. However, when outliers were removed, the prediction intervals suggested the range of true effects that could be reported within an individual study were negative 95% of the time

[1] We also calculated the RVE analyses again for both the generic analyses and emotion-specific analyses, but with the outliers (Plecevic, 2021; Techentin, 2013) removed. When we re-ran the moderator analyses with the outliers removed, we found no change to the significance of the moderating effect of emotion type and valence. See row 5 in Table 2.5 for associated statistics.

(between -0.90 and -0.04). This provided strong support that our reported effect of a VER deficit in ADHD was consistent and robust (Riley et al., 2011).

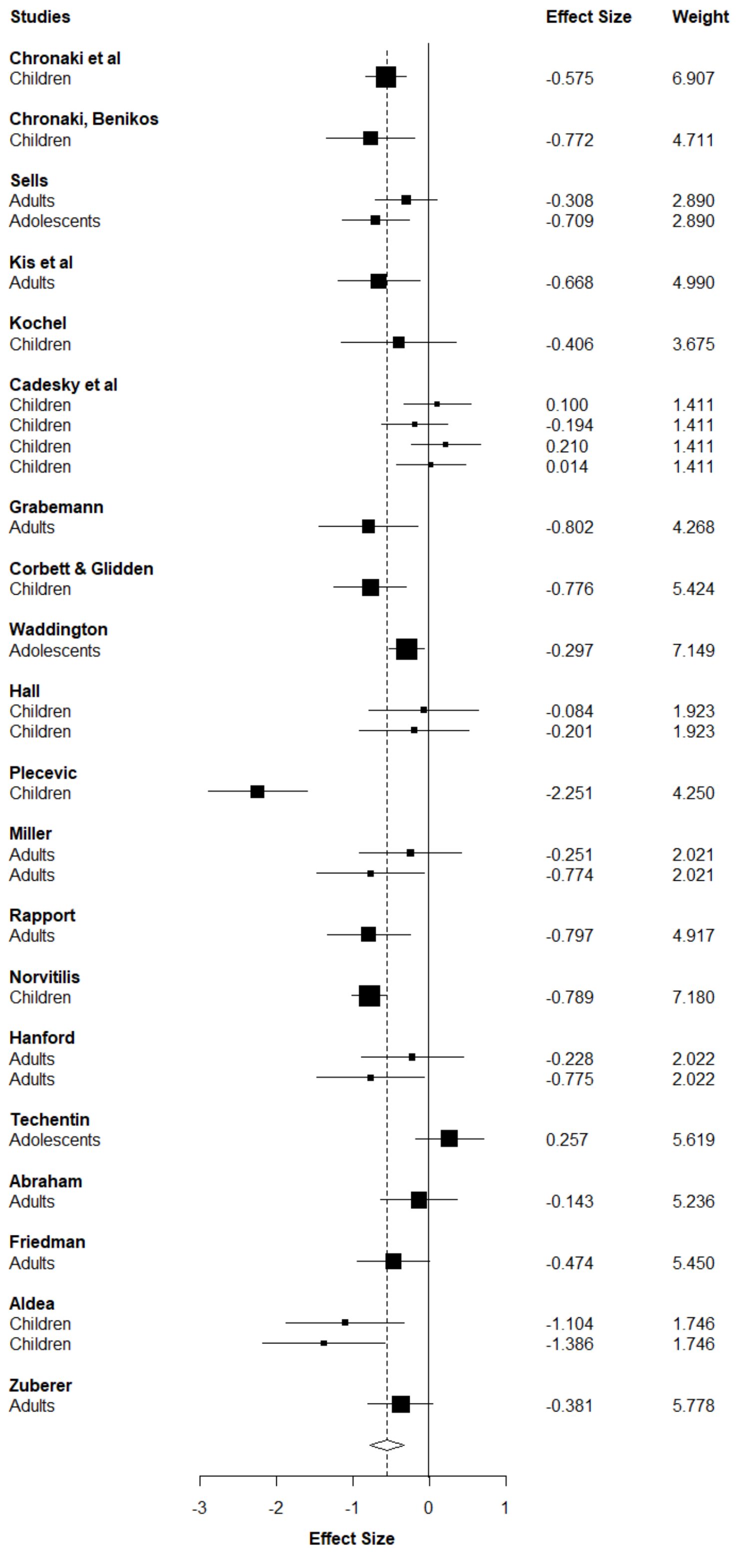

Figure 2.2

Forest Plot Showing SMD (hedges g) of the Generic VER Accuracy Score for the 20 Included Studies.

The figure shows the effect size for each study, and the corresponding weight of each effect size.

Table 2.4

Coefficient Estimates, and Associated T Values, Significance (P Values), and CI of Emotion Type and Valence from Moderator Analyses[2].

Moderator	Coefficient	T	p	CI	df
Emotion Type					
Anger vs Happiness	-0.18	1.39	.194	[-0.11, 0.46]	9.20
Anger vs Fear	-0.18	1.04	.333	[-0.24, 0.60]	6.58
Anger vs Neutral	-0.01	0.06	.951	[-0.42, 0.44]	6.36
Anger vs Sadness	-0.02	0.25	.812	[-0.18, 0.22]	8.20
Valence					
Negative vs Positive	-0.13	1.04	.323	[-0.15, 0.40]	9.32
Negative vs Neutral	0.04	0.22	.836	[-0.44, 0.39]	5.83

[2] Within the 'emotion type' moderator analysis, anger was the intercept for contrasts with the four other emotion types (happiness, fear, neutral and sadness). Within the 'valence' moderator analysis, negative was the intercept for contrasts with two other valence types (positive and neutral). It is important to note, not all 11 studies included in the emotion-specific analyses provided effect sizes for all emotion types (i.e., fearful, sadness, and neutral expressions). There were 36% missing for sadness, 45% missing for fear, 45% missing for neutral expression. As a sensitivity analysis, we re-ran the moderator analysis with only the emotions which were included in all 11 studies (anger and happiness). The mean SMD was -0.369, with 95% CI between -0.574 and -0.164. Moderator analyses (anger vs happiness) remained non-significant. Please see

for numerical trends of all other exploratory variables which cannot reliably be analysed via moderator analysis.

Table 2.5

Effect Sizes (Standardised Mean Difference; SMD) and Associated Statistics for the Sensitivity Analyses, With and Without Outliers (w/o).

	Meta-Analyses									Tests for Publication Bias					
										Eggers Test		Trim and Fill.			Associated Funnel Plot
Distribution	Type	k	No. Studies	SMD	CI (95%)	PI (95%)	Q	I²(%)	Tau	z	p (sig)	No. Inputted Studies (Left or Right Side)	t&f SMD	t&f CI (95%)	
Generic VER (w/o)	RVE	26	18	-0.52	[-0.67, -0.38]	N/A	N/A	40	.03	N/A	N/A	N/A			
Generic VER (No Medication)	RVE	19	13	-0.59	[-0.96, -0.24]	N/A	N/A	77.90	.23	N/A	N/A	N/A			
Generic VER	Naïve	25	20	-0.53	[-0.71, -0.35]	[-1.26, 0.19]	69.83	70.19	.36	0.26	.793	6 (L)	-0.696	[-0.89, -0.50]	Figure 2.3
Generic VER (w/o)	Naïve	23	18	-0.51	[-0.63, -0.39]	[-0.83, -0.19]	30.66	30.05	.15	0.21	.835	1(L)	-0.512	[-0.63, -0.39]	Figure 2.4

Note. Columns 1- 4 display the generic VER distribution of means analysed with or without outliers (w/o), the type of analyses (Robust Variance Estimation; RVE, or a Naïve Random-Effects analyses), the number of effects included (k), and the number of studies included in the analysed distributions respectively. Columns 5 to 9 provide the estimated mean standardised mean difference (SMD) effect size, 95% confidence intervals (CI), 95% prediction intervals (PI), the weighted sum of squared deviations from the mean (Q) and the ratio of true heterogeneity to total variance across observed effect sizes (I^2) and the estimate of between study variance (tau) for each analysed distribution. Columns 10 to 14 report the publication bias results. This includes the eggers test for funnel plot asymmetry statistics (z and p), and the trim and fill (t&f) analyses statistics, including the estimated number of missing studies on the left (L) or right (R) side, the trim and fill adjusted observed standardise mean difference (t&f SMD), and the trim and fill adjusted 95% confidence intervals (t&f CI). Funnel plots for sensitivity analyses are displayed in Figures 2.3 and 2.4.

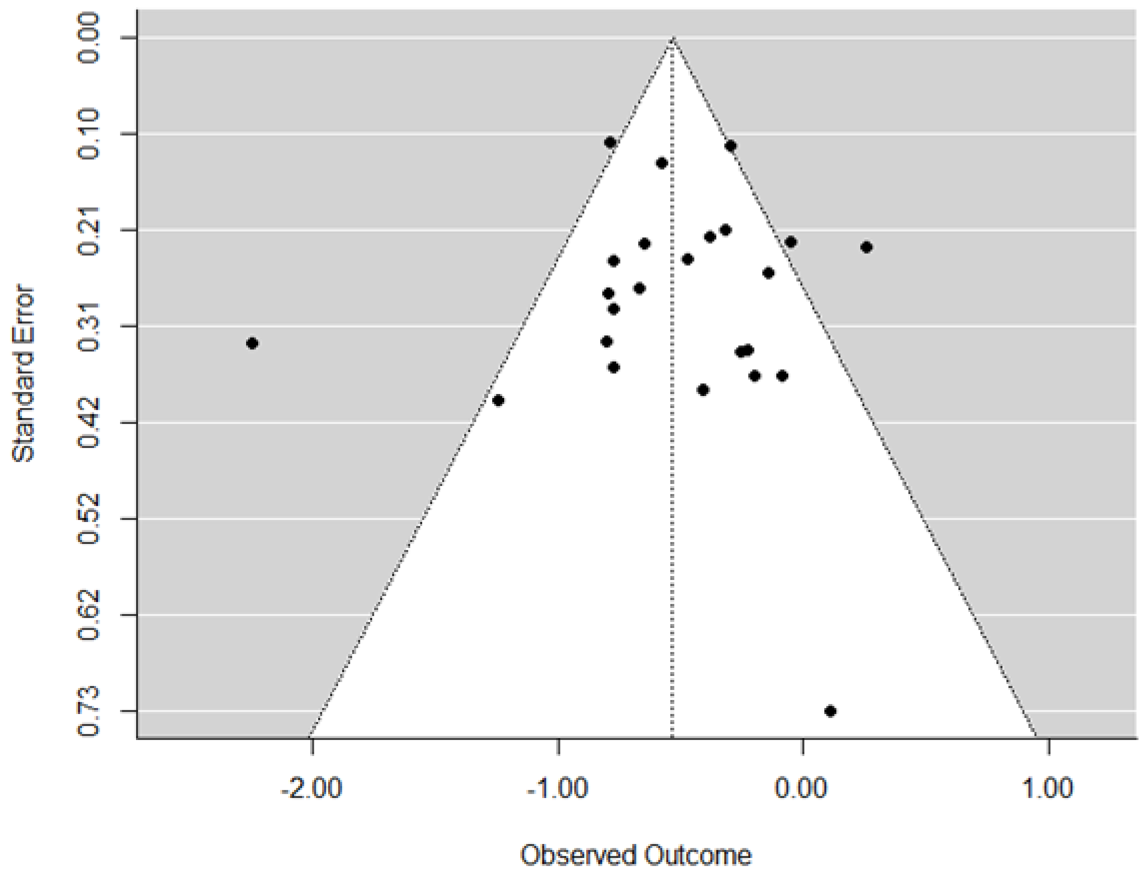

Figure 2.3

Funnel Plot for Generic VER Analysis.

.

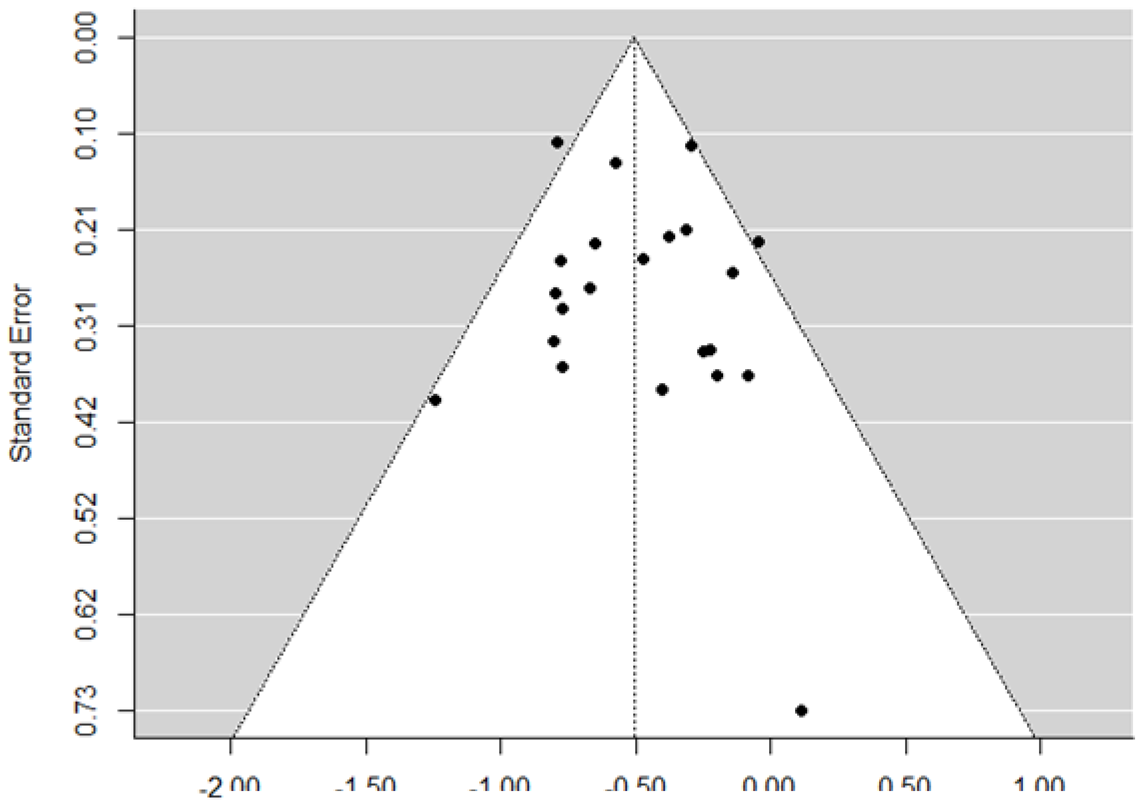

Figure 2.4

Funnel Plot for Generic VER Analysis Without Outliers.

2.6 Discussion

This meta-analysis provided evidence for the presence of VER deficits in ADHD and explored whether they were reflective of a generic attention deviancy and/or atypical emotion processing. Twenty-one published and unpublished studies were included as part of two meta-analyses quantifying VER accuracy scores, and the moderating effect of emotion type and valence on VER deficits in ADHD. First, our results showed a medium effect size for the presence of VER deficits in ADHD (d=0.56). Despite heterogeneity in study methods and characteristics of ADHD participants within the included studies, our results indicated VER deficits in ADHD are found in 95% of studies employing VER tasks, suggesting this effect is reliable and robust. This result is consistent with previous reviews reporting the presence of FER deficits in ADHD (Borhani & Nejati, 2018; Collin et al., 2013; Cooper et al., 2020). This is supportive of the preliminary finding from Bora and Pantelis (2016) who reported an effect of d=0.4 for VER deficits from only six papers prior to 2015. Importantly, the present meta-analyses confirm the effect of VER deficits in ADHD is reliable and consistent due to synthesising effects from a high number of studies (more than 10), in line with best practice. Therefore, our results add weight to the literature stating difficulties to recognise vocal emotional expressions, are present in ADHD.

Second, our meta-analysis showed a lack of moderation to VER deficits present in ADHD by emotion type or valence. There are two possible explanations for this finding. One explanation is that this is a true effect, indicating that the pattern of errors made by individuals with ADHD during VER tasks is random. This suggests VER deficits in ADHD are generic in nature, in support of the view that attentional mechanisms are implicated in emotion recognition difficulties in ADHD. This is consistent with previous research suggesting emotion regulation difficulties in ADHD are associated with deficits in executive function (Groves et al., 2020; Sjowall et al., 2013). These findings are in line with cognitive-behavioural theories (Barkley, 1997; Nigg et al., 2005), which imply individuals with ADHD may have difficulty regulating emotions due to a core deviancy in allocating attention towards emotions. It is also important to consider that attentional and socio-cognitive mechanisms are not mutually exclusive, as perceptual processes, including attention, may be influenced by processes related to motivation and impact subsequent interpretations of emotional information from vocal cues (Crick & Dodge, 1994; Lemerise & Arsenio, 2000).

An alternative explanation for this finding is that this result has arisen because of a lack of power in regard to the moderator analyses, despite including all VER studies to date, and gathering missing emotion-specific data from authors. For instance, most studies in the present meta-analysis only consistently included angry and happy vocal emotion types within the VER tasks. The other emotion types which were included within the VER literature (i.e., fear, sadness, neutral expressions) were not consistent throughout task designs. This inconsistency in emotion types made it difficult to reliably perform moderator analysis with a varied and nuanced selection of emotion types within the present meta-analysis. If this explanation regarding a lack of power is true, it indicates the cumulative evidence within the behavioural VER literature is not statistically robust to provide evidence for emotion-specific VER deficits in ADHD. Therefore, there is insufficient data reported within the behavioural VER literature to provide support for atypicalities in emotional processes implicated by socio-cognitive (Crick & Dodge, 1994) and motivational (Sonuga-Barke, 2005) accounts of ED in ADHD. More thorough research is necessary to determine whether emotion-specific deficits are present in individuals with ADHD, and whether they may be causally implicated in poor clinical and functional outcomes.

Future research might aim to explore whether emotion processing atypicalities contribute to vocal emotion processing deficits in ADHD through use of different experimental designs, employing neuro-imaging techniques, and by considering other important moderator variables, such as ADHD subgroup and comorbidities. This is important as there appear to be some trends within the literature indicating the presence of processing atypicalities at the perceptual processing stage of emotion recognition. For example, neuro-imaging data, such as those from fMRI and EEG studies, have provided support for a specific atypicality to process vocal anger in ADHD when using explicit emotion recognition paradigms (Chronaki, Benikos, et al., 2015; Kochel et al., 2015). Moreover, it has been suggested that exploring biases, rather than deficits, during explicit emotion recognition tasks may be a more sensitive measure of emotion-specific atypicalities at a perceptual level, as deficits are suggested to be a better marker of attention deviance (Cadesky et al., 2000; Uekermann et al., 2010). Alternatively, measuring the preattentive processing of vocal emotion with use of different experimental designs (i.e., when vocal cues are irrelevant to task demands) and EEG techniques may also be a more sensitive measure of emotion-specific atypicalities, as neural markers relating to preattentive, and attentive processing can be isolated with these techniques (See page 35). Within the present meta-analysis, the decision to quantify VER deficits from studies employing explicit emotion recognition tasks was justified due to limited research exploring biases, or the preattentive processing of vocal emotions, within the literature. Additionally, trends within the VER literature suggest it is possible that other variables, such as varying ADHD subtypes, and comorbid CD, may affect the nature of vocal emotion processing in ADHD. For example, Miller et al. (2011) found ADHD-inattentive type was associated with generic VER deficits in adults, whereas hyperactive/impulsive symptoms were associated with a specific deficit for recognising negative emotions. Moreover, children with comorbid ADHD and CD were more likely to make VER errors due to a bias to label ambiguous voices as anger, which was not found in children with ADHD-only (Cadesky et al., 2000). Overall, these trends should be interpreted with caution due to limited evidence. Further research is required to substantiate them within the VER literature.

The finding of robust VER deficits in ADHD complements results from the extensive reports of FER deficits in ADHD (Bora & Pantelis, 2016; Borhani & Nejati, 2018; Collin et al., 2013; Cooper et al., 2020). Together, these findings suggest emotion recognition deficits in ADHD are present across modalities, in both the face and voice. However, there is no evidence from the behavioural literature to support the presence of emotion-specific deficits in the voice, as previously found in the face (Bora & Pantelis, 2016; Borhani & Nejati, 2018). One explanation for this could be due to the dynamic nature of vocal stimuli, compared to static stimuli generally employed in FER studies. For example, perceiving emotions in voices, involves a participant attending to various acoustic parameters which change over time. This activates different areas of the brain compared to interpreting a single static image in FER studies (Schirmer & Adolphs, 2017). It is therefore possible that, due to the nature of the stimuli, FER tasks are more sensitive, compared to VER tasks, in demonstrating the moderating effects of emotion in individuals with ADHD. In support of this, FER tasks which employed dynamic facial stimuli reported deficits for both positive and negative stimuli in children with ADHD compared to controls but did not report any specific deficits for negative emotions (Jusyte et al., 2017; Ludlow et al., 2014). Future research might aim to employ dynamic multimodal presentations of emotional stimuli (e.g., presentation of the voice and face together), which has been shown to increase the accuracy of emotion perception in typical individuals (Klasen et al., 2012). This will enable us to understand whether emotion-specific effects are present when vocal emotional signals are integrated into more holistic emotional expressions, thought to be a more reliable measure of emotion perception (Schirmer & Adolphs, 2017).

In summary, this meta-analysis provides evidence for a robust and reliable effect for the presence of VER deficits in ADHD, in accordance with previous reviews and facial emotion recognition research (Bora & Pantelis, 2016). Secondly, there is a lack of robust evidence in support of the modulation of emotion on VER deficits in ADHD. This suggests VER deficits are generic in nature, supporting the view that ED in ADHD is a secondary problem underpinned, at least in part, by primary attention deficits (Barkley, 1997; Nigg et al., 2005). There is currently not sufficiently robust evidence within the behavioural VER literature to support the definite presence of atypicalities in the perceptual emotion processing mechanisms in ADHD. There is a clear need for further behavioural and neuro-imaging emotion processing studies exploring both biases and deficits during explicit vocal, and multimodal, emotion processing, as well as employing different experimental designs to explore the preattentive processing of vocal emotions ('implicit' emotion recognition), which take into consideration ADHD subtypes and comorbidities. The empirical research within the present thesis (see Chapter Five) aimed to specifically build on the neuroimaging literature by investigating the presence of emotion-specific atypicalities in ADHD with the use of a preattentive, rather than explicit, paradigm as a more sensitive measure of emotion processing. The next Chapters (3 and 4) will discuss how the stimuli and paradigm used were developed and validated in adult participants.

3 VALIDATING A SET OF VOCAL EMOTIONAL STIMULI

3.1 Chapter Overview

As presented in Chapter One (section 1.4); to explore the effect of vocal emotion on VSA with an Emotional Spatial Cueing (ESC) Paradigm, irrelevant vocal emotional stimuli are presented to participants prior to a visual target. To ensure the vocal stimuli accurately reflected a particular emotion (angry, happy or neutral), the present study aimed to validate the stimuli in a sample of 26 heathy adults and 23 typically developing adolescents. The primary goal was to ensure accuracy rates for each stimulus were above chance and in line with previous literature. A secondary goal was to explore the construct validity of the stimuli by investigating if patterns relating to the age, emotion type and behavioural difficulties on VER were consistent with the literature. All vocal stimuli were taken from a pre-existing validated battery (Maurage et al., 2007) and consisted of male and female adult actors producing '/a/' sounds in happy, angry, or neutral prosody. Adults and adolescents were asked to listen to each vocal stimulus and classify the expression by selecting one of three response options. This study was conducted online during the COVID-19 pandemic. Results showed that adults and adolescents recognised all stimuli at above chance levels. A total of six stimuli with high accuracy rates were selected to be included in the ESC paradigm. Item selection was based on a high percentage of inter-judge agreement because these stimuli were recognized by most participants as communicating a particular emotion. The rationale for selecting these six vocal stimuli for the subsequent studies of the thesis is discussed.

3.2 Introduction

To explore vocal emotion processing in children, adolescents and adults, batteries of vocal emotional stimuli have been developed and validated for use in explicit (e.g., vocal emotion recognition) and preattentive (e.g., spatial cueing) tasks. The development of vocal stimuli typically involves actors portraying either linguistic words or phrases (e.g., 'I am going out the room now, and will be back later'; DANVA; Nowicki & Duke, 1994) or non-linguistic interjections (e.g., '/a/'; Maurage et al., 2007) in different emotion tones (e.g., anger, happiness). Next, these stimuli are modified to ensure the standardisation of acoustic properties (e.g., amplitude and duration) whilst ensuring the properties which characterise the emotion within the voice are intact. The standardisation of stimuli, in terms of intensity (dB) and duration, is considered especially important in electrophysiological research to reduce effects of acoustic properties on ERP components (i.e., the N1), thought to be modulated by low-level acoustic features of stimuli processed in the auditory cortex (Schirmer & Kotz, 2006; Engelien et al., 2000). The standardization of vocal stimuli in terms of f0 (in Hz) is less common since fluctuations in pitch are thought to be one of the most definitive features of characterising emotions in voices (Banse & Scherer, 1996; Globerson et al., 2013). Finally, to validate the vocal emotional stimuli, it is necessary to ensure each stimulus portrays an agreed upon emotion type by the majority of participants. This is essential in research exploring socioemotional functioning, as vocal emotions are only considered a valid form of communication if the emotion is mutually perceived by both speaker and listener (Juslin & Laukka, 2003).

Different approaches have been used to validate vocal stimuli. For example, Banse and Scherer (1996) recruited experts from a professional acting school to judge stimuli (pseudo-utterances) portraying 14 different vocal emotions, including anger, fear, sadness, joy, and disgust. The experts rated the authenticity and recognisability of each emotion on a 6-point authenticity scale (1: very good to 6: very poor) and 4-point recognisability scale (1: clearly recognisable, 4: not recognisable). Items were selected as a valid portrayal of the emotion if they had a mean recognisability rating of 2, and a mean authenticity rating of 4 or above. In a sample of 12 non-expert adults, the percentage mean recognition rate for all emotions was above chance, at 55%, suggesting these stimuli represented the correct emotion type (Banse & Scherer, 1996). Morningstar et al (2017, 2018) took the same approach to validate a set of recordings produced by both youth and adult actors, portraying anger, disgust, fear, friendliness, happiness, meanness and sadness. Six adult raters independently judged recordings based on recognisability and authenticity, and intraclass correlation coefficients were computed to access whether raters agreed on recognisability and authenticity of stimuli (Morningstar, Ly, et al., 2018). From these correlation coefficients 140 stimuli were selected to be employed in a VER task with adolescents (aged 13-15 years) and adult participants. Further findings from this study employing the carefully selected 140 vocal stimuli showed percentage accuracy rates for stimuli produced by both adolescent and adult speakers was between 31 and 83%, which was above chance for all stimuli, further validating the stimuli in a younger sample (Morningstar, Ly, et al., 2018).

Another approach is the empirical-normative approach, which has been employed to validate batteries of vocal emotion stimuli in children (Nowicki & Duke, 1994) and adults (Chronaki, Hadwin et al., 2015). Similar to the approach used by Banse and Scherer (1996), judges are instructed to indicate the emotion of the vocal stimulus out of a specified set of emotions (e.g., anger, happiness, fearful, sad) and neutrality, as in VER tasks. However, from these ratings, an item-by-item percent agreement score is calculated for each stimulus by calculating the percentage of judges agreeing on the correct identification of the vocal emotion (Nowiki & Duke, 1994). If the percentage agreements score is above chance, this indicates that the majority of judges agree that the stimulus represents a specific emotion (Chronaki, Hadwin et al., 2015). This approach is thought to ecologically reflect the agreement speakers and listeners automatically make in everyday situations when perceiving emotions from vocal cues (Nowicki & Duke, 1994). When testing the validity of vocal stimuli portraying happy, sad, angry and fearful emotions, from the DANVA vocal stimulus battery using the empirical-normative approach, Nowicki and Duke (1994) reported a high mean percentage accuracy (80%) across a set of vocal emotional stimuli, judged by children aged 6-11 years. This suggests that stimuli from the DANVA battery are a valid set of vocal emotions, which can be used in studies exploring emotion processing in children.

The empirical-normative approach was also employed by Chronaki, Hadwin et al (2015) when validating a set of stimuli vocalising non-linguistic interjection sounds (e.g., '/a/') in angry, happy, sad or neutral tones developed by Maurage et al (2007) and manipulated to be presented at different intensity (dB) levels (100%, 75% and 50%). The empirical-normative validation procedure employed by Chronaki et al. (2015) involved instructing 18 children (aged 4-11 years) and 22 adults to indicate whether they heard an angry, happy, sad or neutral expression within the voice. The mean percentage agreement calculated from children and adults' accuracy scores for the stimuli was 40% or higher for those at mild intensity (angry: 50%, happy:40%, sad: 40%) and 60% or higher for those at moderate intensity (angry:77.50%, happy: 72.50%, sad: 72.50% and neutral: 60%). The finding that higher intensity emotions are better recognised in children and adults in this study is consistent with previous research which similarly demonstrate that high-intensity emotional expressions are better recognised (Banziger et al., 2012; Hess et al., 1997; Juslin & Laukka, 2001; Holz et al., 2021), and thus highlights the importance of employing vocal emotions higher in intensity in preattentive vocal emotion processing studies, in which it is essential for vocal emotions to represent the correct emotion.

The stimuli vocalising short interjection sounds at the highest intensity (100%) have also been validated in adults (Maurage et al., 2007) using a slightly different validation approach with expert and non-expert adults. First, experts in emotional prosody rated the quality of the recording and the recognizability of the emotion on a scale of 1-7, and second, 70 adults (undergraduate students) identified the emotion in the voice using a Likert scale to indicate if the emotion was expressed in the voice. From these ratings, intensity and specificity scores was calculated, which considered how much the vocal stimuli depicted the relevant and irrelevant emotions, respectively. Overall, within this study, the specificity scores (i.e., mean percentage calculated by mean accuracy score for an emotion condition by considering the scoring from irrelevant emotions) were between 75.1% to 97.6%. It is noteworthy that the set of stimuli by Maurage et al (2007) have been validated as having high inter-judge agreement in a set of both adults (Maurage et al., 2007) and children (Chronaki et al., 2015), as these stimuli (presented at 100% intensity) are employed in the ESC in later studies in the thesis with adults (Chapter Four) and adolescents (aged 12-16 years; Chapter Five) participants. However, to date, no studies have validated the vocal emotional stimuli battery developed by Maurage et al (2007) in a sample of adolescent participants to further insure high inter-judge agreement within all age groups. This is necessary due to the continued development of VER in adolescence (Grosbras et al., 2018; Morningstar, Nelson, et al., 2018).

The construct validity of vocal emotional stimuli can additionally be established by exploring whether effects of age, emotion and behavioural difficulties on VER are consistent with patterns reported in previous literature. To explore these effects on VER more reliably, raw accuracy percentage scores are often transformed into 'discrimination accuracy' and 'response bias' scores for each emotion condition (Corwin, 1994; Chronaki, Hadwin et al., 2015). Discrimination accuracy (Pr) scores calculate how accurate participants are at recognising a particular emotion condition. Response bias (Br) scores are also calculated for each emotion condition but portray a participant's tendency to mis-identify a voice as a particular emotion (i.e., a bias to classify voices as angry; Corwin, 1994; Chronaki, Hadwin et al., 2015). Thus, Pr and Br scores are useful when exploring the patterns of VER within emotion conditions, rather than validating each individual stimulus.

Previous literature has assessed the effect of emotion on discrimination accuracy and response bias in children (4-11 years) and adults, when employing female vocal stimuli portraying angry, happy and sad emotions from the Maurage et al (2007) battery in a VER task. Chronaki, Hadwin et al (2015) reported a higher recognition accuracy for angry compared to happy and neutral voices in both adults and children. This is consistent with the literature suggesting a universal advantage to recognise angry vocal emotions compared to other emotion types (Chronaki et al., 2018; Liu & Pell, 2012; Morningstar et al., 2019). Moreover, Chronaki, Hadwin et al (2015) reported a higher response bias to label voices as sad, rather than angry or happy stimuli, suggesting a further effect of emotion on the bias to label voices in children. Together these results are supportive of theories highlighting negative vocal cues as being perceived as highly salient due to their association with potential threat in the environment (Liu & Pell, 2012; Vuilleumier, 2005). There is some mixed evidence regarding the pattern of vocal emotion recognition accuracy and response bias in adolescents.

One study employing stimuli from Montreal Affective Voice (MAV) database report that happiness is recognised with most accuracy in adolescents (Grosbras et al., 2018), which may reflect a heightened sensitivity to rewarding stimuli during this developmental period (Galvan, 2013). However, most other studies employing stimuli from varied batteries report that angry voices are still recognised with the highest accuracy in adolescence compared to happy and neutral voices (Filippa et al., 2022; Morningstar, Ly, et al., 2018; Zupan et al., 2024). It is not yet known if the pattern of VER accuracy and response bias in adolescents (aged 12-16 years), when employing the Maurage et al (2007) stimuli battery, is consistent with previous results, such as a higher accuracy to recognise angry voices (Fillipa et al., 2022; Morningstar, Ly, et al., 2018).

The battery of vocal emotional stimuli developed by Maurage et al (2007) has further been evidenced to demonstrate good construct validity when employed to explore the developmental trajectory of VER accuracy in children and adults (Chronaki et al., 2015). For instance, adults presented with better VER discrimination accuracy compared to 11-year-old children (Chronaki, Hadwin et al., 2015). This is consistent with literature (see Chapter One, page 11) which suggests VER accuracy continues to develop from childhood into adolescence before reaching adult levels during mid to late adolescence (Grosbras et al., 2018; Morningstar, Nelson, et al., 2018). However, it is noteworthy that this effect was driven by the low intensity (dB) vocal emotions, which continued to improve with age throughout late childhood (10-11 years) (Chronaki et al., 2015). For high intensity expressions, older children did not differ in recognition accuracy compared to adults (Chronaki, Hadwin et al., 2015). This is consistent with evidence which suggests recognising subtleties and nuances in emotional expressions takes longer to develop (Fillipa et al., 2022). Previous VER studies employing linguistic stimuli (e.g., ‘why did you do that?’) have further demonstrated that VER accuracy in 13-15 year old adolescents is lower compared to adults but only when recognising vocal cues spoken by youth, and not those spoken by adults (Morningstar, Ly, et al., 2018). No studies have employed the Maurage et al (2007) battery of non-linguistic stimuli in adolescents (12–16-year-olds) to explore the difference in VER accuracy for the present set of stimuli in adolescents compared to adults. If accuracy rates are lower in adolescents compared to adults, this would be consistent with the developmental literature of VER (Grosbras et al., 2018).

Finally, construct validity of vocal emotion stimuli can be established by exploring if patterns in VER accuracy and response bias in individuals with behavioural difficulties is consistent with the literature. For instance, previous research has demonstrated difficulties recognising emotions from vocal expressions are linked with inappropriate responses in social situations and the development of internalising and externalising behaviours (Trentacosta & Fine, 2010). Specifically, adolescents with more depressive symptoms presented with lower accuracy to recognise angry and happy voices during a VER task (Morningstar et al., 2019) and a response bias to misinterpret happy emotional stimuli as angry or fearful in 8 to 20-year-olds (Vidal-Ribas et al., 2018). Moreover, our meta-analysis (Chapter Two) demonstrates that individuals with ADHD have difficulties recognising emotions from voices (Sells et al., 2023). Similarly, a study employing 2 female vocal stimuli per emotion condition (angry, happy and neutral) form the Maurage et al (2007) battery demonstrated that children with ADHD (aged 6-11 years) were less accurate at recognising vocal emotions compared to typically developing children (Chronaki et al., 2015). In addition, children with ADHD displayed a response bias to label angry voices as neutral (Chronaki et al., 2015). This is consistent with the literature which suggests children with ADHD tend to display specific threat-related processing difficulties (Williams et al., 2008; Kochel et al., 2012). This evidence corroborates the validity of the stimuli developed by Maurage et al (2007) to explore vocal emotion processing in children with externalising behaviour problems (e.g., inattention/hyperactivity).

Previous findings exploring patterns of VER in adolescents with externalising behaviours, using stimuli from other batteries, are mixed. For example, Waddington et al. (2018) found levels of inattention/hyperactivity correlated with more errors and slower RTs during a VER task in children and adolescents, aged 7 to 18 years. However, a recent study by Noordermeer et al. (2020) which employed a range of neurocognitive tests, including a VER task, did not find adolescents with ADHD differed in VER accuracy in comparison to adolescents with ADHD and ODD, or typically developing controls. Due to these mixed findings, it is not yet clear whether specific difficulties in recognising vocal anger found in children with traits of ADHD persist into adolescence or decline with development. As of yet, no studies have employed the Maurage et al (2007) battery to explore associations between externalising or internalising problems and VER in adolescents. If emotion processing difficulties in ADHD persist in adolescents, we would expect a specific correlation between the recognition of angry voices and traits of inattention/hyperactivity, consistent with that found in children when employing the same stimuli (Chronaki et al., 2015).

The Present Study

There have not yet been any studies which have validated the battery of stimuli developed by Maurage et al (2007) in an adolescent sample (aged 12-16 years) either through specifically testing the validity of each stimulus with a sample of adolescent judges (e.g., via the empirical-normative approach), or by testing the construct validity of the stimuli battery as a whole, by exploring patterns of emotion, age and behavioural difficulties on discrimination accuracy and response bias. Therefore, the present study aimed to employ an explicit VER task to validate the Maurage et al (2007) battery in adolescents and inform the selection of specific stimuli from the battery for use in later studies in the thesis exploring the preattentive processing of vocal emotion in adults, and adolescents with traits of ADHD (i.e., in Chapters 4 and 5). Aims and predictions are presented below. The emotion types selected from the Maurage et al (2007) battery were angry, happy and neutral voices, for consistency with those included in previous studies exploring vocal emotion processing in children with ADHD (Chronaki et al., 2015). Anger was selected as it is the emotion which is consistently found to be associated with behavioural difficulties, such as ADHD (Chronaki et al., 2015). Happy was selected as an emotion of opposite valence to angry, based on evidence that children with ADHD have similar, albeit milder, difficulties recognising happy vocal cues (Chronaki et al., 2015). Finally, neutral voices were selected to serve as a baseline condition, as in previous VER (Chronaki et al., 2015) and ESC paradigms (Brosch et al., 2009).

Aims

I. The primary aim of this study was to validate each individual stimulus included in the present study by measuring mean percent agreement accuracy rates and perceived intensity for each stimulus to inform selection of specific stimuli for the use in the ESC paradigm.

II. A second aim of the present study was to explore the construct validity of the stimuli by examining whether the pattern of results was consistent with previous literature by investigating effects of emotion, and age on accuracy, response bias and perceived intensity.

III. An exploratory aim was to investigate associations between externalising (inattention, hyperactivity) and internalising (depressive) symptoms, discrimination accuracy, response bias and perceived intensity to further corroborate the validity of the stimuli in an adult and adolescent sample[3].

Predictions

i. It was predicted that mean percent agreement scores for each stimulus would be above chance (i.e., 33.33%). However, only stimuli with the highest percent agreement scores (i.e., above 70%) would be selected for use in the ESC paradigm, consistent with prior literature with children and adults.

ii. It was predicted that adults would score higher on all emotion conditions than adolescents, and that within our adolescent sample, age would positively correlate with emotion recognition accuracy, consistent with literature suggesting VER accuracy continues to improve until mid-late adolescents (Grosbras et al., 2018).

iii. It was predicted that participants (adults and adolescents) would recognise angry voices more accurately compared to happy and neutral voices, based on previous research on VER in children employing the vocal stimuli battery (Maurage et al., 2007) in children and adults (Chronaki et al., 2015).

iv. It was predicted that behavioural difficulties (inattention/hyperactivity) would be negatively associated with vocal anger recognition in both adolescents and adults in our community sample, similar to existing research investigating associations in children (Chronaki et al., 2015). In addition, it was predicted that depressive symptoms would be associated with a negative response bias, as found in previous

[3] Note, the methods and results relating to the third exploratory aim are outlined in Appendix B and Appendix H to avoid taking the focus away from the main aims of validating the stimuli.

studies exploring associations between depressive symptoms and VER (Vidal-Ribas et al., 2018).

3.3 Methods

Participants

In total, 30 adult and 24 adolescent participants were recruited from the community online, through advertisements on social media, emails, and word of mouth. Four adults and one adolescent were identified as outliers, due to having unusually low accuracy scores (i.e., below 40%), and removed from analyses. As these participants' accuracy scores were extremely below the mean, it was possible that they did not understand the task. In total, 26 adults (age range 18-39, mean age = 26.70, SD=4.00, 9 males) and 23 adolescents (age range 12-16, mean age=14.30, SD=1.36, 8 males) were included in analyses. A post-hoc power analysis computed using Gpower suggested that with the current adolescent sample size (23), a power of 0.98, error probability 0.05, was achieved when exploring associations between hyperactivity and anger vocal emotion recognition accuracy scores.

Vocal Stimuli

The vocal stimuli were derived from a battery of vocal emotional expressions (Maurage et al., 2007), including male and female speakers vocalising a non-word vocal 'ah' sound. In the present study 10 happy, 8 angry and 8 neutral prosodic stimuli were utilized due to the number of stimuli available from the original Maurage et al (2007) database of stimuli. In total 26 stimuli were used. Acoustic properties were measured using Pratt software (Boersma & Weenink, 2024). All stimuli were then edited to ensure standardisation of mean intensity (77dB)[4] and rise and fall times (5-20ms) with the software Audacity® (Audacity

[4] Note, since the present study was conducted online due to the COVID-19 pandemic, the intensity with which participants heard the stimuli may have differed depending on the volume set on their computer. However, importantly, participants heard all emotion types / stimuli at the same volume.

Team, n.d.). Vocal stimuli were presented in either the left or right ear via headphones, consistent with later studies employing the ESC paradigm. Table 3.1 displays the acoustic properties of each stimulus included in the present study. Table 3.2 displays the mean acoustic properties of stimuli used in analyses for each emotion condition.

Table 3.1

Acoustic Properties of Each Vocal Stimulus Item Included

Item No.	Duration (ms)	Mean Fundamental Frequency (Hz)	Mean Intensity (dB)
Angry			
5*	693	308.20	77.14
7	686	99.20	77.44
8	622	171.49	77.39
11*	677	132.18	77.03
18	607	335.37	76.75
22	581	295.83	77.38
24	679	228.90	76.76
25	693	308.20	77.14
Happy			
10*	680	250.67	76.96
11	688	289.49	76.79
12	690	159.21	76.77
14	680	183.20	76.87
16	700	388.56	76.94
18	685	350.18	77.05
20*	667	357.97	77.31
21	649	223.31	77.34
25	585	293.41	77.26
26	588	246.28	76.90
Neutral			
6	527	94.45	76.57
11	584	187.78	76.84
13*	698	191.31	76.68
15*	658	183.33	76.78
20	505	101.18	76.87

21	541	207.25	76.77
29	595	126.40	77.39
30	555	101.52	76.95

***** Note, the items with an Asterix were those included in the ESC paradigm. Since, after acoustic analysis, six out of eight neutral stimuli had a shorter duration than expected (under 600ms), these were deemed unsuitable for use in the ESC paradigm.

Table 3.2

Mean Acoustic Properties of Stimuli Used for Each Emotion Condition within the Present Study[5].

	Mean Duration (ms)	Mean Fundamental Frequency (Hz)	Mean Intensity (dB)
Angry	655.62	235.23	77.13
Happy	671.88	260.13	77.03
Neutral	583.38	148.83	76.80

Vocal Emotion Recognition Task

The vocal emotion recognition task was administered online using the platform PSYCHOJS hosted by Pavlovia® (Peirce, 2020). The empirical-normative approach was applied to validate the stimuli by non-expert adult and adolescents (Nowicki & Duke, 1994; Chronaki, Hadwin et al., 2015). Participants were instructed to identify the emotion expressed in each vocal stimulus as either 'Happy', 'Angry' or 'Neutral' by pressing the keys 'H', 'A' and '0' respectively with the respective emotion words. If participants identified an angry or happy voice, they were then asked to rate the intensity of the emotion on a scale of 1 – 4 (e.g., where 1 is not at all angry and 4 is extremely angry). If participants identified a neutral voice, they

[5] See Table C.1 in Appendix C for the mean acoustic properties of the stimuli for each emotion conditions, which were selected for the ESC paradigm.

were asked to rate the intensity of the emotion as 0. Participants first completed three random practice trials. There were three blocks, and each stimulus was played twice in one block in a random order, resulting in 52 trials per block. In total participants heard each stimulus 6 times, leading to a total of 156 trials. After the completion of each block, participants were reminded they could take a break. On average, the task took 15 to 20 minutes to complete. Participants were informed that they could withdraw from the task at any time by exiting the browser. Responses, to completed tasks only, were logged via Pavlovia.

Procedure

Participants were recruited via an online link to a Qualtrics survey in which they read the information sheet and filled in the online consent forms. Participants were then redirected to Pavlovia to complete the vocal emotion recognition task. They were instructed to wear headphones for this task and to complete it on a laptop or desktop computer. Upon completion of the VER task, participants were thanked and reminded to complete the questionnaire measures measuring internalising (depressive symptoms) and externalising (inattention/hyperactivity) symptoms. See Appendix for details of questionnaires employed. Participants were redirected to a separate Qualtrics survey to complete the questionnaire measures. After completion of the questionnaire measures, participants were given the option to be entered into a prize draw to win £50 worth of Amazon vouchers and debriefed in writing on Qualtrics. The study took 30- 40 minutes to complete.

3.4 Data Processing

To explore the validity of each stimulus item by the empirical-normative method, item-by-item percentage agreement was calculated as the mean percentage each stimulus was correctly identified by adolescents and adults. Mean perceived intensity ratings were also calculated for each happy and angry emotion conditions (on a scale 1-4) for adults and adolescents, to ensure stimuli picked for each emotion condition were similar in perceived intensity rating. Mean perceived intensity ratings were not calculated for neutral stimuli.

To explore construct validity by examining effects of age, emotion and behavioural problems on VER, raw data were transformed into discrimination accuracy (Pr) and response bias (Br) scores for each emotion condition, consistent with previous studies exploring VER in children (e.g., Chronaki et al., 2013; 2015). The following formula was used to calculate discrimination accuracy for each emotion condition: Pr = ((number of hits + 0.5)/(number of targets + 1)) - ((number of false alarms + 0.5)/(number of distractors + 1)) (Corwin, 1994). Pr scores closer to 1 suggest accuracy for that emotion condition is better than chance, whereas scores close to 0 suggest accuracy is close to chance and less than 0 suggest accuracy is worse than chance. In addition, tendency to misclassify emotional stimuli with a biased response was calculated using the following formula to calculate response bias, for each emotion condition: Br = ((number of false alarms + 0.5)/(number of distractors + 1))/(1 - Pr). Br scores which are closer to 1 suggest the presence of a systematic bias towards responding with the corresponding emotion condition. Br scores which a closer to 0 suggest the absence of a bias.

Initial Data Treatment

Kolmogorov-Smirnov tests indicated that discrimination accuracy scores in our adolescent sample did not follow a normal distribution due to high accuracy rates for all emotion conditions; angry, D (23) = 0.30, $p < .001$; happy, D (23) = 0.25, $p = .004$; and neutral, D (23) = 0.25, p = .001, probably due to ceiling effects as accuracy scores for all emotion conditions was high (close to 1). Similarly in adults, accuracy scores for angry and neutral voices did not follow a normal distribution, D (26) = 0.21, $p = .006$; D (26) = 0.25, $p < .001$; however, accuracy scores for happy voices were normally distributed, D (26) = 0.17, $p = .056$. Additionally, Kolmogorov-Smirnov tests revealed normally distributed response bias scores for adolescents for happy (D(23) = 0.12, $p = .200$) and angry (D(23)= 0.16, $p = .151$), but not neutral (D(23) = 0.27, $p < .001$) stimuli. In the adult sample, response bias scores were normally distributed for happy (D(26) = 0.13, $p = .200$) and neutral (D(26) = 0.17, $p = .050$) but not angry (D(26) = 0.19, $p = .020$) stimuli. Therefore, due to some conditions being not normally distributed, non-parametric tests were employed to explore the effects emotion and age on accuracy scores and response bias. Finally, Kolmogorov-Smirnov tests revealed normally distributed data for intensity scores, per emotion and age group (adult angry: D(26)=0.12, p=.122; adult happy: D(26)=0.13, p=.200; adolescent angry: D(23)=0.12, p=.200; adolescent happy: D(23)=0.15, p=.194), so parametric tests were employed to explore the effect of emotion and age on intensity scores.

3.5 Results

Item-By-Item Percentage Agreement and Perceived Intensity Ratings

The mean percentage agreement scores and intensity ratings for each item in adults and adolescents are presented in Table 3.3. All stimuli were identified correctly with above chance accuracy (defined as 33.33%) in both adult and adolescent samples. In addition, all stimuli apart from two happy stimuli (items 18 and 26) were correctly identified in over 70% of instances, consistent with previous validation studies in children (Chronaki, Hadwin et al., 2015) and adults (Maurage et al., 2007). Mean perceived intensity ratings for angry stimuli were 3 to 4 (rounded up), suggesting participants perceived angry stimuli as quite to extremely intense, and 2 to 3 (rounded up) for happy stimuli, suggesting participants perceived happy stimuli as a little to quite intense. Intensity ratings are not listed for correctly rated neutral stimuli because participants were not instructed to rate the intensity of the neutral stimuli.

Table 3.3

Means (SDs) of Percentage Recognition Agreement (%) and Perceived Intensity Rating for Each Stimulus Item

Item No.	Adults (N=26)		Adolescents (N=23)	
	Recognition (%)	Intensity	Recognition (%)	Intensity
Angry				
5*	100 (0.00)	3.30 (0.48)	89.13 (20.48)	3.42 (0.43)
7	98.72 (4.53)	3.22 (0.56)	95.65 (10.32)	3.50 (0.50)
8	99.36 (3.27)	3.18 (0.49)	93.47 (13.97)	3.35 (0.50)
11*	100 (0.00)	3.12 (0.62)	92.75 (14.93)	3.23 (0.51)
18	99.36 (3.27)	3.36 (0.49)	89.13 (18.54)	3.50 (0.45)
22	98.08 (5.43)	2.74 (0.47)	86.23 (25.94)	2.86 (0.57)
24	82.05 (30.52)	2.97 (0.78)	82.61 (29.51)	3.25 (0.55)
25	96.15 (10.86)	2.74 (0.44)	91.30 (15.79)	2.71 (0.50)
Happy				
10*	96.15 (13.59)	2.66 (0.51)	88.41 (17.72)	2.88 (0.40)
11	92.31 (15.80)	2.36 (0.54)	85.51 (18.33)	2.47 (0.62)
12	95.51 (12.07)	2.41 (0.58)	86.96 (26.09)	2.49 (0.49)
14	95.51 (8.89)	2.45 (0.49)	89.85 (16.47)	2.46 (0.48)
16	95.51 (11.11)	2.85 (0.50)	86.96 (25.10)	2.92 (0.53)
18	48.08 (41.72)	2.42 (0.67)	49.28 (35.35)	2.45 (0.81)
20*	94.23 (14.10)	2.97 (0.46)	92.75 (17.99)	2.99 (0.42)
21	90.38 (15.76)	2.54 (0.51)	87.68 (18.95)	2.57 (0.41)
25	93.59 (13.40)	2.75 (0.66)	83.33 (25.62)	2.82 (0.63)
26	69.23 (29.70)	2.08 (0.48)	76.81 (23.96)	2.06 (0.45)
Neutral				
6	97.44 (6.13)	-	83.33 (27.52)	-
11	92.95 (13.48)	-	73.91 (30.49)	-
13*	97.44 (7.74)	-	80.43 (28.71)	-
15*	94.87 (10.30)	-	80.43 (30.83)	-
20	95.51 (12.07)	-	84.78 (24.05)	-

21	92.31 (17.15)	-	69.57 (34.32)	-
29	87.82 (17.36)	-	73.91 (22.37)	-
30	98.72 (4.53)	-	84.78 (26.54)	-

Note. Perceived intensity scores were given on a scale from 1-4, where 1 is not at all and 4 is extremely intense. *Vocal stimuli with an asterisk were those selected for use in the ESC paradigm due to high and similar accuracy and perceived intensity for emotion conditions, in both adults and adolescents. See Table C.1 in Appendix C for mean accuracy rates and perceived intensity ratings of the stimuli for each emotion condition selected for use in the ESC paradigm.

The Effect of Emotion on Recognition Accuracy, Perceived Intensity and Response Bias

Means and standard deviations for discrimination accuracy (Pr) and response bias (Br) are presented in Table 3.4. Pr scores for all emotion conditions in adults and adolescents were high (close to 1), suggesting that mean accuracy for each emotion condition was better than chance. Consistent with the literature, Friedman Tests revealed that there was a significant effect of emotion condition on accuracy scores in both adolescents, ($\chi 2(2) = 8.70$, $p = .013$) and adults, ($\chi 2(2) = 7.81$, $p = .020$). Consistent with our prediction, post hoc tests using Wilcoxon signed-rank tests indicated that in adolescents, accuracy for angry voices was significantly higher than for neutral voices ($z = 2.28$, $p = .023$), however not significantly higher than for happy voices, ($z = 1.66$, $p = .097$). In adults, accuracy for angry voices was significantly higher than for happy voices, ($z = 2.20$, $p = .028$), which was consistent with our prediction. However, when Bonferroni corrections were applied, at a significance level set at $p = .025$, this finding was not significant in adults. In addition, a Friedman's test revealed there were no significant effect of emotion on response bias scores in adolescents ($\chi 2(2) = 5.30$, $p = .07$)

or adults ($\chi2(2) = 2.00$, $p = .368$). This suggests there was not an elevated bias towards any specific emotion in the adult or adolescent group.

Not only were angry voices recognised most accurately, but as paired samples t-tests revealed, angry voices were also rated as more intense, compared to happy voices by both adults ($t(25) = 8.19$, $p < .001$) and adolescents ($t(22) = 8.22$, $p < .001$), even after acoustic properties (intensity in dB) were standardised. Table 3.5 displays the means and standard deviations for perceived intensity ratings for happy and angry voices by emotion and age group.

Table 3.4

Mean (SD) of Discrimination Accuracy (Pr) and Response Bias (Br) per Age Group and Emotion Condition

Age groups	Angry		Happy		Neutral	
	Pr	Br	Pr	Br	Pr	Br
Adults	0.93 (0.06)	0.37 (0.23)	0.89 (0.08)	0.34 (0.23)	0.91 (0.08)	0.28 (0.20)
Adolescents	0.81 (0.19)	0.38 (0.25)	0.70 (0.17)	0.37 (0.21)	0.72 (0.26)	0.21 (0.33)

Table 3.5

Mean (SD) of Perceived Intensity Rating for Vocal Expressions per Emotion Condition and Age Group

Age group	Angry	Happy
Adults	3.07 (0.41)	2.56 (0.39)
Adolescents	3.23 (0.34)	2.61 (0.33)

Note. Intensity ratings measured on a scale of 1 to 4 where; 1 = not at all, 2 = a little, 3 = quite, 4 = extremely happy/angry.

The Effect of Age on Recognition Accuracy, Perceived Intensity and Response Bias Scores

To examine the effect of age (adolescents vs adults) on recognition accuracy, Mann Whitney U tests were conducted. In addition, for a more thorough analysis, Spearman's rho correlations examined associations between age (as a continuous variable) and accuracy scores within each age group. Consistent with predictions, a Mann Whitney U test revealed adults presented significantly higher accuracy to recognise angry (U = 199.50, p = .046), happy (U = 181.00, p = .018) and neutral (U = 120.00, p < .001) voices, compared to adolescents. Spearman's rho correlations revealed adults age positively correlated with angry accuracy scores (rs = 0.44, p = .013). Moreover, Spearman's rho correlations revealed age within the adolescent group positively correlated with accuracy rates for happy (rs = 0.44, p = .017) and neutral (rs = 0.43, p = .021) voices. However, the correlation between age in adolescents and angry voices did not reach significance (rs = 0.32, p =.068). Independent t-tests revealed there were no significant differences in intensity ratings between adults and adolescents for either happy (t (47) = 0.56, p = .580) or angry (t(47) = 1.47, p = .147) voices, suggesting that perceived intensity of the vocal stimuli was not influenced by age. Finally, a Mann Whitney U test revealed no significant differences in response bias scores between age group for angry (U=284.50, p=.771), happy (U=276.50, p=.652) and neutral (U=235.50, p=.199) conditions.

Because the primary focus of this study was to validate the vocal stinmuli, the results in relation to the relationship between internalising and externalising behavioural difficulties and VER accuracy and response bias are stated in Appendix H.

3.6 Discussion

The primary aim of this study was to validate a set of vocal stimuli portraying happy, angry, and neutral voices to be used in later behavioural and EEG studies of the thesis. Twenty-six vocal stimuli from the battery created by Maurage et al (2007) were included within an online VER task. First, following an empirical-normative approach to validating the battery of vocal stimuli, results showed the majority of stimuli were recognised with a percentage agreement rate between 80-100%, which was above chance (33.33%). This was consistent with previous accuracy rates reported in children and adults (Chronaki, Hadwin et al., 2015; Maurage et al., 2007). In addition, our findings are consistent with literature demonstrating angry voices are recognised with higher accuracy compared to happy and neutral voices (Chronaki et al., 2018; Pell et al., 2009). Moreover, the expected developmental effects were present in this study. Specifically, VER accuracy was higher in adults compared to adolescents, consistent with prior literature (Morningstar, Ly, et al., 2018; Fillipa et al., 2022; Grosbras et al., 2018). The effects of age and emotion on accuracy were consistent with prior literature which further highlights the construct validity of the vocal stimuli. Finally, our study showed that inattention was negatively associated with VER accuracy (pr) scores for angry voices. This pattern of association between VER and externalising behaviours in adolescents is consistent with the literature which suggests individuals with traits of inattention/hyperactivity have specific difficulties recognising vocal anger (Chronaki et al., 2015; Chronaki et al., 2013; Manassis et al., 2007; Plećević et al., 2020), further corroborating the validity of the vocal stimuli.

This is the first study to validate the set of stimuli developed and standardised by Maurage et al (2007) in a sample of adolescent and adult judges using the empirical normative approach. Results demonstrate high percentage accuracy rates (above 80%) for the majority of stimuli employed in the present study in both the non-expert adult and adolescent group. This is consistent with previous research employing the same battery of stimuli which reported accuracy rates between 60-80% in children, when using a similar empirical-normative approach (Chronaki, Hadwin, et al., 2015). These findings add to the validation of stimuli from the same battery in adults, in a study which demonstrated specificity scores of between 75-97% (Maurage et al., 2007). Our findings suggests that the stimuli employed in the present study have high inter-judge agreement in both adults and adolescents, and since mutual agreement of vocal emotions is considered essential when defining emotions in voices (Juslin & Laukka, 2003), this confirms the validity of the present stimuli for use in the ESC paradigm with adolescent and adult participants. Importantly, these findings informed the selection of stimuli for the ESC paradigm, based on high accuracy rates (above 80%) in accordance with previous literature (Chronaki et al., 2015).

Considering the effect of emotion on discrimination accuracy, our findings showed a consistent specific advantage to recognise angry voices, compared to happy and neutral voices. This effect has been demonstrated previously in adults and children (Chronaki et al., 2018; Liu & Pell, 2012; Pell et al., 2009), and adolescents (Morningstar et al., 2019; Davis et al., 2020), who also report angry voices are recognised more accurately compared to other expressions (i.e., sadness, disgust, fear, friendliness, happiness, and meanness). The ability to recognise negative compared to positive emotions more accurately is in line with theories suggesting recognising threatening vocal cues, such as anger, is necessary to ensure human survival, and therefore has evolved to be perceived as highly salient (Liu & Pell, 2012; Vuilleumier, 2005). Consistent with this, the present study demonstrated that the perceived intensity as rated by adults and adolescents was significantly larger for angry (i.e., quite to extremely intense) compared to happy voices (a little to quite intense). This is interesting, as the objective acoustic intensity levels (dB) were standardised between emotion conditions, suggesting the perceived intensity levels were specific to the perception of emotion, possibly due to angry vocal cues capturing attention more automatically. This finding is consistent with a study in adults which also instructed participants to rate the perceived intensity of the emotion and reported that intensity ratings were higher for negative compared to positive expressions (Holz et al., 2021). Moreover the combination of both high accuracy and high perceived intensity scores for angry voices in the present study is consistent with research by Morningstar et al (2021), who similarly reported that when angry voices were rated as higher in perceived intensity, they were also better recognised in terms of accuracy in adults (aged 18 to 38). The fact that effects of emotion on accuracy and intensity within the present study are consistent with previous literature, corroborates the validity of the present stimuli. For the purpose of selecting stimuli for the ESC paradigm, happy and angry stimuli were selected if they were categorised with a mean intensity of 3 (quite intense) to ensure stimuli for happy and angry stimuli were as equal as possible in terms of perceived intensity.

In terms of the effect of age on VER, as predicted, our results showed that adults presented with higher accuracy for all emotions compared to adolescents. These results support existing literature suggesting VER does not reach adult levels in late childhood (i.e. 11 years), as opposed to facial emotion recognition, but continues to develop throughout adolescence (Chronaki, Hadwin, et al., 2015; Grosbras et al., 2018; Morningstar, Ly, et al., 2018). Morningstar, Ly, et al. (2018) showed that the performance of adolescents (aged 13 to 15 years) in a VER task was less accurate than that of adults only when hearing vocal emotional expressions (meaningful linguistic sentences) produced by other adolescents, and not adults. The findings from the present study add to this by suggesting that this effect holds true when adolescents listened to non-linguistic vocal emotions produced by adults. Therefore, our results provide further evidence that recognition of basic vocal emotions follows a developmental trajectory that extends throughout adolescence. In relation to the effect of age on accuracy of the individual emotion conditions, our findings showed accuracy increased with age in adolescents for happy and neutral voices, but not for angry voices. This may be because vocal anger was also perceived as more intense than happy voices, and the recognition of more intense vocal emotions have been reported to mature earlier in children (Chronaki, Hadwin, et al. 2015). Additionally, it is possible the association between age and accuracy for vocal anger was not as strong due to ceiling effects from high accuracy rates. Overall, it is a strength of the present study that predicted age and emotion effects were found in relation to VER accuracy, especially considering the study was conducted entirely online due to the COVID-19 pandemic.

To further highlight the construct validity of the vocal stimuli employed in the present study, the results exploring the relationship between externalising behaviours (inattention/hyperactivity) and VER accuracy demonstrate that adults who reported higher levels of inattention were less accurate at recognising angry voices. This is consistent with previous research which similarly showed that adults with ADHD displayed greater difficulties when perceiving angry voices compared to adults without ADHD (Kis et al., 2017). Our study expands on this finding by establishing difficulties to recognising angry voices are specifically associated with higher levels of inattention in a community sample of adults. Replication of this effect found in a community sample highlights the importance of recruiting non-clinical samples in future studies. Moreover, the same association between traits of inattention/hyperactivity and anger recognition accuracy in adolescents was close to significance. This association most likely did not reach significance because of a low sample size recruited due to this being only an exploratory aim within the present study. However, importantly, the pattern of results in adolescents is consistent with previous literature in children (Chronaki et al., 2013; Manassis et al., 2007; Plećević et al., 2020). In addition, this result is partially consistent with research which has found that adolescents with more symptoms of hyperactivity and inattention performed with slower reaction times and more generic errors in a VER task (Waddington et al., 2018). Therefore, overall, the pattern of results pointing to an association between angry VER accuracy scores and traits of inattention/hyperactivity in adults and adolescents further corroborates the validity of the stimuli and highlights the important role of anger in underlying ADHD traits.

Finally, no significant effects of emotion, age, or behavioural characteristics, were found on response bias scores within the present validation study. This is consistent with research by Chronaki, Hadwin et al (2015) who did not report any differences in response bias scores for angry and happy voices in older children or adults. Specifically, only preschoolers were found to display a response bias to label voices as happy (Chronaki, Hadwin et al., 2015). In addition, our results did not indicate any significant associations between a change in depressive symptoms and response biases in VER. This contrasts to previous literature which robustly reported associations between more depressive symptoms, such as in depressive disorders, and a negative response bias (i.e. to angry voices) in VER tasks (Kan et al., 2004; Vidal-Ribas et al., 2018). However, since negative response biases have previously been shown most effectively in ambiguous emotional contexts (Schultz et al., 2000), it is likely that vocal stimuli employed in the present study were not ambiguous enough in order to capture predicted effects relating to bias. This is demonstrated by the high accuracy rates reported in the present study. Although, this was a strength of this study when validating the stimuli, and when exploring associations between externalising behaviours and response accuracy, it appears limiting in terms of exploring response biases. Overall, the finding of no significant effect of emotion on response bias in the present study suggests that response bias may not be as useful as accuracy in validating and selecting stimuli for use in the ESC paradigm.

In conclusion, the present study validated a standardised set of stimuli produced by actors vocalising an /’ah/ sound (Maurage et al., 2007) for use in vocal emotion processing tasks, such as the ESC paradigm, with adults and adolescent participants. First, the present study demonstrated accuracy item-by-item percentage agreement rates were high, above chance and consistent with previous literature. Second, the present study demonstrated good construct validity of the stimuli as predicted effects of emotion and age on VER accuracy were present in adult and adolescents. Interestingly, despite all stimuli being well standardised, angry voices were recognised with the highest accuracy rates and perceived as most intense, supporting the idea that angry vocal cues are perceived as more salient, perhaps due to capturing attention. The following ERP studies (Chapters 4 and 5) will further test this idea more thoroughly by measuring the P1 component, reflecting the capture of attention by vocal emotion in an ESC paradigm. Importantly, results related to accuracy, as well as characteristics of acoustic properties from this validation study informed the selection of six stimuli (two per emotion condition) for use in the ESC paradigm. Stimuli were selected based on high accuracy rates (above 80%) and a moderate (a mean rating of 3) perceived intensity rate[6], for all emotion types in accordance with previous developmental literature (Chronaki et al., 2015).

[6] See Table C.1 in the appendices for mean acoustic characteristics, accuracy rates (%) and perceived intensity of the stimuli selected for each emotion condition in the ESC paradigm

4 VALIDATION OF A NOVEL EMOTIONAL SPATIAL CUEING PARADIGM IN HEALTHY ADULTS: A BEHAVIOURAL AND ELECTROPHYSIOLOGICAL PILOT STUDY.

4.1 Chapter Overview

Chapter Three successfully validated a set of vocal emotional stimuli to be employed in a novel emotional spatial cueing (ESC) paradigm measuring preattentive and attentive VEP atypicalities in adolescents with inattentive/hyperactive traits. Since most previous studies exploring these processes via behavioural and electrophysiological techniques have been conducted in adults, and not adolescents[7], the present study aimed to validate the ESC paradigm as a measure of preattentive and attentive VEP in healthy adults. Four experiments were conducted to validate the ESC paradigm on a behavioural and neural level in a total sample of 49 adults. Experiments One to Three optimised and validated the ESC paradigm on a behavioural level by exploring its effectiveness at demonstrating attentional capture by irrelevant vocal cues by measuring RTs to the visual target. Experiment Four further validated the task on a neural level by investigating if the ESC is effective at isolating a) the P2 component to irrelevant vocal emotions, as a measure of the preattentive processing of vocal emotion; and b) the P1 component to a visual target, as a measure of the capture of visual spatial attention (VSA) by irrelevant vocal cues, consistent with the literature in healthy adults when employing dot-probe paradigms (Brosch et al., 2009). An exploratory aim was to

[7] The introduction includes a review of the literature employing spatial cueing paradigms (and similar tasks) in both adults and adolescents with traits of inattention/hyperactivity, to provide the rationale for the development of a paradigm which would be suitable to use with adolescents in a later ERP study (Chapter Five).

examine the validity of the paradigm at measuring the effects of vocal emotion on later attentional mechanisms. This was examined by isolating the P3 component to the visual target, consistent with the literature in healthy adults when employing other similar spatial cueing paradigms (Wang et al., 2022). Findings from all four experiments demonstrated the ESC paradigm had good construct validity when a) measuring RTs as a behavioural measure of the capture of attention by irrelevant vocal emotion and b) isolating the P2 component as a neural marker of the preattentive processing of vocal emotion. However, it is unclear whether the ESC paradigm had good construct validity when isolating the P1 as a neural measure of the capture of attention by irrelevant vocal cues. Finally, there is some evidence that the Target-P3 component within the ESC paradigm reflects the effects of vocal cues on mechanisms of later attentional allocation to a visual target. Implications for the use of the ESC to explore preattentive and attentive emotion processing in adolescents with traits of inattention/hyperactivity are discussed.

4.2 Introduction

The present thesis aimed to develop a paradigm to reliably isolate electrophysiological markers of the preattentive processing of vocal emotion. To explore the preattentive processing of vocal emotion (i.e., the perceptual processes which occur prior to the attentional capture of emotion; Jaaskelainen et al., 2004; Chronaki & Marsh, 2024), previous studies have incorporated task-irrelevant vocal emotional stimuli in experimental tasks (Mauchand & Zhang, 2022). For example, within previous ERP studies investigating the preattentive processing of vocal emotion adults were asked to identify the gender of vocal stimuli rather than the emotion (Lin et al., 2022), or to attend to the number of syllables in pseudowords and detect the deviant stimuli depending on the number of syllables (Gadeke et al., 2013). These studies reported a larger P2 response (120-160ms) in the frontocentral electrode sites, in response to happy and angry compared to neutral voices, suggesting the P2 is a neural marker of the preattentive processing of vocal emotion (Schirmer & Kotz, 2006; Lin et al., 2022; Gadeke et al., 2013). Alternatively, studies employing oddball paradigms typically present a standard stimulus type serially, interspersed with occasional deviant stimuli (Charpentier et al., 2018). Within these oddball paradigms, participants are instructed to passively view or listen to a vocal stimulus whilst ignoring the steam of standard and deviant stimuli. For example, Charpentier et al (2018) instructed adults and children to watch a silent movie, whilst passively hearing neutral voice standards and neutral and angry voices as deviants. The results revealed that angry deviants elicited shorter MMN (200-300ms) latencies compared to neutral deviants (Charpentier et al., 2018). This was interpreted as reflecting a rapid preattentive processing of angry compared to neutral voices (Charpentier et al., 2018). Similarly, other oddball studies have demonstrated an enhanced MMN response

(200-300ms) to angry compared to happy and sad vocal deviants in adults, indicating that the MMN in oddball paradigms may be a neural marker of the preattentive processing of vocal emotion (Deng et al., 2022; Kao & Zhang, 2023). In summary, paradigms instructing participants to ignore the emotion in the voice and focus on other aspect of the stimulus (i.e., gender; Lin et al., 2022; Gadeke et al., 2013), have isolated ERP components (e.g., MMN, P2) at around 200ms, suggested to reflect a preattentive processing stage of vocal emotion in healthy adults. To date, there have not been any ERP studies which have explored the neural processing of vocal emotion using preattentive or oddball paradigms in adolescents. This is important as prior to exploring atypicalities in the preattentive processing of vocal emotion in adolescents with ADHD, it is necessary to understand the neural markers of preattentive VEP in healthy adolescents.

The paradigm aimed to measure the involuntary capture of attention by vocal emotion. The involuntary capture of certain stimuli (e.g., emotional over neutral stimuli) is thought to be due to a preattentive gating mechanism which determines stimuli salience, and the degree to which those stimuli capture attention (Jaaskelainen et al., 2004). Previous behavioural studies have shown that salient auditory stimuli involuntarilyy capture visual attention over relevant visual stimuli, using spatial cueing paradigms (Spence &Driver, 1997; McDonald et al., 2000; Mazza et al., 2007). For example, Spence and Driver (1997) first developed the orthogonal cueing paradigm (OCP), to explore the effects of irrelevant visual and auditory stimuli on spatial attention in adults. The procedure of one trial of the OCP comprised of a central fixation (usually a light or cross), followed by a random delay of 400-650ms. Following this, an auditory (i.e., 2000Hz of pure tone) or visual (led light) cue was presented in either the left or right hemisphere for around 100ms. Next, after a delay dependent on the SOA (stimulus onset asynchrony), a visual (i.e., red LED lights) or auditory (i.e., white noise) target was then presented in the left or right hemisphere in one of four up or down positions (see Figure 4.1 for a schematic view of the OCP). Adult participants were instructed to ignore the cue stimulus and respond to the target stimulus by pressing a button corresponding to its up or down position. Trials could either be congruent (in which the location and target were presented in the same left or right hemisphere), or incongruent (in which the location of cue and target were presented in opposite hemispheres). Congruent and incongruent trials were presented equally and randomly to ensure there were no predictive elements of the cueing stimuli, and they were irrelevant to task demands (Spence & Driver, 1997). Across six experiments, in which cues and targets varied in terms of modality, the most striking finding Spence and Driver (1997) reported was a congruency effect on reaction times (RTs) when auditory cues preceded visual targets; response times were faster and more accurate in congruent, as opposed to incongruent, trials. Therefore, the authors suggested irrelevant auditory stimuli involuntarily capture and facilitate visual spatial attention (VSA; Spence & Driver, 1997). This is in line with theories which suggest the auditory attentional system acts as an 'early warning system' which automatically prioritises task-irrelevant auditory stimuli over task-relevant visual stimuli, helping us to respond quickly to potentially threatening events outside the visual field (Murphy et al., 2017).

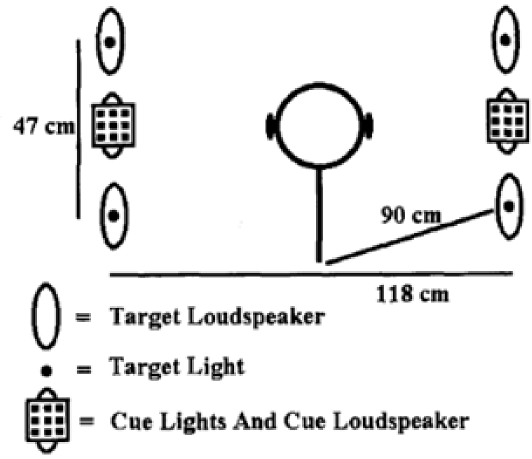

Figure 4.1

Schematic Diagram of the Original Orthogonal Cueing Paradigm (Spence & Driver, 1997). Auditory and visual stimuli were presented in one of four positions specified on the figure. Reproduced from Perception & Psychophysics, Vol 59/Issue 1, Spence & Driver, "Audiovisual links in exogenous covert spatial attention", Copyright 1996, with permission from Springer Nature

The congruency effect on reaction times, as reported by Spence and Driver (1997), is robust within the spatial cueing paradigm literature. In a similar cueing task, an irrelevant auditory cue was presented to the left or right side of a fixation prior to a visual mask presented at the congruent or incongruent location (McDonald et al., 2000). On half of the trials, a visual target was presented at the masked location, but on the other half of the trials the visual target did not appear. Participants were asked to respond via a button press to indicate whether the target was present. Results showed adults responded 20ms quicker to the target on congruent compared to incongruent trials, supporting the presence of a congruency effect due to irrelevant auditory cues involuntarily capturing VSA (McDonald et al., 2000). Moreover, a study which incorporated a similar spatial cueing paradigm, but incorporated predictive knowledge of the targets location (i.e., written information signalling which visual field the target would appear before each block), similarly reported a congruency effect on performance (Mazza et al., 2007). Participants were still faster to respond to targets in the congruent as opposed to the incongruent location, even when they were aware the target was going to appear in the incongruent location, suggesting that irrelevant auditory cues involuntarily captured visual spatial attention (Mazza et al., 2007).

In addition, behavioural studies have shown that preattentively processed emotional stimuli involuntarily capture attention more rapidly than neutral stimuli (Vuilleumier & Pourtois, 2007; Lima et al., 2019; Brosch et al., 2008). However, there are mixed findings within the literature as to whether the involuntary attentional capture of irrelevant vocal emotion has a facilitating (Brosch et al., 2009; Zimmer et al., 2022) or distracting effect (Kattner & Ellermeier, 2018; Kattner et al., 2024; Peschard et al., 2017; Sander et al., 2005) on task-relevant attentional processing. For example, the involuntary capture of vocal anger has been reported in a dot-probe paradigm which incorporated vocal emotions as cueing stimuli with adult participants (Brosch et al., 2008). In the dot-probe paradigm (See Figure 4.2), auditory stimuli consisting of angry and neutral prosody in the form of pseudo-words (64.5dB) were presented simultaneously in the left and right ear (e.g., an angry voice in the left ear, and a neutral voice in the right) prior to a visual target, which was congruent or incongruent to the location of the angry vocal cue (Brosch et al., 2008). Participants were instructed to press either the left or right arrow key button to indicate the location of the visual target. Results demonstrated that adults were faster to respond to targets preceded by a spatially congruent compared to incongruent angry vocal cue (Brosch et al., 2008). As both neutral and angry vocal cues were presented simultaneously, these results suggest angry voices selectively captured attention over neutral voices. This is consistent with previous dot-probe paradigms which included irrelevant visual threatening stimuli such as angry facial expressions (Bar-Haim et al., 2007; Kappenman et al., 2014). However, it is important to note that these effects were only demonstrated in the first two blocks of the task when the visual target was presented in the right visual field (Brosch et al., 2008). The lack of significant findings in the third block possibly suggests the presence of habituation effects within the dot-probe paradigm due to the repetition of threat-related stimuli without any related consequences (Brosch et al., 2008). In addition, the presence of lateralisation effects within the dot-probe paradigm are common, in that RTs are often faster when targets are presented on the right compared to left visual hemifield (Bradley, Mogg & Millar, 2010), possibly due to participants adopting strategies to attend to one side of the screen over the other (Bradley, Mogg et al., 2010). Overall, the behavioural results by Brosch et al. (2008) support the theories which suggest that angry voices are preattentively perceived with higher salience than neutral voices, and therefore selectively and involuntarily capture attention and facilitate task-relevant performance (Dalton & Lavie, 2004; Hansen & Hansen, 1988; Ohman & Mineka, 2001).

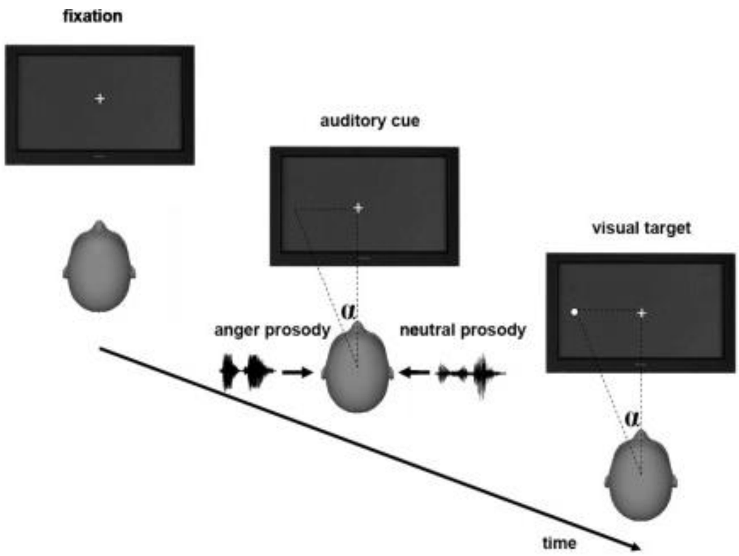

Figure 4.2

A Graphical Representation of the Emotional Dot-Probe Paradigm Used By Brosch Et Al (2008). Reprinted from Cognition, Volume 106/Issue 3, Brosch, Grandjean, Sander & Scherer, “Behold the voice of wrath: Cross-modal modulation of visual attention by anger prosody”, Pages 1497-1503, Copyright (2008), with permission from Elsevier.

There have not been any studies which have employed similar dot-probe paradigms with vocal emotional cues in adolescents. This is important because it is essential to understand the effect of vocal emotion on attentional mechanisms in healthy adolescents before exploring atypicalities in these processes in adolescents with inattentive/hyperactive traits (see Chapter Five). Existing studies using dot-probe paradigms in adolescents have utilised facial expressions (e.g., happy, angry and neutral) as cueing stimuli, and reported that congruent trials, in which happy or angry faces were spatially congruent to the target, consistently resulted in faster RTs compared to incongruent trials, in adolescents (Monk et al., 2006; Telzer et al., 2008). This finding is consistent with the pattern of findings in the adult literature (Brosch et al., 2008). These results from dot-probe paradigms recruiting adolescent and adult participants suggest that facial emotional cues (happy and angry) are effective at capturing and facilitating VSA in younger and older participants. It would therefore be reasonable to hypothesise that vocal emotional cues (happy and angry) may also be effective at capturing VSA in adolescents, in consistency with findings in adults (Brosch et al., 2008).

There are, however, some limitations with behavioural studies employing dot-probe tasks to examine the effect of vocal emotion on VSA. Firstly, it is not fully understood which mechanisms account for the effect of emotion on RTs in dot-probe tasks (Brosch et al., 2008). There are several possible mechanisms which may account for this effect. First, within congruent trials, it is possible that faster RTs occur because the salient (i.e. threatening) cues involuntarily capture attention and thus, facilitate attention towards the visual target. This is consistent with theories which suggest threatening stimuli facilitate the allocation of attention to aid survival (Dalton & Lavie, 2004; Hansen & Hansen, 1988; Ohman & Mineka, 2001). An additional explanation is that during incongruent trials slower RTs are observed because salient threatening cues similarly capture attention, but instead distract attention from the visual target, due to difficulties to disengage from the spatial location of the incongruent threatening cue stimulus (Koster et al, 2004; Van Rooijen et al., 2017).

The distracting effect of vocal emotion has been evidenced in studies employing auditory distraction tasks (Kattner & Ellermeier, 2018; Kattner et al., 2024) and dichotic listening tasks (Peschard et al., 2017; Sander et al., 2005). For example, Kattner & Ellermeier (2018) employed an auditory distraction task in which adults were asked to recall a list of words whilst listening to irrelevant vocal emotional stimuli. Kattner & Ellermeier (2018) reported the poorer recall of words when an angry voice, as opposed to a neutral voice, was presented as a distractor. Moreover, in dichotic listening tasks, adults were presented with vocal cues (neutral and angry) simultaneously in the left and right ear and asked to discriminate the gender of voice in the specified attended ear (Peschard et al., 2017; Sander et al., 2005). These studies, employing dichotic listening tasks, found adults were slower to discriminate the gender of an attended angry voice compared to an attended neutral voice (Peschard et al., 2017; Sander et al., 2005). This is thought to reflect the involuntary capture of vocal anger which holds limited attentional resources later in processing (Hughes, 2014; Sander et al., 2005). Taken together, behavioural studies provide evidence for the involuntary capture of preattentively processed vocal emotion (specifically anger) in adults. In addition, the literature indicates that irrelevant vocal emotion cues can have a facilitating or distracting effect on task performance (i.e., RTs), depending on task design.

To more reliably measure the involuntary capture of irrelevant vocal emotion, prior to effects of vocal emotion on later attention allocation to a target, ERP studies using spatial cueing paradigms have explored the neural underpinnings of the capture of VSA by irrelevant auditory (Störmer et al., 2009) and vocal emotional (Brosch et al., 2009) cues by targeting the P1 component. As discussed in Chapter One (Page 35), the P1 component (90-190ms) is generated in the visual cortex and is argued to reflect early spatial attention allocation to a visual target (Luck et al., 2000; Mangun et al., 1998; Mangun & Hillyard, 1991). The P1 has been found to be enhanced when a visual stimulus appears in an already attended location (Luck et al., 2000; Mangun et al., 1998; Mangun & Hillyard, 1991). For example, one variation of a cueing paradigm presented two Gabor patches simultaneously as visual targets on the left and right hemisphere, which varied in contrast and orientation (vertical/horizontal; Störmer et al., 2009). On each trial an auditory cue (83ms of pink noise played in either the left or right ear) preceded the presentation of the Gabor patches by 150ms. Participants were instructed to indicate the orientation of the Gabor patch highest in contrast by pressing the corresponding button. ERPs were time-locked to the onset of the target stimuli, and were analysed at sites PO7/ PO8, in the ventral occipital-temporal visual cortex (Störmer et al., 2009). The authors reported that when participants made errors in judging a visual stimulus as being higher in contrast, this response corresponded to a larger P1 amplitude to the visual target in congruent trials (Störmer et al., 2009). This research was presented as evidence that irrelevant auditory cues can alter early perceptual processing in the visual cortex, relating to altered perceptual judgements of visual appearance (Störmer et al., 2009) in accordance with theories which suggest auditory cues automatically capture VSA (Murphy et al., 2017).

Similarly, the ERP study by Brosch et al (2009) explored the capture of VSA by vocal emotion in adults in a dot-probe paradigm, by targeting the P1 component. This dot-probe paradigm was similar to that in the behavioural study discussed earlier in this chapter (See Figure 4.2; Brosch et al., 2008), but the task was optimised for an EEG study. For example, the authors included an upward or downward pointing triangle as the target stimulus instead of left/right button press response (Brosch et al., 2009). Participants were instructed to press the 'B' key response button only when the orientation of the triangle was pointed in either the upward or downward position. These were GO conditions which included in only 10% of trials. The other 90% of trials were 'no-go' trials, which did not require a button-press response to avoid motor preparation influencing the EEG data (Brosch et al., 2009). Results showed adults displayed a larger P1 response to trials preceded by angry vocal cues which appeared in a spatial location congruent to the visual target as opposed to trials preceded by an angry vocal in a spatial location incongruent to the target (Brosch et al., 2009). However, interestingly these effects were driven by targets presented in the left visual field, suggesting the presence of lateralisation effects on the P1 response. Since targets in the left visual field are processed in the right hemisphere, this is in line with evidence which suggests some acoustical features of emotional prosody involve the right hemisphere more so than left hemisphere (Brosch et al., 2009; Ross & Monnot, 2008). Overall, the effect of anger on an enhanced P1 response supports the idea that angry vocal cues rapidly and automatically capture attention (Dalton & Lavie, 2004; Hansen & Hansen, 1988; Ohman & Mineka, 2001).

In support, a more recent ERP study incorporated vocal emotions as cueing stimuli in a spatial cueing paradigm in adults (Zimmer et al., 2022). Within this paradigm, the authors incorporated bi-modal (faces and voices together) emotional cueing stimuli (fearful and neutral) for 1000ms to just one hemifield (left or right) on each trial followed by a visual target (a white up/downward pointing triangle). Participants were instructed to specify the upward or downward point of the triangle. Behavioural results showed faster RTs to targets preceded by fearful compared to neutral bi-modal cues. In addition, participants displayed faster reaction times in congruent trials preceded by fearful cues compared to incongruent trials preceded by fearful cues (Zimmer et al., 2022). Moreover, in line with Brosch et al (2009), results showed a larger P1 component in response to congruent compared to incongruent trials (Zimmer et al., 2022). Finally in congruent trials only, the P1 was enhanced after fearful compared to neutral cues, suggesting that negative emotional cues facilitate the capture of attention (Zimmer et al., 2022). However, as the cues were bimodal, it is not known whether this effect was due to emotion being presented within the face or voice separately or a combination of both modalities.

There have not yet been any studies which have incorporated other vocal emotion types (i.e., happy) in spatial cueing paradigms to explore the effects of vocal emotion on the capture of VSA. Research which has explored the effect of happy emotions in other modalities (i.e., facial stimuli) on VSA, have not reported a robust effect of happiness on the P1 component. For example, in dot-probe studies using happy facial expressions, the congruency effect on the P1 was not found when happy faces were used in valid trials (Pourtois et al., 2004; Santesso et al., 2008). In contrast, Pfabigan and colleagues (2014) found enhanced P1 amplitudes in response to congruent trials preceded by happy compared to neutral facial expressions in female adults. The same finding was not reported in male participants, possibly suggesting that happy stimuli are more likely to capture attention in females due to evolutionary or socio-cultural influences (Pfabigan et al., 2014). It is important to explore the effect happy vocal cues may have on the involuntary capture of VSA, as there may be different mechanisms associated with the preattentive processing of negative (i.e. angry) and positive (i.e. happy) stimuli. For instance, there has been some evidence that the striatum, a subcortical area of the brain within the basal ganglia, may be specifically involved in the prioritisation of positive stimuli relating to reward-seeking behaviours (Ernst, 2014; Bogert et al., 2016). The role of the striatum in processing positive stimuli is particularly evident during adolescence, a period characterised by a heightened sensitivity to rewards (Galván, 2013). In support, fMRI studies using monetary reward tasks have demonstrated enhanced activation in the striatum in response to reward-related outcomes (Ernst et al., 2005; Galván et al., 2013; Silverman et al., 2015) and happy faces (Mueller et al., 2017) in adolescents compared to adults.

Limited research has employed spatial cueing paradigms to explore the involuntary capture of VSA by vocal emotion (angry and happy) by measuring the P1 response in children or adolescents. Although the P1 amplitude, as a measure of early visual attention, does not change significantly with age (Perchet et al., 2007), it is not known whether vocal emotion enhances the P1 response in children and adolescents, in the same way as it does in adults. Understanding this is important because it serves as a basis for exploring atypicalities in the capture of attention by vocal emotion in children and adolescents with ADHD (see Chapter Five). Moreover, thus far, no studies have employed spatial cueing paradigms to explore differences in the capture of irrelevant auditory cues in individuals with ADHD compared to healthy controls. However, Perchet et al. (2001) employed a unimodal visual cueing task, known as the Posner Paradigm (see Figure 1.2 on page 38) in 24 children with ADHD and 13 controls (aged 6-11-years) to explore atypicalities in the capture of irrelevant visual cues on VSA in ADHD. Within the Posner paradigm, a visual target (a red star) was presented in either the left or right hemifield for 400ms, preceded by a visual cue (a yellow rectangle) which appeared for 500ms in the same potential left or right locations as the target. In congruent and incongruent trials, the cue remained on the screen until after the target disappeared. Children were instructed to press the left or right buttons corresponding to the hemifield the target was presented in as quickly as possible. Children were also rewarded if they made less than three errors in the task. When exploring ERP components time-locked to the target stimuli, healthy children (without ADHD) exhibited a significantly larger P1 component to congruent trials compared to incongruent trials, consistent with research using spatial cueing paradigms in adults (Störmer et al., 2009). Compared to controls, children with ADHD showed a reduced P1 component in response to congruent compared to incongruent trials (Perchet et al., 2007). This finding suggests an atypical neural response to the use of cues to prime attention in children with ADHD. However, as of yet, no studies have employed a spatial cueing paradigm with vocal emotions to explore atypicalities in the capture of VSA in ADHD.

In parallel to exploring the involuntary capture of attention by vocal emotion by measuring effects of emotion on early visual processing (i.e., the P1), previous research has additionally reported effects of vocal emotion on a later P3 component in spatial cueing paradigms. As discussed in Chapter One (page 38) the P3 is typically observed between 300 and 600ms at midline centro-parietal sites to task-relevant stimuli in adults and is argued to reflect the allocation of attentional resources (Wang et al., 2022; Li et al., 2018; Broyd et al., 2012). Previous dot-probe and emotional spatial cueing paradigms have measured the P3 component to a subsequent target (Li et al., 2018; Wang et al., 2022). For example, in a dot-probe task which incorporated happy-neutral and sad-neutral facial expressions as cueing stimuli, the P3 amplitude was larger to the target when it was preceded by congruent sad faces compared to congruent neutral faces in adults with Major Depressive Disorder (Li et al., 2018). The authors suggested this enhanced P3 response reflects difficulties in disengaging attention from the irrelevant negative stimuli in individuals with depression, after the initial capture of attention (Li et al., 2018). In support, in a spatial cueing paradigm with environmental sounds as cueing stimuli, a larger P3 amplitude was reported to congruent trials preceded by positive cues (e.g., baby laugh/bird song) compared to congruent trials preceded by neutral cues (e.g., clock ticking; Wang et al., 2022). The authors suggested more attentional resources were allocated and maintained to the irrelevant positive emotional sounds compared to the irrelevant neutral sounds (Wang et al., 2022). Additionally, the P3 was larger to incongruent compared to congruent trials suggesting that more attentional resources were employed to disengage attention from the irrelevant (as opposed to the congruent) spatial location in order to attend to the relevant visual target (Wang et al., 2022). Taken together, the P3 component appears to be a neural marker of the capture of attention by vocal emotion, specifically reflecting the greater recruitment of attentional resources to disengage from the irrelevant emotional stimulus and engage to the target. However, as the P3 occurs later in processing, it reflects downstream effects of the capture of attention.

The Present Study

Since the studies by Brosch et al (2008, 2009) are the only studies which have measured the effects of vocal emotion on VSA, the design of the novel ESC within the present thesis was based on these dot-probe paradigms (Brosch et al; 2008, 2009). However, the task was modified to present one vocal cue at a time to make comparisons between the effect of vocal cue types (e.g., happy, angry, neutral) on preattentive processing stages and VSA. The inclusion of one cueing stimulus at a time in our task is consistent with the Spatial Cueing Paradigms by Spence and Driver (1997), McDonald et al., (2000), and Zimmer et al., (2022). In addition, as we aimed to recruit a sample of adolescents with traits of ADHD in a later ERP study (see Chapter Five), the task was designed to sustain the attention of participants who are likely to present with difficulties in sustained attention. Finally, the task was optimised for the purpose of an ERP study as our aim was to isolate neural markers of both the preattentive processing of vocal emotion (Lin et al., 2022; Gädeke et al., 2013) and VSA as in previous cueing paradigms (Brosch et al., 2009; Perchet et al., 2001). With these goals in mind, a more thorough explanation of the rationale for the inclusion of all methodological design features of the ESC paradigm is presented in Appendix D.

The primary aim of the present study was to optimise and validate the ESC task on a behavioural and neural level in adults, for use in a later ERP study exploring preattentive and attentive vocal emotion processing in adolescents with ADHD (Chapter Five).

Therefore, the study aims were as follows:

I) To investigate the effects of vocal emotion on the involuntary capture of attention in adults on a behavioural level by measuring RTs to the visual target.

II) To isolate the P2 component to the vocal cue as a neural marker of preattentive vocal emotion processing in adults within the ESC paradigm, by exploring the effect of emotion on the P2.

III) To isolate the P1 component as a neural marker of the early visual processing of the target within the ESC paradigm in adults, by exploring the effect of emotion and congruency on the P1.

IV) To isolate the P3 component as neural marker of the later attentional allocation to the target within the ESC paradigm in adults, by exploring the effect of emotion on and congruency on the P3.

Consistent with findings from previous dot-probe studies (Brosch et al., 2008; 2009), spatial cueing paradigms (Spence & Driver, 1997; Zimmer et al.., 2022), and ERP preattentive paradigms in adults (Gadeke et al., 2013), the following predictions were made:

i) Reaction Times:

a) There will be a main effect of congruency on RTs. Adults would be faster to respond to congruent compared to incongruent trials.

b) There will be a main effect of emotion on RTs. Specifically, participants will respond faster to trials preceded by emotional (angry and happy) compared to neutral voices. There will be an interaction effect between congruency and emotion on RTs. Specifically, adults will respond faster to congruent trials preceded by angry voices compared to congruent trials preceded by neutral voices.

ii) Cue-locked ERPs.

a) There will be a main effect of emotion on the P2 amplitude to the vocal cue. The P2 to the vocal cue would be larger to angry and happy voices compared to neutral voices.

iii) Target-locked ERPs

a) There will be a main effect of congruency on the P1 and P3 amplitude to the visual target. The P1 amplitude to the visual target will be larger to congruent compared to incongruent trials. The P3 amplitude will be larger to incongruent compared to congruent trials.

b) There will be a main effect of emotion on the P1 and P3 amplitude to the visual target. The P1 and P3 amplitude will be larger to targets preceded by angry and happy vocal cues compared to neutral vocal cues.

c) There will be an interaction effect between emotion and congruency on the P1 to the visual target. The P1 and P3 amplitude would be larger to congruently cued visual targets preceded by angry and happy vocal cues compared to congruently cued targets preceded by neutral vocal cues.

A first exploratory aim of the present study was to explore the presence of lateralisation and habituation effects on target processing (e.g., RTs, P1 and P3). A second exploratory aim was to investigate associations between RTs and ERP components with behavioural and psychopathology characteristics (ADHD traits, depressive symptoms) for consistency with our later ERP study employing the ESC in adolescents[8].

---

[8] Methods and results relating to this aim are presented in Appendix B and H as they are not the focus of this study.

To meet these aims, the present study consisted of four pilot experiments. Experiment one to three were preliminary pilot experiments, which aimed to test the paradigm behaviourally in the first instance in a small sample of adults. This aimed to validate the task in terms of producing the expected effects of congruency and emotion on reaction times as reported in previous spatial cueing paradigms (Spence & Driver, 1997; Brosch et al., 2008). The goal was to modify the task if necessary if the expected effects were not present. Next, once the task design was optimised and validated based on the results from the three behavioural pilot experiments, experiment four aimed to validate the final version of the paradigm on a neural level.

4.3 General Method

Participants.

In Experiment One, 12 adults aged between 18 and 35 (3 males) were recruited from an opportunity sample of students and staff at the University of Central Lancashire. A different set of 12 adults (aged 18-35 years) were recruited for Experiment Two. Datasets from Experiment One and Two were combined in Experiment Three, resulting in a total of 24 participants. This meets a priori power calculations (GPower software) which indicated that a sample size of 17 adults was required to detect these effects, with an effect size of 0.44 (α error probability of .05). Due to the pilot nature of experiments one to three, data on behavioural characteristics of participants were not collected.

In Experiment Four, thirty-six adults were recruited from the University of Central Lancashire through word of mouth and the participant recruitment pool. One participant was excluded due to being an outlier (i.e., errors and RT outliers exceeded 25% of the participants data). Participants with a minimum of 70% correct and artefact free trials were included in ERP analyses (see page 151 for details). Complete behavioural and ERP data were available from 25 adult participants (mean age= 24 years, SD= 5.73, age range 18-35, 8 males). This meets apriori power analyses using Gpower and Pangea software (Westfall, 2015) which indicated that a sample size of a) 16 participants was necessary to achieve effects of congruency on the P1 in adults, with a power of 0.947, and b) 25 participants was necessary to achieve predicted effects of emotion on the P2 component, with a power of 0.979. All participants had normal hearing (self-reported), no neurological or psychiatric diagnoses, and no visual impairments.

The Vocal Stimuli.

The study utilised six non-linguistic vocal stimuli (happy, angry and neutral) developed by Maurage et al (2007) vocalising an '/a/' sound by a female speaker (see Table C.1 in Appendix C for acoustic properties and accuracy rates from the validation study). All stimuli were standardized on mean intensity (77dB), duration (700ms), and rise and fall time (5-20ms) as measured by Praat software (Boersma & Weenink, 2024), and were validated to represent happy, angry, and neutral vocal expressions (See Chapter Three). In Experiment One, the maximum perceived intensity (dB) of these stimuli as measured by a Decibel Meter (Max Measure MM-SM01), with a frequency weighting of A, was between 64 and 68dB. In Experiment Two and four, the intensity increased to 77dB. A plastic cup was used to simulate the pinna against the left or right headphone, whilst taking these readings for each stimulus.

The Emotional Spatial Cueing Paradigm Design.

In all experiments, adults took part in the Emotional Spatial Cueing (ESC) Paradigm (See Figure 1.4, page 48). There were three blocks with 96 trials each, making a total of 288 trials. A total of 50% (144) of trials were congruent and 50% were incongruent. An even number (96) of angry, happy, and neutral vocal cues were presented randomly within all trials. Participants were sat 60cm away from the screen and were given the following instructions, verbally and in writing: *'Your task is to catch the stars which will appear in the left or right side of the screen. You will also hear some voices – try to ignore these. You should respond to the star using the index and middle finger of your dominant hand, by pressing the left or right arrow key button. For example, if the star appears on the left-hand side of the screen, you should press the left arrow key button, and if the star appears on the right-hand side of the screen, you should press the right arrow key button. Make sure your eyes stay fixated on the* *cross and try to respond as fast as you can. Try not to make any errors.'* Each block took roughly 5 minutes to complete. Participants were given the option to take a break after each block. Participants took part in 8 practice trials at the beginning of the task to check their understanding of the instructions. Participants were given the opportunity to ask any questions after the practice trials. After checking participants understood the instructions, adults proceeded to the first main experimental block. Button press responses were collected via Psychopy software (Pierce et al., 2019).

The procedure of one trial of the ESC paradigm comprised of the following: (1) a fixation cross appeared on the screen for variable amount of time between 500 and 1000 milliseconds, (2) a vocal cue stimulus portraying an angry, happy or a neutral voice was then presented in one ear (left or right) for a duration of 700ms, (3) after 750ms from the onset of the vocal stimulus (SOA), the target (a yellow star) appeared on the left or the right side of the screen for 100ms, (4) participants responded using their index and middle fingers on their dominant hand by pressing the left and right buttons, respectively, on the keyboard to indicate the location of the target (5) the interstimulus interval (ITI) consisted of 2000ms seconds in Experiment One and 1000ms in Experiment Two and Four. The visual target was presented in either the congruent or incongruent hemifield to the preceding vocal cue. Therefore, a congruent trial occurred when participants heard the voice and saw the target in the same hemifield. An incongruent trial occurred when participants heard the voice and saw the target in the opposite hemifield. The presentation of congruent and incongruent trials, and the presentation of angry, happy and neutral trials was random but equal. The vocal cues were presented to either the left or right ear via headphones (Sennheisser, HD 380 pro). The within subject independent variables consisted of congruency (congruent and incongruent) and emotion (angry, happy, and neutral) in the vocal stimuli. The dependent variable was the reaction time to correctly respond to the visual target in all experiments. In Experiment Four only, additional dependent variables included ERP components: Cue-P2, the Target-P1 and the Target-P3 (see page 152 for details)

Procedure.

In all four experiments**,** informed written consent was gathered from all participants. The researcher informed participants at the start of the study that they could withdraw their participation at any time. In all experiments, participants completed the ESC Task. However, in Experiment Four only, the researcher set up the EEG equipment prior to the participants completing the ESC task. Participants were encouraged to sit as still as possible and keep eye movements to a minimum throughout the experimental task. After completion of the experimental task, participants completed the questionnaire measures for behavioural and psychopathology characteristics[9]. At the end of the study, all participants were debriefed verbally and in writing. This study was approved by the UCLan Psychology ethics committee (Science 0037).

Behavioural Data Processing.

First, in all experiments, all reaction time data were pre-processed to include only correct responses within the range of 200-1000ms. Second, because Experiments One and Two employed a small sample size, further outliers in these experiments were identified via boxplots created for each participants dataset. These outliers were identified as correct

[9] See Appendix B for details of these measures employed to meet our exploratory aim. Because the purpose of this experiment was to validate the task on a neural level, and behavioural and psychopathology characteristics were not entered as covariates in analyses in the present experiment, details of these measures and analyses are included in Appendix H.

responses which were 1.5 times faster or slower than correct response which fell in the interquartile range (i.e., responses which fell above the third quartile or below the first quartile) The total percentage of data removed due to errors and RT outliers from the task was 7.44% in Experiment One, 5.17% in Experiment Two. In Experiment Four, a more lenient approach was taken to deal with outliers due to including a larger sample size. Specifically, outliers were only removed if they fell outside the range of 200-1000ms, consistent with the literature (Brosch et al., 2008). A small percentage of data (2.16%) was removed in Experiment Four due to errors or outliers in the RT data.

Electrophysiological Recording and Processing

In Experiment Four only, EEG data were recorded using an electrode cap (Quick cap) from 64 electrodes based on the 10-20 system (see E.1 in Appendix E for the EEG montage) using Curry 7 software. Data were analysed using Neuroscan (Version 4.5). The EEG data were sampled at 1000 Hz with an online band pass filter at 0.01 to 100Hz using AC procedure. The reference and ground electrodes were embedded in the electrode cap. Vertical electro-oculogram (vEOG) was recorded, with one electrode placed directly beneath the left eye. Horizontal electro-oculogram (hEOG) were recorded from two electrodes placed next to the left and right eyes. Impedance for vEOG, hEOG, reference and 64 electrodes were kept below 5KΩ during recording.

Two ERP epochs were defined: 1) from the onset of the vocal cue stimuli and 2) from after the onset of the visual target stimuli within the ESC paradigm. The vocal cue epoch was defined as -100ms pre-stimulus to 700ms post stimulus. The target epoch was defined as -50ms pre-stimulus to 500ms post stimulus. Both epochs were filtered offline with high-pass band edge frequency of 0.2Hz and low-pass band frequency of 15Hz. An ocular artefact reduction procedure (Semlitsch et al., 1989) based on vEOG and hEOG activity was used to remove the influence of blinks and other eye movements. Epochs were rejected for both vocal cue and target if amplitudes exceeded ±150 µV at any EOG or scalp sites included in analyses. Epochs for the target were additionally rejected if participants responded incorrectly or outside of a window between 200 to 1000ms. For the vocal cue epoch, average ERPs were calculated for each emotion type (angry, happy, neutral). A baseline-to-peak mean amplitude method was used to calculate the ERP components (Picton et al., 2000). This method involved averaging the amplitude over a specified time window and sites for a component, and then measuring the difference in voltage between the peak and the baseline level, which is equal to zero (Picton et al., 2000). The following components were targeted. The P2 component to the vocal cue was targeted with a time window between 130 and 250ms at fronto-central sites (F1, FZ, F2, FC1, FCZ, FC2, C1, CZ, C1). The P1 to the visual target was observed between a window of 100-170ms at parietal-occipital sites (PO7,PO8,PO3,PO4,POZ,O1,O2,OZ), and the P3 to the visual target was observed between 290 and 450ms at parietal sites (PZ,P1,P2). Mean amplitude was initially calculated for each individual site. Following this, the mean amplitude for each ERP component was calculated as a combined score for a number of groups of electrode sites (scalp regions), in order to increase the reliability of the measurement (Dien & Santuzzi, 2005). The inclusion of these sites and time windows were informed from the literature (cue-P2: Gadeke et al., 2013; Jessen & Kotz, 2011; Target-P1: Störmer et al., 2009; Brosch et al., 2009; Target-P3: Wang et al., 2022; Li et al., 2018; Broyd et al., 2012) and visual inspection of the individual averages.

A minimum of 70% artefact free epochs for each condition per participant were used for calculating ERP averages. If a participant presented with less than 70% clean epochs, they were excluded from analyses. The mean and SD of the number of epochs included in the analyses for each vocal cue condition were as follows: angry (M=90.80, SD=5.61), happy (M=90.56, SD=5.98), neutral (M=89.80, SD=6.42). The mean and SD of the number of epochs included in the analyses for each target condition were as follows: angry (M=90.40, SD=4.65), happy (M=91.16, SD=3.53), neutral (M=91.40, SD=3.38), congruent (M=136.08, SD=5.63), incongruent (M=136.88, SD=5.63).

4.4 Experiment One

Results

Table 4.1 shows descriptive statistics of RTs for correct response (after outliers removed) for trials across congruency and emotion conditions in Experiment One. A 2 x 3 repeated measures ANOVA was conducted to explore the effect of emotion and congruency on mean RTs. There was no significant main effect of emotion ($F(2,22) = 0.94$, $p = .403$) or congruency ($F(1,11) = 4.60$, $p = .055$) on RTs. However, as the main effect of congruency was close to significance, this effect may have been due to a lack of power from the small sample size. There was no emotion x congruency interaction effect on RTs ($F(2,22) = 0.04$, $p = .958$). Figure 4.3 presents a bar graph showing mean RTs for targets preceded by angry, happy, and neutral vocal cues in congruent and incongruent trials.

When repeating the analysis with the data from block one of the task only, results showed that the congruency effect reached significance ($F(1,11) = 5.74$, $p = .035$). Within block one, participants were slower to respond to incongruent compared to congruent trials, consistent with the findings reported by Brosch et al (2008). There was no significant main effect of emotion ($F(2,22) = 0.14$, $p = .872$) or emotion x congruency interaction effect on RTs ($F(2,22) = 0.02$, $p = .980$) within block one. When repeating the analysis with data from block 2 only, there were no significant main effects of congruency ($F(1,11) = 3.26$, $p = .098$), emotion ($F(2,22) = .983$, $p = .390$) or an emotion x congruency interaction effect on RTs ($F(2,22) = 0.81$, $p = .459$). This pattern of results did not change when repeating the analyses with data from block 3 of the task only (all $ps > .05$).

Table 4.1

Mean (SD) RTs for Correct Responses to Targets Preceded by Vocal Cues Differing by Emotion (Angry, Happy, Neutral) and Congruency (Congruent, Incongruent) Conditions in Experiment One

	Congruent	Incongruent	Total
Angry Voice	319 (38.20)	323 (40.29)	321 (11.24)
Happy Voice	316 (33.34)	322 (39.25)	319 (10.26)
Neutral Voice	316 (33.03)	322 (26.57)	319 (9.94)
Total M(SE)	317 (9.96)	322 (11.05)	320 (8.09)

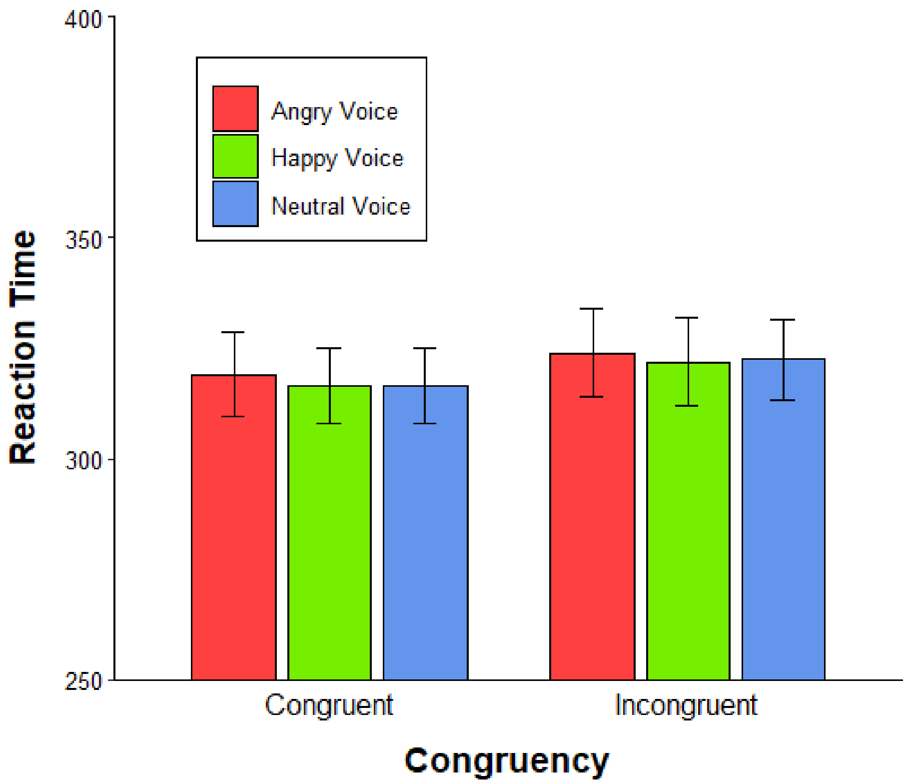

Figure 4.3

Bar Graph Showing Mean RTs (in ms) to Targets in Emotion (Angry, Happy And Neutral) and Congruency (Congruent, Incongruent) Conditions in Experiment One.

Interim Discussion

First, based on data from block one in Experiment One, the ESC paradigm produced the expected congruency effects on RTs, consistent with that reported by Brosch et al (2008). This suggests the first block of the ESC paradigm is a valid measure of the capture of VSA by irrelevant auditory cues, as in prior cueing paradigms (Spence & Driver, 1997). A more thorough interpretation of this finding is discussed on page 180. Second, the present paradigm did not demonstrate the expected emotion, or congruency by emotion interaction effects on RTs. One possible reason that the ESC paradigm did not produce the expected effects of emotion may have been due to the saliency of the vocal stimuli employed. Although the vocal stimuli were measured as 77dB via Praat software, in line with previous literature

(Brosch et al., 2008, 2009; Chronaki et al., 2015), the volume participants objectively perceived the stimuli was between 64 and 68 decibels as measured by a decibel metre. Previous research reported that the loudness (dB) of an auditory stimulus can affect effects on visual cognitive tasks. For example, within a Stroop paradigm, when auditory stimuli were played at 70db, this significantly reduced cognitive performance compared to when the same auditory stimuli were played at 50db (Schlittmeier et al., 2014). Therefore, it is possible that a disadvantage of this version of the task was due to the loudness (dB) at which the vocal cueing stimuli were objectively perceived. To address this, in Experiment Two, a simple change to the task was made to increase the loudness of the vocal cueing stimuli.

Another possible reason the paradigm did not produce the expected effects of emotion could be related to the error rate. Specifically, as the error rate within the present study was smaller (1.2%) than in previous studies (1.4-1.6%), this may indicate that the present paradigm is too easy. Therefore, one way to increase the difficulty of the task without significantly modifying key elements affecting congruency effects or de-optimising the task for EEG analysis was to reduce the ITI from 2000ms to 1000ms. This modification to the ITI in the paradigm was deemed appropriate for the subsequent experiments in this study.

In conclusion, as effects of emotion and interaction effects were not evident in Experiment One, two adaptations to the task design were made to optimise the task. First, the intensity (dB) of the vocal stimuli increased to 77dB. Second, the ITI decreased to 1000ms. Experiment Two aimed to pilot the optimised ESC paradigm on a behavioural level with a different sample of 12 adults.

4.5 Experiment Two

Results

Table 4.2 shows descriptive statistics of RTs for correct trials across congruency and emotion conditions. A 2 x 3 repeated measures ANOVA was conducted to explore the effect of emotion and congruency on mean RT scores. The results showed no significant main effect of either emotion ($F(2,22) = 2.53$, $p = .103$) or congruency ($F(1,11) = 1.19$, $p = .298$). There was also no significant emotion x congruency interaction effect on RTs ($F(2,22) = 1.13$, $p = .340$). When repeating the analyses with the data from block one only, results revealed there was no significant main effect of congruency ($F(1,11) = 1.89$, $p = .196$), but there was a significant main effect of emotion on RTs ($F(2,22) = 6.07$, $p = .008$). Post hoc tests revealed that participants were slower to respond to trials preceded by angry compared to happy voices ($MD = 9.57$, $p = .007$). Finally, there was a close to significance trend for an emotion x congruency interaction effect on RT: ($F(2,22) = 3.09$, $p = .066$). When repeating the analyses with the data from block 2 only, there was no significant main effect of congruency ($F(1,11) = 0.59$, $p = .457$) or main effect of emotion ($F(2,22) = 0.05$, $p = .940$) or emotion x congruency interaction effect ($F(2,22) = 0.77$, $p = .476$) on RTs. When repeating the analyses with data from block 3 only, there was no significant main effect of congruency ($F(1,11) = 1.38$, $p = .266$) or main effect of emotion ($F(2,22) = 0.81$, $p = .459$) or interaction effect ($F(2,22) = 1.37$, $p = .273$) on RTs. Figure 4.4 presents a bar graph showing mean RTs for correct responses for emotion by congruency conditions.

Table 4.2

Means (SD) RTs to Targets Preceded by Vocal Cues Differing by Emotion (Angry, Happy, Neutral) and Congruency (Congruent, Incongruent) Conditions in Experiment Two.

	Congruent	Incongruent	Total
Angry Voice	321 (46.88)	328 (60.23)	324 (15.86)
Happy Voice	317 (47.79)	323 (61.88)	320 (15.74)
Neutral Voice	322 (50.82)	324 (55.59)	323 (15.27)
Total M(SE)	320 (13.96)	325 (17.03)	320 (8.09)

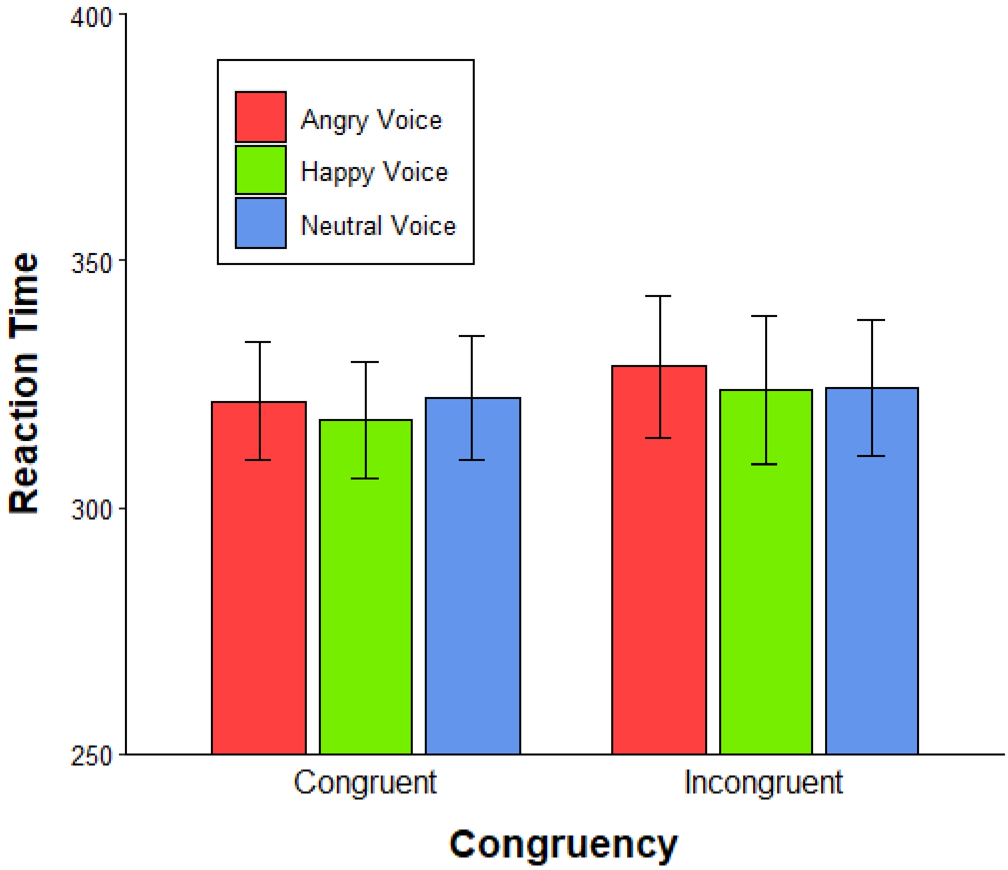

Figure 4.4

Bar Graph Showing Mean RTs (in ms) to Targets Differing by Emotion (Angry, Happy, Neutral) and Congruency (Congruent, Incongruent) in Experiment Two

Interim Discussion

The aim of Experiment Two was to examine whether the small adaptations to the design of the ESC paradigm (i.e., increasing the dB to 77dB and decreasing the ITI to 1000ms) increased the validity of the task in terms of emotion and interaction effects on RT scores, whilst keeping congruency effects reported in Experiment One in-tact.

Firstly, results demonstrated a significant effect of emotion on RT scores in block one of the task, possibly suggesting the design adaptations made to the task resulted in increased sensitivity of the task to detect emotion effects. Specifically, adults presented with slower RTs to trials preceded by angry compared to happy vocal cues. The direction of the emotion effects was not in line with previous studies employing dot-probe tasks, which found that participants responded faster to targets preceded by spatially congruent angry voices compared to targets preceded by spatially congruent neutral voices (Brosch et al., 2008). However, the present findings are consistent with results from auditory distraction tasks (Kattner & Ellermeier, 2018; Kattner, 2023) and dichotic listening (Peschard et al., 2017; Sander et al., 2005) tasks, which similarly report irrelevant anger voices have a distracting effect on performance. A thorough interpretation of this finding is discussed on page 180.

Secondly, the present study surprisingly did not reveal any significant effects of congruency on RTs. As no changes were made in terms of increasing or decreasing possible congruency effects in experiment two, it is possible that this is due to a lack of power from recruiting a small sample of adults. Finally, the present study revealed a close to significance emotion by congruency interaction effect on RTs. Similarly, it is possible this did not reach significance due to a lack of power from a small sample size. Due to the small number of participants tested within experiment two, further analyses with data from all 24 participants from both Experiment One and Two was necessary to more thoroughly test the validity of the ESC task. This further work is presented in the next experiment.

4.6 Experiment Three

Results

Table 4.3 presents descriptive statistics of RTs for correct responses to targets for emotion and congruency conditions from all 24 adults who took part in Experiment One and Two. Descriptive statistics are presented for the overall task (blocks 1-3), and for just block one. A mixed 2 x 3 x 2 ANOVA, with emotion and congruency as within-subject variables, and task version (I.e., one and two) as the between-subject variable, was conducted. The dependent measure was RT scores. Results revealed there were no significant main effects of emotion (F (2,44) = 2.85, p = .069), or congruency (F (1,22) = 2.05, p=.058) on RTs, however both were close to significance. In addition, there was no significant main effect of version on reaction times (F (1,22) = 0.03, p=.877). Finally, results showed there were no significant interaction effects between emotion and congruency and version on RTs (all ps >.05).

When the same analyses were repeated with the data from block one only, there were significant effects of both congruency (F (1,22) = 6.37, p=.019) and emotion (F (2,44) = 4.09, p=.024) on RTs. Participants were faster to respond to congruent compared to incongruent trials. Additionally, post hoc tests revealed participants were significantly slower to respond to trials preceded by angry compared to happy voices (MD=5.51, p=.043). There was no significant main effect of version (F (1,22) = 0.15, p=.703) on RTs, nor any further significant interaction effects between congruency, emotion and version (all ps>.05). Finally, to explore lateralisation and habituation effects, a 2 (congruency) x 3 (emotion) x 2 (target hemifield) x 3 (block) ANOVA was conducted with RT scores as the dependent measure. There was no significant main effect of hemifield (F (1,23) =2.70, p=.114) or block (F (2,46) =0.04, p=.961) on RT scores. There were no other significant interaction effects (all ps >.05).

Table 4.3

Means (SD) RTs to Targets Preceded by Vocal Cues in Emotion (Angry, Happy, Neutral) and Congruency (Congruent, Incongruent) Conditions in Experiment Three.

	Congruent	Incongruent	Total
Overall task			
Angry	320.23	326.19	323.21
Happy	317.20	322.85	320.02
Neutral	319.26	323.35	321.31
Total	318.90	324.13	
Block One			
Angry	318.93	327.73	323.33
Happy	314.78	320.87	317.83
Neutral	320.70	322.92	321.81
Total	318.14	323.84	

Note. RTs (in ms) are presented for the overall task (blocks 1-3) and for block one (n=24).

Interim Discussion

Experiment Three aimed to combine datasets from Experiments One and Two to examine the validity of the task in terms of emotion, congruency and interaction effects in a sample of 24 adults. Results from this analysis suggest our task design features were mostly in line with established spatial cueing paradigms and dot-probe paradigms incorporating both non-emotional (Spence & Driver, 1997) and emotional (Brosch et al., 2008) auditory cueing stimuli in adults. However, the overall task did not produce the expected effects. The predicted effects were limited to block one of the task. Specifically, findings from block one revealed adults were faster at responding to congruent compared to incongruent trials, in accordance with our predictions. Second, the paradigm produced a significant effect of emotion on RTs. Participants were slower at responding to trials preceded by angry compared to happy voices. Finally, the analyses from all three experiments did not demonstrate the predicted interaction effects between emotion and congruency.

The finding of faster RTs to congruent compared to incongruent trials in block one of the task is important because it demonstrates that the present spatial cueing paradigm produces robust congruency effect in accordance with previous studies (Spence & Driver, 1997; Brosch et al., 2008). Since there were no predictive elements of the auditory stimuli (i.e., vocal emotions) in the present paradigm, the presence of congruency effects on RTs are thought to reflect the automatic attentional capture of irrelevant auditory stimuli, which facilitate performance (Spence & Driver, 1997). This is consistent with theories which suggest the auditory attentional systems prioritises task irrelevant auditory stimuli, to enable us to respond quickly to events outside of our visual field (Murphy et al., 2017). Since the finding of a congruency effects on RTs is consistent with a wide body of literature employing varied spatial cueing paradigms with emotional and neutral auditory stimuli, this finding corroborates the validity of the task as a spatial cueing paradigm.

The finding that adults were slower to respond to trials preceded by angry vocal cues contrasts with previous findings from studies employing dot-probe paradigms (e.g., Brosch et al., 2008). Previous research has shown that spatially congruent angry vocal cues result in faster response times to trials preceded by angry compared neutral voices, thus suggesting that angry vocal cues capture and facilitate VSA in dot-probe paradigms (Brosch et al., 2008). The contrasting findings between the present ESC paradigm and the dot-probe paradigm employed by Brosch et al (2008) may be due to differences in the task design. Specifically, the present ESC paradigm differed from the dot-probe paradigm used by Brosch and colleagues (2008) in that only one vocal cueing stimulus was presented unilaterally, rather than two competitive vocal cueing stimuli presented bimodally in the dot-probe paradigm (Brosch et al.,2008). This is important as different mechanisms may underlie the emotion effects in the two paradigms. For example, within the dot-probe paradigm, neutral and angry vocal cues competed for attention, and vocal anger was thought to facilitate RTs due to a bias in attention towards the angry over the neutral vocal cue, possibly due to vocal anger being perceived as more salient than the neutral voice (Brosch et al., 2008). However, within the present paradigm there was no competition for attention when only one irrelevant vocal cue stimulus was presented at a time unilaterally.

Taking these design differences in consideration, it is possible that within the present ESC paradigm, angry vocal cues were still perceived with more salience and captured attention, however rather than competing for and thus facilitating RTs, the perceived salience had an alternative overall distracting effect on RTs. This explanation aligns with previous research employing slightly different paradigms to explore the effect of vocal emotion on attentional mechanisms. For example, the hindering effects of irrelevant vocal anger have been reported in auditory distraction (Kattner & Ellermeier, 2018; Kattner, 2023) and dichotic listening (Peschard et al., 2017; Sander et al., 2005) tasks. Similarly, in paradigms which incorporate other types of negative stimuli, such as aversive or threatening visual stimuli, those stimuli have been found to slow down RTs when these stimuli are irrelevant to task demands (Hart et al., 2010; Melcher et al., 2011). For example, Hart et al (2010) employed a Stroop-like paradigm, in which participants were required to indicate how many visual digits were presented in an array of congruent (e.g., four digits in an array of the number 4's) or incongruent (e.g., three digits in an array of the number 4's) trials. Each trial was preceded by an irrelevant picture depicting either aversive or neutral content. Results showed participants displayed slower RTs in incongruent trials when they were primed by aversive compared to neutral stimuli (Hart et al., 2010). Theoretically, it is thought that negative or threatening stimuli may hinder cognitive performance due to holding limited attentional resources, and therefore creating a difficulty to engage on the task-relevant stimulus (Hughes, 2014). Therefore, within the present ESC task, it is possible that vocal anger initially captured VSA, but then maintained attention, creating a difficulty for participants to engage attention on the visual target, resulting in slower RTs.

The finding of no significant interaction effects between emotion and congruency is surprising. There are several possible reasons for this. First, it is possible that as the task was optimised for ERP analyses, it was not sufficiently sensitive on a behavioural level to reveal interaction effects. Second, it is possible that the lack of an interaction effect is a true null finding, perhaps caused by the emotion, specially anger, within the voice having an overarching distracting effect on performance overall. The Additive Factors Model (AFM) can be applied in this context to help explain the presence of main effects (i.e., congruency and emotion), but a lack of an interaction effect on RTs. The AFM focuses on how quickly participants make RTs when errors are rare and there are two or more factors (Sternberg, 1998), as in the present ESC paradigm. The AFM model suggests that combination of main effects, but not interaction effects, produces an additive effect on RTs. This suggests that the mental process in the ESC paradigm can be split into two stages, in that emotion influences one stage of attention allocation, and congruency influences a different stage in attention allocation. Moreover, the lack of an interaction effect suggests that emotion and congruency influence no stage in common. This is consistent with the idea that congruency appears to have a facilitating effect on performance, whereas emotion appears to have a distracting effect on performance in the present ESC paradigm. This is also consistent with theories of perception and attention which suggest there are separate mechanisms for processing emotional and non-emotional perceptual and attentional features of stimuli, and both can occur in parallel, to produce additive effects on responses (Pourtois et al., 2013). Further exploration of the underlying mechanisms relating to potential additive effects of congruency and emotion at specific stages in processing with the ESC paradigm can be more thoroughly investigated with ERP techniques. As reviewed on page 129, the cue-P2, has been thought to reflect the preattentive processing of vocal emotion prior to attentional capture, next the Target-P1 is considered as an index of the automatic capture of VSA by vocal cues, and finally the Target-P3 has been argued to reflect the effect of vocal emotion on later attention allocation mechanisms when responding to relevant target stimuli. Therefore, Experiment Four aimed to validate the ESC paradigm at producing the expected congruency and emotion effects on these ERP components, consistent with previous literature (see page 129).

It is important to note that the significant findings relating to effects of congruency and emotion on RTs in Experiment Three were limited to the first block of the task. This possibly suggests effects of habituation as noted by Brosch et al (2008), who similarly only reported significant effects on RTs within the first two blocks of the paradigm, and not the third block. Similar habituation effects have been reported in ERP studies which found that emotion effects on the N1 component were only evident within the first half (180 trials) of a VER task in healthy 6-11-year-old children (Chronaki et al., 2015). Despite this, other studies employing emotional auditory stimuli in spatial cueing paradigms did not report habituation effects in adults when including a larger number of total trials (e.g., 600; Zimmer et al., 2022; 360; Wang et al., 2022). Moreover, Spence & Driver (1997) reported overall congruency effects in a spatial cueing paradigm with eight blocks of trials. It is noteworthy that when data from all 24 adult participants was combined in Experiment Three, both effects of emotion and congruency were close to significance in the overall task. Therefore, it is possible that effects were only significant in block one of the present pilot study due to the small size of the adult sample or due the behavioural characteristics (e.g., age, IQ) of the adult sample, which were not noted due to the preliminary nature of the pilot experiments. Experiment Four aimed to measure behavioural (i.e., age, IQ, gender, hyperacuity) and psychopathology characteristics (i.e.., hyperactivity, inattention, depressive symptoms) of adult participants. These characteristics may influence effects of emotion and congruency on RTs and ERPs, In addition these measures were added for consistency with a later ERP study with adolescent participants. See Chapter Five for the rationale of the inclusion of these measures. Moreover, a full sample of 25 participants were recruited in the next experiment to more thoroughly validate the final version of the ESC paradigm.

In conclusion, three pilot experiments were conducted to inform the design of the ESC. The ESC was optimised and validated behaviourally in a sample of 24 adults with two slightly different versions of the paradigm. Based on the results from Experiment Three, the ESC paradigm demonstrated high validity, at least in the first block of the task, for implementation in ERP studies. First, this paradigm produced the expected congruency effects in the first block of the task, indicating that irrelevant spatially located vocal cues captured VSA and facilitated performance, consistent with previous studies employing spatial cueing paradigms (Spence & Driver, 1997). Second, when making small adaptations to the paradigm, additional emotion effects on RT scores were evident in the first block of the task (see Experiment Two). However, it is not clear whether these emotion effects specifically reflect the capture of VSA. The next experiment employed ERP techniques, with the aim to elucidate the underlying cognitive mechanisms of emotion effects on VSA.

4.7 Experiment Four

Results

Initial RT Data Treatment. Kolmogrov Smirnoff tests indicated that all RT variables were normally distributed (Angry Congruent: D=0.12, p=.200, Angry Incongruent: D=0.08, p=.200, Happy Congruent: D=0.14, p=.200, Happy Incongruent: D=0.14, p=.197, Neutral Congruent: D=0.13, p=.200, Neutral Incongruent: D=0.13, p-.200). Pearsons' correlations revealed no significant correlations between behavioural (IQ, hyperacuity, age) and RTs (r values from -.20 to .24). In addition, no significant correlations were reported between psychopathology[10] (depressive symptoms, inattentive and hyperactive symptoms) characteristics and RT scores (r values from -.20 to .24). Therefore, no behavioural or psychopathology characteristics were added as covariates in subsequent analyses.

RT Main Analyses. Table 4.4 shows descriptive statistics of RTs for correct responses to the visual target by congruency and emotion conditions. A 2 x 3 repeated measures ANOVA was conducted to explore the effect of emotion and congruency on mean RTs. The results showed a significant main effect of congruency (F (1,24) = 5.56, p = .027) and a main effect of emotion (F (1,24) = 4.42, p =.017) on RTs. Consistent with our predictions and findings from the previous pilot experiments, participants were faster to respond to congruent compared to incongruent trials. This effect was demonstrated within the whole task (blocks 1-3), rather

---

[10] Note, because the aim of the present study was to validate the paradigm at a neural level, more detailed supplementary analyses relating to how psychopathology affected RTs and ERPs to cue, and target are presented in Appendix H, when conducting additional analyses comparing adult and adolescent data.

than just block one (as reported in Experiment Three). Two post hoc t-tests revealed adults were slower to respond to trials preceded by an angry compared to a happy (t (24) = 2.94, p=.004) or neutral (t (24) =2.28, p=.016) voice, in accordance with results from Experiment Three. The alpha value was adjusted to p=.025 to control for multiple comparisons. In line with findings from Experiment Three, there was no significant congruency by emotion interaction effect on RT scores (F (2,48) = 0.52, p=.595). Figure 4.5 presents a bar graph showing RTs for the emotion and congruency conditions.

Table 4.4

Means (SD) RTs to Targets Preceded by Vocal Cues in Emotion (Angry, Happy, Neutral) and Congruency (Congruent, Incongruent) Conditions in Adults (n=25) in Experiment Four

	Congruent	Incongruent	Total
Angry	327 (51)	331 (50)	329 (9.88)
Happy	320 (47)	326 (49)	323 (9.49)
Neutral	321 (52)	328 (49)	324 (10.11)
Total	323 (9.93)	328 (9.70)	

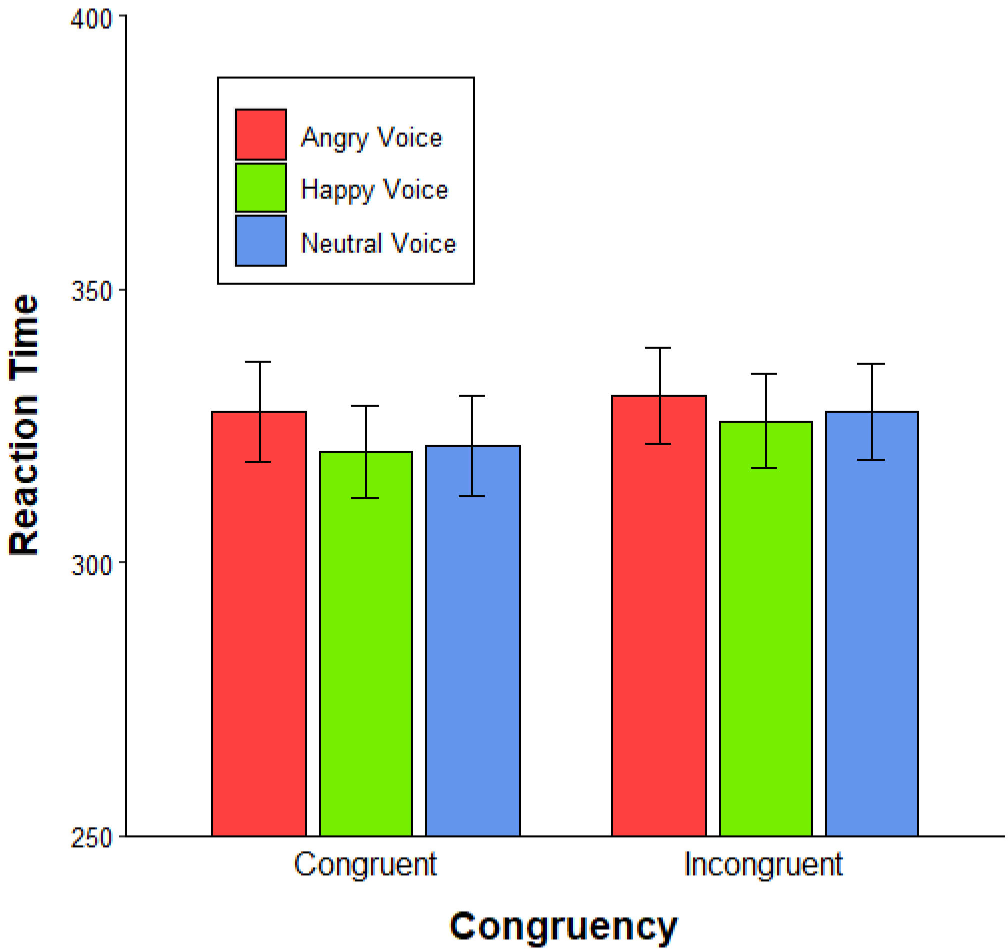

Figure 4.5

Bar Graph Showing Mean RTs (in ms) to Targets Differing by Emotion (Angry, Happy, And Neutral) and Congruency (Congruent And Incongruent) Conditions in Adults (n=25) in Experiment Four.

Initial ERP Data Treatment. Kolmogorov-Smirnov tests indicated that all ERP data for the Vocal Cue and Visual Target were normally distributed for all emotion and congruency conditions (ps >.05; See Table E.1 in Appendix E for complete statistics), apart from the Target-P3 in angry incongruent conditions (d=0.21, p=.005). Pearsons' correlations revealed no significant correlations between behavioural (IQ, hyperacuity, age) and ERP amplitudes for components P2, P1, and P3 (r values from -.20 to .24). In addition, no significant correlations between psychopathology (depressive symptoms, inattentive and hyperactive symptoms) and ERP amplitudes were found (r values from -.20 to .24). Therefore, no behavioural or psychopathology measures were added as covariates in the ERP analyses.

ERP Main Analyses

Vocal Cue-P2. In accordance with our predictions, a one-way repeated measures ANOVA revealed a main effect of emotion on P2 amplitudes ($F(2,48) = 19.12$, $p<.001$). Specifically, adults presented with larger P2 amplitudes to happy and angry compared to neutral voices (neutral and happy: $MD = 1.22$, $p<.001$; neutral and angry: $MD = -0.78$, $p=.005$). These results suggest that within the ESC paradigm, the P2 component may reflect the preattentive processing of vocal emotion in adults. Figure 4.6 presents the grand averages of the P2 component to the vocal cue. Table E.2 in Appendix E presents the mean P2 amplitude values for angry, happy and neutral vocal expressions.

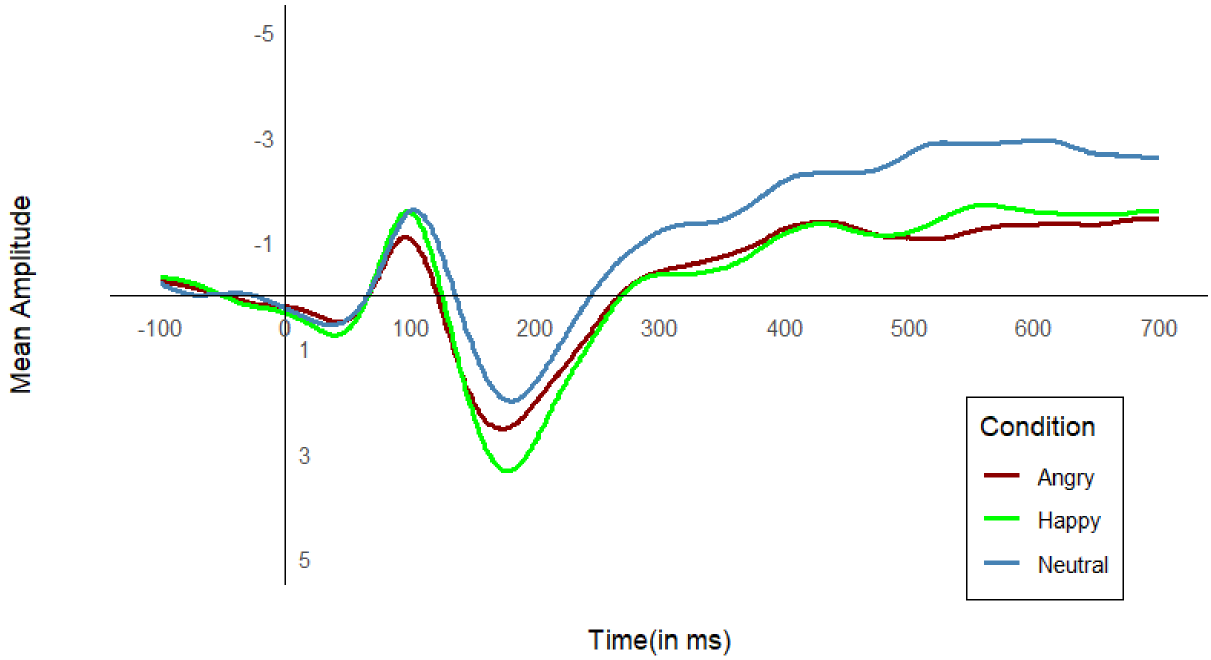

Figure 4.6

Grand Average of the Frontocentral P2 Component (130 – 250 ms) to Angry, Happy and Neutral Vocal Expressions in Adults (n=25). Scale Is -5 To + 5 Mv.

Target-P1. In contrast with our predictions, a 2 x 3 repeated measures ANOVA revealed no main effects of congruency ($F(1,24) = 0.60$, $p = .446$) or emotion ($F(2,48) = 2.75$, $p = .074$) on Target-P1 amplitudes. In addition, there was no significant interaction effect between emotion and congruency on P1 amplitudes ($F(2,48) = 0.61$, $p = .941$). Figure 4.7 presents the grand averages of the P1 to visual target in congruent and emotion conditions. As a secondary analysis, a 2 (congruency) x 3(emotion) x 2(target hemifield) repeated measures ANOVA was conducted to explore if congruency effects on the Target-P1 amplitude were evident when the target was presented to the left hemisphere, as reported by Brosch et al (2009). Results revealed there was no significant main effect of target hemifield on P1 amplitudes ($F(1,24) = 1.85$, $p = .186$). Additionally, there was no significant hemifield x congruency interaction effects on the Target-P1 ($F(1,24) = 1.01$, $p = .326$). In addition, there was no significant hemifield x emotion interaction ($F(2,48) = 1.55$, $p = .224$), or congruency x emotion x hemifield interaction effect on P1 amplitudes ($F(2,48) = 0.50$, $p = .612$). Table E.3 in Appendix E presents the mean P1 amplitudes time-locked to the visual target for all emotion and congruency conditions.

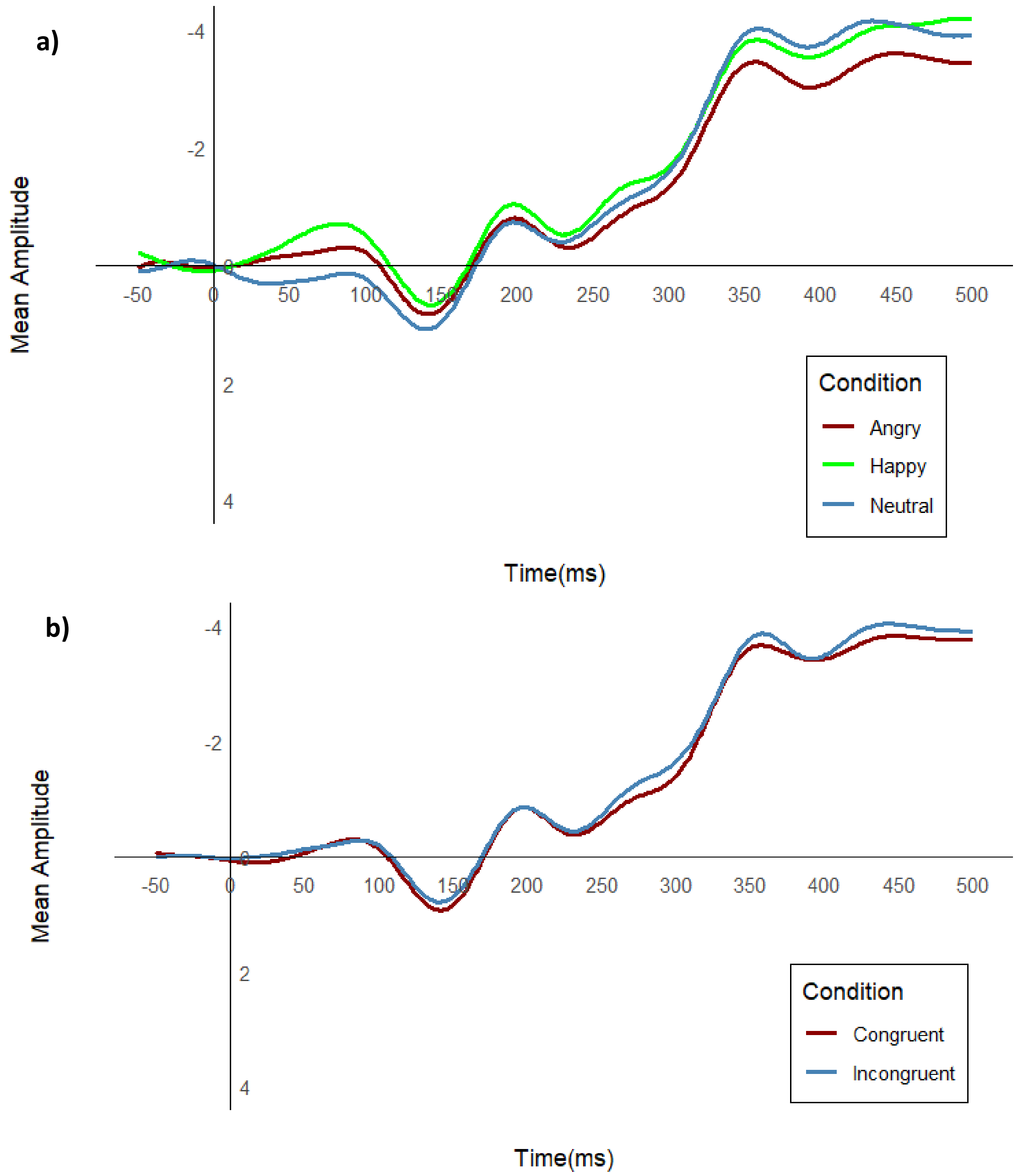

Figure 4.7

Grand Averages Showing the Parietal Occipital P1 Component (100-170ms) to the Target Differing by a) Emotion (Angry, Happy, Neutral) and b) Congruency (Congruent, Incongruent) Conditions in Adults (n=25). Scale is -4 to + 4 μV.

Target-P3. Table E.4 in Appendix E presents the mean P3 amplitudes time-locked to the visual target for all emotion and congruency conditions when presented in the left and right hemifield. First, a 2 (congruency) x 3 (emotion) repeated measures ANOVA was conducted. Results revealed no main effect of emotion ($F(2,48) = 0.08$, $p = .920$) or congruency ($F(1,24) = 0.49$, $p = .488$) on the Target-P3 amplitude. Additionally, there was no congruency x emotion interaction effect on the Target-P3 amplitude ($F(2,48) = 0.82$, $p = .449$). Next, to explore effects of laterality on the Target-P3 amplitude, a 2 (congruency) x 3 (emotion) x 2 (hemifield) repeated measures ANOVA was conducted. There was no main effect of congruency ($F(1,24) = 0.04$, $p = .846$), emotion ($F(2,48) = 0.43$, $p = .656$) or hemifield ($F(1,24) = 0.22$, $p = .644$) on the Target-P3 amplitude. However, results revealed a significant interaction effect between congruency and hemifield ($F(1,24) = 11.23$, $p = .003$). Post hoc tests (Bonferroni corrected) revealed that when targets were presented in the right visual field, participants displayed a marginally significant larger P3 to incongruent compared to congruent trials ($MD = 0.47$, $p = .050$). Figure 4.8 presents the grand averages of the P3 to visual target by congruency and emotion in the right and left visual field. Finally, there was no significant hemifield by congruency interaction effects on the Target-P3 amplitude ($F(2,48) = 0.12$, $p = .898$). There was no significant congruency by emotion by hemifield interaction effect on the Target-P3 amplitude ($F(2,48) = 0.87$, $p = .424$).

a)

b)

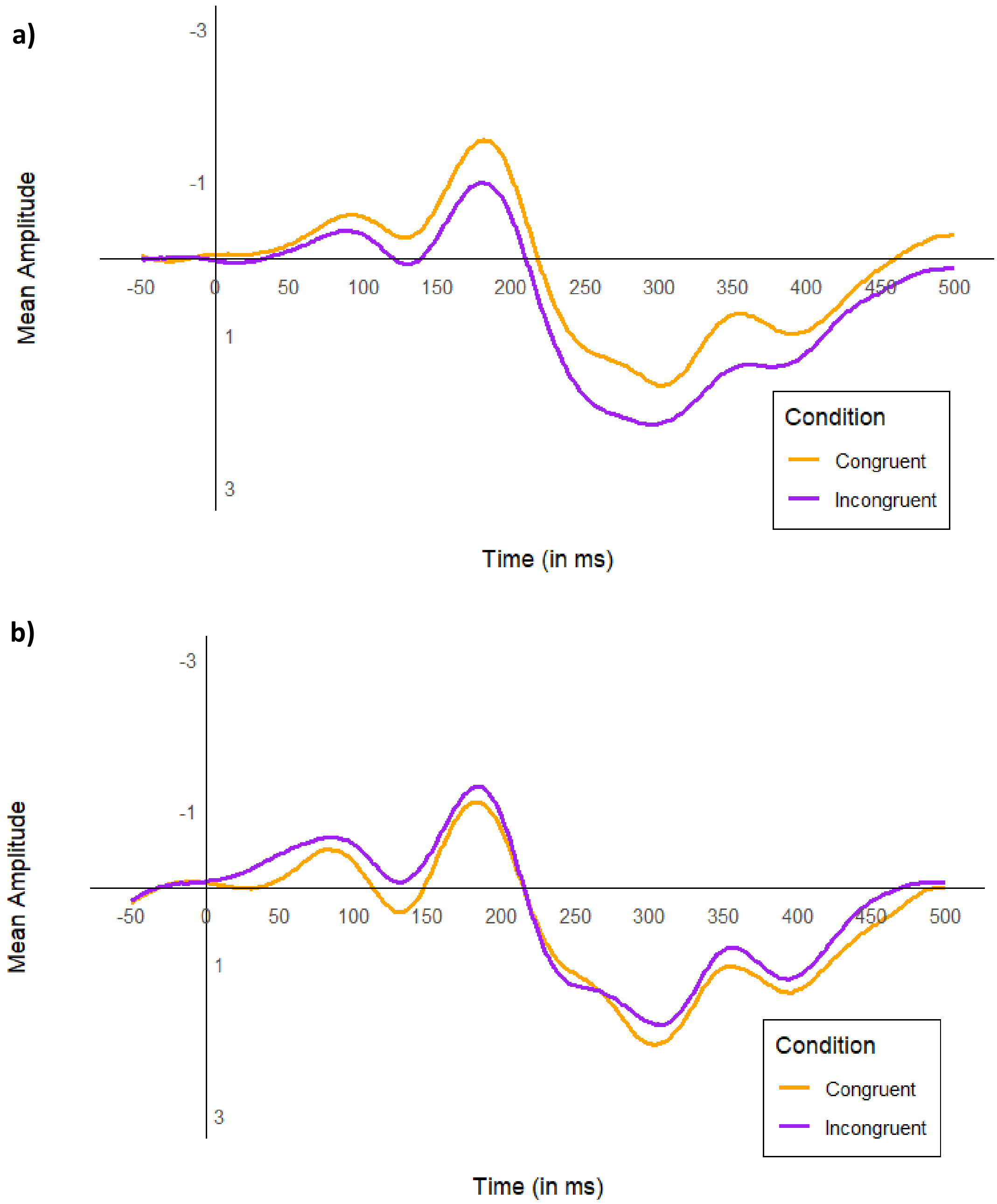

Figure 4.8

Grand Averages of the Parietal P3 Component (290-450ms) in Response to the Visual Target Following Congruent and Incongruent Trials when the Target was Presented in the a) Right Visual Field and b) Left Visual Field in Adults (n=25). Scale is -3 to + 3 µV.

4.8 Discussion

Experiment Four adds to the findings from Experiments One to Three to provide robust evidence that the ESC paradigm is effective at producing reliable additive, rather than interactive, congruency and emotion effects on RTs in adults. Second, Experiment Four validated the ESC paradigm on a neural level. Specifically, results demonstrated the ESC paradigm is a valid measure of the preattentive processing of vocal emotion, as reflected by larger P2 amplitude to irrelevant emotional (i.e., angry and happy) compared to neutral voices, meeting the first objective of the study. However, it is unclear whether the ESC paradigm is a valid measure of downstream effects of the involuntary capture of attention by vocal emotion on a neural level, as the expected effects of emotion and congruency on the Target-P1 and Target-P3 components were not evident in adults.

First, the present results replicated the congruency and emotion effects on RTs reported in the pilot experiments (1-3). The presence of these main effects, and the lack of emotion by congruency interaction effects suggests that within the present paradigm, emotion and congruency have additive rather than interactive effects on reaction times. Importantly these effects were not limited to the first block of the task, thus strengthening the findings reported in Experiment Three. Specifically, participants were faster to respond to congruent compared to incongruent trials. The congruency effect on RTs is a robust finding, consistent with findings from previous cueing paradigms (Spence & Driver, 1997; McDonald et al., 2000; Mazza et al., 2007). The congruency effect on RTs is therefore thought to reflect the involuntary capture of VSA by irrelevant spatially located vocal cues (Spence & Driver, 1997). This pattern of results is in accordance with theories suggesting irrelevant auditory stimuli are prioritised in our attentional systems (Murphy et al., 2017). Additionally, the main effect of emotion on RTs indicated that participants were slower to respond to trials preceded by angry compared to happy and neutral vocal cues. This replicates findings from Experiment Three. Interestingly, in Experiment Four effects were not limited to the first block of the task. This finding is consistent with studies employing auditory distraction or dichotic listening paradigms which similarly reported a distracting effect of task-irrelevant vocal anger on performance, as demonstrated by slower reaction times (Peschard et al., 2017; Sander et al., 2005) and poorer recall of words (Kattner & Ellermeier (2018) in healthy adults. Within the ESC paradigm, the distracting effect of vocal anger on RTs probably reflects the involuntary capture of attention by vocal anger, resulting in later difficulties to disengage from the emotional stimulus, and engage in a task-relevant visual target, as suggested previously (Koser et al., 2004; Rooijen, Ploeger & Kret, 2017). This finding is in accordance with theories which suggest task-irrelevant negative or threatening stimuli may hinder performance by capturing and holding limited attentional resources (Huges, 2014). The next study aims to examine whether this is also the case in younger (adolescent) participants (see page 186).

Second, the present study extends existing literature by validating the ESC paradigm at a neural level. The present study isolated the P2 component as a neural marker of the preattentive processing of vocal emotion in adults in accordance with our predictions. Specifically, adults displayed a larger P2 amplitude to irrelevant angry and happy compared to neutral voices. This is in accordance with previous studies employing preattentive paradigms which similarly reported an emotion effect on the P2 amplitude to irrelevant vocal cues (Lin et al., 2022; Gadeke et al., 2013). Results from the present study support the model of vocal emotion processing proposed by Schirmer & Kotz (2006). This model suggests the P2 reflects a preattentive stage of processing vocal emotion, when the emotional significance of the voice is processed (i.e., via the STS and amygdala) prior to the capture of attention (see Figure 1.3, page 41; Schirmer & Kotz, 2006). This suggests adults preattentively perceive happy and angry voices with greater salience than neutral voices. The effect of emotion on the P2 is robust in the present study, and consistent with the literature. This finding corroborates the validity of the ESC paradigm in measuring the preattentive processing of vocal emotion in adults. This is an important finding, since the main aim of the thesis is to disentangle preattentive vocal emotion processing, from attentive processing in adolescents with traits of ADHD. The next study (Chapter Five) aims to explore whether the P2 is also a neural marker of preattentive vocal emotion processing in adolescents.

It is surprising that the present study did not reveal any significant effects of emotion and congruency on the Target-P1 component, thought to reflect the involuntary capture of VSA by irrelevant auditory cues, as reported in previous spatial cueing paradigms (Störmer et al., 2009). This contrasts with results from previous research employing a dot-probe paradigm, which reported a larger P1 amplitude in response to targets preceded by spatially congruent angry vocal cues (Brosch et al., 2009). Based on theory suggesting the P1 amplitude is enhanced when a visual stimulus appears in an already attended location in healthy adults (Luck et al., 2000; Mangun et al., 1998), the lack of congruency and emotion effects on the P1 amplitude in present study may suggest the vocal cue did not effectively influence early VSA to the target within the ESC paradigm. Alternatively, it is possible that the capture of VSA by the vocal cues occurred earlier than the onset of the visual target within the ESC paradigm. Supporting this, a previous study which employed a spatial cueing paradigm to explore the stages of visual processing of a target modulated by the involuntary orienting of attention to an auditory stimulus similarly did not report effects of congruency on the P1 component in response to the target (McDonald & Ward, 2000). In their study, McDonald and Ward (2000) similarly suggested that the lack of congruency effects on the P1 amplitude may have been due to the capture of VSA by the auditory cue occurring after the initial sensory processing of the target due to incorporating an auditory stimulus short in duration (70ms) and a short SOA (100-300ms). In the present study, a longer SOA (750ms) was incorporated in the ESC paradigm, which raises the possibility that the rapid capture of attention by the auditory cue occurred prior to the initial sensory processing of the target. Since 750ms in the present study was the shortest possible SOA, necessary to incorporate the duration of the vocal emotional stimuli (700ms) and a 50ms baseline prior to the visual target, it was not deemed suitable to make any modifications to the SOA included in the ESC paradigm.

An alternative explanation for the lack of congruency or emotion effects on the P1 could be related to insufficient statistical power within the present study, resulting in a small P1 component to the visual target. This is plausible, as the overall P1 effects observed were small (0.46μV average amplitude) compared to those reported in other studies (e.g., 3μV amplitude; Brosch et al., 2009). Since design features relating to the visual target (i.e., size, colour, duration) were consistent with those in previous spatial cueing paradigms with adults (Brosch et al., 2009) and children (Perchet et al., 2007), it is possible that a larger number of trials (i.e., 400-600) are necessary to produce a more reliable P1 response in the ESC paradigm, as recommended by Luck (2005). However, because the next study aimed to recruit adolescents with traits of inattention and hyperactivity, it was not deemed appropriate to increase task length due to potential difficulties participants with inattention and hyperactivity may have in sustaining attention on the task. Moreover, considering that both the present study (i.e., Experiment 3) and previous studies (i.e., Brosch et al., 2009) report habituation effects on the P1 component and RTs increasing the number of trials might increase the strength of these effects. An alternative approach to increase the power to reach significant effects of congruency and emotion on the P1 component is to increase the sample size. Therefore, potential effects of congruency and emotion on the P1 amplitude to the visual target in the ESC paradigm are explored in more detail in Chapter Five, which recruited a larger number (60) of adolescent participants (See page 237).

Finally, the present study did not report any significant effects of emotion on the P3 component. This finding contrasts with our predictions and previous studies, which reported a larger P3 component in response to targets preceded by emotional auditory stimuli compared to neutral stimuli. These previous findings have been interpreted as reflecting a mechanism of later maintenance of attention on the spatial location of the irrelevant emotional stimuli, stemming from the early capture of attention by irrelevant emotional stimuli (Li et al., 2018; Wang et al., 2022). Interestingly, however, the present study did reveal that the P3 was larger in incongruent compared to congruent trials, but only when the visual target was presented in the right visual field. These results provide partial evidence for a congruency effect on the later P3 component. This finding is consistent with previous research which reported a similar effect in an emotional spatial cueing paradigm, where the P3 component was thought to reflect the recruitment of additional attentional resources to disengage attention from spatial location of the prior cueing stimulus and engage on the relevant visual target (Wang et al., 2022). Overall, the finding of congruency effects on the P3 further highlights the validity of the ESC paradigm, at producing the expected target related effects, consistent with the reaction time data. Moreover, this pattern of findings highlights the importance of further exploring emotion and congruency effects on the later P3 components in the subsequent study with adolescents with varying levels of ADHD traits, as it is possible that the P3 component is a more sensitive measure than the early P1 component at capturing the effects of congruency and emotion on attention mechanisms during target processing within the ESC paradigm. The next study which recruits a larger number of adolescent participants will further explore this possibility.

In summary, based on behavioural and ERP results from adult participants, the present study provides evidence that the ESC paradigm is a valid measure of the preattentive processing of vocal emotion at a neural level, as demonstrated by a larger P2 response to emotional compared to neutral voices. This is an important finding and in accordance with our first objective when developing the ESC paradigm. Second, the presence study provides evidence that the ESC paradigm is a valid measure of the involuntary capture of VSA by irrelevant vocal emotion at a behavioural level, as demonstrated by the presence of a congruency effect on RTs, and a distracting effect of vocal anger on RTs. It is unclear from the present ERP findings whether the ESC paradigm can isolate electrophysiological correlates (e.g., P1) of the effects of vocal emotion on early visual attentional mechanisms during target processing stages. However, the present study provided some initial evidence that the P3 component captures effects of congruency on the later attentional mechanisms underlying target processing, and thus may be a more sensitive neural measure of emotion and congruency effects on attention allocation to a target. The next chapter presents the experimental work which employed the present ESC paradigm to primarily explore the preattentive processing of vocal emotion, and its effect on visual spatial attention in adolescents with varying levels of ADHD traits.

5 THE EFFECT OF VOCAL EMOTION ON VISUAL SPATIAL ATTENTION IN ADOLESCENTS WITH ADHD TRAITS: A BEHAVIOURAL AND ERP STUDY

5.1 Chapter Overview

This chapter presents a behavioural and ERP study which employs the ESC paradigm, designed and validated in adults in Chapter Four. The present study aims to investigate a) whether adolescents with ADHD traits present with a preattentive hypersensitivity to vocal emotion (e.g. anger, happiness), and b) the effect of vocal emotion on the allocation of attention during target processing in adolescents with ADHD traits. Sixty-three adolescents with varying levels of inattention and hyperactivity were recruited from the community in the present study. Three types of analyses were employed. The first analysis aimed to isolate a neural marker of preattentive vocal emotion processing (reflected by the Cue-P2) in adolescents. In addition, the first analysis explored the involuntary capture of VSA by vocal emotion, by measuring the effects of emotion on the early visual processing of the target (reflected by the Target-P1) and the later attentional allocation to the target (reflected by the Target-P3) within the whole sample of adolescents. Second, correlational analyses explored associations between a) ADHD traits and preattentive vocal emotion processing, and b) ADHD traits and attentional allocation during target processing. Finally, analysis of a subset of participants with extreme (high vs low) ADHD traits was performed to more thoroughly examine the effects found in the correlational analyses. Results from present study revealed a larger P2 amplitude to angry and happy compared to neutral voices in adolescents with high compared to low traits of inattention and hyperactivity. This builds on studies employing explicit VER paradigms, suggesting a hyper-orientation towards vocal anger in children with ADHD (Chronaki et al., 2015), by more conclusively providing evidence for emotion-specific atypicalities during the preattentive processing of vocal emotion in ADHD. Findings of the present study support socio-cognitive (Lemerise & Arsenio, 2000) and motivational dysfunction theories of ADHD (Sonuga-Barke, 2005, 2012) and have important implications for clinical practice.

5.2 Introduction

High levels of inattention and hyperactivity are present in up to 23% of children and adolescents who do not meet the threshold for an ADHD diagnosis, but present with similar difficulties as those with a clinical diagnosis (Balázs & Keresztény, 2014; McLennan, 2016). This is consistent with dimensional approaches to ADHD (Sonuga-Barke., 2023). Individuals with high levels of ADHD traits often present with emotional difficulties, such as emotion dysregulation (i.e., problems with recognising, attending and regulating emotions; Shaw et al., 2014; Faraone et al., 2019), as well as comorbidities such as depression (Park & Chang, 2022) and conduct problems, including irritability and anger problems (Gustafsson et al., 2018; Taylor et al., 1996). Experimental research has confirmed that ADHD is linked with difficulties in recognising emotions (e.g., anger, happiness, sadness, fear and disgust) in vocal emotion recognition (VER) tasks, consistent with our meta-analysis in Chapter Two (Sells et al., 2023). However, it is debated whether VER difficulties in ADHD reflect generic attention deficits, in support of executive function (EF) theories (Barkley, 1997) or atypicalities in perceptual emotion processes in support of motivational dysfunction (Sonuga-Barke, 2005, 2012) and socio-cognitive theories of ADHD (Lemerise & Arsenio, 2000).

To contribute to the theoretical debate surrounding the underlying neurocognitive atypicalities of vocal emotion difficulties in ADHD, previous research has relied on studies employing explicit vocal emotion processing tasks. For example, behavioural studies report individuals with ADHD display generic attention difficulties during vocal emotion processing in support of EF theories (Sells et al., 2023), whilst neuro-imaging research reports anger-specific atypicalities during vocal emotion processing (Kochel et al., 2015; Chronaki et al., 2015), in support of perceptual processing theories, such as the socio-cognitive models (Lemerise & Arsenio, 2000) and motivational dysfunction theories (Sonuga-Barke, 2012). For example, Kochel et al (2015) conducted a near-infrared spectroscopy study with children (aged 10 years) with and without ADHD who completed a vocal emotion recognition task. Findings showed that during the processing of vocal anger, boys with ADHD presented with reduced activation of the right STG, thought to reflect atypicalities in attention allocation to emotionally relevant stimuli in explicit VER tasks (Kochel et al., 2015). Additionally, ERP research reported children with ADHD present with an enhanced N1 to vocal anger when completing an explicit emotion recognition task, leading the authors to hypothesise that the N1 may be a marker for an automatic hyper-orientation to vocal threat in ADHD (Chronaki et al., 2015). This is consistent with findings from the facial emotion processing literature, which show that when fearful faces are presented subliminally (i.e., for 30ms) followed by a neutral expression for 470ms, adolescents (aged 12-14 years) with ADHD showed greater amygdala activation to fearful faces compared to healthy controls (Posner et al., 2012). In summary, neuro-imaging results from studies employing explicit VER paradigms provide some initial evidence for an automatic hyper-orientation to threat-related (e.g., vocal anger) stimuli in ADHD (Chronaki et al., 2015; Kochel et al., 2015).

In contrast to the hypothesis of a hyper-orientation to specifically threat-related stimuli in ADHD, there has been some research which suggests adolescents with ADHD present with attention allocation atypicalities in both positive and negative emotionally salient contexts (Chronaki et al., 2017). For instance, in an ERP study, adolescents (aged 10-16) took part in a Monetary Incentive Delay task, in which participants were instructed to respond to a visual target stimulus as fast as they could. A cueing stimulus indicated whether the speed of participants' responses resulted in monetary gains (i.e., positive reinforcement) or losses (i.e., negative reinforcement; Chronaki et al., 2017). Results demonstrated that adolescents with ADHD presented with enhanced P3 and CNV amplitudes, thought to reflect attention allocation and response preparation to the visual target, in both positive and negative reinforcement conditions (Chronaki et al., 2017). This provides some evidence that both negative and positive motivationally salient cueing stimuli enhances attention allocation to a target stimulus in adolescents with ADHD. This is consistent with motivational-dysfunction theories, such as the delay-aversion hypothesis, which argue atypicalities in perceptual processing of emotions may lead to behaviours associated with seeking out immediate rewards, rather than experiencing negative emotion associated with waiting for a delayed outcome (Sonuga-Barke, 2005, 2012). However, as of yet, no studies have reported a hyper-orientation to positive vocal emotion (e.g., happy voices) in ADHD, as demonstrated for threat-related (angry voices) expressions (Chronaki et al., 2015). Moreover, it is not known from studies employing explicit paradigms (e.g., VER or MID tasks) whether enhancements in attention allocation in emotionally salient contexts (both negative and positive) stem from atypicalities in preattentive or attentive emotion processing. Exploring the preattentive processing of vocal emotion in ADHD is necessary to understand more conclusively if there are atypicalities in the perceptual stage of emotion processing, independently from later attentional processing stages in ADHD, as proposed by socio-cognitive (Lemerise & Arsenio, 2000) and motivational-dysfunction theories (Sonuga-Barke, 2005, 2012).

To investigate atypicalities in the involuntary capture of preattentively processed vocal emotion in ADHD, a behavioural study explored the effect of irrelevant vocal emotion on attentional mechanisms in ADHD using a cross-modal oddball paradigm (Chronaki & Marsh, 2024). Within this study, a community sample of 93 children (aged 7-10 years), 51 adolescents (aged 11-15 years) and 50 adults (aged 18-48 years) indicated the colour of stepping-stones on a screen, whilst ignoring irrelevant emotional (angry, happy) and neutral voices (Chronaki & Marsh, 2024). The results demonstrated that children who displayed higher levels of hyperactivity, as measured by the Strengths and Difficulties Questionnaire (SDQ), were slower to respond to visual targets (i.e. stepping-stones) preceded by irrelevant angry vocal cues. This study provided evidence for a threat-specific involuntary capture of attention by preattentively processed vocal emotion in ADHD, resulting in a distracting effect on task performance (Chronaki & Marsh, 2024). Therefore, this study provided support that a hyper-orientation to threat-related stimuli in ADHD may stem from atypicalities in preattentive processing of negative expressions, such as vocal anger. However, no studies so far have provided direct evidence for specific preattentive atypicalities in vocal emotion processing prior to attentional capture.

To disentangle the preattentive processing of vocal emotion, from the involuntary capture of attention in preattentive paradigms, electrophysiological techniques can be used due to their high temporal resolution. The present thesis validated an ESC paradigm to measure the preattentive and attentive processing of vocal emotion with electrophysiological techniques in healthy adults in Chapter Four. Within the ESC paradigm irrelevant angry, happy and neutral vocal cues are presented in the left or right ear prior to a spatially congruent or incongruent visual target. Participants are asked to ignore the vocal emotion and indicate the location of the visual target. The study presented in Chapter Four provided evidence that the ESC paradigm is a valid measure of the preattentive processing of vocal emotion in adults, as reflected by a larger frontocentral P2 response (130-250ms) to emotional compared to neutral voices. Second, Chapter Four demonstrated that the ESC paradigm is a valid measure of the involuntary capture of attention by vocal emotion as reflected by slower responses to targets preceded by angry compared to neutral voices. This was thought to reflect a threat-specific involuntary capture of attention by vocal emotion in adults, resulting in a distracting effect on task performance consistent with findings from dichotic listening paradigms and auditory distraction paradigms in adults (Peschard et al., 2017; Sander et al., 2005; Kattner & Ellermeier, 2018 – see Chapter Four for a more detailed discussion of these studies).

Chapter Four also presented evidence for the effectiveness of the ESC paradigm at measuring the involuntary capture of attention by vocal emotion on a neural level in healthy adults. For instance, adults presented with a larger parietal P3 response (290-450ms) to targets preceded by spatially incongruent compared to congruent vocal cues when targets were presented in the right visual field. A larger Target-P3 response to incongruent compared to congruent trials within spatial cueing paradigm is thought to reflect the use of greater cognitive resources to engage attention on the relevant target, due to the involuntary capture of attention by the irrelevant cueing stimuli (Wang et al., 2022). In addition, a larger P3 amplitude to targets preceded by congruently located positive compared to neutral irrelevant cues has similarly been suggested to reflect the greater use of attentional resources recruited to maintain attention on the location of the irrelevant auditory emotional cueing stimulus (Wang et al., 2022). Stronger evidence for the involuntary capture of vocal emotion in the ESC paradigm would derive from observing a larger parietal-occipital Target-P1 (100-170ms) component to targets preceded by spatially congruent emotional (happy and angry) voices compared to spatially congruent neutral voices. The P1 component reflects the early visual processing of the target stimuli (Hillyard et al., 1998). A larger P1 to targets preceded by spatially congruent vocal anger has been evidenced to specifically reflect the rapid capture of visual spatial attention by threat-related stimuli in previous dot-probe and spatial cueing paradigms in healthy adults (Brosch et al., 2009; Zimmer et al., 2022 – see Chapter Four for a more detailed discussion of this literature). Thus far, no studies have employed a similar ESC paradigm to isolate the neural markers of preattentive (i.e., the Cue-P2) and attentive (i.e., the Target-P1 and P3) vocal emotion processing in adolescents. This is necessary prior to exploring atypicalities in the components reflecting preattentive and attentive vocal emotion processing in adolescents with high levels of inattention and hyperactivity.

A separate body of literature has provided evidence for atypicalities in the P2 component to non-emotional auditory stimuli in ADHD, in both preattentive (Lazzaro et al., 2001) and attentive (Barry et al., 2009; Oades et al, 1997) contexts. For example, adolescents with ADHD (aged 11-17 years) presented with a larger P2 amplitude at midline electrodes (Fz, Cz, Pz) to irrelevant background auditory tones (50ms, 1000Hz, 60dB) as well as target auditory tones (50ms, 1500Hz, 60dB; Lazzaro et al., 2001). The finding that the P2 was larger to irrelevant auditory stimuli may suggest that the P2 reflects a preattentive sensory hypersensitivity to auditory stimuli in ADHD. This is consistent with studies employing oddball tasks which instruct participants to silently count the number of auditory target tones whilst ignoring visual stimuli, and similarly reported larger P2 amplitudes to the auditory targets in children (Oades et al., 1997) and adults (Barry et al., 2009) with ADHD compared to healthy controls. The interpretation that a larger P2 amplitude is reflective of a sensory hypersensitivity to auditory stimuli in ADHD is consistent with reports that ADHD is associated with higher-than-normal levels of auditory hyperacuity (Rali et al., 2020; Paulin et al., 2016); a perceived sensory hypersensitivity to sounds within the environment leading to intolerance of sounds in noisy situations. Consistent with this, auditory hyperacuity, as measured by a hyperacuity questionnaire, in adults without an ADHD diagnosis was associated with longer P2 latency and larger P2 amplitude in response to changes in frequency (e.g., 250 Hz to 275 Hz) and intensity (e.g., 60 dB to 68 dB) of auditory tones (Zlotnik et al., 2018). In addition, auditory hyperacuity is often associated with other psychopathology comorbid with ADHD, such as depression and anxiety (Paulin et al., 2016). To date, there are no studies which have explored the effect of emotion on the hypersensitivity to auditory stimuli in ADHD. If there is a specific effect of emotion (i.e., angry, happy) on the P2 amplitude in ADHD, this would suggest a specific preattentive hypersensitivity to vocal emotion in individuals with ADHD, in support of perceptual processing theories of ADHD (Sonuga-Barke, 2005; 2012).

Finally, previous research has reported ERP atypicalities relating to the effect of irrelevant (i.e., preattentively processed) emotional stimuli on attention allocation (i.e., as reflected by the P3 component) in children with ADHD (López-Martín et al., 2015). For example, in an ERP study, López-Martín et al. (2015) employed a go/no-go task with emotional (positive, negative) and neutral pictures used as background contexts. Letters (A, B and X) were embedded within these pictures, which acted as go/no-go stimuli, thus, the emotional content was irrelevant to task demands. Twenty-four children (8 to 13 years) with

ADHD and a control group of 24 healthy children participated. Results revealed the frontocentral P3 component to no-go trials was larger during negative and positive background contexts in children with ADHD, but not controls. The authors suggested that this may reflect the greater use of cognitive resources recruited in ADHD to disengage from and inhibit responses in emotional situations (López-Martín et al., 2015). This is thought to reflect the involuntary capture of attention by irrelevant emotions causing later difficulties to disengage from irrelevant stimuli and focus on task demands. This is consistent with the behavioural literature suggesting children with hyperactivity are more distracted by irrelevant angry vocal cues (Chronaki & Marsh, 2024). As of yet, no studies have explored ERP atypicalities reflecting the involuntary capture of attention by preattentive vocal emotion in adolescents with traits of inattention and hyperactivity. In the ESC paradigm, this would be reflected by differences in the parietal occipital P1 and parietal P3 components to targets preceded by spatially congruent and incongruent emotional (e.g., angry and happy) compared to neutral voices, consistent with the adult literature (Chapter Four; Wang et al., 2022; Brosch et al., 2009; Zimmer et al. 2022). A larger P1 to targets preceded by spatially congruent angry voices is expected to reflect the threat-specific involuntary capture of early visual spatial attention in adolescents with inattention and hyperactivity, whilst a larger P3 to targets preceded by spatially congruent and incongruent emotional voices (angry and happy) is expected to reflect the greater cognitive resources recruited in individuals with ADHD in emotional contexts to engage in task-relevant target processing.

The Present Study

The present study primarily aimed to employ the ESC paradigm with electrophysiological and behavioural measures, to investigate the preattentive processing of vocal emotion, and the involuntary capture of vocal emotion in adolescents with varying levels of inattention and hyperactivity.

First the present study aimed to isolate the behavioural and neural markers of preattentive and attentive vocal emotion processes in a community sample of adolescents. Second, the present study aimed to explore associations between preattentive and attentive processes and levels of inattention and hyperactivity in adolescents. Therefore, the following set of aims were made in relation to behavioural and neural markers of the preattentive processing and involuntary capture of vocal emotion in a community sample of adolescents:

I. To investigate the effects of vocal emotion on the involuntary capture of attention in adolescents on a behavioural level by measuring RTs to the visual target.
II. To isolate the neural markers of preattentive vocal emotion processing in adolescents (i.e., the P2 component to the vocal cue).
III. To investigate the P1 and P3 components to the visual target as neural markers of the involuntary capture of attention by vocal emotion in adolescents. Specifically, the study aimed to measure the effect of vocal emotion on a) the early visual processing of the target by measuring the P1 component, and b) the later attentional allocation to the target by measuring the P3 component.

First, consistent with findings from the adult literature, and previous research employing similar spatial cueing paradigms (Brosch et al., 2009; Zimmer et al., 2022 – see Chapter Four for a detailed discussion of these studies), the following predictions were made in relation to these first set of aims.

i. Reaction Time:

a) There will be a main effect of congruency on RTs. Adolescents will present with faster RTs to congruent compared to incongruent trials.

b) There will be a main effect of emotion on RTs. Adolescents will respond slower to trials preceded by angry compared to neutral voices.

ii. Cue-Locked ERPs:

a) There will be a main effect of emotion on the P2 amplitudes. Adolescents will present with larger P2 amplitudes to emotional (angry and happy) compared to neutral voices.

iii. Target-Locked ERPs:

a) There will be a main effect of congruency on the P1[11] and P3 amplitudes to the visual target. Adolescents will present with a larger P1 amplitude to congruent compared to incongruent trials. Adolescents will present with a larger P3 amplitude to incongruent compared to congruent trials.

b) There will be a main effect of emotion on the P1 and P3 amplitudes to the visual target. Adolescents will present with larger P1 and P3 amplitudes to visual targets preceded by emotional compared to neutral voices.

c) There will be an emotion x congruence interaction effect on the Target-P1 and P3. The P1 and P3 amplitudes will be larger to targets preceded by spatially congruent emotional (angry and happy) compared to neutral voices.

[11] We additionally explored effects of emotion and congruency on an earlier C1-like component to the visual target, based on prior literature and visual inspection of the individual averages for the target data. The rationale, analyses and discussion relating to this component are presented in Appendix F.

Second, to investigate the presence of a preattentive hypersensitivity to vocal emotion in adolescents with varying levels of inattention and hyperactivity, the following set of aims were formulated:

I. To investigate associations between inattention and hyperactivity and the involuntary capture of vocal emotion on a behavioural level (i.e., RTs).

II. To investigate associations between inattention and hyperactivity and the preattentive processing of vocal emotion (i.e., Cue-P2).

III. To investigate associations between inattention and hyperactivity and the effects of vocal emotion on downstream target processing. Specifically, the present study aimed to explore associations with the effect of vocal emotion on a) the early visual processing (i.e., P1) of the target, and b) later attention allocation (i.e,. P3) to the target.

Although recent developmental research suggests there are threat-specific perceptual processing atypicalities in ADHD (Chronaki et al., 2015; Chronaki & Marsh, 2024), predictions were made based on previous research which suggests adolescents with ADHD present atypicalities in perceptual processing in both positive and negative contexts (Chronaki et al., 2019). Therefore, the following predictions were made:

i. Inattention and hyperactivity will be positively associated with reaction times to trials preceded by emotional (angry and happy) compared to neutral voices.

ii. Inattention and hyperactivity will be positively associated with the P2 amplitude to emotional (angry and happy) compared to neutral voices.

iii. Inattention and hyperactivity will be positively associated with the P1 amplitude to targets preceded by spatially congruent emotional (angry and happy) compared to neutral vocal cues.

iv. Inattention and hyperactivity will be positively associated with the P3 amplitude to targets preceded by emotional (angry and happy) compared to neutral vocal cues.

5.3 Methods.

Participants.

Sixty-Eight adolescents (age range: 12-16) were recruited from the community through schools in the local area, social media and word of mouth. Participants with a minimum of 60% correct and artefact free trials were included in ERP analyses (see page 206 for details). Behavioural and ERP data were available from 63 adolescent participants. However, in the vocal cue ERP analyses, an additional 2 participants were excluded due to a low percentage of clean epochs (below 60%). In the target ERP analyses, 3 participants were excluded due to poor data quality following visual inspection of the individual participant averages. In total, complete vocal cue ERP data were available from 61 participants, and complete target ERP data were available from 60 participants. This was in accordance with power calculations (G*Power software) which indicated that a sample size of 59 individuals can detect an effect size of 0.32 (a priori calculation) with an error probability of 0.05 (Correlational design: Bivariate normal model) when investigating associations between reaction time and ERP measures and traits of inattention/hyperactivity in adolescents. All participants had normal hearing (under 25dB), as measured by an audiometer, no neurological diagnoses, or visual impairment, and an IQ greater than 70. Seven participants had a formal diagnosis of ADHD. Table 5.1 presents the sample characteristics.

Table 5.1

Descriptive Statistics (Means, SDs) of Sample Characteristics (n=63).

	Adolescents
Age	13.62 (1.34)
Gender (% Male)	57%
Handedness (% Right)	93.7%
IQ	103.05 (12.99)
Hyperacuity	11.81 (8.16)
Hearing Threshold (dB)	11.86 (5.28)
Psychopathology	
SDQ	
Inattention/Hyperactivity	4.95 (2.75)
Hyperactivity	2.79 (1.67)
Inattention	2.16 (1.23)
Conduct Problems	1.98 (1.76)
GHQ Depressive Symptoms	1.71 (2.54)
ERQ	
Emotion Reappraisal	18.14 (4.60)
Emotion Suppression	12.30 (3.26)

Note, psychopathology characteristics were measured using the Strengths and Difficulties Questionnaire (SDQ), General Health Questionnaire (GHQ) and Emotion Regulation Questionnaire (ERQ). See pages 203-204 for full details of these measures.

Measures

The Emotional Spatial Cueing (ESC) Task. Adolescent participants completed the ESC paradigm as presented in the Pilot Study (Version 2 employed in Experiments 2 and 4; see page 148), which included the same vocal stimuli from the Maurage et al. (2007) battery used in the ESC paradigm in Chapter Four (See page 148) and validated in Chapter Three (see page 108).

Weschler Abbreviated Scale of Intelligence (WASI-II). The WASI-II (Weschler, 2011) is the same measure of intelligence used with adult participants in Chapter Four. Performance on the WASI sub-tests has been reported to be close to the normative average in clinical cases of ADHD, however typically scores are slightly lower in ADHD compared to matched controls (Weschler, 2011). Adolescents completed the Matrix Reasoning and Vocabulary sub-tests (See page 203 for details). Raw scores from the matrix reasoning and vocabulary sub-test were converted to an IQ score based on the age-appropriate standardised scores according to the WASI manual (Weschler, 2011) The WASI has good reliability (r=0.86 to 0.91, for the vocab; r=0.86-0.91, for the matrix) and good test-retest stability (r=0.87, for the vocab; r=0.77, for the matrix) in adolescents aged 12 to 16.

Hyperacusis Questionnaire (HQ). The HQ (Khalfa et al., 2002) is the same 14-item self-report questionnaire which measures hyperacuity as employed with adults in Chapter Four (See page 284). The HQ has also previously been used with adolescent participants in a study exploring hyperacusis in ASD (Danesh et al., 2015). In the present study, the HQ had good internal reliability in adolescents (Chronbach's alpha = 0.90).

Puretone Audiometric Assessment. To ensure participants' hearing threshold was within the average range of 0-25dB following the British Society of Audiology Recommended

Procedures (2018), we conducted audiometric testing using a clinical audiometer (Type 4 tone, Interacoustics, AS608). The following instructions were given to participants before testing:

"I am going to test your hearing by measuring the quietest sounds that you can hear. As soon as you hear a sound (tone), press the button. You might hear the sound from either your left ear or your right ear. You should press the button no matter which ear you hear it in. Release the button as soon as you think you no longer hear the sound."

After ensuring participants understood the instructions, and the headphones were correctly placed on the head, participants were asked to focus their gaze in the opposite direction to the researcher and the audiometer. Participants response to the tone was indicated on the audiometer by a light which only the researcher could see. Tones were presented to the adolescents starting at a frequency of 1500Hz, followed by 1000 and 500Hz. These were first presented to the right ear and then to the left ear. The hearing threshold for each ear and frequency was obtained by presenting the tone first at 30dB, and following a satisfactory response from the participants, the decibel of the tone was reduced in 10dB steps (i.e., from 30 to 20 to 10dB) until no further response occurred. Next, the decibel of the tone was increased by 5dB until a response was made. To ensure reliability, the decibel was reduced by 10dB and increased by 5dB again until the participant responded. The researcher ensured the duration of the tone and the interval between the tones varied between 1 and 3 second so the timing of the tone was not predictable to the participant. The hearing threshold for each frequency level was defined as the lowest decibel the participants responded to in at least half of all trials. The overall hearing threshold was calculated as the mean of the six thresholds obtained across both ears (three frequencies in each ear).

Measures of Adolescent Psychopathology

The Strengths and Difficulties Questionnaire (SDQ). The SDQ (Goodman, 1997) screens for behavioural and emotional difficulties in children and adolescents, aged between 3 and 16 years old. In the present study the parent, teacher and self-report version was used. Both parent/teacher and self-report versions of the SDQ consist of 25-items scored on a 3-point Likert-scale including 0 (not true), 1 (somewhat true), and 2 (certainly true). The SDQ includes five subscales measuring hyperactivity/inattention, conduct problems, emotional symptoms, peer problems, and prosocial behaviours. A total score for each subscale can be obtained by adding up the scores for each of the items relating to each independent sub-scale (Goodman & Goodman, 2009). Because our study focused on externalising behaviour, we analysed only 3 subscales: inattention/hyperactivity (5 items), conduct problems (5 items) and emotional problems (5 items) scales for each informant (parent, teacher, self). The three subscales employed in the current study had good internal reliability for the self-report (Cronbach's alpha =0.77), parent-reported (Cronbach's alpha = 0.84), and teacher-reported (Cronbach's alpha = 0.87) measure. The SDQ scores can be used as a continuous scale, or participants can be categorised into four categories for each sub-scale based on cut-off scores. For example, within the parent-rated inattention/hyperactivity scale, a score of 0-5 is classified as 'Close-to-average' in which 80% of the population should fall into, a score of 6-7 as 'Slightly raised' (10% of the population), 8 as 'High' (5% of the population) and 9-10 as 'Very High' (5% of the population; Goodman & Goodman, 2009). For the purposes of the present theses, as our primary aim was to explore how traits of inattention and hyperactivity were associated with measures on the Emotional Spatial Cueing paradigm, we treated SDQ scores on a continuous scale. However, to ensure we had a normal distribution of inattention/hyperactivity scores, we employed the parent/teacher version as a pre-screening measure during the recruitment of adolescents with a range of scores from 0-10. In order to obtain a representation of high and low ADHD traits in the overall sample, we recruited adolescents so that 50% of participants had a score of 0-5 and 50% of participants had a score of 6-10 on the inattention/hyperactivity scale.

The General Health Questionnaire (GHQ- 12). We employed the same self-report GHQ (Goldberg and Williams., 2000) in adolescents as we did in adults (See page 284) to measure depressive symptoms. This questionnaire has been validated for use in adolescents (aged 11 onwards) to measure psychological wellbeing (Tait et al., 2003) and depressive traits (Lundin et al., 2016). A higher score reflects greater depressive symptoms. The GHQ had good internal reliability in adolescents within the present study (Cronbach's alpha = 0.85). This is consistent with reliability scores in adults in Chapter Four (0.84)

The Emotion Regulation Questionnaire for Children and Adolescents (ERQ-CA). This is a self-report questionnaire for children and adolescents (aged 9 -18) consisting of 10 items scored on a 5-point Likert scale (Gullone & Taffe, 2012). This is a revised version of the Emotion Regulation Questionnaire previously designed for adults (Gross & John, 2003). This questionnaire consists of two sub-scales: emotion suppression and cognitive reappraisal. Emotion suppression is a measure of emotion dysregulation, which involves inhibiting or restraining emotional expressions and is measured via four items. Example questions measuring emotion suppression include 'I keep my feelings to myself' or 'I control my feelings by not showing them'. Cognitive reappraisal is an emotion regulation strategy, which involves thinking about a situation differently, and is measured via six items. Example questions measuring emotion reappraisal include 'When I want to feel happier, I think about something different' or 'When I want to feel less bad (e.g., sad, angry or worried), I think about something different.' In this study, the cognitive reappraisal scale had good internal reliability (Cronbach's alpha = 0.82), which is consistent with previous reports in a similar age group (Cronbach's alpha = 0.83; Gullone & Taffe, 2012). The internal reliability of the emotion suppression scale in the present study (Cronbach's alpha = 0.57) was slightly weaker than previous reports (Cronbach's alpha = 0.75; Gullone & Taffe, 2012).

Procedure

Informed written consent was gathered from parents or guardians of adolescents. Adolescent participants provided written assent. The researcher informed participants at the start of the study that they could withdraw their participation at any time. First, participants' hearing was measured with the clinical audiometer. Next, the researcher set up the EEG equipment and the participants completed the ESC Task. After completion of the experimental task, adolescent participants completed the relevant questionnaire measures for behavioural (e.g., IQ, Hyperacuity) and psychopathology characteristics. Parents also completed the SDQ at the time of the study[12]. At the end of the study, all participants and their parents were debriefed verbally and in writing. Adolescents were offered a £10 amazon voucher as reimbursement for their time, and parents were offered a £10 amazon voucher to reimburse their travel costs. This study was approved by the UCLan Psychology Ethics Committee (Science 0037).

[12] Note, this was in addition to parents and teachers answering the SDQ pre-screening question via a telephone call prior to completion of the study. Only answers from the SDQ completed at the time of the study were included in analysis.

Behavioural Data Processing

All reaction time data were pre-processed to include only correct responses within the range of 200-1000ms. A small percentage of data (6.34%) was removed due to errors or outliers in the RT data.

Electrophysiological Recording and Processing

Electrophysiological data were recorded using the same 66-electrode cap (Quick cap), Neuroscan and Curry 7 software and online filter settings (0.01 to 100Hz) as in the Experiment 4, Chapter Four (See page 151). Impedance for vEOG, hEOG, references and 64 electrodes were kept below 5KΩ during recording. Consistent with Study 4, the ERP epochs for the Vocal Cue and the Visual Target were defined from 100 to 700ms, and from 50 to 500ms, respectively. The offline filters, ocular artefact reduction procedure (Semlitsch et al., 1989), artefact rejection, and baseline-to-peak mean amplitude (Picton et al., 2000) techniques were applied to the present study as outlined on page 151.

The following components were targeted. The P2 component to the vocal cue was targeted between 130 and 250ms at fronto-central sites (F1, FZ, F2, FC1, FCZ, FC2, C1, CZ, C1). The P1 to the visual target was targeted between 100 and 170ms at parietal-occipital sites (PO7,PO8,PO3,PO4,POZ,O1,O2,OZ). Finally, the P3 to the visual target was targeted between 290 and 350ms[13] at parietal sites (PZ,P1,P2). The inclusion of these sites and time windows were informed by the ERP literature in emotion processing (P2: Jessen & Kotz, 2011; P1: Brosch et al., 2009; P3: Broyd et al., 2012) and visual inspection of the individual participant

[13] This time window differs slightly to the P3 time window used in Chapter 4 due to visual inspection of the time window. In the present study, the P3 appeared to be captured completely prior to 350ms. This differs to previous studies which reported the P3 in the time window of 250-450ms (Broyd et al., 2012), which was used in Chapter 4.

averages. Mean amplitude was initially calculated for each individual site. Following this, the mean amplitude for each ERP component was calculated as a combined score for a number of groups of electrode sites (scalp regions), in order to increase the reliability of the measurement (Dien & Santuzzi, 2005). Based on visual inspection of the data showing a C1-like component, we additionally explored an early C1 component in adolescents targeted between 50 and 100ms at parietal-occipital sites (PZ, POZ, PO3 and PO4; see 0 for details).

A minimum of 60% artefact-free epochs for each condition per participant were used for calculating ERP averages. This criterion was more flexible than those in the previous analysis with adult participants (Study 4) to ensure the inclusion of adolescents who presented with higher levels of inattentive and hyperactive traits. However, this still ensured a minimum of 20 trials per condition per participant were used to calculate ERP averages per condition. The mean (M), standard deviation (SD) and the minimum and maximum (MM) number of epochs included in analyses per condition was as follows, VOCAL CUE: Angry (M: 86.74, SD: 8.59, MM: 63-96); Happy: (M: 87.02, SD: 7.75, MM: 60-96); Neutral: (M: 87.08, SD: 8.20, MM: 62-96); TARGET: Angry Congruent (M: 42.00, SD: 5.91, MM: 23-48); Happy Congruent (M: 42.28, SD: 6.01, MM: 22-48); Neutral Congruent (M: 42.25, SD: 5.13, MM: 22-48); Angry Incongruent (M: 41.98, SD: 5.78, MM: 24-47); Happy Incongruent (M: 41.60, SD: 6.05, MM: 24-48); Neutral Incongruent (M: 41.73, SD: 5.62, MM: 22-48). Note, there was a total possible number of 96 trials per emotion condition for the cue analysis, and a total possible number of 48 trials per condition for the target analysis. There was no difference in the number of artefact-free trials between conditions for the cue ($p=.798$) and the target ($p>.100$).

5.4 Analytic Strategy

Three sets of analyses were conducted to meet the aims of the present study.

In Analysis One (page 210) we examined whether the ESC paradigm produced the expected effects on a behavioural and neural level in adolescents. Therefore, variables relating to adolescent psychopathology and emotion regulation were not included in this analysis. In accordance with our aims and predictions, emotion (angry, happy, neutral) and congruency (congruent, incongruent) were included as IVs in analyses. However, as we found an effect of hemifield on the Target-P3 in our adult data, and Brosch et al (2009) reported an effect of hemifield on Target-P1 data, we included target hemifield as an independent variable in a secondary exploratory analysis of the Target-P1 and P3 data. Consistent with Chapter Four, IQ and behavioural characteristics (e.g., age, hyperacuity, hearing threshold) were explored as covariates in this analysis because they may influence task performance For example, hyperacuity has been found to be associated with larger a larger P2 component to voices (Zlotnik et al., 2018), and hearing threshold might impact performance as those with a higher hearing threshold may more easily ignore voices than those with a low hearing threshold. To make comparisons between our adult and adolescent data, additional exploratory analyses comparing the effects of age group are presented in Appendix G. In addition, since expected effects of congruency were not found on the P1 component in adults, exploratory analyses were conducted on an earlier target C1-like component to investigate if congruency effects appear earlier than expected within the ESC paradigm (see 0).

In Analysis Two (page 220), ADHD was defined as a dimensional variable consistent with recent theory (Sonuga-Barke, 2023), to first examine associations between traits of inattention/hyperactivity and preattentive vocal emotion processing (i.e., via the P2). In addition, Analysis Two examined associations between inattention/hyperactivity and the downstream effects of vocal emotion on target processing (i.e., via RTs, the P1 and P3 response). To more robustly explore associations between effects of emotion (e.g., angry, happy) on the preattentive and attentive processing of vocal emotion, and traits of inattention/hyperactivity, difference scores were calculated prior to conducting correlational analyses, with the neutral condition as a baseline. Calculations of difference scores for RTs, Cue-P2, Target-P1, and Target-P3 are outlined on pages 220 and 221. Emotion regulation, conduct problems and depressive symptoms were explored as covariates, along with IQ and behavioural characteristics, because they are all highly correlated with ADHD and have been found to be associated with ADHD traits and vocal emotion recognition in previous studies (Chronaki et al., 2015). To control for IQ, behavioural characteristics (age, hyperacuity, hearing threshold), emotion regulation, and comorbidities of ADHD (conduct problems, depressive symptoms), preliminary Spearman's Rho[14] correlations were conducted between these variables, inattention/hyperactivity and all dependent measures (i.e., Cue-P2, Target-P1 and Target-P3) prior to the main analysis. Any variables which significantly correlated with either inattention and hyperactivity or the dependent variables were entered as covariates in the main analyses. These preliminary analyses are outlined in section 5.6.

Finally, in Analysis Three (page 232), ADHD was defined as a discrete variable to conduct analyses investigating the effect of ADHD traits (high vs low) on the Cue-P2, Target-P1 and Target-P3 components. This is consistent with the categorical approach to ADHD, as

[14] Non-parametric spearman's rho correlations were conducted because Kolmogorov-Smirnov tests revealed that behavioural characteristics (age, hearing threshold, IQ, hyperacuity), emotion suppression and comorbidities of ADHD (conduct problems, depressive symptoms) were not normally distributed. See Appendix F for the KS statistics.

outlined in diagnostic manuals (e.g., DSM-IV) and previous research exploring emotion processing in ADHD (Chronaki et al., 2015). Within Analysis Three, a subset of participants, presenting with very low or very high ADHD traits, was included in analyses (sees page 232 for details).

5.5 Analysis One

Initial Data Treatment

Behavioural Data Treatment. A Kolmogorov-Smirnov test indicated that RT scores for the happy congruent and neutral incongruent trials did not follow a normal distribution (p=.007; p=.004), however all other RT scores did follow a normal distribution (p=.058 to .096). Additionally, Kolmogorov-Smirnov tests indicated that behavioural characteristics (age, IQ, hyperacuity, hearing threshold) were not normally distributed (p<.05). Spearman's rho correlations revealed a significant negative correlation between age and RTs (r=-0.47, p<.001). The younger the adolescent, the slower they responded to the target. There were no other significant correlations between RTs and behavioural characteristics (p>.05; See Table F.3 in the appendices for complete statistics). In addition, an independent sample t-test revealed no significant differences in RTs between males and females (t(60) = 1.54, p=.130).

ERP Data Treatment. Kolmogorov-Smirnov tests indicated that all amplitude values for the Cue-P2, Target-P1 and Target-P3 were normally distributed for most of the emotion and congruency conditions (p=.069 to .200), apart from the amplitude values for the Target-P3 to happy congruent and incongruent conditions (p=.027; p=.022). Table F.1 in the appendices presents the relevant statistics. However, because behavioural characteristics were nonparametric (see Table F.2 in appendices), Spearman's rho correlations were conducted to explore preliminary correlations between these variables and ERP component. Spearman's

Rho correlations revealed a significant a negative correlation between the Cue-P2 and age (r=-0.24, p<.05). The younger the adolescent, the larger their overall P2 amplitudes. There were no other significant correlations between behavioural characteristics and ERP component amplitude. Table F.3 in the appendices present statistics (r and p values) for all preliminary correlations. In addition, independent samples t-tests revealed no significant differences in the Cue-P2, Target-P1 and P3 amplitudes between males and females (all ps>.05).

Results

Reaction Times. Table 5.2 shows descriptive statistics of RTs for trials per congruency and emotion in 63 adolescents. As most RTs were normally distributed, parametric tests were conducted to explore the effects of emotion and congruency on reaction times. Consistent with results in adults (see page 171), a 2 (congruency) x 3 (emotion) repeated measures ANOVA revealed significant main effects of congruency (F(1,62) = 20.39, p<.001) and emotion (F(2,124) = 3.26, p= .042) on RT scores. Consistent with our predictions, adolescents were faster to respond to targets preceded by spatially congruent compared to incongruent vocal cues. Additionally, post hoc tests (Bonferroni corrected) revealed adolescents were slower to react to targets preceded by angry compared to neutral voices (MD=4.46, p=.049). Non-parametric tests did not change the pattern of results. However, when age was added as a covariate in analyses, main effects of emotion and congruency on RTs were no longer significant (ps>.05). Figure 5.1 presents a bar graph showing mean RTs for correct responses for emotion by congruency conditions.

Table 5.2.

Means (SD) RTs to Targets Preceded by Vocal Cues Differing by Emotion (Angry, Happy, Neutral) and Congruency (Congruent, Incongruent) Conditions in Adolescents (n=63).

	Congruent	Incongruent	Total
Angry	360.11 (52.20)	365.88 (57.49)	362.99 (54.15)
Happy	356.22 (50.68)	365.49 (55.09)	360.85 (52.03)
Neutral	356.18 (50.75)	360.89 (51.73)	358.54 (50.53)
Total	357.50 (50.20)	364.09 (53.67)	

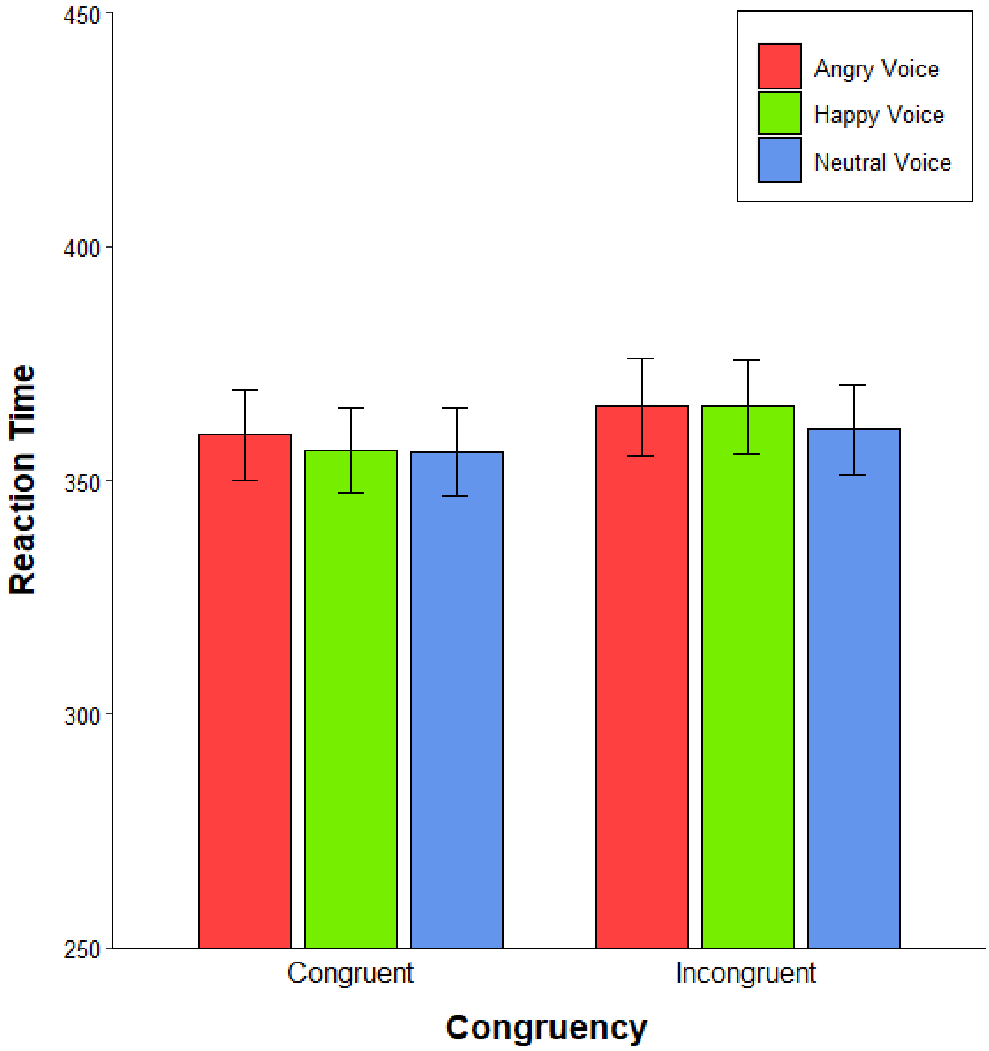

Figure 5.1

Bar Graph (with Standard Error Bars) Showing Mean RTs (in ms) to Targets Differing by Emotion (Angry, Happy, And Neutral) and Congruency (Congruent And Incongruent) Conditions in the ESC Task in Adolescents (n=63)

Vocal Cue-P2. Consistent with our predictions, a one-way repeated measures ANOVA revealed a main effect of emotion on P2 amplitudes ($F (2,120) = 58.91$, $p<.001$). Post-hoc tests (Bonferroni-corrected) revealed adolescents presented with larger P2 amplitudes to happy and angry compared to neutral voices (neutral and happy: $MD = 1.64$, $p<.001$; neutral and angry: $MD = 1.31$, $p=.005$). Additionally, adolescents presented with a larger P2 amplitude to happy compared to angry voices ($MD=0.33$, $p=.048$). These findings suggest the P2 is a neural marker of preattentive vocal emotion processing in adolescents. However, when age was added as a covariate, the main effect of emotion on the P2 amplitude did not reach significance ($F (2,118) = 2.69$, $p=.072$). Figure 5.2 presents the grand averages of the Cue-P2 by emotion condition. Table F.4 in Appendix F presents the means and SDs for the P2 component to the vocal cue for each emotion condition.

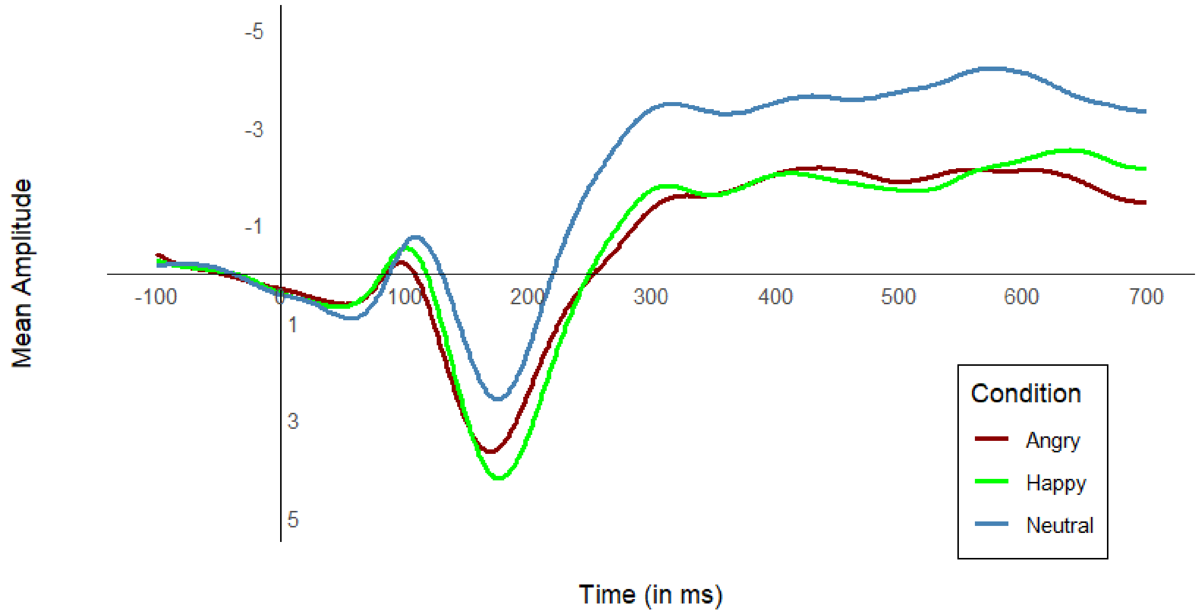

Figure 5.2

Grand Average Showing the Frontocentral P2 Component (130-250 ms) to Angry, Happy and Neutral Vocal Expressions in Adolescents (n=61). Scale is -5 to + 5 μV.

Target-P1. Figure 5.3Figure 5.2 presents the grand averages of the Target-P1 by congruency and emotion. Consistent with findings in the previous study with adults, a 2 (congrurncy) x 3 (emotion) repeated measures ANOVA revealed no main effect of congruency on Target-P1 amplitudes (F (1,59) =0.53, p=.469). However, results revealed a main effect of emotion on Target-P1 amplitudes (F (2,118) = 7.83, p=<.001). Specifically, the P1 amplitudes were larger to targets preceded by angry compared to neutral (MD=0.61, p=.001) and happy (MD=0.64, p=.003) voices. Finally, there was no interaction effect between emotion and congruency on Target-P1 amplitudes (F (2,118) = 0.50, p=.608). Interestingly, the effects of emotion on the P1 amplitude continued to survive correction when controlling for the C1 component (50-100ms) in analyses (F (2,116) = 7.94, p<.001). Table F.5 in Appendix F presents the mean P1 amplitudes time-locked to the visual target for all emotion and congruency conditions.

As a secondary exploratory analysis, a 2 (congruency) x 3 (emotion) x 2 (target hemifield) repeated measures ANOVA was conducted to explore if congruency effects on the Target-P1 amplitude were present when the target was presented to the left hemisphere, as reported by Brosch et al (2009). Consistent with our findings in adults, results revealed there was no significant main effect of hemifield ($F(1,59) = 2.06$, $p = .157$), or any significant hemifield x congruency, hemifield x emotion or hemifield x congruency x emotion interaction effects on the Target-P1 amplitude ($p > .05$).

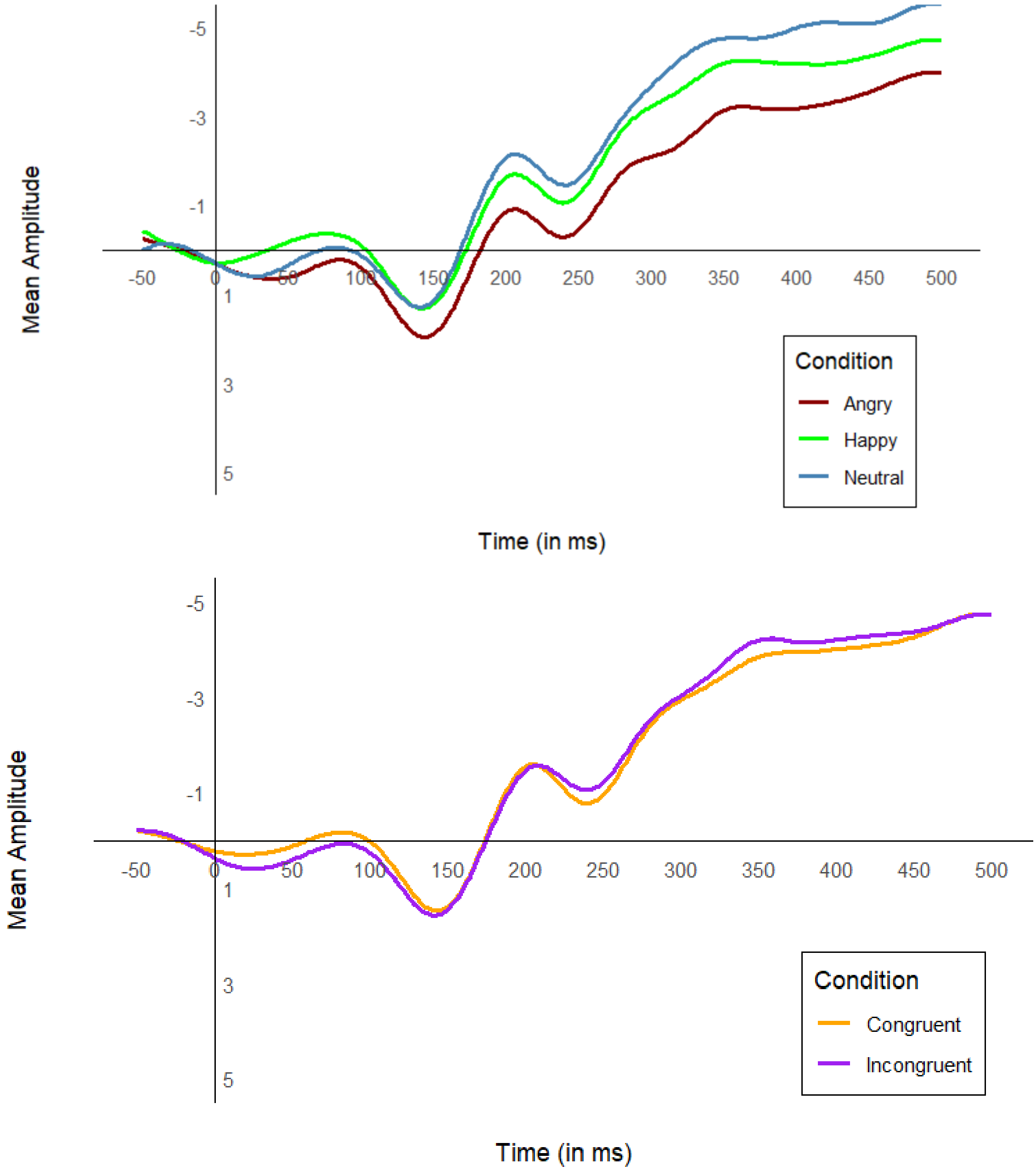

Figure 5.3

Grand Average Showing the Parietal Occipital P1 Component (100-170ms) to the Target in a) Emotion (Angry, Happy, Neutral) and b) Congruency (Congruent, Incongruent) Conditions in Adolescents (n=60. Scale is -5 to 5 mv

Target-P3. A 2 (congruency) x 3 (emotion) repeated measures ANOVA was conducted. There was a main effect of congruency on the Target-P3 amplitude (F (1,59) =4.25, p=.044). Consistent with our predictions, adolescents presented with a larger P3 amplitude to targets in incongruent compared to congruent trials. There was also a main effect of emotion on the Target-P3 amplitude (F (2,118) = 4.84, p=.010). Specifically, post hoc results using t-tests (Bonferroni corrected, alpha value =.025) revealed the P3 was significantly larger to targets preceded by angry compared to neutral voices (t (59) =3.00, p=.002), consistent with predictions. The difference between the P3 amplitudes to targets preceded by happy compared to neutral voices was marginally significant (t (59) = 1.89, p=.032), however did not survive correction for multiple comparisons. There was no significant congruency by emotion interaction effect on the Target-P3 amplitude (F (2,118) = 0.20, p=.738). Figure 5.4 presents the grand averages of the Target-P3 by congruency and emotion. Table F.6 in Appendix F presents the mean P3 amplitudes time-locked to the visual target for all emotion by congruency conditions.

As a secondary exploratory analyses, target hemifield was added as an IV in a 2 (congruency) x 3 (emotion) x 2 (hemifield) analyses. There was no main effect of hemifield (F (1,59) = 0.14, p=.708) on the Target-P3 amplitude. In addition, there were no congruency x hemifield (F (1,59) = 2.52, p=.118), emotion x hemifield (F(2,118) = 1.88, p=.161 or congruency x emotion x hemifield (F (2,118) = 1.62, p=.201) interaction effects on the Target-P3 amplitude.

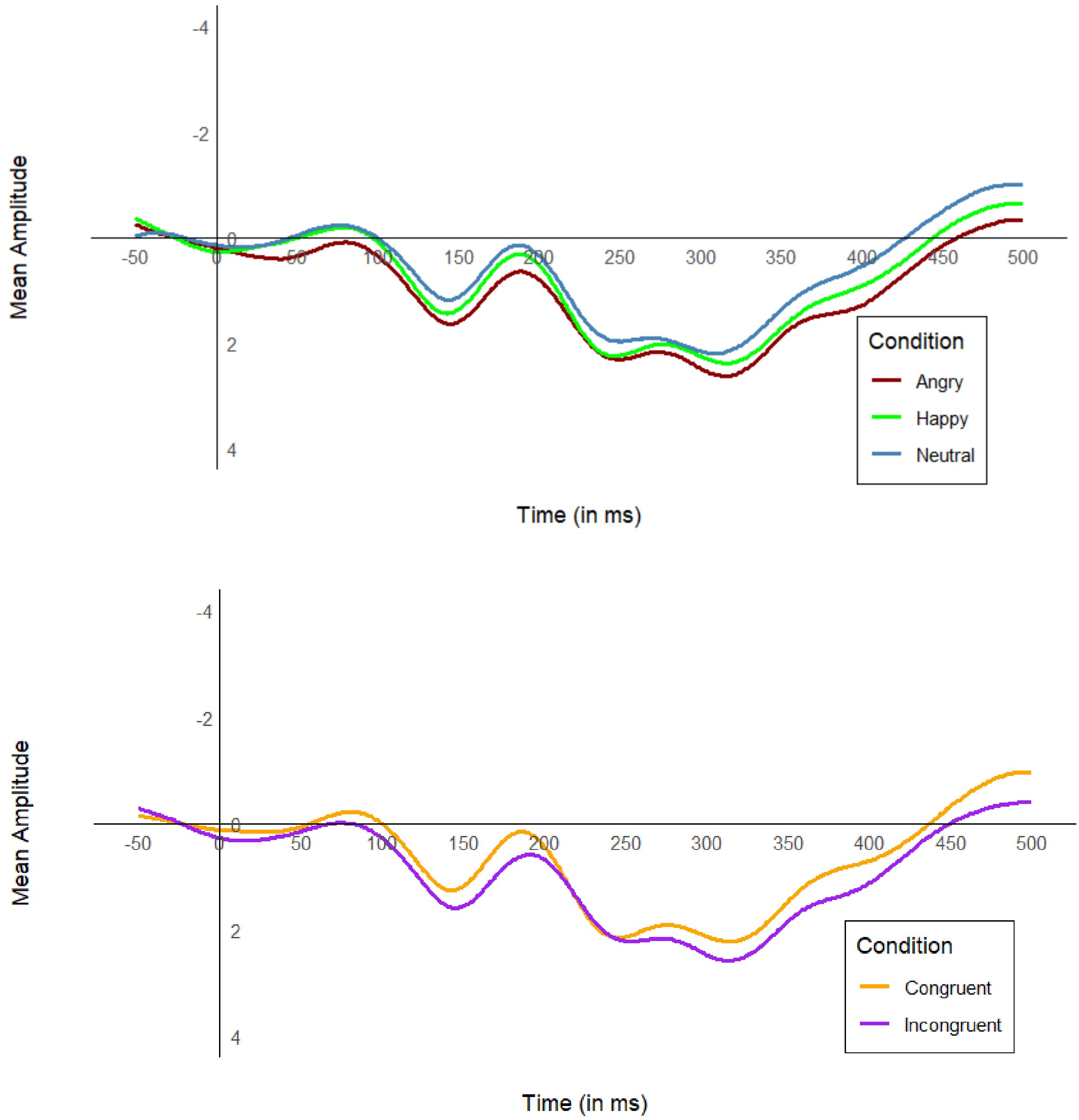

Figure 5.4

Grand Averages Showing the Parietal P3 Component (290-350 ms) to the Target in a) Emotion (Angry, Happy, Neutral) and b) Congruency (Congruent, Incongruent) Conditions in Adolescents (n=60). Scale is -4 to + 4 µV.

5.6 Analysis Two. The relationship between ERP components and traits of inattention and hyperactivity

Behavioural Data Treatment.

RT difference scores were calculated to more accurately measure anger-specific and happiness-specific effects within the ESC paradigm. For example, an angry-neutral RT difference score was calculated. A positive Angry-Neutral RT difference score would reflect slower RTs to targets preceded by angry compared to neutral voices. The same was applied to happy voices. Kolmogorov-Smirnov tests indicated that all RT difference scores followed a normal distribution (p>.05).

Preliminary Spearman's rho correlations between psychopathology (inattention/hyperactivity, conduct problems, depressive symptoms), emotion regulation scores, IQ, and behavioural characteristics (age, auditory hyperacuity, hearing threshold) are presented in Table F.3 in the appendices. These correlations revealed no evidence of collinearity between our variables. However, there was a strong positive correlation between hyperactivity and inattention and conduct problems (r=0.66, p<.001), consistent with the literature (Gustafsson et al., 2018; Taylor et al., 1996). In addition, IQ, and hearing threshold were negatively associated with inattention and hyperactivity. Auditory hyperacuity was positively associated with all psychopathology measures (inattention/hyperactivity, depressive symptoms, and conduct problems). Therefore, conduct problems, hyperacuity, IQ and hearing threshold were entered as covariates in analyses.

In addition, Spearmans' rho correlations between RTs, RT difference scores, behavioural characteristics, emotion regulation scores, and psychopathology measures (conduct problems, depressive symptoms) are presented in Table F.8. Results revealed significant correlations between RTs and age ($p<.001$). Therefore, age was included as an additional covariate in RT correlational analyses.

ERP Data Treatment

Vocal Cue ERPs. Difference scores were calculated for the Cue-P2 amplitude. Specifically, an Angry-Neutral P2 amplitude score was calculated to capture anger-specific effects. The same was repeated with happy with a Happy-Neutral P2 amplitude difference score. Therefore, a greater Angry-Neutral and Happy-Neutral P2 difference scores are considered to reflect a greater difference in the preattentive perception of emotional compared to neutral voices. Kolmogorov-Smirnov tests indicated that all P2 amplitude difference scores were normally distributed ($p>.05$). Preliminary Spearmans' rho correlations between the Cue-P2 difference scores, behavioural characteristics, emotion regulation scores, and adolescent psychopathology measures (conduct problems, depressive symptoms) are presented in Table F.9. Results revealed significant correlations between the Cue-P2 amplitude to happy voices and age ($r=-0.26$, $p=.043$) and IQ ($r=-0.25$, $p=.048$). In addition, there was a significant correlation between hearing threshold and Angry-Neutral ($r=-0.32$, $p=.026$) and Happy-Neutral ($r=-0.36$, $p=.010$) P2 amplitude difference scores. Finally conduct problems significantly correlated with the Cue-P2 amplitude to happy voices ($r=.035$, $p=.006$) and the Happy-Neutral P2 amplitude difference score ($r=.028$, $p=.028$). Therefore, IQ, age, hearing threshold and conduct problems were explored as covariates in the P2 analyses.

Visual Target ERPs. A Kolmogorov-Smirnov test revealed all target ERP-components amplitudes were normally distributed ($p>.05$), apart from the P1 amplitude scores to targets preceded by happy voices ($p=.035$). As above, difference scores were calculated for the

Target-P1 amplitudes. For instance, difference scores were calculated for only the congruent trials, in line with predictions and previous research, to more accurately reflect the involuntary capture of attention by emotional voices compared to neutral voices (e.g., Angry Congruent – Neutral Congruent; Happy Congruent – Neutral Congruent P1 amplitude). In addition, for consistency with the effects of emotion on early visual processing reported in the previous analysis (page 211), difference scores were calculated as Angry–Neutral and Happy–Neutral P1 amplitude difference scores respectively. In the same way, difference scores were calculated for the Target-P3 amplitudes. Kolmogrov-Smirnov tests indicated that the Target-P1 and P3 amplitude difference scores were normally distributed ($p>.05$).

To explore the presence of covariates prior to data analysis, Spearmans' rho correlations were conducted between Target-P1 and P3 amplitudes, difference scores, behavioural characteristics, emotion regulation scores, and other adolescent psychopathology characteristics (conduct problems, depressive symptoms). Results revealed a significant correlation between the P1 amplitude to targets preceded by spatially congruent happy voices and emotion reappraisal ($r=-0.37$, $p=.004$), and the Angry Congruent- Neutral Congruent P1 difference score ($r=-0.33$, $p=.019$). The Angry Congruent - Neutral Congruent P1 difference score also correlated with hyperacuity and depressive symptoms. Therefore, emotion reappraisal, hyperacuity and depressive symptoms were included as a covariate in the Target-P1 analyses. In addition, emotion suppression significantly positively correlated with the Angry-Neutral and Happy-Neutral P3 amplitude difference scores ($r=.034$, $p=.016$; $r=0.29$, $p=.040$). Emotion reappraisal significantly negatively correlated with the P3 amplitudes to targets following angry ($r=-0.38$, $p=.008$) and happy ($r=-0.32$, $p=.023$) vocal cues. Finally, conduct problems significantly positively correlated with the Happy-Neutral P3 difference score (r=0.33, p=.011). Therefore, emotion suppression, emotion reappraisal and conduct problems were included as covariates in Target-P3 analyses.

Results

Reaction Times. Since the majority of RTs were normally distributed, Pearsons's correlations were conducted to explore the relationships between RTs (for all congruency and emotion conditions) and inattention/hyperactivity. Pearson's correlations also explored the relationship between RT difference scores (Angry-Neutral, Happy-Neutral) and inattention/hyperactivity. In contrast to our predictions, results showed no significant correlations (ps>.05). Additionally, when age, hyperacuity, IQ, hearing threshold and conduct problems were added as covariates, no significant correlations were observed (p>.05). Due to multiple comparisons, the alpha value was Bonferroni adjusted to p=.01 (5 correlations). Results are summarised in Table F.12 in Appendix F.

Cue-P2. Pearson's correlations were conducted to explore the relationships between the Cue-P2 amplitudes to emotional (happy, angry) and neutral voices and hyperactivity and inattention. Due to multiple comparisons, the alpha value was Bonferroni adjusted to p=.01 (5 correlations). In accordance with our predictions, the P2 amplitude to happy voices significantly correlated with inattention/hyperactivity (p<.001). However, surprisingly, there was no significant correlations between the P2 amplitude to angry voices and inattention/hyperactivity. When controlling for age, IQ, hearing threshold and hyperacuity, this did not change the pattern of the results. When controlling for conduct problems, the correlation between inattention/hyperactivity and the P2 amplitude to happy voices did not survive correction for multiple comparisons (r=0.27, p=.019).

Additionally, Pearson's correlations explored the relationship between the P2 difference scores (Angry-Neutral, Happy-Neutral) and inattention/hyperactivity. Results showed the Happy-Neutral P2 amplitude difference score significantly positively correlated with inattention and hyperactivity; however, this did not survive correction for multiple comparisons (p=.013). When controlling for hyperacuity, IQ and age, the pattern of results did not change. When controlling for hearing threshold, the correlation between inattention/hyperactivity and the Happy-Neutral P2 difference score remained significant and survived Bonferroni correction for multiple comparisons (p=.003). When controlling for conduct problems, the correlation between inattention/hyperactivity and the Happy - Neutral P2 difference score did not remain significant (r=0.17, p=.095). Surprisingly, there were no significant correlations between the Angry-Neutral P2 difference score and inattention/hyperactivity. Table 5.3.presents all correlations between the P2 amplitude for all emotion conditions and difference scores with inattention/hyperactivity. Figure F.1 (Appendix F) presents a scatter plot of the correlation between the Happy-Neutral P2 difference score and inattention and hyperactivity.

Table 5.3.

Pearson's Correlations (One-Tailed) Between the Cue-P2 and Inattention And Hyperactivity With and Without Controlling for Conduct Problems (CP) and Hearing Threshold (HT; n=61).

	Inattention/ Hyperactivity	Inattention/ Hyperactivity (controlling for CP)	Inattention/ Hyperactivity (controlling for HT)
Angry	0.28*	0.14	0.19
Happy	**0.41*** **	0.27*	**0.36** **
Neutral	0.22*	0.14	0.11
Difference Scores			
Angry – Neutral	0.11	0.01	0.13
Happy – Neutral	0.28*	0.17	**0.39** **

*Note, *p<.05,* **p<.01, ***p<.001. Correlations in bold indicate correlations which survived correction for multiple comparisons when no other variables were controlled for in analysis. When controlling for both HT and CP at the same time, only the correlation with the Happy-Neutral P2 difference score was significant (r=0.29, p=.020), but did not survive correction.

Target-P1. Pearson's correlations were conducted to examine the relationship between the Target-P1 amplitudes and inattention/hyperactivity. Due to multiple comparisons, the alpha value was Bonferroni adjusted to $p=.007$ (7 correlations). In contrast with our predictions, there were no significant correlations between the Target-P1 amplitude in spatially congruent trials and inattention/hyperactivity (all $ps > .05$). In addition, Pearson's correlations examined the relationship between the Angry Congruent-Neutral Congruent P1 difference score and the Happy Congruent-Neutral Congruent P1 difference score and inattention/hyperactivity. Results revealed no significant correlations between the P1 amplitude difference scores and hyperactivity and inattention. The pattern of results did not change when controlling for hearing threshold, IQ, hyperacuity, emotion reappraisal or depressive symptoms (all $ps >.05$).

Interestingly, when controlling for conduct problems, there was a significant positive association between inattention/hyperactivity and the P1 amplitude to targets preceded by spatially congruent neutral voices ($r=0.33$, $p=.006$), which survived correction. In addition, there was a negative correlation between inattention/hyperactivity and the Angry Congruent – Neutral Congruent P1 difference score ($r=-0.26$, $p=.020$) and the Happy Congruent – Neutral Congruent P1 difference score ($r=-0.31$, $p=.008$). Only the latter survived correction, however the direction of the effect (negative) was inconsistent with the predicted (positive) direction. Specifically, the Happy Congruent-Neutral Congruent P1 difference score was smaller in adolescents with higher compared to lower inattentive/hyperactive traits. Table 5.4 presents all Pearson's correlations between inattention/hyperactivity and P1 amplitude scores in spatially congruent trials. Table F.13 in Appendix F presents additional correlations conducted between the P1 amplitude to targets preceded by emotional voices in all trials (congruent and incongruent) to reflect the effect of emotion found in Analysis One This shows that the pattern of results did not change.

Table 5.4

Pearson's Correlations (One-Tailed) Between the Target-P1 in Congruent Trials and Inattention/Hyperactivity With and Without Controlling for Conduct Problems (CP).

	Inattention/ Hyperactivity	Inattention/ Hyperactivity (Controlling for CP)
Angry Congruent	0.07	0.03
Happy Congruent	0.01	-0.06
Neutral Congruent	0.14	**0.32****
Difference Scores		
Angry Congruent - Neutral Congruent	-0.06	-0.27*
Happy Congruent - Neutral Congruent	-0.11	**-0.31****

*Note, *p<.05,* **p<.01, ***p<.001. Only those in bold survived correction for multiple comparisons.

Target-P3. Pearson's r correlations were conducted to explore the relationship between Target-P3 amplitudes and inattention/hyperactivity (see Table 5.5) Due to multiple comparisons, the alpha value was Bonferroni adjusted to $p=.01$ (5 correlations). In contrast with our predictions, no significant correlations were observed between the Target-P3 and inattention/hyperactivity. When analyses were repeated with the Target-P3 difference scores, results showed no significant correlation between the P3 difference scores and inattention and hyperactivity. The pattern of results did not change when adding emotion reappraisal, emotion suppression, hyperacuity or IQ as covariates in analyses ($p>.05$). However, when controlling for hearing threshold, there was a positive association between inattention/hyperactivity and the P3 amplitude to targets preceded by happy voices ($r=0.31$, $p=.018$). However, this did not survive corrections for multiple comparisons. In addition, when controlling for conduct problems, there was a negative correlation between the Happy - Neutral P3 difference score and inattention/hyperactivity ($r=-0.27$, $p=.019$). However, this also did not survive correction for multiple comparisons.

Table 5.5

Pearson's Correlations (One-Tailed) Between the Target-P3 and Inattention/Hyperactivity With and Without Controlling for Conduct Problems (CP) and Hearing Threshold (HT).

	Inattention/ Hyperactivity	Inattention/ Hyperactivity (controlling for CP)	Inattention/ Hyperactivity (controlling for HT)
Angry	0.03	0.05	0.17
Happy	0.09	-0.04	0.31*
Neutral	0.05	0.13	0.19
Difference Scores			
Angry-Neutral	-0.04	-0.14	-0.04
Happy-Neutral	0.06	-0.27*	0.14

*Note, *p<.05,* **p<.01, ***p<.001. However, none of these survived corrections for multiple texting.

5.7 Analysis Three: Categorical Analyses

To more thoroughly examine the effects of inattention/hyperactivity on the preattentive processing of vocal emotion, a further set of analyses examined if the above effects on Cue-P2 differed between a subset of adolescents with very high (present within 10% of the population; Goodman & Goodman, 2009) symptoms of ADHD compared to adolescents with very low symptoms of ADHD. The primary aims of these analyses were to investigate the effects ADHD traits (High vs Low) on reaction times and the neural markers of pre-attentive vocal emotion processing (Cue-P2). It was predicted that there will be a main effect of Group on the Cue-P2. We predicted adolescents with high levels of inattention/hyperactivity would present with a larger P2 amplitudes compared to adolescents with low levels of inattention/hyperactivity. Second, we predicted a Group x Emotion interaction effect on the Cue-P2. Adolescents with high levels of inattention/hyperactivity would present with a larger P2 amplitude to emotional (angry and happy) compared to neutral voices. In addition, effects of ADHD group on the target processing (RTs, Target-P1 and P3) were explored, but due to the lack of power in this set of analyses, these results are presented in Appendix F.

In total, 26 participants were included in the categorical analyses. Only adolescent participants with extreme low or high levels of inattention and hyperactivity (IH) were included in the present analyses. These two groups (Low IH, High IH) were determined based on the distribution of SDQ Inattention/Hyperactivity scores in the sample (Goodman & Goodman, 2009). Adolescents who scored between '0-2' were categorised into the 'Low IH' group; and adolescents who scored between '8-10' were categorised into the 'High IH' group. Based on the number of adolescents who presented with scores within these ranges from the

61 adolescents included in the previous analysis, a total of 13 adolescents were included in the 'Low IH' group and 12 adolescents were included in the 'High IH' group for the Cue-P2 analyses. Sample characteristics for all adolescents included in the 'Low' and 'High' IH group are presented in Appendix F (Table F.14). In the 'High IH' group, 5 adolescents had a clinical diagnosis of ADHD. No adolescents had a clinical diagnosis of ADHD in the 'Low IH' group.

An a priori power analyses based on the study by Chronaki et al (2015) revealed a power of 0.999 with an effect size of 0.59, when testing differences in early ERP components (i.e, N1) to vocal stimuli in 25 clinically diagnosed children with ADHD compared to 25 healthy controls. However, since the present analysis included a smaller number of participants in each group (e.g., 13), a further post hoc power analysis was conducted using Gpower, which revealed a power of 1 with an effect size of 1.85 (α error probability of .05). when exploring effects of ADHD group and emotion on the Cue-P2 amplitude.

Data Pre-Processing

EEG Data Treatment. EEG data were processed with the same techniques as in Analysis One and Two. The mean (SD) number of artefact-free epochs included in the cue analysis for each group x emotion condition are as follows: HIGH TRAITS (Angry: 85.83 (8.37), Happy: 85.75 (6.64), Neutral: 87.25 (7.58)); LOW TRAITS (Angry: 88.54 (6.71), Happy: 88.38 (7.46), Neutral: 88.15 (7.16)). There was no significant difference in the number of artefact-free trials between groups for the cue (p=.468).

Kolmogrov-Smirnov tests indicated that the Cue-P2 data were normally distributed for all emotion conditions in the Low ADHD traits group but were not normally distributed for the Cue-P2 to neutral (D=0.26, p=.021) and angry (D=0.33, p<.001) voices in the High ADHD Traits group. Spearman's rho correlations revealed no significant correlations between behavioural characteristics (hearing threshold, IQ, age, hyperacuity) and the Cue-P2 (ps>.05). There was a significant correlation between conduct problems and the Cue-P2 to happy (r=.58, p=.002) and angry (r=.56, p=.004) voices. Therefore, conduct problems were included as a covariate in analyses.

Results

Cue-P2. Although Kolmogorov Smirnov tests indicated that Cue-P2 data was not normally distributed for angry and neutral voices in the High traits group, since the majority of data was normally distributed parametric tests were conducted. When conducting non-parametric tests, the pattern of results did not change. A 3 (emotion) x 2 (group) ANOVA was conducted on the Cue-P2 amplitude. Consistent with our predictions, there was a main effect of group on the P2 amplitude (F (1,23) = 11.13, p=.003). Adolescents with high traits of ADHD presented with significantly larger P2 amplitudes compared to adolescents with low traits of ADHD. In addition, there was a significant emotion by group interaction effect on the P2 amplitudes (F(2,46) = 4.23, p=.021). Adolescents with high traits of ADHD presented with significantly larger Cue-P2 amplitudes to angry (MD=1.98, p=.006) and happy (MD= 2.56, p<.001) voices compared to adolescents with low traits of ADHD. The same difference was not found for neutral voices (MD=1.13, p=.089). This supports the hypothesis that adolescents with high levels of ADHD present with a preattentive sensitivity to emotional (angry and happy) vocal stimuli compared to adolescents with low levels of ADHD. Figure 5.5 presents the grand averages of the Cue-P2 per group. Table F.7 in Appendix F presents the mean Cue-P2 amplitudes per emotion and ADHD group. When conduct problems was added as a covariate in this analysis, the main effect of group on the Cue-P2 amplitude remained significant ($F(1,22) = 7.24$, $p=.013$), but the interaction between emotion and group did not remain significant ($F(2,44) = 2.65$, $p=.082$).

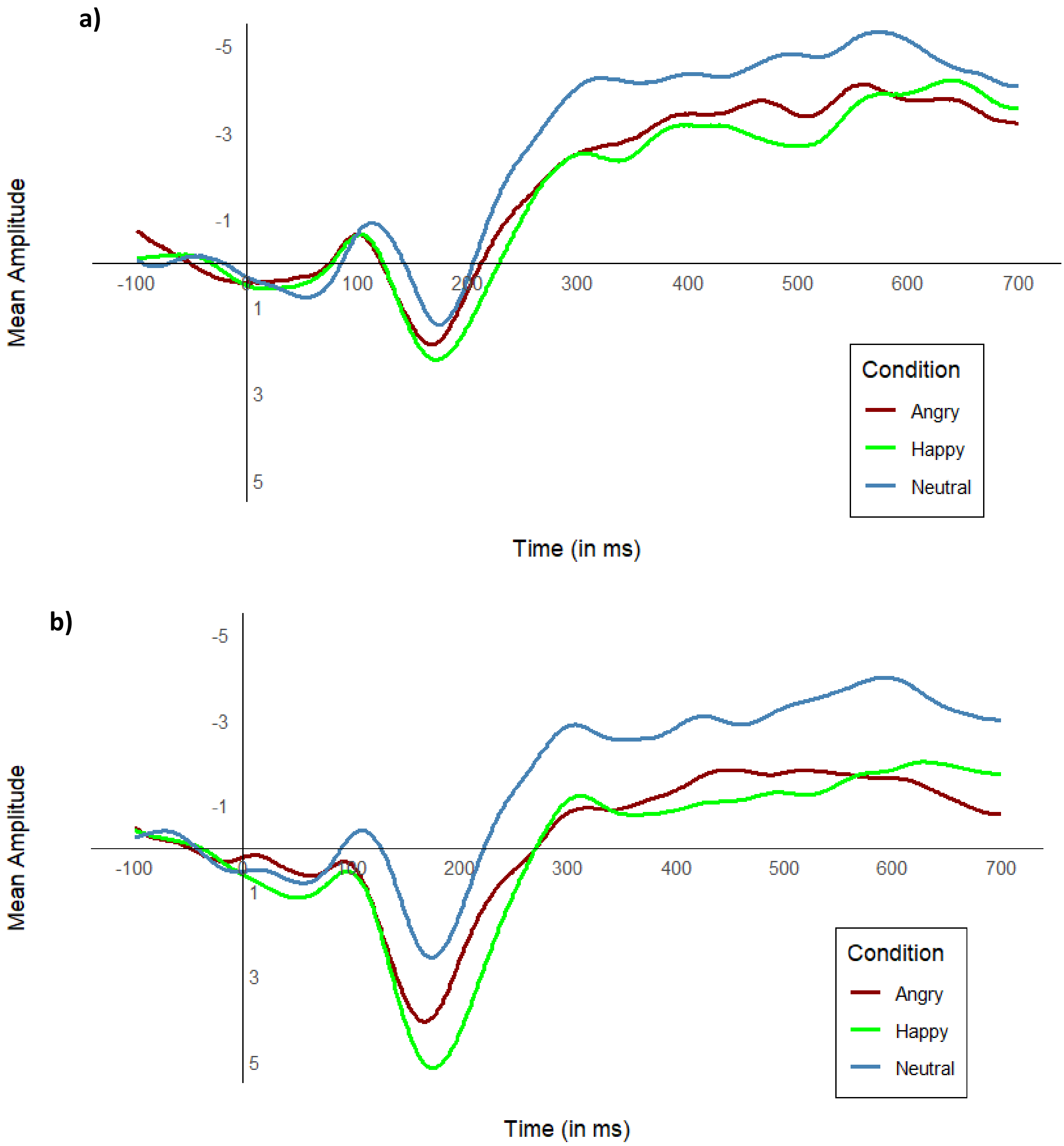

Figure 5.5

Grand Averages Showing the Frontocentral P2 Component (130-250ms) to the Angry, Happy and Neutral Vocal Cue in Adolescents with a) Low (n=13) and b) High (n=12) Traits of Inattention and Hyperactivity. Scale is -5 to + 5 µV.

5.8 Discussion

The present study employed an ESC paradigm to disentangle atypicalities in preattentive and attentional processing in adolescents with inattention and hyperactivity. First, the present study provided evidence for a distracting effect of vocal anger on performance (RTs) in the ESC paradigm in adolescents, providing support for a threat-specific involuntary capture of attention on a behavioural level. Second, the present study isolated the Cue-P2 component as a neural marker of preattentive vocal emotion processing in adolescents. Third, the present study provided further evidence for a threat-specific involuntary capture of attention by vocal anger in adolescents on a neural level, as reflected by larger Target-P1 and P3 amplitudes in conditions of angry compared to neutral voices. Fourth, this study provided novel evidence for the presence of a preattentive hypersensitivity to vocal emotion (happy and angry) in adolescents with high levels of inattention and hyperactivity. This finding was confirmed by categorical analyses. Finally, overall, there were no robust associations between inattention/hyperactivity and attention capture by irrelevant emotional voices (angry and happy) in adolescents, as reflected by RTs, the Target-P1 and P3 amplitudes. Each of these findings will be discussed in turn.

First, behaviourally, the present study reported congruency and emotion effects on reaction times within the ESC paradigm in adolescents. Specifically, adolescents were faster to respond to targets preceded by spatially congruent compared to incongruent vocal cues. Additionally, adolescents were slower to respond to targets preceded by angry compared to neutral voices. These findings are consistent with previous findings in adults employing the ESC paradigm (Chapter Four). The congruency effect on RTs is thought to reflect the involuntary capture of visual spatial attention by irrelevant vocal cues and is consistent with findings employing other spatial cueing paradigms with auditory stimuli in adults (Spence & Driver, 1997; Wang et al., 2022). The additive effect of slower reaction times to targets preceded by vocal anger is consistent with studies using dichotic listening and auditory distraction paradigms (Sander et al., 2005; Kattner & Ellermeier, 2018), and is thought to reflect the involuntary capture of attention by vocal anger followed by later difficulties to disengage from the irrelevant vocal stimulus and engage on the relevant target stimulus (Koster et al., 2004; Van Rooijen et al., 2017). Therefore, within the ESC paradigm, slower RTs to targets preceded by vocal anger is thought to reflect a threat-specific involuntary capture of attention by vocal emotion and downstream distracting effects on performance (i.e., RTs). This is consistent with theories which suggest negative stimuli capture attention due to adaptive advantages for survival (Hansen & Hansen, 1988; Ohman & Manika, 2001). Importantly, findings are consistent with the adult literature. This validates RTs as a behavioural measure of the involuntary capture of attention by vocal emotion in adolescents in the ESC paradigm.

Second, a clear Cue-P2 component was observed at Frontocentral sites in adolescents consistent with the literature (Mahajan & McArthur, 2012). In addition, the Cue-P2 was modulated by the emotional significance of the vocal stimuli, with larger amplitudes to emotional (angry and happy) compared to neutral voices in adolescents. However, it should be noted that this effect did not hold when adding age as a covariate in the analysis. This is surprising, however since the emotion effect was still close to significance when controlling for age, and it was consistent with adults, using the same ESC paradigm (Chapter Four), and other preattentive paradigms (Gadeke et al., 2014), it is reasonable to assume that the effect of emotion in adolescence is a true effect. Therefore, it is likely that the effect of emotion on the Cue-P2 component in adolescents reflects a preattentive stage of vocal emotion processing, prior to the capture of attention. This is a novel finding which confirms the Cue-P2 component within the ESC paradigm as a neural marker of preattentive vocal emotion processing in adolescents. However, in contrast to what was found in adults in Chapter Four, adolescents presented with a larger Cue-P2 amplitude to happy compared to angry voices. A greater sensitivity to positive compared to negative stimuli in adolescence is not surprising as adolescence is a period of greater sensitivity towards rewarding stimuli (Galvan, 2013). For example, fMRI studies have reported that adolescents present with significantly more ventral striatum activity when they earned money in monetary-reward tasks compared to adults (Ernst et al., 2005; Galvan et al., 2006). Moreover, previous VER research has reported that adolescents display higher accuracy when recognising happy voices in a recognition task (Grosbras et al., 2018), which differs from children and adults, who consistently recognise negative emotions with higher accuracy (Pell et al., 2009; Chronaki et al., 2018; Zupan, 2015). However, it is important to note that age (adults vs adolescents) by emotion interaction effects on the Cue-P2 component were not confirmed in additional comparative analysis (Appendix G). Therefore, it is not known whether a greater preattentive sensitivity to happy compared to angry voices in the present sample is due to age effects. Taking into consideration both the role of age as a covariate in the Cue-P2 analysis, and a potential adolescence-specific sensitivity towards happy voices, further developmental research is required to explore the effect of age on the preattentive processing of vocal emotion.

Third, the results indicated that preattentively processed vocal anger affected the early visual processing of the target, as reflected by the P1 component. Specifically, adolescents exhibited a larger P1 amplitude to targets preceded by angry compared to neutral voices. The effect of negative vocal emotion on the P1 component has previously been evidenced in ESC paradigms in adults (Brosch et al., 2009; Zimmer et al., 2022). For example, Brosch et al (2009) reported a larger P1 response to targets preceded by spatially congruent vocal anger compared to spatially congruent neutral voices in a dot-probe paradigm in adults. Similarly, Zimmer et al (2022) reported a larger P1 response to targets preceded by spatially congruent bi-modal (faces and voices) fearful expressions in adults in an ESC paradigm. In these cases, a larger P1 response to targets preceded by spatially congruent emotional expressions was thought to indicate the rapid involuntary capture of VSA by the irrelevant emotion. In the present study, a similar congruency by emotion effect on the P1 was not found. Therefore, it cannot be concluded with certainty that the P1 in the present study specifically reflects the capture of VSA by vocal emotion, as reported in previous studies (Brosch et al., 2009; Zimmer et al., 2022). However, the robust effect of vocal anger on the P1 amplitude suggests that preattentively processed angry voices must have had an effect (in some form) on the early visual processing of the target.

Regarding the lack of congruency effects on the P1 amplitude in adolescents, as discussed in Chapter Four, it is possible that this may be due to a lack of power. Despite the increased power with a larger sample size (60) in the present study, there was no evidence of congruency effects on the P1. It could be argued that increasing the number of trials in the ESC paradigm may have further increased the power to achieve congruency effects on the P1. However, this was not suitable in the present study due to recruiting a large number of participants with inattention/hyperactivity. In addition, in Chapter Four, it was suggested that the involuntary capture of VSA by spatially congruent vocal cues might occur earlier in the ESC paradigm, and thus might be reflected in an earlier component. Within the present study, this possibility was explored by analysing the effects of emotion and congruency on an earlier C1-like component, thought to reflect the earliest activations in the visual cortex (Foxe & Simpson, 2002; Rauss et al., 2011). Although similar anger-specific effects on the C1 component were demonstrated, there was no evidence of congruency or interaction effects on the C1 component. This provides further support that although early visual components to the target (e.g., C1, P1) might not specifically reflect the capture of VSA within the ESC paradigm, they do reflect an effect of emotion on the early visual processing of the target, which occurred as early as 50ms after the target onset (e.g., C1). Moreover, when controlling for the C1 component in the P1 analyses, the effects of vocal anger on the P1 component held. This strengthens the reliability of the finding relating to an anger-specific effect on the early visual processing of the target (i.e., P1) in the present study. Therefore, overall, due to the robust findings of emotion effects on early visual components to the target, it is reasonable to suggest the P1 reflects the involuntary capture of early visual attention by vocal cues, which is enhanced by vocal anger. In addition, it is possible that within the ESC paradigm, a congruency effect on early visual attention is a true effect, but there was not sufficient power within the present study to capture potential congruency effects, as there was to capture emotion effects, on the P1 component. If this is the case, the lack of emotion by congruency effects on the P1 possibly suggest that the effect of emotion on visual attention is independent from the effect of spatial congruency on visual attention. However, if the lack of a congruency effect on the P1 within the present ESC paradigm is a true effect, this may suggest that effects of emotion override spatial congruency effects on visual attention. If this is the case, this suggests that emotion and spatial congruency do interact during early visual processing. In either case, it appears emotion is prioritised over spatial congruency when competing for visual attention.

Although congruency effects on the P1 component were not observed in the ESC paradigm, downstream congruency effects were observed on the Target-P3 component. Adolescents presented with larger P3 amplitudes to targets preceded by incongruent compared to congruent vocal cues, consistent with previous findings in adults (Chapter Four; Wang et al., 2022). In addition, not only did irrelevant angry voices influence the early visual processing of the target, but also influenced the later attentional processing of the target, as reflected by larger P3 amplitudes for targets following angry compared to neutral voices. This is consistent with previous ESC studies in adults which also reported a larger P3 component to targets preceded by emotional (e.g., positive) compared to neutral emotional cueing stimuli (e.g., environmental sounds; Wang et al., 2022). A larger P3 suggests greater use of cognitive resources to attend to the visual target, due to the involuntary capture of VSA by the emotional cueing stimulus (Wang et al., 2022). Therefore, within the present study a larger P3 to targets preceded by angry vocal cues supports the hypothesis of a threat-specific involuntary capture of attention by irrelevant angry voices. Taken together, larger P1 and P3 amplitude to targets preceded by angry voices are consistent with theories which suggest threat-related stimuli rapidly capture attention due to adaptive advantages (Hansen & Hansen, 1988; Ohman & Minika, 2001). Moreover, evidence of emotion effects on components reflecting attentive processing, is in accordance with socio-cognitive models which suggest perceptual processes, such emotional information may influence attentional processes (Lemerise & Arsenio, 2000).

Interestingly, this is the first study to demonstrate a preattentive hypersensitivity to vocal emotion in adolescents with high traits of ADHD. First, the present findings demonstrate a positive relationship between the Cue-P2 to happy voices and inattention/ hyperactivity.

This suggests adolescents with higher levels of inattention/hyperactivity perceive happy voices as more salient. In addition, the difference in Cue-P2 amplitude to happy compared to neutral voices significantly correlated with inattention and hyperactivity. This corroborates the finding that there is a greater difference in the perception of salience for happy compared to neutral stimuli in adolescents with higher levels of inattention/hyperactivity. It was, however, surprising that there was no evidence of an association between the Cue-P2 to angry voices and inattention/hyperactivity. Taken together, these results suggest there is a reward-specific hypersensitivity to vocal emotional stimuli in adolescents with high levels of inattention/hyperactivity. This contrasts with previous reports of a hyper-orientation to threat-related vocal stimuli in 6-11-year-old children with ADHD (Chronaki et al., 2015). It is possible that the present findings of a preattentive hypersensitivity to specifically happy voices in ADHD are specific to the period of adolescence, in which rewarding stimuli are more important for critical socio-emotional development during this period (Galvan et al., 2013; Blakemore & Mills, 2014). This is also consistent with a study which employed a Monetary Incentive Delay task with adolescents with ADHD and controls, and found that ADHD was associated with enhanced attentional ERP components (i.e., the CNV) to positive reinforcement, which similarly suggests an attentional hypersensitivity to rewards in ADHD (Chronaki et al., 2017). A second possibility is that anger-specific effects were not evident in the present study due to methodological differences with previous research. For example, this study recruited a community sample of adolescents and measured ADHD on a dimensional scale, while previous research recruited a clinical sample of children with ADHD (Chronaki et al., 2015). It is possible that threat-specific atypicalities are present individuals with a clinical diagnosis of ADHD, who also present with higher levels of depression (Barkley & Fischer, 2010; Bunford et al., 2015). In our study there were no strong associations between emotional problems (e.g., depressive symptoms, emotion dysregulation) and ADHD traits, in contrast with prior research (Brocki et al., 2019; Sjowall et al., 2017). Moreover, the study by Chronaki et al. (2015) employed an explicit emotion recognition paradigm, and thus atypicalities in attentional mechanisms may have resulted in anger-specific 'hyper-orientation' effects and thus, may not reliably reflect preattentive hypersensitivity effects.

To overcome the methodological limitations relating to the dimensional approach to ADHD within the present study, categorical analyses confirmed the pattern of findings from the correlational analyses. Specifically, the Cue-P2 amplitude to both happy and angry voices was enhanced in adolescents with high traits of ADHD compared to low traits of ADHD. First, this finding extends research which has found a larger P2 amplitude to irrelevant non-emotional auditory stimuli (50ms tones) in adolescents with ADHD compared to controls (Lazzaro et al., 2001). The present study supports the conclusion that a larger P2 amplitude is reflective of a hypersensitivity to auditory stimuli in ADHD (Lazzaro et al., 2001; Barry et al., 2009) and extends research by showing that this hypersensitivity is enhanced by the presence of emotionally significant (e.g., angry and happy) stimuli in ADHD. Second, the finding that the Cue-P2 amplitude was enhanced for angry voices in adolescents with extreme levels of inattention/hyperactivity is consistent with previous neuro-imaging research which showed enhanced amygdala activation when subliminally processing fearful faces in adolescents with ADHD compared to healthy controls (Posner et al., 2012). As the P2 reflects the preattentive processing of the emotional significance of the voice in the early perceptual pathways (i.e., via the STS and amygdala; Schirmer & Kotz, 2006), the present findings add to this body of literature by providing support for an emotion-specific preattentive sensitivity to vocal cues (both happy and angry) in adolescents.

Taken together findings from both the correlational and categorical analysis on the Cue-P2 support the hypothesis that individuals with ADHD present with a preattentive hypersensitivity to vocal emotion (Chronaki et al., 2015). Moreover, these findings are consistent with socio-cognitive (Lemerise & Arsenio, 2000) and motivational dysfunction theories of ADHD (Sonuga-Barke, 2005; 2012). These theories suggest that atypicalities in the lower-level perceptual processing pathways contribute to emotion processing difficulties in ADHD, independently from the influence of attentional mechanisms. It is, however, important to note, that the association with the Cue-P2 to happy voices was also significant with conduct problems in preliminary analyses, and when conduct problems were controlled for in the Cue-P2 analyses, the association with inattention/hyperactivity was no longer significant. This suggests that a preattentive hypersensitivity to happy voices might not be specific to inattention and hyperactivity symptoms but be an underlying neurocognitive atypicality of other externalising problems in adolescents, such as conduct problems. This is not surprising as conduct problems is a common comorbidity with clinical ADHD (Gustafsson et al., 2018; Taylor et al., 1996), and within the present study inattention and hyperactivity and conduct problems were highly correlated.

Next, conduct problems appeared to play a significant role when investigating associations between inattention/hyperactivity and the P1 amplitude to the visual target. Without controlling for conduct problems, no significant correlations were found, but controlling for them revealed that adolescents with higher levels of inattention/hyperactivity presented with larger P1 amplitudes to targets preceded by spatially congruent neutral voices. This suggests the higher the levels of hyperactivity and inattention, the higher the early visual processing of the target. Previous research has demonstrated an atypical P1 response to visual targets preceded by neutral cues in children with ADHD (Perchet et al., 2007; Nazari et al., 2010). For example, children (aged 6-11 years) with ADHD displayed smaller P1 amplitudes to targets preceded by visual cues in a spatial cueing paradigm, reflecting atypical attention allocation to a relevant visual target (Perchet et al., 2007). In contrast, the present findings revealed larger P1 amplitude to visual targets preceded by neutral vocal cues, possibly due to differences in cueing modality (visual vs auditory). Since auditory stimuli capture VSA more automatically and effectively than visual stimuli in adults (Spence & Driver, 1997), it is possible that visual and auditory cueing stimuli result in differential attentional atypicalities in ADHD. Further research is required to substantiate this claim.

In addition, there was a negative correlation between inattention/hyperactivity and the Happy-Neutral P1 difference score. This suggests that the higher the inattention/hyperactivity, the smaller the difference in the P1 amplitude to targets preceded by spatially congruent happy compared to neutral voices. Combined with the result demonstrating an enhanced P1 amplitude following neutral voices, this suggest that adolescents with higher levels of inattention/hyperactivity present with enhanced early visual target processing, but only when targets are preceded by neutral voices. When irrelevant happy or angry voices are present, this enhancement is not present. In contrast to our predictions, and previous findings (Chronaki et al., 2015), this may reflect an enhanced involuntary orientation to neutral, rather than emotional, voices in adolescents with higher levels of inattention/hyperactivity, when the voices are congruent to task demands. If this is the case, this finding has clinical implications. For example, it suggests that the use of neutral tone by clinicians, parents and educators may be beneficial to help adolescents with inattention/hyperactivity focus on a task. Additionally, contrasting results with previous studies (e.g., Chronaki et al., 2015) likely reflect differences in task design and the ERP components. For instance, Chronaki et al (2015) demonstrated a larger N1 amplitude to angry voices in children with ADHD when completing an explicit vocal emotion recognition task, hypothesised to reflect a preattentive hypersensitivity and hyper-orientation to vocal anger in children with ADHD (Chronaki et al., 2015). Moreover, a study which employed a go/no-go task with emotional (happy and angry) faces as Go stimuli and neutral faces as No-Go stimuli demonstrated that adolescents with ADHD displayed enhanced P1 amplitudes to the emotional (positive and negative) compared to the neutral facial stimuli (Karalunas et al., 2020). However, as emotional faces in the go/no task were relevant to task demands, the P1 in this context likely reflected both preattentive and attentive facial emotion processing, similar to the N1 component to explicitly processed vocal stimuli in the study by Chronaki et al. (2015). Within the present study, differential associations between inattention/hyperactivity and the Cue-P2 and Target-P1 suggest that although adolescents with inattention/hyperactivity present with a preattentive hypersensitivity to vocal emotion, this does not appear to lead to downstream enhancement effects of early visual processing to a relevant target. This suggests that irrelevant emotional voices do not enhance the capture of visual spatial attention in adolescents with inattentive/hyperactive traits. These findings come in contrast to previous findings from explicit emotion recognition tasks (Chronaki et al., 2015; Karalunas et al., 2020). These findings need to be considered in relation to the task design in the present study which aimed to disentangle components related to the preattentive processing of vocal emotion (i.e., the Cue-P2) from visual attentional capture (i.e., the Target-P1).

The present study did not report any robust associations between RTs, and the Target-P3 and inattention/hyperactivity. Although there was some evidence of a negative association between the Happy-Neutral Target-P3 difference score and inattention/hyperactivity, when controlling for conduct problems, this association did not survive correction for multiple comparisons. Moreover, the direction of this effect was consistent with findings on the earlier P1 component, and similarly suggests no association between the modulation of the P3 by irrelevant emotional voices and inattention/hyperactivity. Since a larger P3 reflects the greater use of cognitive resources to attend to the visual target (Wang et al., 2022), this finding suggests that adolescents with higher levels of inattention/hyperactivity did not display atypicalities in later attentional processing to the target preceded by irrelevant emotional voices. This contrasts with previous studies employing go/no-go tasks in children with ADHD, who presented with larger P3 amplitude in no-go trials when negative and positive emotional backgrounds were present (Lopez-Martin et al., 2015), as well as studies which reported enhanced P3 amplitudes in adolescents with ADHD during conditions of negative and positive reinforcement in a monetary incentive task (Chronaki et al., 2017). However, the present findings are consistent with a study employing an oddball task in adults, which reported no differences in the Target-P3 to auditory neutral stimuli in ADHD, despite the presence of early sensory atypicalities (i.e., via the P2) to the auditory stimuli (Barry et al., 2009). The authors suggested this reflects early sensory atypicalities being overcome through effortful processing in ADHD (Barry et al., 2009). It is possible that a similar interpretation applies to the present study in that atypicalities in preattentive vocal emotion processing (i.e., via the Cue-P2) do not result in later attention allocation atypicalities downstream in adolescents with higher levels of inattention/hyperactivity. Alternatively, the lack of association between inattention/hyperactivity and later attentional mechanisms (i.e., Target-

P3 and RTs) could be due to the simplicity of the ESC task compared to other more complex tasks, such as go/no-go tasks, which required participants to inhibit responses, and thus required higher levels of attentional control.

It is interesting to note that some preliminary findings demonstrated emotion regulation and conduct problems were associated with the P1 and P3 amplitude to emotional compared to neutral voices. These results suggest that atypicalities in the downstream effects of irrelevant vocal emotions on attentional mechanisms may be specifically associated with emotion dysregulation and conduct problems, both of which are common in individuals with high levels of inattention/hyperactivity (Shaw et al., 2014; Faraone et al., 2019; Gustafsson et al., 2018). However, as the present study did not aim to disentangle the differential roles of conduct problems, emotion dysregulation and inattention/hyperactivity symptoms in preattentive and attentive vocal emotion processing, further research is required to substantiate this hypothesis.

Considering the findings from the present study, future research should employ the ESC paradigm to further investigate differences in preattentive and attentive vocal emotion processing in children and adolescents. Specifically, given potential developmental differences, future research should explore emotion-specific differences in the preattentive processing of vocal emotions in children, adolescents and adults with ADHD. This is important because threat-specific atypicalities observed in children with ADHD (Chronaki et al., 2015), were not observed in the present adolescent sample, possibly suggesting a shift towards reward-specific atypicalities in older individuals with inattention/hyperactivity. Second, future research should more closely explore the link between emotion dysregulation in ADHD and a threat-specific preattentive hypersensitivity or hyper-orientation to vocal emotions.

This is crucial as ED is strongly associated with higher rates of a clinical diagnosis of ADHD (Shaw et al., 2014), and greater social problems (McKay et al., 2023). Understanding whether specific neurocognitive atypicalities, within preattentive or attentive emotion processing, differentiate severe presentations of ADHD from milder symptoms of ADHD would support research showing that ED is an important clinical feature of ADHD, and inform clinical interventions. Finally, considering the role of conduct problems in the emotion processing atypicalities in ADHD, further research is required to explore preattentive and attentive vocal emotion processing in adolescents with a range of externalising behaviours (i.e., conduct problems and hyperactivity). Future studies should recruit a large sample size of adolescents with a range of externalising behaviour problems to disentangle the role of inattention/hyperactivity from conduct problems.

In summary, the present study employed the ESC paradigm to disentangle preattentive from attentive processing in adolescents with traits of inattention and hyperactivity. First, the findings support the use of the ESC to measure the preattentive processing of vocal emotions (reflected by the Cue-P2), and downstream effects on taret processing in adolescents (reflected by the Target-P1, and P3). Specifically, angry voices appear to have a greater effect on the involuntary capture of attention in adolescents, compared to neutral voices. This study provided novel evidence for a preattentive hypersensitivity to vocal emotions (happy and anger) in adolescents with high levels of inattention and hyperactivity, in support of motivational dysfunction theories (Sonuga-Barke, 2012) and socio-cognitive models (Lemerise & Arsenio, 2000). However, target-related (attentive) mechanisms do not appear to be affected by the presence of irrelevant emotional voices (i.e., happy and angry) in adolescents with higher compared to lower levels of inattention/hyperactivity. On the other hand, adolescents with higher levels of inattention/hyperactivity appear to present with attentional atypicalities to the target preceded by neutral voices. This additionally supports the role of emotion in contributing to variations in the allocation of attention in support of perceptual processing theories (Sonuga-Barke, 2005; Lemerise & Arsenio, 2000). Finally, there is some preliminary evidence that atypicalities relating to the effect of irrelevant emotional voices on target-related (attentive) mechanisms might be specific to adolescents who present difficulties, such as emotion dysregulation and conduct problems. Further research is required to disentangle the role of conduct problems, emotion dysregulation and inattention/hyperactivity in atypical preattentive and attentive emotion processing in children and adolescents with a clinical diagnosis of ADHD.

6 GENERAL DISCUSSION

6.1 Chapter Overview

First, this chapter presents a summary of the motivation for conducting the thesis before summarising the key findings of the thesis. Second, the theoretical and clinical implications of the main findings are discussed. Third, the limitations of this work are described and ideas for future research are suggested. Finally, the key contributions of the thesis are summarised in a concluding paragraph.

6.2 Overview of Thesis Findings

It is debated whether emotion processing difficulties in ADHD stem from general executive function atypicalities, consistent with the cognitive-behavioural theory of ADHD (Barkley, 1997, 2015), or whether they additionally reflect atypicalities in the neuro-cognitive pathways, which are implicated in emotion processing, consistent with motivational-dysfunction models of ADHD (Sonuga-Barke, 2005, 2012). To contribute to this debate, previous ERP research employing explicit emotion processing tasks (e.g., VER and MID tasks) in ADHD have revealed enhancements in attention allocation in emotionally salient (both positive and negative) contexts (Chronaki et al., 2017; Chronaki et al., 2015). It was hypothesised that a preattentive hypersensitivity to vocal emotions may contribute to these emotion effects on attentional mechanisms in ADHD, consistent with motivational dysfunction theories (Sonuga-Barke, 2012). In support, Chronaki et al. (2015) revealed a larger N1 component, reflecting the early sensory processing of vocal anger stimuli in children with ADHD compared to controls (Chronaki et al., 2015). However, due to employing an explicit VER task, when the vocal emotion was relevant to task demands, it was not clear whether preattentive or attentive processing contributed to the effect of anger on the N1 component in that study (Chronaki et al., 2015). To disentangle preattentive from attentive mechanisms underlying vocal emotion processing, recent research has begun to explore the effect of preattentively processed (i.e., irrelevant) vocal emotions on attention in ADHD (Chronaki & Marsh, 2024). For example, Chronaki & Marsh (2024) reported a distracting effect of irrelevant vocal anger on performance (RTs) in children with ADHD. This supports the hypothesis that a hyper-orientation to threat-related stimuli stems from atypical preattentive processing of negative vocal emotions (Chronaki & Marsh, 2024). However, no studies so far have confirmed the presence of specific preattentive vocal emotion processing atypicalities prior to attentional capture using temporally sensitive ERP techniques.

To address this gap in the literature, the present thesis developed a novel ESC paradigm to disentangle preattentive and attentive vocal emotion processing in ADHD employing behavioural and ERP techniques. The present thesis aimed to investigate a) if adolescents with high levels of inattentive and hyperactive traits present with a preattentive hypersensitivity to vocal emotion and b) if irrelevant vocal emotion influences attention allocation in adolescents with high levels of inattentive and hyperactive traits. To meet these aims, this thesis conducted four studies which revealed several key findings. Each of these will be summarised in turn before discussing the theoretical and clinical implications of the main findings.

Chapter Two consisted of a meta-analysis which thoroughly reviewed the literature of vocal emotion recognition in ADHD. This was the first study to synthesise effect sizes for vocal emotion recognition difficulties in ADHD. VER deficits in ADHD were found in 95% of studies employing explicit VER tasks. This confirmed that the presence of VER deficits in ADHD was a reliable and consistent finding, which was in line with that previously reported in the facial emotion recognition literature (Bora & Pantelis, 2016; Borhani & Nagati, 2018). In addition, moderator analysis showed that VER deficits were not modulated by emotion. This finding demonstrated that there is insufficient data in the behavioural VER literature to conclusively identify the presence of atypicalities in the emotion processing pathways in ADHD. Therefore, Chapter Two confirmed that explicit VER tasks are limited in their ability to isolate emotion-specific atypicalities in ADHD due to the role of attentional processing necessary to identify task-relevant vocal emotions. Overall, findings from this meta-analysis confirmed the gap in the literature to explore 'implicit' vocal emotion processing, independent from attentional capture, by employing preattentive, rather than explicit emotion recognition paradigms.

Chapter Three aimed to validate a set of vocal emotional stimuli, originally developed by Maurage et al. (2007) in adults and adolescents. These stimuli were incorporated into a novel 'preattentive' ESC paradigm in later studies. This study was novel because these stimuli had not previously been validated in an adolescent population. Importantly, the study reported high accuracy rates (above 80%) in both adults and adolescents. Since reciprocal understanding of vocal emotions is essential in the use of vocal emotions for effective communication (Jusline & Laukka, 2004), this confirmed the validity of the stimuli for use in the later ESC paradigm. In addition, Chapter Three demonstrated similar age and emotion effects on VER accuracy as reported in previous literature. For example, adolescents and adolescents not only recognised angry voices more accurately compared to happy and neutral voices, but also perceived angry voices as more intense than happy voices, which was consistent with previous VER studies (Chronaki et al., 2018; Liu & Pell, 2012; Morningstar et al., 2019). This supported the advantage of employing angry vocal stimuli in explicit VER tasks

(Hansen & Hansen, 1988; Ohman & Manika, 2001), perhaps due to these stimuli being preattentively perceived as more salient, or due to later attentional mechanisms prioritising vocal anger when it is relevant to task demands.

Chapter Four presented the first study employing a novel ESC paradigm combining behavioural and ERP techniques in adults, to more conclusively explore the time course of preattentive and attentive vocal emotion processing. Although previous ERP studies in adults have isolated the P2 component as a neural marker of preattentive vocal emotion processing (Gadeke et al., 2013; Lin et al., 2022), Chapter 4 presented the first ERP study to disentangle the neural markers of both preattentive and attentive vocal emotion processing. First, Chapter Four similarly reported a larger Cue-P2 component to emotional (angry and happy) compared to neutral voices in adults, validating the Cue-P2 as a neural marker of preattentive vocal emotion processing in the ESC paradigm. Second, Chapter Four demonstrated that behaviourally, participants responded slower to targets preceded by irrelevant angry compared to neutral voices. This is consistent with findings from studies employing dichotic listening (Peschard et al., 2017; Sander et al., 2005), auditory distraction tasks (Kattner & Ellermeier., 2018) and cross-modal oddball tasks (Chronaki & Marsch, 2024), which have demonstrated a distracting effect of irrelevant vocal anger on task performance in adults. In addition, an anger-specific effect on RTs is consistent with the anger-specific effect on recognition accuracy reported in the validation study (Chapter Three). Finally, there was a lack of evidence relating to the downstream effects of emotion on ERP components reflecting early visual target processing (i.e., the P1) and later attentional target processing (i.e., the P3), both of which were hypothesised to reflect the capture of attention by the vocal cues. However, because robust emotion and congruency effects were demonstrated on RTs, it was concluded that the ESC paradigm was a valid behavioural measure of the involuntary capture of attention by vocal emotions, and the lack of corresponding neural correlates likely reflected a lack of power.

Finally in Chapter Five, a behavioural and ERP study employing the ESC paradigm was conducted to explore preattentive and attentive vocal emotion processing in 60 adolescents with varying levels of inattention/hyperactivity. This study generated a number of novel and interesting findings. First, consistent with our findings in adults (Chapter Four), adolescents responded slower to targets preceded by angry compared to neutral voices, further demonstrating that irrelevant vocal anger has a distracting effect on performance (RTs) in adolescents. Second, adolescents presented with a larger P2 amplitude to angry and happy voices compared to neutral voices, consistent with findings in adults (Chapter Four). This was a novel finding and suggests that the Cue-P2 is a neural marker of the preattentive emotion processing in adolescents, consistent with the adult literature (Gadeke et al., 2013). However, surprisingly, when controlling for the effect of age, the effect of emotion on the P2 amplitude was not significant. Moreover, although we expected, based on the validation study, that angry voices would be preattentively perceived more saliently than happy voices, the Cue-P2 component was larger to happy compared to angry voices in adolescents. This provided evidence of a reward-specific preattentive sensitivity in adolescents, not found in adults. Interestingly, novel findings from Chapter Five highlighted an anger-specific effect on the early visual processing, as reflected by the P1 component, of the target and a downstream anger-specific effect on later attentional mechanisms to the target, as reflected by the P3 component. This is consistent with the anger-specific effects on RTs found in adults and adolescents in the present thesis, and further suggests that angry voices involuntarily capture attention in adolescents.

Consistent with our predictions, correlational analysis in Chapter Five revealed that adolescents with higher levels of inattention/hyperactivity presented a larger Cue-P2 amplitude to happy voices. In addition, categorical analysis revealed that adolescents with high levels of inattention/hyperactivity presented with a larger Cue-P2 amplitude to both emotional (angry and happy) voices compared to adolescents with low levels of inattention/hyperactivity. Taken together, since the P2 component is suggested to reflect the preattentive processing of vocal emotion (Gadeke et al., 2013; Schirmer & Kotz, 2006), these results provide novel evidence in support of a preattentive hypersensitivity to vocal emotion in ADHD. It is interesting to note, however, that these effects were no longer significant when conduct problems were controlled for in analysis, possibly suggesting that conduct problems may mediate a preattentive hypersensitivity to vocal emotion in ADHD.

Finally, the associations between emotion effects and inattention/hyperactivity appeared to be specific to the vocal cue processing, as there were no significant associations between emotion effects on target-locked ERPs and RTs and inattention/hyperactivity. Specifically, there were no significant associations between inattention/hyperactivity and a) the effect of anger on RTs, b) the effect of anger on early visual processing of the target (as reflected by the P1) and c) the effect of anger on later attention allocation to the target (as reflected by the P3). Taken together, these results suggest that there are no atypicalities in how irrelevant vocal emotion (anger and happiness) influence attention allocation mechanisms in adolescents with higher levels of inattention/hyperactivity. However, interestingly, after controlling for conduct problems, there was some novel evidence that the presence of irrelevant neutral voices enhances early visual processing of a target (i.e., the P1) in adolescents with higher levels of inattention/hyperactivity.

Overall, Chapter Five presented the first study to a) identify a preattentive hypersensitivity to vocal emotion (happy and angry) in adolescents with high levels of inattention/hyperactivity and b) identify a robust ‘distracting’ effect of irrelevant vocal anger on attentional mechanisms in adolescents with varying (high and low) levels of inattention/hyperactivity. The theoretical and clinical implications of these findings are summarised in the following section.

6.3 Theoretical Implications of Main Findings

Findings Relating to the Preattentive Processing of Vocal Emotion (Cue P2)

First, the finding that both happy and angry voices resulted in a larger Cue-P2 amplitude compared to neutral voices in the ESC paradigm in both adults and adolescents suggests that emotional (angry and happy) voices are preattentively perceived as more salient than neutral voices. This supports the model of vocal emotion processing (Schirmer & Kotz, 2006), which argues that the emotional significance of voices is processed in the perceptual pathways of the brain (e.g., the STS) via subcortical regions, such as the amygdala, prior to the capture of attention. Second, the finding of a greater Cue-P2 amplitude for happy compared to angry voices in adolescents is supportive of theories which view adolescence as a period of greater sensitivity towards rewarding stimuli stemming from enhanced activation of the striatum, a subcortical region of the brain in the basal ganglia (Galván et al., 2013). This finding is consistent with fMRI research employing monetary reward tasks which has shown enhanced activation of the striatum in adolescents, during the processing of stimuli in response to reward-related outcomes (Ernst, 2014; Bogert et al., 2016). In addition, the finding of an enhanced Cue-P2 to happy voices suggests that reward-sensitivity in adolescence extends to voices, consistent with fMRI findings of enhanced striatum activity in the face processing literature in adolescents (Mueller et al., 2017). Moreover, it is interesting that the effect of emotion on the Cue-P2 amplitude was rendered non-significant when controlling for the effect of age. This suggests that effect of emotion on the Cue-P2 amplitude in adolescents is not stable across development (from 12 to 16 years). As the effect of age was not found in adults, this suggests that in adolescents the preattentive processing of vocal emotions is still developing, consistent with the developmental trajectory of explicit vocal emotion recognition in adolescents, which suggests vocal emotion processing continues to develop into late adolescence (Grosbras et al., 2019; Fillipa et al., 2022).

Second, the finding that high levels inattention/hyperactivity were significantly associated with a larger Cue-P2 amplitude to happy and angry voices provides robust support for the hypothesis of a preattentive hypersensitivity to vocal emotion in ADHD (Chronaki et al., 2015; Chronaki & Marsh, 2023). This finding is consistent with motivational dysfunction theories of ADHD suggesting that socio-emotional difficulties are characteristic of ADHD, and that these difficulties may stem from specific atypicalities in the regions of the brain involved in emotion and reward processing (e.g., amygdala, striatum; Sonuga-Barke, 2012). Previous ERP research has supported motivational dysfunction theories by demonstrating that relevant motivational stimuli (i.e., negative and positive reinforcement) versus neutral stimuli enhance attention allocation (i.e., larger P3 amplitudes) to a target in a monetary incentive task in 10-16-year-old adolescents with ADHD (Chronaki et al., 2017). Moreover, previous MRI research employing an Escape-Delay Incentive task has shown that 10-18 year-old adolescents with ADHD present with reduced volumes in the amygdala when processing stimuli in the 'certain delay' condition compared to the 'no delay' condition (Van Dessel et al., 2020). It was suggested that this reflects atypicalities in subcortical processing (e.g., amygdala) when processing the negative emotions associated with delay in ADHD, in support of delay-aversion theories of ADHD (Van Dessel et al., 2020). However, in both studies, since the motivational stimuli included were relevant to task demands, it was not known if atypicalities reflected emotion processing atypicalities independently from attentional difficulties. The present findings reporting a preattentive hypersensitivity to vocal emotions in ADHD add to this literature by providing the first ERP evidence of atypicalities in the preattentive processing pathways of vocal emotions.

It is interesting that the present thesis reported only a preattentive hypersensitivity to happy (and not angry) voices in adolescents with higher levels of inattention/hyperactivity in correlational analysis in Chapter Five. This contrasts with previous research which has shown anger-specific effects on the N1 amplitude to explicitly processed emotional voices in 6-11-year-old children with a clinical diagnosis of ADHD. One explanation for the presence of happy-specific effects during the preattentive processing of vocal emotion in the present study could be that this effect is specific to adolescents, and not children, with ADHD. This hypothesis is consistent with theories which suggest adolescents present with a greater sensitivity to rewards (Galván, 2013). Moreover, this finding suggests that reward sensitivity is heightened in adolescents with higher levels of inattention/hyperactivity. This is in accordance with motivational dysfunction theories such as the delay aversion hypothesis, which suggest individuals with ADHD are more likely to seek out immediate salient rewards in their environment, rather than wait for a delayed (positive or negative) outcome (Sonuga-Barke, 2005). For example, if adolescents with ADHD perceive the immediate positive stimuli (i.e., messages on phone) with higher saliency than later negative consequences (i.e., punishment for not doing homework), this might make it particularly difficult for adolescents with ADHD to ignore potentially distracting positive stimuli.

An alternative explanation for the presence of happy-specific effects compared to anger-specific effects reported in previous studies (Chronaki et al., 2015) could be due to the recruitment of a non-clinical sample in the present study. Thus, anger-specific effects might be associated with clinical presentations of ADHD. Within the present thesis, this argument was supported by the fact that in Analysis Three in Chapter Five, when a subsample of adolescents with extreme levels of inattention/hyperactivity (i.e., high SDQ scores) were included in categorical analyses, an additional anger-specific effect on the Cue-P2 emerged. That is, adolescents with high levels of inattention/hyperactivity presented with a larger Cue-P2 amplitude to both angry and happy voices compared to adolescents with low levels of inattention/hyperactivity. Further research is necessary to explore both developmental and clinical reasons for the presence of a reward-specific preattentive hypersensitivity to vocal emotions in adolescents with higher levels of inattention/hyperactivity (See section 6.6).

In addition, within the present study, effects relating to the preattentive hypersensitivity to vocal emotion (anger and happiness) were rendered non-significant when conduct problems were controlled for in both correlational and categorical analyses. The effect of conduct problems on emotion processing in ADHD is also consistent with a trend in the VER literature which suggests that comborbid CD affects the nature of vocal emotion recognition atypicalities in ADHD (Miller et al., 2011; Cadesky et al., 2000; Sells et al., 2023). The present findings have theoretical implications as they suggest that a preattentive hypersensitivity to vocal emotion might be an underlying neurocognitive atypicality which drives the presence of conduct problems in ADHD (or vice versa; conduct problems might drive preattentive emotion processing atypicalities). These findings raise questions as to whether preattentive emotion processing atypicalities are present in adolescents who present with ADHD only (without comorbidities, such as conduct problems). This possibility is in accordance with theories which suggest that the combination of ADHD and emotion difficulties, such as ED, is a risk factor for developing comorbidities in ADHD (Steinberg & Drabick, 2015; Seymour et al., 2014). If this is the case, it is possible that motivational dysfunction theories may not be specific to ADHD but extend to the neurocognitive underpinnings of other externalising disorders, such as conduct disorder. In addition, the view that neurocognitive atypicalities are not categorically specific to one disorder, but the same atypicality (i.e., a preattentive hypersensitivity) and associated symptoms (i.e., emotion difficulties) may present at varying levels in other disorders, is consistent with the dimensional view of ADHD, as a form of neuro-divergence (Sonuga-Barke et al., 2023; Astle et al., 2022).

Findings Relating to the Effect of Vocal Emotion on Target Processing.

In both samples of adults and adolescents, the present thesis reported an anger-specific distracting effect of irrelevant cues on performance (RTs) to a visual target in the ESC task. This is consistent with previous behavioural studies which report that irrelevant vocal anger has a distracting effect on performance in adults (Kattner & Ellermeier, 2018, Sander et al., 2005). These findings support theories which suggest that when vocal anger is irrelevant to task demands, the voice captures attention involuntarily and holds limited attentional resources later in processing, and thus has a distracting effect on task performance (Huges, 2014, Sander et al., 2005). Consistent with this theory, the present thesis additionally reported an anger-specific effect on a neural level by reporting enhanced P1 and P3 components to targets preceded by angry compared to neutral voices in the overall adolescent sample. The presence of an enhanced P1 to targets preceded by angry voices supports prior research which suggests angry voices involuntarily captures early visual attention (Brosch et al., 2009). The presence of an enhanced P3 to targets preceded by angry voices is consistent with previous research which has shown that the involuntary capture of visual attention by vocal emotion has downstream effects on later attentional processing (Li et al., 2018, Wang et al., 2022). Moreover, this finding supports theory which argues that irrelevant emotional voices hold limited attentional resources and thus, may cause difficulties to engage on a target (Huges, 2014). Overall, findings of anger-specific effects on target processing in adults and adolescents are consistent with research which has shown that angry voices are recognised with higher accuracy in adults (Zupan, 2015; Maurage et al., 2007), adolescents (Morningstar et al., 2019) and children (Chronaki et al., 2018). Findings that angry voices are recognised with higher accuracy in adults and adolescents are also confirmed in Chapter 3 of the present thesis. However, it is important to note that the present study only found effects of anger on the Target P1 and Target P3 in adolescents but not adults. It is possible that effects of emotion on the involuntary capture of vocal emotion are stronger in adolescents compared to adults on a neural level, perhaps because adolescence is associated with heightened sensitivities to emotional stimuli (Blakemore, 2008). However, further developmental research is necessary to confirm this (See section 6.6). Overall, anger-specific effects on target processing support theories which suggest the ability to recognise and attend to negative (i.e., threatening) compared to positive emotions is integral to human survival (Liu & Pell, 2005; Vuilleumier, 2005). Moreover, these findings suggest that the involuntary capture of visual attention by irrelevant angry vocal cues occurs typically in healthy adults and adolescents.

The present study reported preattentive emotion processing atypicalities but not downstream emotion effects on target processing in adolescents with higher levels of inattention/hyperactivity. The finding that inattention/hyperactivity was not associated with emotion-specific effects on target processing contrasts with previous research which reported that when motivationally salient stimuli (e.g., monetary rewards) were relevant to task demands, this enhanced attention allocation to a target in adolescents with ADHD (Chronaki et al., 2017). It is possible that the absence of downstream emotion effects on target processing in the present thesis is explained by the fact that vocal emotions were irrelevant to task demands, compared to previous research which included emotional stimuli which were relevant to task demands (Chronaki et al., 2015; Chronaki et al., 2017). If this is the case, this would suggest that a preattentive hypersensitivity to vocal emotion has a greater effect on the allocation of attention in ADHD when emotional stimuli are task relevant. This hypothesis is consistent with motivational dysfunction and social-cognitive theories which suggest that attention (i.e., related to task relevance) and emotion processing atypicalities interact to contribute to variations in the allocation of attention in ADHD (Sonuga-Barke, 2005, 2012; Lemerise & Arsenio, 2000).

Next, in Chapter Five, we reported a significant positive association between inattention/hyperactivity and the Target-P1 component to targets preceded by neutral voices. This suggests that irrelevant neutral voices have an enhancing effect on early visual processing of the target and are more likely to involuntarily capture attention in adolescents with higher levels of inattention/hyperactivity. Similarly, it possible that individuals with lower levels of inattention/hyperactivity find it easier to ignore irrelevant neutral voices. An enhanced P1 to targets following irrelevant neutral auditory stimuli contrasts to studies which find a smaller P1 to visual targets preceded by irrelevant neutral visual stimuli in children with ADHD (Perchet et al., 2007). These contrasting findings could be due to differences in how cueing modality (visual vs auditory) influences attention allocation in ADHD. A greater P1 response to targets preceded by auditory compared to visual cueing stimuli may suggest that irrelevant auditory stimuli involuntarily capture and distract individuals with ADHD more compared to visual cueing stimuli. This hypothesis is consistent with theories which suggest irrelevant auditory stimuli are more likely to capture attention and orientate attention to potentially threatening events outside of the visual field for adaptive purposes (Murphy et al., 2014). Moreover, this finding is consistent with research which has shown that individuals with higher levels of inattention/hyperactivity present with higher sensitivity (reflected by the P2 component) to auditory stimuli (Oades et al., 1997; Barry et al., 2009) and research which suggests there is a high prevalence of auditory hyperacuity ADHD (Rali et al., 2020; Paulin et al.,2016).

In addition, the smaller amplitude difference in the effect of emotional and neutral voices on early visual processing (i.e., the P1) in adolescents with higher levels of inattention/hyperactivity suggests neutral voices may capture attention similarly to emotional voices in ADHD. If this is the case, this may explain greater difficulties in recognising vocal emotions in ADHD (Sells et al., 2023). For instance, if emotional and neutral voices are perceived during attentional processing as more similar, this may lead to errors or biases in perceiving neutral voices as emotional or vice versa. This hypothesis is consistent with social-cognitive theories which argue that atypicalities in the encoding of emotional stimuli may lead to incorrectly recognising emotional stimuli, which in turn leads to incorrect responses in social situations (Lemerise & Arsenio, 2000). For example, if a neutral tone used by a friend captures attention to a similar degree as an angry tone, the receiver might be more likely to perceive the tone as anger and this may lead to a more hostile response. Similarly, if irrelevant neutral voices capture attention to a similar degree as emotional voices in individuals with traits of ADHD, this might explain why students with ADHD struggle to focus in classroom environments (e.g. when students and teachers are talking in a neutral tone in group settings).

Overall, findings of a larger Cue-P2 to irrelevant vocal emotion in adolescents with higher levels of inattention/hyperactivity provide evidence for atypicalities in the neuro-cognitive pathways which preattentively process emotions. This provides evidence in favour of motivational dysfunction theories to explain the presence of emotional difficulties in ADHD (Sonuga-Barke, 2012). Moreover, as there is some evidence that conduct problems play a role in the preattentive processing of vocal emotions. This finding supports a dimensional view of ADHD and suggests that motivational dysfunction theories might be specific to the neuro-cognitive atypicalities in a subgroup of individuals with comorbid ADHD and CD. Finally, it is unclear from the present findings how a preattentive hypersensitivity to vocal emotion interacts with visual processing to contribute to variations in the allocation of attention in ADHD. However, findings of a smaller Happy-Neutral Target-P1 difference score in adolescents with higher levels of inattention/hyperactivity correspond to the presence of VER errors and biases found in ADHD. Together, these findings further support theories which suggest that individuals with ADHD present with difficulties during the encoding of emotional expressions which contribute to socio-emotional difficulties (Lemerise & Arsenio, 2000; Sonuga-Barke, 2012).

6.4 Clinical Implications

The findings of the present thesis have important implications for clinical practice. Clinical implications deriving from the present research are suggested below.

First, the present findings highlight the clinical importance of examining preattentive processes during emotion perception in ADHD. Adolescents with high levels of inattention/hyperactivity preattentively perceive emotional voices as more salient than those with low levels of inattention/hyperactivity. This has implications for clinical practice as it suggests that emotion processing difficulties in ADHD exist at a preattentive level in adolescents with ADHD traits. If emotion processing difficulties exist at a preattentive level, then a goal of therapeutic intervention might be to increase emotional self-awareness and emotional regulation in individuals with traits of ADHD. Clinicians might aim to incorporate therapeutic techniques which tap onto preattentive emotion processes. For instance, there has been some evidence that play and creative art therapies increase emotional self-awareness and reduce externalising behaviour problems in children and young people (Parker et al. 2021; Ray, Schottelkorb, & Tsai 2007). In addition, findings which suggest that conduct problems might drive a preattentive hypersensitivity to vocal emotions in ADHD suggest that clinicians should carefully assess the presence of conduct problems in adolescents with ADHD. It is possible that in cases in which conduct problems are high, treatments could be tailored to consider potentially greater levels of emotional difficulties in adolescents with ADHD. Similarly, in adolescents with high levels of inattention/hyperactivity only (without comorbidities or emotion difficulties), an emotion-specific treatment approach may not be suitable. In summary, findings of a preattentive hypersensitivity to emotional stimuli in ADHD can inform prevention programs to support emotional development in adolescents with traits of inattention/hyperactivity and conduct problems in the community.

Specifically, the present findings support the presence of a preattentive hypersensitivity to happy vocal stimuli in adolescents with high levels of inattention/hyperactivity. Similarly, previous research reported that relevant reward-specific motivational stimuli, such as stimuli associated with negative and positive reinforcement, can aid task performance in monetary-incentive tasks in adolescents with ADHD (Chronaki et al., 2017). Findings from the present study suggest that when a happy (e.g. encouraging) tone is used, this may be beneficial to task performance. However, it must be noted that the present study only employed task-irrelevant vocal emotions to more robustly explore the preattentive processing of vocal emotions. Therefore, findings from the ESC paradigm appeared to support distracting (not facilitating) effects of vocal emotions, which were specific to vocal anger, and it was not clear how happy voices interact with, and potentially facilitate, attention in task-relevant situations. Thus, further research is required to explore how the preattentive processing of vocal emotions may interact with attention to potentially facilitate performance in other contexts (See section 6.6).

Next, as the present thesis presented evidence that a preattentive hypersensitivity to angry voices was only present in adolescents with extremely high levels of inattention/hyperactivity (i.e., very high SDQ scores), this suggests that a preattentive hypersensitivity to angry voices may be specific to a clinical diagnosis of ADHD. This has implications regarding the clinical significance of ED in ADHD (Shaw et al., 2014). For example, it is possible that ED in ADHD specifically stems from a preattentive hypersensitivity to vocal anger. However, this is speculative because the present study did not recruit a clinical sample of ADHD, nor were there high levels of ED present within the community sample of adolescents. Therefore, further research with clinical populations should be conducted to test this hypothesis (see section 6.6.).

The present findings highlighted that irrelevant vocal anger has a distracting effect on task performance in adolescents with varying levels of inattention/hyperactivity. Specifically, adolescents with both high and low levels of inattention/hyperactivity were slower to respond to targets preceded by irrelevant angry compared to neutral voices. This finding has implications for educators and parents. Parents and teachers should aim to avoid potentially distracting emotional voices, especially anger, in situations which require focus and attention to a task. Importantly, this applies not only to adolescents with high levels of inattention/hyperactivity, but also to typically developing adolescents.

Next, the present findings revealed that adolescents with higher levels of inattention/hyperactivity presented with a smaller Happy-Neutral Cue-P2 amplitude difference score, suggesting that adolescents with ADHD might find it more difficult to differentiate between neutral and happy stimuli. This is in accordance with findings which suggest individuals with ADHD display difficulties to recognise emotional (e.g., happy) voices (Sells et al., 2023). Taken together, these findings have clinical implications, as they suggest that parents and educators should more explicitly express the emotional significance and value of vocal stimuli. This might help children and adolescents with ADHD to better understand the emotional significance of a social situation and respond appropriately (Lemerise & Arsenio, 2000).

6.5 Limitations

Despite the present thesis making important contributions to theory and clinical practice, it also has a number of limitations. There are discussed in detail below.

First, a limitation of the present thesis was that we did not explore thoroughly the role of conduct problems in the preattentive and attentive vocal emotion processing stages. For example, although the present thesis provided some evidence for the role of conduct problems in possibly driving a preattentive hypersensitivity to vocal emotions, we did not provide confirmatory evidence for this. However, because conduct problems and inattention/hyperactivity were highly correlated in the present study, it was difficult to disentangle the contribution of conduct problems from that of hyperactivity. Moreover, because the focus of the present thesis was on traits of inattention/hyperactivity, we did not aim to recruit a high number of adolescents with conduct problems. Therefore, given our small sample size, our study did not have the statistical power to compare two groups of participants (with and without conduct problems) to examine whether our findings are independent of conduct problems.

Second, within the present thesis, adult participants were only recruited for the purpose of validating the task design, and we did not aim to conduct a developmental study. Considering this, the present study recruited 25 adults compared to a later sample of 60 adolescents. This could be seen as a limitation of the present study as we were not able to robustly compare preattentive and attentive vocal emotion processing across the two age groups. Despite this, it is interesting that the present study reported a reward-specific effect on preattentive vocal emotion processing in adolescents which was not reported in adults, and there was a significant effect of age on the Cue-P2 amplitude in the adolescent sample.

Taken together, these preliminary findings relating to age effects on preattentive vocal emotion processing suggest that further developmental research is necessary to explore these effects further (see section 6.6).

Next, it could be argued that including two vocal stimuli per emotion condition in the ESC paradigm was limiting. For instance, if a greater number of stimuli per emotion condition were included, this could potentially have increased the reliability of the emotion effects reported in Chapters 4 and 5. It is not known whether emotion effects on RTs, the Cue-P2 and the Target-P1 and Target-P3 would be replicated when incorporating a larger number of vocal emotional stimuli. However, the inclusion of only two stimuli per emotion condition is consistent with previous studies exploring vocal emotion processing in ADHD (Chronaki et al., 2015).

In addition, by including only two emotion types (happy and angry) in the present ESC paradigm, it is unclear whether emotion-effects on the preattentive and attentive processes are present when including more varied emotion types, such as disgust, surprise, fear. However, in the present study, the incorporation of happy and angry stimuli were justified because they had been evidenced to be recognised with high accuracy in children and adolescents, compared to other emotion types (e.g., disgust, surprise; Sauter et al., 2013). Moreover, including more emotion types would have increased the length of the task, which may not have been suitable with adolescents with high levels of inattention/hyperactivity.

Another limitation of the present thesis was that it was not conclusively clear why congruency effects on the Target-P1 component were not observed in adults and adolescents in Chapters Four and Five. This finding was not in line with previous studies incorporating spatial cueing paradigms, showing consistent effects of spatial congruency on the P1 component (Spence & Driver, 1997, Brosch et al., 2009, Wang et al., 2022). It was concluded that since we found effects of emotion on the Target-P1 and P3 in adolescents, the task was still an effective measure of the capture of attention by vocal emotions and it was possible that emotion effects are prioritised in early visual processing systems, and thus override congruency effects within the present ESC paradigm (as discussed in Chapter Five). However, since there were no interaction effects between congruency or emotion, overall, the present thesis did not report any evidence to support an enhanced effect of vocal emotion on the allocation of attention in ADHD, as demonstrated in previous studies (Chronaki et al., 2017, 2019; Chronaki & Marsh, 2024). This is possibly due to including irrelevant emotions in the ESC paradigm, which had a distracting effect on attention overall, compared to a facilitating effect, as reported in previous studies using different paradigms (i.e., the Monetary-Incentive paradigm; Chronaki et al., 2017; the dot-probe task; Brosch et al., 2009). Section 6.6 outlines some suggestions on how to strengthen the sensitivity of the task to detect emotion-attention interaction atypicalities in ADHD.

6.6 Directions for Future Research

Taking the key findings, theoretical contributions and strengths and limitations of the present thesis into consideration, this section offers some suggestions for future research.

First, to build upon motivation-dysfunction theories of ADHD, and to more thoroughly investigate the clinical significance of the present findings, future research might aim to conduct a case control study to investigate the preattentive processing of vocal emotions in clinical cases of ADHD compared to controls. For example, future research could recruit a group of adolescents with a clinical diagnosis of ADHD, a group of adolescents with sub-threshold levels of ADHD, and a control group of adolescents. This would examine whether anger-specific effects on the Cue-P2 amplitude are more pronounced in clinical cases of ADHD. Similarly, future research could explore the role of conduct problems in a preattentive hypersensitivity to vocal emotion in ADHD by including separate groups of participants with conduct problems such as a group with 'ADHD only', a group with 'ADHD and CD' and a group with 'CD only'. This design would more robustly examine questions surrounding the mediating effect of conduct problems on a preattentive hypersensitivity to vocal emotions in ADHD, preliminary shown in the present thesis.

Second, to address developmental questions, such as the age at which children begin to preattentively process vocal emotions, and whether preattentive vocal emotion processing continues to develop in adolescence (at a similar rate as explicit vocal emotion processing), future studies might aim to employ the ESC paradigm to examine the developmental trajectory of preattentive vocal emotion processing. For example, researchers might aim to recruit a group of children, adolescents and adults in a future study to conduct more robust categorical analysis across age groups. This would help to establish whether adolescents present with greater preattentive sensitivity to positive vocal emotions compared to adults and children. This research would contribute to reward processing theories in adolescence (Galvan, 2013). Similarly, further ERP research might aim to examine the developmental trajectory of the neural markers of the preattentive vocal emotion processing during childhood and adolescence. Finally, further fMRI research might aim to explore if there is an enhanced striatum activity to irrelevant happy voices in adolescents compared to adults to test theories on reward sensitivity in adolescence.

Next, it is important to note that a strength of the present thesis was the development of a novel ESC paradigm, which included irrelevant vocal emotions, to effectively measure the preattentive processing of vocal emotions via the Cue-P2 component. Therefore, future research might aim to employ the ESC paradigm or similar paradigms to further explore the preattentive processing of vocal emotion. For example, future research might aim to explore if a range of vocal emotion types are preattentively processed to the same degree as happy and angry voices, or if a larger P2 to emotional compared to neutral voices is specific to more intense vocal emotions (e.g., anger), which are recognised with high accuracy (as in the present thesis). Similarly, to increase the reliability of the present findings on preattentive vocal emotion processing, future studies might aim to include a greater number of items per emotion type.

In addition, a limitation of the paradigm was that it was not sensitive at measuring interaction effects between attention and emotion, or at measuring the facilitating effects of vocal emotion on attention (See section 6.5), future studies might aim to adapt the present task design to better explore atypicalities in the interaction between emotion and attention in ADHD. For instance, to better explore atypicalities in the capture of attention by irrelevant vocal emotions in ADHD, future research may alter the task design, in order for the Target-P1 component to be a more sensitive measure of the capture of attention by vocal emotion. For example, previous research employing similar spatial cueing paradigms demonstrated stronger emotion effects on the P1 response when the visual target was presented as a white triangle which pointed either upward or downward. Participants responded only when the white triangle pointed upward or downward, which possibly increased the difficulty of the task (Brosch et al., 2009; Wang et al., 2022). In addition, a larger effect of irrelevant vocal anger on the P1 response was found when incorporating a dot-probe study, when the vocal stimuli (anger and neutrality) were presented simultaneously in adults (Brosch et al., 2009). The studies by Brosch and colleagues (2008, 2009) also reported a facilitating effect of emotion on RTs. Therefore, dot-probe tasks, (when irrelevant vocal stimuli compete for attention) might be more effective than the ESC paradigm in exploring potential facilitating effects of preattentively processed vocal emotions on performance.

Finally, the present study reported a larger P1 response to targets preceded by neutral vocal stimuli in ADHD. This contrasts to previous studies which reported a smaller P1 response to non-emotional targets preceded by neutral visual stimuli in ADHD (Perchet et al., 2007). Therefore, future research could explore differences in the way auditory and visual distractors influence visual attention in ADHD. For instance, similar ESC paradigms might be effective at measuring the effects of bi-modal stimuli (e.g., facial and vocal cues together) and unimodal (either visual or auditory cues) on preattentive and attentive processing. This would help explore differences in how auditory and visual emotional stimuli distract or facilitate attention in ADHD.

6.7 Conclusion

In summary, the present thesis presented four studies which reviewed the vocal emotion processing literature in ADHD and developed a novel a paradigm to address the gap in the literature regarding the preattentive processing of vocal emotion processing in ADHD. Overall, findings from the present PhD thesis support the hypothesis that individuals with ADHD present with a preattentive hypersensitivity to vocal emotion (Chronaki et al., 2015). These findings are consistent with motivational dysfunction theories of ADHD (Sonuga-Barke, 2005, 2012). Interestingly, within the present thesis, effects relating to the preattentive processing of vocal emotion in ADHD appear to be specific to happy voices. This is consistent with theories suggesting adolescence is a period of a heightened sensitivity to rewarding stimuli (Galván, 2013). Moreover, conduct problems appear to play a role in both the preattentive and attentive emotion processing atypicalities observed in ADHD within the present study. Further research is necessary to confirm if anger-specific effects are specific to clinical levels of ADHD symptoms, and to more thoroughly explore the role of conduct problems in potentially mediating emotion-specific vocal emotion processing atypicalities in ADHD.

Appendix A. Chapter Two (Meta-Analysis) Supplementary Material

Full Search Strategy for Study 1

The Boolean expression used for the literature search was:

(ADHD OR “attention-deficit hyperactivity disorder” OR “attention-deficit/ hyperactivity disorder” OR “externalising behaviour*” OR “externalizing behavior*” OR “externalising problems” OR “externalizing problems” OR hyperact* OR inattent* OR impuls* OR ‘attention problem*’ OR ‘hyperkinetic disorder) AND (emotion* OR affect* OR ang* OR happ* OR joy OR sad* OR surprise* OR fear* OR disgust* OR fright*)) AND (prosod* OR voice* OR speech OR vocal* OR auditory OR paralanguage OR nonverbal) AND (recogni* OR percept* OR decod* OR identify* OR process* OR naming OR detect* OR match* OR interpret* OR understand*).

The asterisk on some terms allowed for different possible endings e.g., processing, identifying. We formulated our search expression based on terms included in the titles of a list of pre-collected studies exploring associations between vocal emotion recognition and ADHD. Most searches were limited to title and/or abstracts for most terms, apart from words similar to ‘vocal’, which were searched throughout the whole text. This was because some studies which measured emotion recognition in different modalities do not specify the vocal modality in the title or abstract. To identify additional studies, we searched the reference lists of all included studies, and we searched (via google scholar) for any papers we had missed which had cited one key and most cited paper (Cadesky et al, 2000). See the concept table (Table A.1) used to form the search strategy.

Table A.1

Concept Table of Boolean Search Strategy.

Vocal (voc*)	Emotion*	Recognition (recogn*)	ADHD
Voice*	Affect*	Perception / perceive (Perce*)	Attention-deficit/hyperactivity disorder
Prosody (prosod*)	Anger (Ang*)	Decode (decod*)	Attention-deficit hyperactivity disorder
auditory	Happy (Happ*)	Identify*	“Externalising behaviour*”
Speech	Joy*	Process*	“Externalizing behavior*”
Paralanguage	Sad*	Naming (nam*)	“Externalising problems”
Nonverbal	Surprise (Surpris*)	Detect*	“Externalising problems”
	Fear*	Matching (Match*)	Hyperactivity (Hyperact*)
	Disgust*	Interpret *	Inattention (Inattent*)
	Frightened (fright*)	Understand*	Impulsivity (Impuls*) “Attention problem*” “hyperkinetic disorder”

PubMed

Search modes – Boolean search

14/04/2022

((((((((((((ADHD*[Title/Abstract]) OR “attention-deficit hyperactivity disorder”[Title/Abstract]) OR “attention-deficit/ hyperactivity disorder”[Title/Abstract]) OR “externalising behaviour*”[Title/Abstract]) OR “externalizing behavior*”[Title/Abstract]) OR “externalising problems*”[Title/Abstract]) OR “externalizing problems*”[Title/Abstract]) OR hyperact*[Title/Abstract]) OR inattent*[Title/Abstract]) OR impuls*[Title/Abstract]) OR “attention problems”[Title/Abstract]) OR “hyperkinetic disorder” [Title/Abstract])) AND ((((((((((((emotion*[Title/abstract]) OR affect*[Title/Abstract]) OR anger[Title/Abstract]) OR angry[Title/Abstract]) OR happ*[Title/Abstract]) OR joy[Title/Abstract]) OR sad[Title/Abstract]) OR sadness[Title/Abstract]) OR surprise*[Title/Abstract]) OR fear*[Title/Abstract]) OR disgust*[Title/Abstract]) OR fright*[Title/Abstract])) AND ((((((((prosod*[tw]) OR voice[tw]) OR speech[tw]) OR vocal[tw]) OR auditory[tw]) OR paralanguage[tw]) OR nonverbal[tw])) AND (((((((((((recogni*[Title/Abstract]) OR perce*[Title/Abstract]) OR decod*[Title/Abstract]) OR identify*[Title/Abstract]) OR process*[Title/Abstract]) OR naming[Title/Abstract]) OR detect*[Title/Abstract]) OR match*[Title/Abstract]) OR interpret*[Title/Abstract]) OR understand*[Title/Abstract]))

585 results

557 deleted due to ineligibility from titles/ abstracts

28 articles screened in more detail

PsychoInfo and Psych articles

Search modes – Boolean Search

14/04/2022

((((((((((((AB ADHD*) OR AB "attention-deficit hyperactivity disorder") OR AB "attention-deficit/ hyperactivity disorder") OR AB "externalising behaviour*") OR AB "externalizing behavior*") OR AB "externalising problems") OR AB "externalizing problems") OR AB hyperact*) OR AB inattent*) OR AB impuls*) OR AB "attention problems") OR AB "hyperkinetic disorder")) AND ((((((((((((AB emotion*) OR AB affect*) OR AB anger) OR AB angry) OR AB happ*) OR AB joy) OR AB sad) OR AB sadness) OR AB surprise*) OR AB fear*) OR AB disgust*) OR AB fright*)) AND ((((((((TX prosod*) OR TX voice*) OR TX speech) OR TX vocal) OR AB auditory) OR AB paralanguage) OR AB nonverbal)) AND ((((((((((AB recogni*) OR AB perce*) OR AB decod*) OR AB identify*) OR AB process*) OR AB naming) OR AB detect*) OR AB match*) OR AB interpret*) OR AB understand*))

642 results

222 deleted due to duplication

411 deleted due to ineligibility from titles/ abstracts

9 articles retrieved to be screened in more detail

ProQuest Central (Psychology) and ProQuest Dissertations and Theses Global

14/04/2022

(TI,AB(ADHD) OR TI,AB("attention-deficit hyperactivity disorder") OR TI,AB("attention-deficit/ hyperactivity disorder") OR TI,AB("externalising behaviour*") OR TI,AB("externalizing behavior*”) OR TI,AB(“externalising problems”) OR TI,AB(“externalizing problems”) OR TI,AB(hyperact*) OR TI,AB(inattent*) OR TI(impuls*) OR TI,AB(“attention problems”) OR TI,AB(“hyperkinetic disorder”)) AND (TI(emotion*) OR TI(affect*) OR TI,AB(anger) OR TI,AB(angry) OR TI,AB(happ*) OR TI,AB(joy) OR TI,AB(sad) OR TI,AB(sadness) OR TI,AB(surprise*) OR TI,AB(fear*) OR TI,AB(disgust*) OR TI,AB(fright*)) AND (FT(prosod*) OR FT(voice*) OR TI,AB(speech) OR FT(vocal) OR TI,AB(auditory) OR TI,AB(paralanguage) OR TI,AB(nonverbal)) AND (TI,AB(recogni*) OR TI,AB(perce*) TI,AB(decod*) OR TI,AB(identify*) OR TI,AB(process*) OR TI,AB(naming) OR TI,AB(detect*) OR TI,AB(match*) OR TI,AB(interpret*) OR TI,AB(understand*))

389 results

53 deleted due to duplication

328 deleted due to ineligibility

8 articles retrieved to be screened

Further identification of articles

From Citation searching: 5 articles retrieved to be screened.

From reference screening: 4 articles retrieved to be screened.

Authors unpublished work: 1 article retrieved to be screened

Appendix B. Additional Adult Questionnaire Measures

This section details the questionnaire measures included to measure adult behavioural and psychopathology characteristics in Chapters 3 and 4.

Assessment of Behavioural Characteristics in Adults

Weschler Abbreviated Scale of Intelligence (WASI). The *WASI* (Weschler, 2011) is a measure of cognitive ability (i.e., IQ) for individuals between the ages of 6 and 90 years. The full-scale version of the WASI includes four subtests: Vocabulary, Similarities, Block Design and Matrix Reasoning. Due to time limitations, adults were asked to complete two subtests, the Matrix Reasoning and Vocabulary tests. The combination of these two tests only is a valid measure of intelligence, and recommended to use together, as stated in the WASI handbook (Weschler, 2011). The Matrix Reasoning is a measure of non-verbal fluid reasoning and instructs participants to view a series of patterns and use abstract reasoning and visual processing skills to choose the correct pattern in the sequence. During this task the participant is instructed to state the number of the pattern which completes a series of 4 grided patterns out of 5 possible choices. There are 35 items. Participants receive a score of 0 for incorrect responses and a score of 1 for correct response. The maximum score participants can receive is 35. The vocabulary subtest is a measure of verbal intelligence, which also taps into memory, learning ability and language development. Participants are asked to verbally define up to 37 words, which are presented orally and visually. For each item, participants receive a score of 0,1 or 2 depending on the content of the definition based on set criteria in the WASI manual. Adults can receive a maximum of 80 points for this test. Raw scores from the matrix reasoning and vocabulary test are converted to an IQ score based on the standardised scores presented in the WASI manual. The reliability coefficients in adults for the WASI (as noted in the handbook) range from 0.90 – 0.98 for the vocabulary test and 0.88 to 0.96 for the matrix reasoning test. The reliability 'test-retest' coefficients was 0.88 for the vocabulary tests and 0.72 for the matrix reasoning test in adults aged 17 to 54 (see handbook; Weschler, 2011).

Hyperacusis Questionnaire (HQ). The HQ (Khalfa et al., 2001*)* is a 14-item self-report questionnaire which measures hyperacuity, defined as perceived hypersensitivity to sound. Answers to items are measured on a 4-point scale consisting of 'no', 'yes, a little', 'yes, quite a lot' and 'yes, a lot' scoring 0 to 4 points, respectively. Examples of questions include 'Do you find it harder to ignore sounds around you in everyday situations?' and 'Do you have trouble reading in a noisy or loud environment?'. The HQ has been validated and normalised for use in community samples of adults (Khalfa et al., 2001). In the present study, the HQ had good internal reliability (Chronbach's alpha = 0.89).

Self-Report Measures of Adult Psychopathology

Attention Deficit Hyperactivity Disorder-Current Behaviour Scale (CBS). The CBS (Barkley & Murphy, 1998) questionnaire is used to measure symptom dimensions of ADHD including inattention, and hyperactivity/impulsivity in adults based on the DSM-IV criteria. All 18 items were scored on a 4-point Likert scale including 'rarely', 'sometimes' 'often' and 'very often'. Exemplar items which measure inattention include 'I don't listen when spoken to directly' and exemplar items which measure hyperactivity include 'I fidget with hands or feet or squirm in seat.' In the present study, this questionnaire was scored by marking symptomatic answers (i.e. 'often' or 'very often') with a 1 and asymptomatic answers (i.e. 'rarely' or 'sometimes') with a 0. It includes subscales measuring inattention, hyperactivity/impulsivity, and combined-type ADHD (combined inattention and hyperactivity/impulsivity). A score of more than 5 on either the inattentive or hyperactivity/impulsivity scale is thought to represent impairment above the threshold for clinical diagnosis (Barkley, 2011). In the current study, the combined scale in the current study had good internal consistency and reliability (Chronbach's alpha = 0.86).

General Health Questionnaire (GHQ). The GHQ -12 is a self-report measure of depressive symptoms. It consists of 12 items scored on a 1- 4-point Likert scale (e.g., from 'better than usual' to 'much less than usual'). The GHQ is a valid measure of self-reported internalising symptoms, such as recent depressive symptoms in adults (Ozdemir & Rezaki, 2007). The questionnaire instructs participants to compare how they are currently feeling compared to how they generally felt two weeks prior. An example of one question from the GHQ is "Have you recently been able to concentrate on whatever you are doing?". The GHQ had good internal reliability in the current study (Cronbach's alpha = 0.84), which is consistent with high rates (Cronbach's alpha =0.82 to 0.93) reported in previous studies (Goldberg & Williams, 1998; Goldberg, 1978; Makowska et al., 2002).

Appendix C. Chapter Three Supplementary Materials and Analysis

Table C.1

Mean Acoustic Characteristics, Accuracy Rates and Perceived Intensity of Stimuli Included in Emotion Condition in the ESC Paradigm.

Emotion	Stimuli (Item numbers)	Actor Gender	Duration	f0	dB	Adult accuracy rate (%)	Adolescent accuracy rate (%)	Adult Perceived Intensity	Adolescent Perceived Intensity
Neutral	15, 13	Female	678ms	187.32	76.73	96.05	80.43	-	-
Happy	20. 10	Female	673ms	304.32	77.14	95.19	90.58	2.81	2.94
Angry	5, 11	Female	685ms	219.69	77.09	100	90.94	3.21	3.33

Note, this information is taken from Study Three when validating a larger selection of vocal stimuli from Maurage et al (2007) battery.

Additional Analyses: Correlations Between Internalising and Externalising Behavioural Difficulties and Vocal Emotion Recognition Accuracy and Response Bias Scores

Tables C.2 and C.3 display partial correlations between questionnaire measures with accuracy (Pr) and response bias (Br) scores in adults and adolescents, respectively. Age was used as a covariate in these analyses because it was significantly associated with accuracy scores. Consistent with our hypotheses, there was a significant negative correlation (1-tailed) between inattention and accuracy scores for angry voices in adults, (p = .002). Those with higher inattention scores exhibited lower accuracy to recognise angry voices. Adults' response bias towards angry voices also negatively correlated with inattention scores (p = .046). Moreover, consistent with what we hypothesised, hyperactivity and accuracy scores for angry voices were negatively correlated in adolescents (p = .019). Adolescents with higher hyperactivity scores presented with lower accuracy to recognise angry voices. Adolescents' inattention scores also marginally correlated with accuracy scores for angry voices (p = .050). However, despite a clear trend in our data demonstrating a consistent association between inattention and hyperactivity and accuracy scores for angry voices, when Bonferroni corrections were applied at a significance level of p =.005, only the correlation between inattention and vocal anger recognition in adults survived correction.

Contrary to expectation, there was a negative correlation (1-tailed) between self-perceived change in depressive symptoms and an angry response bias in adults (p = .008), and a positive correlation between depressive symptoms and a happy response bias (p=.016). Those with fewer depressive symptoms had a higher response bias to recognise voices as angry. Additionally, in adults, depressive symptoms positively correlated with accuracy to recognise neutral (p = .007) and happy voices (p = .019). In adolescents, depressive symptoms positively correlated with a response bias to recognise voices as neutral (p = .009). However, none of these findings survived multiple testing when Bonferroni corrections were applied at a significance level of p = .005.

Table C.2

Partial Correlations Between Emotion Accuracy Scores (Pr) and Response Bias Scores (Br) with Traits Of ADHD in Adults

		Hyperactivity	Inattention
Angry	Pr	-0.24	-0.56**
	Br	-0.27	-0.34*
Happy	Pr	-0.04	-0.06
	Br	-0.16	-0.26
Neutral	Pr	0.17	0.10
	Br	-0.04	0.10

**p<.05, **p<.01, ***p<.001*

Table C.3

Partial Correlations Between Emotion Accuracy Scores (Pr) and Response Bias Scores (Br) with Traits Of ADHD in Adolescents.

		Hyperactivity	Inattention
Angry	Pr	-0.44*	-0.35*
	Br	-0.21	-0.15
Happy	Pr	-0.28	-0.20
	Br	-0.08	-0.05
Neutral	Pr	-0.22	-0.06
	Br	0.28	0.32

**p<.05, **p<.01, ***p<.001*

Appendix D. Rationale for Methodological Features of the ESC Paradigm

This section provides a rationale for all methodological design features of the paradigm. See on page 48 for a graphical representation of one trial within the ESC paradigm. As in Chapter One, each phase of the experimental paradigm is discussed in the order it appears.

First phase: Fixation Cross. At the start of each trial, participants were presented with a fixation cross to signify the start of a new trial, which stayed on screen until the end of each trial. Participants' visual fixation on the cross throughout the trial ensured their eye movements were limited. A visual fixation cross was used as previous studies have shown that healthy adults can maintain visual fixation on a cross during visual orienting tasks and during auditory orienting tasks (Mayer et al., 2007; Spence & Driver, 1994). Between the presentation of the cross and the onset of the vocal cue, a random and variable time interval (i.e., the interstimulus interval; ISI) between 500 to 1000ms was implemented, allowing for a 100ms baseline prior to the onset of the vocal cue stimulus. Within ERP studies, a variable ISI is essential to reduce artefacts from anticipation effects, overt attention, alpha oscillations, and helps filter out overlapping neural activity from previous trials (Luck, 2014).

Second phase: Vocal Cue. In the second phase of the task, vocal cue stimuli were randomly presented to each ear, with an equal probability on any trial. In this respect, these cues did not predict the target's spatial location, consistent with previous spatial cueing paradigms (Spence & Driver, 1997). Two stimuli were selected for each emotion condition to control for stimulus-specific variations. Non-linguistic vocalisations were chosen to minimise semantic or language-specific effects on attentional capture. All vocal stimuli were standardised in terms of intensity, duration and rise and fall times to ensure that any effects on ERP components were due to the emotional significance and not acoustic characteristics.

The stimulus onset asynchrony (SOA) between the voice and target was fixed at 750ms, consisting of the 700ms vocal cue and a 50ms baseline interval prior to the target. This slightly differed from a variable SOA (550ms -750ms), consisting of 750ms of vocal cue stimulus, used in a previous dot-probe task (e.g., Brosch al., 2009). Despite overlap effects in ERP analyses in paradigms with SOAs of 1 second or less (Luck, 2014), rather than including a variable SOA to reduce these effects, we opted for a fixed SOA, and a 50ms target-baseline. By including a baseline prior to the visual target, this helped to reduce the overlap effects. Although a 50ms baseline is slightly shorter than the recommended time for a baseline period (i.e., 20% of the epoch length; Luck, 2014), previous studies measuring early ERP components (e.g., N1; Joutsiniemi et al, 1998; Lange, 2010; Raij et al., 1997) have used a 50ms baseline and reported significant effects, such as relating to predictable vs unpredictable timing prior to a tone target on N1 amplitude (Lange, 2010). Moreover, a shorter 50ms baseline ensured a sequential presentation of vocal cues and visual targets, whilst avoiding inhibition of return (IOR) effects occurring when participants respond slower to congruent, compared to incongruent, trials because the target appears after the initial reflexive orienting response (Mayer et al., 2007). Previous spatial cueing paradigms incorporating auditory cues have reported IOR effects when the SOA was greater than 800ms (Mayer et al., 2007; McDonald & Ward, 2000), so it was vital to ensure the SOA was less than 800ms (i.e.,750ms) by incorporating a shorter (50ms) baseline period. Overall, the fixed 750ms SOA and 50ms baseline were chosen to optimise the task for ERP analysis.

Third phase: Visual Target. In the third phase of the emotional spatial cueing paradigm, a visual target (yellow star) was presented for 100ms on either the left or right side of the screen at a 24-degree angle from the eye, by instructing participants to sit 60cm from the computer screen. A yellow star was chosen because it is salient and child-friendly, consistent with previous developmental studies (Pollak & Tolley-Schell, 2003; Perchet et al., 2001). The angle and duration of the target were consistent with previous spatial cueing ERP paradigms with auditory stimuli in adults (Brosch et al., 2009; McDonald & Ward, 2000). Additionally, the presentation of a visual stimulus was suitable for use with EEG techniques, because when a visual target is presented for 100ms, this is short enough for offset times not to affect ERPs (Luck, 2014). Finally, as adolescents have been shown to display adult levels of visual processing speed (Croker & Maratos, 2011), a duration of 100ms was considered suitable for use with adolescents. The spatial location of the visual target on either the left or right side of the screen was random, with equal probability of congruent and incongruent trials, consistent with previous spatial cueing paradigms (Spence & Driver, 1997; Brosch et al., 2009; Feng et al., 2014), to ensure no predictive elements were added to the task.

Fourth Phase: Button Press Response. In the fourth phase of the ESC, participants were required to press the left or right arrow key button to indicate the location of the visual target, with their dominant hand. Participants had infinite time to respond, after which the next trial begun. This allowed participants to blink and re-focus before the onset of the next trial and helped reduce ocular artefacts from blinking and eye movements. Incorporating a button-press response allowed the measurement of reaction times, as a behavioural measure of visual spatial attention (Brosch et al., 2008; Spence & Driver., 1997). As button press responses could introduce artefacts in the ERP analysis (Brosch et al., 2009), ERP studies have used an upward or downward facing triangle as a visual target and have instructed participants to respond only on ‘go’ trials (e.g., when the triangle was upward), which were subsequently removed from the EEG analysis (Brosch et al., 2009). Because this approach removes potentially informative trials from the analysis and increases the number of trials and length of the task, it was not considered appropriate when testing adolescents with potentially limited attentional capacity in the present thesis. Instead, the present paradigm ensured consistent motor preparation via a button press for all trials, a method shown not to introduce artefacts in vocal emotion recognition ERP studies with younger participants (Chronaki et al., 2015).

Number of Trials. The present task included a shorter number of total trials (288) compared to 640 trials included in the EEG study by Brosch et al (2009), reducing the length of our paradigm to approximately 15 minutes, which was considered suitable for adolescents with high traits of inattention and hyperactivity. This is justified as although Luck (2014) recommends including 400 to 600 trials in adults to reliably explore early components, such as the P1 to a visual target, Brosch et al (2009) found significant effects of emotion on the P1 component in a dot-probe paradigm in only the first block which included 160 trials in total (i.e. 80 trials per emotion condition), possibly due to habituation effects. Moreover, the total number of trials (96) per emotion condition within the present paradigm is supported by previous developmental ERP research showing habituation effects on early ERP components (e.g., significant emotion effects on the N1 only in the first half of the task) in vocal emotion recognition tasks with a smaller number of trials (60 trials per emotion; Chronaki et al. 2015).

Appendix E. Chapter Four Supplementary Analysis and Materials

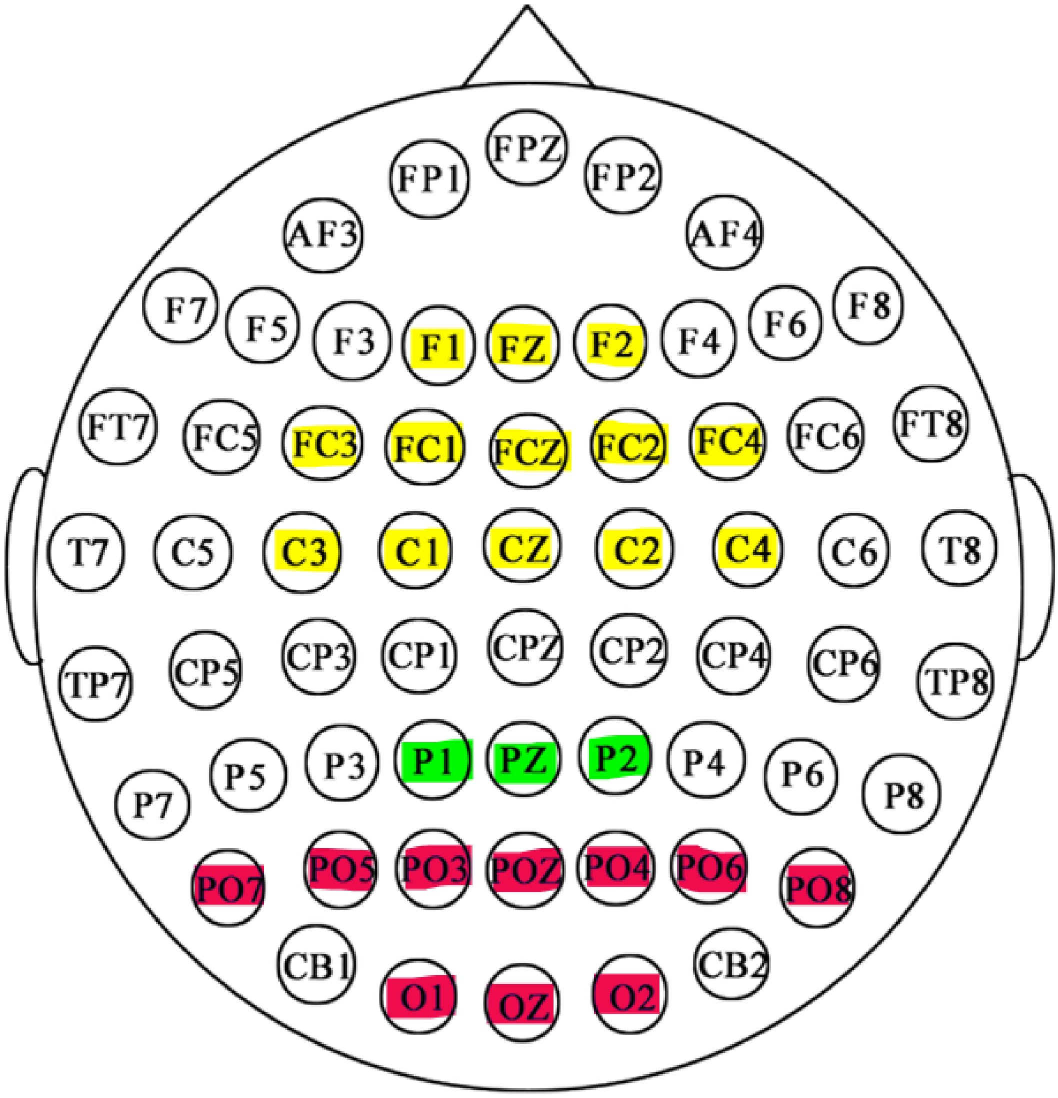

Figure E.1

ERP Montage of Sites Included in Analyses. Frontocentral Sites highlighted in yellow were included in Cue-P2 analysis. Mid-line parietal sites highlighted in green were included in Target-P3 analysis. Parietal Occipital sites highlighted in red were included in Target-P1 analysis.

Table E.1

Kolmogorov Smirnov Statistics, D (P), Testing for Normality of Data in All ERP Analyses in Adults (n=25).

	Adults (n=25)
Cue-P2	
Neutral	0.13 (.200)
Happy	0.11 (.200)
Angry	0.12 (.200)
Target-P1	
Angry Congruent	0.12 (.200)
Angry Incongruent	0.14 (.200)
Happy Congruent	0.09 (.200)
Happy Incongruent	0.09 (.200)
Neutral Congruent	0.15 (.148)
Neutral Incongruent	0.17 (.120)
Target-P3	
Angry Congruent	0.13 (.200)
Angry Incongruent	0.21 (.005)
Happy Congruent	0.16 (.103)
Happy Incongruent	0.17 (.073)
Neutral Congruent	0.15 (.168)
Neutral Incongruent	0.10 (.200)

Table E.2

Mean (SD) Amplitudes of the P2 Component to the Vocal Cue for Each Emotion Condition in Adults (n=25)

FrontoCentral P2	
Angry	1.88 (2.21)
Happy	2.32 (2.00)
Neutral	1.11 (2.23)

Table E.3

Mean (SD) Amplitudes of the P1 Component to the Visual Target by Emotion and Congruency Conditions in Adults (n=25)

Parietal Occipital P1			
	Congruent	Incongruent	Total
Angry	0.45 (1.32)	0.41 (1.52)	0.43 (0.26)
Happy	0.33 (1.25)	0.16 (1.43)	0.25 (0.23)
Neutral	0.77 (1.40)	0.66 (1.43)	0.72 (0.22)
Total	0.52 (0.22)	0.41 (0.22)	

Table E.4

Mean (SD) Amplitudes of the P3 Component to the Visual Target by Emotion and Congruency Conditions in the Right and Left Visual Field (RVF, LVF) in Adults (n=25)

Parietal Occipital P1			
	Congruent	Incongruent	Total
Right Visual Field			
Angry	0.66 (3.44)	1.49 (2.40)	1.08 (2.82)
Happy	0.75 (3.27)	1.36 (2.93)	1.06 (2.89)
Neutral	1.35 (3.07)	1.30 (3.55)	1.33 (3.04)
Total	0.93 (2.89)	1.38 (2.69)	
Left Visual Field			
Angry	1.12 (2.52)	1.21 (3.33)	1.16 (2.77)
Happy	1.75 (3.34)	0.85 (3.16)	1.29 (2.94)
Neutral	1.57 (1.99)	1.21 (2.81)	1.39 (2.23)
Total	1.31 (2.41)	1.08 (2.75)	

Appendix F. Chapter Five Supplementary Analysis

Table F.1

Kolmogorov Smirnov Statistics, D (p), of Dependent Variables, Testing for Normality of Data

	Adolescents (n=63)	Low Traits (n=13)	High Traits (n=13)
Reaction Times			
Angry Congruent	0.12 (.067)	0.16 (.200)	0.13 (.200)
Angry Incongruent	0.12 (.064)	0.24 (.047)	0.14 (.200)
Happy Congruent	0.13 (.007)	0.14 (.200)	0.13 (.200)
Happy Incongruent	0.10 (.096)	0.19 (.197)	0.18 (.200)
Neutral Congruent	0.14 (.004)	0.15 (.200)	0.18 (.200)
Neutral Incongruent	0.12 (.058)	0.17 (.200)	0.12 (.200)
Cue-P2			
Neutral	0.11 (.069)	0.15 (.200)	0.26 (.021)
Happy	0.11 (.087)	0.22 (.085)	0.19 (.200)
Angry	0.09 (.200)	0.16 (.200)	0.33 (<.001)
Target-P1			
Angry Congruent	0.08 (.200)	0.18 (.200)	0.14 (.200)
Angry Incongruent	0.06 (.200)	0.22 (.149)	0.17 (.200)
Happy Congruent	0.08 (.200)	0.15 (.200)	0.17 (.200)
Happy Incongruent	0.08 (.200)	0.16 (.200)	0.16 (.200)
Neutral Congruent	0.05 (.200)	0.20 (.200)	0.17 (.200)
Neutral Incongruent	0.06 (.200)	0.17 (.200)	0.12 (.200)
Target-P3			
Angry Congruent	0.07 (.200)	0.22 (.131)	0.14 (.200)
Angry Incongruent	0.07 (.200)	0.32 (.003)	0.16 (.200)
Happy Congruent	0.12 (.027)	0.21 (.200)	0.15 (.200)
Happy Incongruent	0.12 (.022)	0.17 (.200)	0.24 (.045)
Neutral Congruent	0.08 (.200)	0.15 (.200)	0.11 (.200)
Neutral Incongruent	0.10 (.200)	0.22 (.148)	0.17 (.200)

Table F.2

Kolmogorov-Smirnov Statistics D (p) of Inattention/Hyperactivity, Behavioural Characteristics, Emotion Regulation and ADHD Comorbidities

Variables	KS Statistics
Inattention/Hyperactivity	0.09 (p=.200)
Behavioural Characteristics	
Age	0.17 (p<.001)
Hearing Threshold	0.25 (p<.001)
Hyperacuity	0.16 (p<.001)
IQ	0.12 (p=.033)
Emotion Regulation	
Emotion Reappraisal	0.13 (p=.039)
Emotion Suppression	0.09 (p=.200)
ADHD Comorbidities	
Conduct Problems	0.20 (p<.001)
Depressive Symptoms	0.26 (p<.001)

Table F.3

Spearman's Rho (Two-Tailed) Correlations Between Dependent Measures, Adolescent Psychopathology and Behavioural Characteristic (n=63).

	RTs	Cue-P2	Target-P1	Target-P3	Inattention/ Hyperactivity	Conduct Problems	Depressive Symptoms	Emotion Reappraisal	Emotion Suppression	IQ	Age	Hyperacuity	Hearing Threshold
RTs	-												
Cue-P2	0.13	-											
Target-P1	0.24	0.22	-										
Target-P3	0.08	-0.08	**0.35****	-									
SDQ-Hyperactivity/ Inattention	0.12	**0.38****	0.07	0.07	-								
SDQ-Conduct Problems	0.09	**0.24***	-0.01	0.05	**0.66*****	-							
Depressive Symptoms	0.13	-0.06	0.16	0.04	0.09	0.19	-						
Emotion Reappraisal	-0.02	-0.12	-0.26	**-0.33***	-0.17	-0.09	**-0.32***	-					
Emotion Suppression	-0.13	0.23	-0.02	0.02	0.14	**0.27***	0.11	0.03	-				
IQ	-0.02	-0.18	-0.02	0.04	**-0.31****	**-0.40*****	-0.12	-0.08	-0.13	-			
Age	**-0.47*****	**-0.24***	0.04	0.02	-0.13	-0.10	0.07	0.08	-0.06	-0.09	-		
Hyperacuity	0.15	0.12	0.16	0.13	**0.24***	**0.37****	**0.42*****	-0.16	**0.24***	-0.18	0.02	-	
Hearing Threshold	0.02	-0.12	0.18	-0.03	**-0.25***	-0.20	-0.09	0.11	0.04	-0.08	0.08	0.05	-

*Note:*p<.05, **p<.01, ***p<.001*

Table F.4

Mean (SD) Amplitudes for the Cue-P2 Component per Emotion Condition in Adolescents (n=61)

Emotion	Mean (SD)
Angry	2.21 (2.11)
Happy	2.55 (1.97)
Neutral	0.91 (1.99)

Table F.5

Mean (SD) Amplitudes for the Target-P1 for Emotion by Congruency Conditions in Adolescents (n=60)

Emotion	Congruent	Incongruent	Total
Angry	1.22 (1.96)	1.52 (1.97)	1.37 (1.72)
Happy	0.79 (1.98)	0.74 (2.10)	0.77 (1.83)
Neutral	0.69 (1.65)	0.78 (1.97)	0.74 (1.46)
Total	0.91 (1.49)	1.02 (1.66)	

Table F.6

Mean (SD) Amplitudes for the Target-P3 for Emotion by Congruency Conditions in Adolescents (n=60).

Emotion	Congruent	Incongruent	Total
Angry	1.87 (3.80)	2.40 (2.56)	2.16 (2.57)
Happy	1.79 (2.39)	2.07 (2.79)	1.95 (2.43)
Neutral	1.62 (2.91)	1.66 (3.24)	1.52 (2.69)
Total	1.69 (2.41)	2.06 (2.55)	

Table F.7

Mean (SD) Amplitudes for Cue-P2 by Emotion Condition and ADHD Group

FrontoCentral P2	Low Traits (n=13)	High Traits (n=12)
Angry	0.58 (1.62)	2.57 (1.64)
Happy	1.08 (1.52)	3.64 (1.56)
Neutral	-0.11 (1.80)	1.01 (1.31)

Table F.8

Spearman's Rho Correlations Between RTs and Behavioural Characteristics, Psychopathology and Emotion Regulation (n=63)

RTs	Age	IQ	Hyperacuity	Hearing Threshold	Emotion Suppression	Emotion Reappraisal	Conduct Problems	Depressive Symptoms
Angry	-0.46***	0.04	0.17	0.08	-0.12	-0.02	0.12	0.13
Happy	-0.47***	-0.01	0.11	0.07	-0.09	-0.02	0.12	0.15
Neutral	-0.51***	0.03	0.13	0.06	-0.16	-0.02	0.08	0.13
Difference Scores								
Angry – Neutral	0.18	0.08	0.24	0.18	0.22	0.01	0.18	0.05
Happy-Neutral	0.09	-0.21	0.08	0.10	0.26	0.05	0.19	0.05

Note: *p<.05, **p<.01, ***p<.001

Table F.9

Spearman's Rho, (2 Tailed) Correlations Between the Cue-P2 and Behavioural Characteristics, Emotion Regulation and Psychopathology (n=61).

P2 Amplitude	Age	IQ	Hyperacuity	Hearing Threshold	Emotion Suppression	Emotion Reappraisal	Conduct Problems	Depressive Symptoms
Angry	-0.23	-0.15	0.12	-0.16	0.23	-0.04	0.22	0.01
Happy	-0.26*	-0.25*	0.16	-0.18	0.28	-0.17	0.35**	0.05
Neutral	-.19	-0.07	0.06	0.04	0.20	-0.11	0.14	-0.09
Difference Scores								
Angry – Neutral	-0.14	-12	0.09	-0.32*	0.13	0.05	0.12	0.09
Happy – Neutral	-0.12	-0.14	0.1	-0.36**	0.05	-0.05	0.28*	0.18

Note: *p<.05, **p<.01, ***p<.001

Table F.10

Spearman's Rho, (Two-Tailed) Correlations Between the Target-P1 and Behavioural Characteristics, Emotion Regulation and Psychopathology (n=60)

P1 Amplitude	Age	IQ	Hearing Threshold	Hyperacuity	Emotion Suppression	Emotion Reappraisal	Conduct Problems	Depressive Symptoms
Angry Congruent	-0.02	0.05	0.06	0.23	0.17	-0.25	0.11	0.29*
Happy Congruent	0.04	0.14	0.14	0.05	-0.20	-0.37**	-0.03	0.01
Neutral Congruent	0.05	-0.03	-0.12	-0.24	0.08	-0.02	-0.08	-0.01
Difference Scores								
Angry– Neutral Congruent	-0.07	0.04	0.15	0.30*	0.07	-0.33*	0.24	0.30*
Happy –Neutral Congruent	0.02	0.16	0.24	0.09	-0.24	-0.25	0.03	0.02

Note: *p<.05, **p<.01, ***p<.001

Table F.11

Spearman's Rho, (Two-Tailed) Correlations Between the Target-P3, Behavioural Characteristics, Emotion Regulation and Psychopathology (n=60).

	Age	IQ	Hearing Threshold	Hyperacuity	Emotion Suppression	Emotion Reappraisal	Conduct Problems	Depressive Symptoms
Angry	0.02	0.11	-0.03	0.09	0.15	-0.38**	-0.04	0.01
Happy	0.12	-0.07	0.02	0.25	0.12	-0.32*	0.12	0.20
Neutral	-0.01	0.09	0.02	0.06	-0.14	-0.16	-0.08	-0.06
Difference Scores								
Angry-Neutral	0.13	-0.04	-0.01	-0.06	0.34*	-0.23	0.14	0.03
Happy-Neutral	0.15	-0.20	0.03	0.15	0.29*	-0.17	0.33*	0.24

Note: *p<.05, **p<.01, ***p<.001

Table F.12

Pearsons Correlations (One-Tailed) Between RTs snd Inattention/Hyperactivity (n=63).

RTs	Inattention/ Hyperactivity
Angry	0.13
Happy	0.14
Neutral	0.14
Difference Scores	
Angry – Neutral	-0.01
Happy - Neutral	0.01

*Note, *p<.05,* **p<.01, ***p<.001.

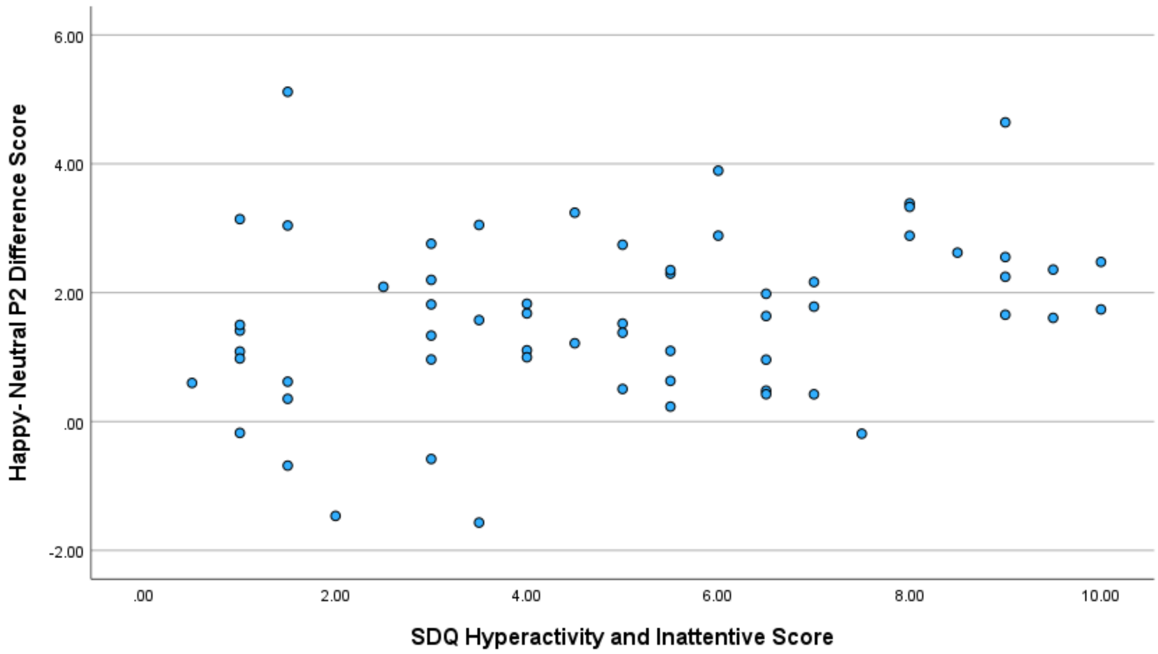

Figure F.1

Scatter Graph Showing Difference Scores for the Cue-P2 Amplitude for Happy Voices with the Cue-P2 Amplitude for The Neutral Voices as Baseline

Table F.13

Pearsons's Correlations (One-Tailed) Between the Target-P1 and Inattention/Hyperactivity With and Without Controlling for Conduct Problems (CP).

	Inattention/ Hyperactivity	Inattention/Hyperactivity (controlling for CP)
Angry	0.12	0.08
Happy	0.04	0.01
Neutral	0.08	0.26*
Difference Scores		
Angry-Neutral	0.07	-0.18
Happy-Neutral	-0.03	-0.24*

*Note, *p<.05,* **p<.01, ***p<.001. However, none of these correlations survived corrections for multiple testing.

Target C1 Literature Review and Analysis

Literature Review

An exploratory analysis investigated possible effects of congruency and emotion earlier than the P1 component to the visual target. For these exploratory analyses we focused on the C1 component, which appears between 50 and 100ms in midline parietal sites within the present data, consistent with previous literature (Foxe & Simpson, 2002; Rauss et al.., 2011). The C1 is thought to reflect the earliest activation of the visual cortex (V1) prior to the Target-P1 component, which reflects further occipital cortex activity (Foxe & Simpson, 2002; Rauss et al.., 2011). Findings from intracranial recordings in monkeys support the view that early activation of the visual cortex occurs as early as 40-60ms (Raiguel et al. 1989; Givre et al. 1994). To explore the topography and latency of the C1, Foxe and Simpson (2002) employed a selective-attention paradigm in which they presented bilateral red disks centred to the left and right of a fixation cross for 280ms. These target stimuli were preceded by a central word cue which instructed participants to either attend to the visual target or attend instead to an auditory target. Results revealed an early C1 component at approximately 50-62 ms as a negativity at POZ, this was followed by the next phase of the C1 between 60-80ms, which was a positive current at PO3 and PO4. The C1 component continued between 80 and 100ms.

There is debate within the literature as to whether an early C1 component may be influenced by attentional modulations, similar to the P1 component (Hillyard et al., 1998). For instance, Rauss and colleuges (2011) reviewed the effect of top-down mechanisms on the C1 component and reported evidence in favour of both sides of the debate. In favour of attentional mechanisms influencing the C1 amplitude, studies with monkeys found that when their attentional focus was manipulated, very early modulations of visual processing have been reported, as early as 50ms in the visual cortex (Motter, 1994; Ito & Gilbert, 1999). Additionally, Kelly et al (2009) reported effects of endogenous attention on the C1 in adult participants, in a spatial cueing paradigm which presented a centrally presented arrowhead which covertly directed attention to the left or right side of the screen. Findings revealed enhanced C1 amplitudes when actively attending to the location of a subsequently presented grating target stimulus (Kelly et al., 2009). Despite this, few studies have reported similar congruency effects on the C1 (as found on the P1) with prior auditory cues in spatial cueing paradigms. For example, McDonald et al (2005) presented a nonpredictive spatial auditory cue prior to visual targets in a spatial cueing paradigm, and although a C1 was evident in the waveforms (mean latency 72ms) in parietal occipital site, it was not altered by congruency of the auditory cue.

To add to the debate on congruency effects on the C1 component, Slotnick (2020) reviewed the literature on spatial attention effects on the C1 component. Specifically, Slotnick (2020) suggested what experimental conditions are most likely to produce spatial attention effects on the C1. For example, Slotnick (2020) concluded that exogenous orienting is more likely to elicit C1 attentional modulation with SOA between 69 – 300ms. Moreover, some studies only reported effects of attention when the target was presented in the upper visual field. Additionally, Slotnick (2017) suggested the C1 has its peak magnitude at midline parietal-occipital sites, and were specifically strong at sites POZ, Pz and CPZ.

In addition, Rauss et al (2011) reviewed some studies which explored effects of emotion on the C1 component. Notably, in a dot-probe study which presented fearful, happy and neutral faces prior to a visual target, Pourtois et al (2004) reported a larger C1 amplitude to fearful and happy facial cueing stimuli at 80-100ms at parietal sites. These C1 modulations correlated with greater P1 effects on the subsequent target stimuli. However, no effects of emotion on the C1 were observed to the target stimulus. In support of emotion effects modulating the C1 component, Stolarova et al (2006) conditioned neutral stimuli (small black and white squares) with unpleasant or arousing photographs, and other neutral stimuli with low arousing pictures. Findings revealed the C1 component (65-90ms), was significantly more negative to the conditioned stimuli compared to baseline at the posterior sites. This provides evidence that the emotional significance relating to a stimulus can lead to changes in early visual processing at the amplitude of the C1 component.

Overall, congruency (attention) and emotion effects could influence initial visual processing, reflected in an early C1-like component within our paradigm. This would appear in parietal-occipital sites (centrally presented – POZ, PO3, PO4, O1, OZ, O2, PZ, CPZ) between 50 and 100ms. However, the C1 according to Slotnick (2017) component may not be as robust as in other studies because stimuli were presented centrally (not upper visual field), and perceptual load was low in our study. In summary, we identified a C1-like component in the visual target ERP data and investigated effects of congruency and emotion on the C1. Results and a brief discussion are presented below.

Results

A 2 x 3 repeated measures ANOVA revealed there was no significant main effect of congruency on the C1 amplitude ($F(1,59)= 3.55$, $p=.065$). There was a significant main effect of emotion on the C1 amplitude ($F(2,118) = 7.57$, $p<.001$). There was no congruency x emotion interaction ($F(2,118) = 0.42$, $p=.657$) on the C1. Post-hoc tests revealed a significantly larger C1 (i.e., more positive) amplitude for targets preceded by happy compared to angry voices

(MD=0.60, p<.001). There was no significant difference between the C1 amplitude to targets preceded by angry compared to neutral voices (MD=0.39, p=.084), or between targets preceded by happy compared to neutral voices (MD=0.21, p=.596).

Discussion

Overall, there were no significant effects of congruency on the C1, although these effects were close to significance and in the right direction. Interestingly, there was a main effect of emotion on the C1. Participants displayed a more positive C1 amplitude to happy compared to angry voices. It is interesting that emotion effects in the present ESC paradigm appear as early as the C1 component. Angry voices elicited a more negative amplitude as early as 50ms, which is consistent with a more negative amplitude compared to neutral voices observed at the P1 component. This suggests the effect of vocal anger on early visual processing occurs as early as 50ms after the onset of the visual target in the ESC paradigm. However, overall, the analysis of the C1 did not provide evidence for congruency or emotion x congruence interaction effects during target processing.

Target Categorical Analysis

Table F.14

Sample Characteristics for Adolescents with High and Low Levels of ADHD Traits

	Low ADHD Traits (n=13)	High ADHD Traits (n=12)
Age	13.69 (1.31)	13.67 (1.23)
Gender (% Male)	46%	69%
Handedness (N Left)	0 Left	1 Left
IQ	110.31 (10.70)	97.33 (11.48)
Hyperacuity	10.31 (8.90)	17.50 (9.99)
Hearing Threshold	12.86 (6.36)	10.91 (6.25)
Psychopathology		
SDQ		
Inattentive and hyperactivity	1.23 (0.38)	8.96 (0.70)
Conduct Problems	0.81 (0.73)	4.17 (1.77)
GHQ Depressive Symptoms	1.61 (2.36)	2.47 (2.46)
ERQ		
Emotion Reappraisal	19.67 (4.18)	16.36 (4.20)
Emotion Suppression	11.00 (2.65)	11.63 (4.21)

Data Processing

There were fewer artefact-free epochs for adolescents with High trait compared to the Low trait group for the target (p=.028). There were no other differences in the number of artefact-free trials between conditions for either group (p>.20). The mean (SD) number of artefact-free epochs were as follows: HIGH TRAITS: Angry Congruent: 40.54 (6.45), Happy Congruent: 40.46 (6.19), Neutral Congruent: 40.62 (4.48), Angry Incongruent: 41.08 (5.84), Happy Incongruency: 40.15 (6.45), Neutral Incongruent: 39.85 (5.71); LOW TRAITS: Angry Congruent: 45.00 (2.23), Happy Congruent: 45.27 (1.79), Neutral Congruent: 44.18 (3.54), Angry Incongruent: 44.91 (2.55), Happy Incongruent: 44.18 (2.96), Neutral Incongruent: 43.91 (3.24).

Behavioural Data Treatment

A Kolmogorov-Smirnov test indicated that most RT variables for both groups of participants were normally distributed (See Table F.1). However, the RT scores for the angry incongruent condition for the low traits group was not normally distributed (D=0.24, p=.047). Additionally, Kolmogorov-Smirnov tests indicated that within the subset sample, behavioural characteristics (hearing threshold, IQ, age, hyperacuity) and psychopathology characteristics (conduct problems, depressive symptoms) were not normally distributed (p<.05). Spearman's rho correlations revealed that age significantly correlated with RTs (r=-0.62, p=<.001), therefore age was explored as a covariate in analyses. No other associations reached significance (all ps >.05).

Kolmogorov-Smirnov tests indicated that the Target-P1 and P3 were normally distributed for all emotion x congruency conditions in both Low and High ADHD groups. Spearmans' rho correlations revealed no significant correlations between behavioural (IQ.

hyperacuity, age, hearing threshold) and P1 amplitudes (ps>.05). In addition, Spearman's rho correlations revealed a significant correlation between the Target-P3 to incongruent trials preceded by happy voices and hyperacuity (r=0.39, p=.018). Therefore, hyperacuity was added as a covariate in Target-P3 analyses.

Results

Reaction Times. A 2 (ADHD group) x 2 (congruency) x 3 (emotion) ANOVA revealed no main effect of group on RTs (F(1,24)= 0.55,p=.467). There was no congruency x group interaction (F(1,24) = 0.29, p=.590), or emotion x group interaction (F(2,48) =0.01, p=.995) or emotion x congruency x group interaction (F(2,48) = 1.88, p=.164). When age was added in as a covariate, the pattern of results did not change (p>.05).

Target-P1. A 2 (group) x 3 (emotion) x 2(congruency) mixed ANOVA was performed on the Target-P1 amplitudes values. Results revealed no additional effects of group on the Target-P1 (F(1,22)=1.32, p=.263). There were no group by emotion, group by congruency or group by emotion by congruency interaction effects on the Target-P1 (all ps > .05).

Target-P3. A 2 (congruency) x 3 (emotion) x 2 (ADHD group) ANOVA was performed on the Target-P3 amplitudes values. There was no main effect of group on the Target-P3 (F(1,22) = 0.46, p=.507). Finally, there were no congruency by group, emotion by group or congruency by emotion by group interaction effects on the Target-P3 (all ps>.05).

Appendix G. Comparative Analysis between Adolescents and Adults

Participants

To conduct comparative analysis between adult and adolescent participants, we applied the same inclusion criteria to both sets of data. Therefore, we excluded seven participants from the adolescent dataset who had a clinical diagnosis of ADHD or other diagnoses. Additionally, to ensure dataset met criteria to conduct parametric tests, we ensured equal sample sizes in both adult and adolescent groups by only including the first 25 adolescent participants who met the inclusion criteria. In total, 25 adolescent and 25 adults were included in this analysis. Sample characteristics are provided in Table F.15.

Table F.15

Sample Characteristics for Adolescents (n=25) and Adults (n=25)

	Adolescents (n=25)	Adults (n=25)
Age	13.84 (1.43)	24 (5.73)
Gender (% Male)	40	32
Handedness (% Right)	92	96
IQ	108.52 (11.48)	110.68 (10.00)
Hyperacuity	12.28 (9.51)	15.29 (8.61)
Psychopathology		
CBS		
Adult Inattentiveness	-	2.12 (2.57)
Adult Hyperactivity	-	1.84 (1.52)
SDQ		
Adolescent Inattention and Hyperactivity	3.04 (2.31)	-
Adolescent Conduct Problems	1.06 (1.08)	-
GHQ Depressive Symptoms	2.20 (2.80)	2.92 (3.17)
ERQ		
Emotion Reappraisal	18.73 (4.76)	-
Emotion Suppression	12.75 (2.79)	-

Results

RTs. A 2 (congruency) x 3 (emotion) x 2 (Age group) ANOVA was conducted. Results revealed a main effect of emotion ($F(2,96) = 4.69$, $p=.011$), a main effect of congruency ($F(1,48) = 15.62$, $p<.001$) on RTs. Additionally, there was a main effect of age group on RT ($F(1,48) = 6.65$, $p=.013$). Adults presented with faster reaction times to the visual target than adolescents. There were no emotion x age group, congruency x age group or emotion x congruency x age group interaction effects on RTs ($p>.05$). Post hoc tests revealed slower RTS to angry compared to both happy ($p=.039$) and neutral ($p=.018$) voices.

Cue-P2. A 3 (emotion) x 2 (age group) ANOVA revealed a main effect of emotion ($F(2,96) =40.59$, $p<.001$) on the P2 amplitude to the vocal cue. There was no main effect of age on the P2 amplitude ($F(1,48) = 0.31$, $p=.574$). In addition there was no emotion x age group interaction effect on the P2 amplitude ($F(2,96) = 1.43$, $p=.246$).

Target-P1. A 3 (emotion) x 2 (congruency) x 2 (age group) revealed a main effect of emotion on the P1 amplitude to the visual target ($F (2,96) = 3.45$, $p=.036$). There was no main effect of congruency on the P1 amplitude ($F (1,48) = 0.01$, $p=.918$). There was no main effect of age group on the P1 ($F (1,48) = 0.59$, $p=.445$). There was a close to significant emotion x age group interaction effect on the P1 amplitude ($F (2,96) = 2.99$, $p=.055$). Post hoc tests confirmed that the larger P1 to angry compared to happy and neutral voices was found only in adolescents ($p=.014$, $p=.053$) but not adults. There was no other significant congruency x age group or emotion by congruency x age group interaction effects on the P1 amplitude (all $ps>.05$).

Target-P3. A 3 (emotion) x 2(congruency) x 2 (age group) revealed no main effect of emotion on the P3 amplitude to the visual target ($F (2,96) = 0.69$, $p=.502$. There was also no main effect of congruency (F (1,48) = 0.05, p=.834, or emotion by congruency interaction effect on the P3 amplitude (F (2,96) = 1.52, p=.224). There was no main effect of age group on the P3 amplitude (F (1,48) = 0.01, p=.969). Finally, there were no other significant congruency x age group, or age group by emotion, or age group by emotion by congruency interaction effects on the P3 amplitude (all ps>.05).

Appendix H. Exploratory Associations Between Dependent Variables (RTs and ERPs) and Adult Inattention and Hyperactivity

Data Processing and Initial Data Treatment

A Kolmogorov-Smirnov test indicated that adult hyperactivity and inattention were not normally distributed ($D=0.19$, $p=.021$; $D=0.27$, $p<.001$). Therefore, non-parametric tests were conducted. Preliminary Spearman's rho correlations identified significant positive correlations between depressive symptoms and adult inattention ($r=0.50$, $p=.012$) and adult hyperactivity ($r=0.43$, $p=.043$). In addition, hyperacuity positively correlated with adult inattention ($r=0.62$, $p=.001$) and adult hyperactivity ($r=0.49$, $p=.013$). Therefore, both depressive symptoms and hyperacuity were added as covariates in analysis. No other significant correlations were found between adult hyperactivity and inattention and age or IQ ($ps>.05$).

RT Results

For consistency with the adolescent correlational analyses, correlations were conducted to explore the relationships between RTs (for all emotion conditions), the angry-neutral RT difference score, and adult inattention and hyperactivity. When conducting Spearman's rho correlations, no significant correlations between RTs and inattention or hyperactivity were found (all $ps>.05$). When conducting Pearsons's correlations and controlling for hyperacuity and depressive symptoms, this did not change the pattern of results (all $ps>.05$).

Cue-P2 Results

For consistency with the adolescent correlational analyses, correlations were conducted to explore the relationships between the Cue-P2 (for all emotion conditions), the P2 difference scores (angry-neutral; happy-neutral), and inattention and hyperactivity. Spearman's rho correlations revealed no significant correlations between the P2 amplitude and inattention or hyperactivity (all ps>,05). When conducting Pearson's correlations and controlling for depressive symptoms, the correlation between the P2 amplitude for neutral voices and hyperactivity was close to significance (r=0.40, p=.058), but no other correlations were significant (all ps>.05). Similarly, no correlations were significant when controlling for hyperacuity (all ps>.05).

Target-P1 Results

For consistency with the adolescent correlational analyses, correlations were conducted to explore the relationships between the Target-P1 (for all spatially congruent emotion conditions), the P1 difference scores (angry congruent -neutral congruent; happy congruent-neutral congruent), and inattention and hyperactivity. When conducting Spearman's rho correlations, no significant correlations between the Target-P1 and inattention or hyperactivity were found (all ps>.05). When conducting Pearsons's correlations and controlling for hyperacuity and depressive symptoms, this did not change the pattern of results (all ps>.05).

Target-P3 Results

For consistency with the adolescent correlational analyses, correlations were conducted to explore the relationships between the Target-P3 (for all emotion conditions), the P3 difference scores (angry-neutral congruent; happy-neutral), and inattention and hyperactivity. When conducting Spearman's rho correlations, there was a close to significance correlation between adult inattentiveness and the angry-neutral P3 amplitude ($r=0.37$, $p=.069$), but no other significant correlations between the Target-P3 and inattention or hyperactivity were found (all $ps>.05$). When conducting Pearsons's correlations and controlling for hyperacuity and depressive symptoms, this did not change the pattern of results (all $ps>.05$).

REFERENCES

Abraham, J. (2004). *Emotional intensity and affect recognition among adults with ADHD*, Southeastern Louisiana University].

Aguert, M., Laval, V., Lacroix, A., Gil, S., & Le Bigot, L. (2013). Inferring emotions from speech prosody: not so easy at age five. *PLoS One*, *8*(12), e83657.

Aldea, R. F. (2013). *Emotion Recognition and Social Functioning in Children with and Without Attention Deficit Hyperactivity Disorder.*, University of Kentucky].

Alho, K., Woods, D. L., & Algazi, A. (1994). Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials. *Psychophysiology*, *31*(5), 469-479.

Anderson, B. A. (2016). Social reward shapes attentional biases. *Cogn Neurosci*, *7*(1-4), 30-36. https://doi.org/10.1080/17588928.2015.1047823

APA. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)* (5th Edition ed.). American Psychiatric Publishing.

Astle, D. E., Holmes, J., Kievit, R., & Gathercole, S. E. (2022). Annual Research Review: The transdiagnostic revolution in neurodevelopmental disorders. *Journal of Child Psychology and Psychiatry*, *63*(4), 397-417.

Audacity Team. (n.d.). [Computer software]. Retrieved March 2020, from https://www.audacityteam.org

Balázs, J., & Keresztény, Á. (2014). Subthreshold attention deficit hyperactivity in children and adolescents: a systematic review. *European child & adolescent psychiatry*, *23*, 393-408.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology*, *70*(3), 614.

Banziger, T., Hosoya, G., & Scherer, K. R. (2015). Path Models of Vocal Emotion Communication. *PLoS One*, *10*(9), e0136675.

Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the Geneva Multimodal expression corpus for experimental research on emotion perception. *Emotion*, *12*(5), 1161.

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.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65-94. https://doi.org/Doi.10.1037/0033-2909.121.1.65

Barkley, R. A. (2015). Emotional dysregulation is a core component of ADHD. In R. A. Barkley (Ed.), Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment (4th ed., pp. 81–115). The Guilford Press.

Barkley, R. A., & Barkley, R. A. (1998). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (Vol. 2). New York: Guilford press.

Barkley, R. A., & Fischer, M. (2010). The unique contribution of emotional impulsiveness to impairment in major life activities in hyperactive children as adults. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(5), 503-513.

Barkley, R. A., & Murphy, K. R. (1998). *Attention-Deficit Hyperactivity Disorder: A Clinical Workbook., 2nd edn*.

Barry, R. J., Clarke, A. R., McCarthy, R., Selikowitz, M., Brown, C. R., & Heaven, P. C. (2009). Event-related potentials in adults with Attention-Deficit/Hyperactivity Disorder: an investigation using an inter-modal auditory/visual oddball task. *International Journal of Psychophysiology*, *71*(2), 124-131.

Barry, R. J., Johnstone, S. J., & Clarke, A. R. (2003). A review of electrophysiology in attention-deficit/hyperactivity disorder: II. Event-related potentials. *Clinical neurophysiology*, *114*(2), 184-198.

Bisch, J., Kreifelts, B., Bretscher, J., Wildgruber, D., Fallgatter, A., & Ethofer, T. (2016). Emotion perception in adult attention-deficit hyperactivity disorder. *J Neural Transm (Vienna)*, *123*(8), 961-970. https://doi.org/10.1007/s00702-016-1513-x

Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*, 187-207.

Boersma & Weenink (2024). Praat: doing phonetics by computer [Computer program]. Version 6.4.16, retrieved 29 July 2024 from http://www.praat.org/

Bogert, B., Numminen-Kontti, T., Gold, B., Sams, M., Numminen, J., Burunat, I., . . . Brattico, E. (2016). Hidden sources of joy, fear, and sadness: Explicit versus implicit neural processing of musical emotions. *Neuropsychologia*, *89*, 393-402. https://doi.org/10.1016/j.neuropsychologia.2016.07.005

Bora, E., & Pantelis, C. (2016). Meta-analysis of social cognition in attention-deficit/hyperactivity disorder (ADHD): comparison with healthy controls and autistic spectrum disorder. *Psychol Med*, *46*(4), 699-716. https://doi.org/10.1017/S0033291715002573

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods*, *1*(2), 97-111. https://doi.org/10.1002/jrsm.12

Borhani, K., & Nejati, V. (2018). Emotional face recognition in individuals withattention-deficit/hyperactivity disorder: a review article. *Developmental Neuropsychology*, *43*(3), 256-277. https://doi.org/10.1080/87565641.2018.1440295

Bostanov, V., & Kotchoubey, B. (2004). Recognition of affective prosody: Continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology*, *41*(2), 259-268.

Bradley, B. P., Mogg, K., & Millar, N. H. (2000). Covert and overt orienting of attention to emotional faces in anxiety. *Cognition & Emotion*, *14*(6), 789-808.

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.

Brennan, A. R., & Arnsten, A. F. (2008). Neuronal mechanisms underlying attention deficit hyperactivity disorder: the influence of arousal on prefrontal cortical function. *Annals of the New York Academy of Sciences*, *1129*(1), 236-245.

British Society of Audiology (2018). *Recommended Procedure. Pure tone air conduction and bone conduction threshold audiometry with and without masking.* Received from https://www.thebsa.org.uk/wp-content/uploads/2023/10/OD104-32-Recommended-Procedure-Pure-Tone-Audiometry-August-2018-FINAL-1.pdf

Brocki, K. C., Eninger, L., Thorell, L. B., & Bohlin, G. (2010). Interrelations between executive function and symptoms of hyperactivity/impulsivity and inattention in preschoolers: A two year longitudinal study. *Journal of abnormal child psychology*, *38*, 163-171.

Brocki, K. C., Forslund, T., Frick, M., & Bohlin, G. (2019). Do individual differences in early affective and cognitive self-regulation predict developmental change in ADHD symptoms from preschool to adolescence? *Journal of Attention Disorders*, *23*(13), 1656-1666.

Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2008). Behold the voice of wrath: cross-modal modulation of visual attention by anger prosody. *Cognition*, *106*(3), 1497-1503. https://doi.org/10.1016/j.cognition.2007.05.011

Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2009). Cross-modal emotional attention: emotional voices modulate early stages of visual processing. *J Cogn Neurosci*, *21*(9), 1670-1679. https://doi.org/10.1162/jocn.2009.21110

Broyd, S. J., Richards, H. J., Helps, S. K., Chronaki, G., Bamford, S., & Sonuga-Barke, E. J. (2012). An electrophysiological monetary incentive delay (e-MID) task: a way to decompose the different components of neural response to positive and negative monetary reinforcement. *Journal of neuroscience methods*, *209*(1), 40-49.

Broyd, S. J., Richards, H. J., Helps, S. K., Chronaki, G., Bamford, S., & Sonuga-Barke, E. J. (2012). An electrophysiological monetary incentive delay (e-MID) task: a way to decompose the different components of neural response to positive and negative monetary reinforcement. *Journal of neuroscience methods*, *209*(1), 40-49.

Bruin, K. J., Kenemans, J. L., Verbaten, M. N., & Van der Heijden, A. H. C. (1998). Localization of spatial attention processes with the aid of a probe technique. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *108*(2), 110-122.

Brunia, C. H., Hackley, S. A., van Boxtel, G. J., Kotani, Y., & Ohgami, Y. (2011). Waiting to perceive: reward or punishment?. *Clinical Neurophysiology*, *122*(5), 858-868.

Bunford, N., Evans, S. W., & Langberg, J. M. (2018). Emotion dysregulation is associated with social impairment among young adolescents with ADHD. *Journal of Attention Disorders, 22*(1), 66-82. https://doi.org/10.1177/1087054714527793

Bunford, N., Evans, S. W., & Wymbs, F. (2015). ADHD and emotion dysregulation among children and adolescents. *Clinical child and family psychology review*, *18*, 185-217.

Burra, N., Kerzel, D., Munoz Tord, D., Grandjean, D., & Ceravolo, L. (2019). Early spatial attention deployment toward and away from aggressive voices. *Soc Cogn Affect Neurosci*, *14*(1), 73-80. https://doi.org/10.1093/scan/nsy100

Bush, G., Valera, E. M., & Seidman, L. J. (2005). Functional neuroimaging of attention-deficit/hyperactivity disorder: a review and suggested future directions. *Biological psychiatry*, *57*(11), 1273-1284.

Cadesky, E. B., Mota, V. L., Mota, V. L., & Schachar, R. J. (2000). Beyond words: how do children with ADHD and/or conduct problems process nonverbal information about affect?. Journal of the American Academy of Child & Adolescent Psychiatry. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(9), 1160-1167.

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal Decoding Skills and Relationship Well-being in Adults. *Journal of Nonverbal behavior*, *23*(1), 91-100.

Cartwright, A. (2024). *Vocal emotional recognition in children with behavioural problems*. Unpublished BSc thesis, University of Central Lancashire.

Charpentier, J., Kovarski, K., Roux, S., Houy-Durand, E., Saby, A., Bonnet-Brilhault, F., . . . Gomot, M. (2018). Brain mechanisms involved in angry prosody change detection in school-age children and adults, revealed by electrophysiology. *Cogn Affect Behav Neurosci*, *18*(4), 748-763. https://doi.org/10.3758/s13415-018-0602-8

Chronaki, G., & Marsh, J. E. (2024). Distraction by vocal anger in children and adolescents with hyperactivity. *Journal of Cognitive Psychology*, 1-16.

Chronaki, G., Benikos, N., Fairchild, G., & Sonuga-Barke, E. J. (2015). Atypical neural responses to vocal anger in attention-deficit/hyperactivity disorder. *J Child Psychol Psychiatry*, *56*(4), 477-487. https://doi.org/10.1111/jcpp.12312.

Chronaki, G., Benikos, N., Soltesz, F., & Sonuga-Barke, E. J. (2019). The reinforcing value of delay escape in attention deficit/hyperactivity disorder: An electrophysiological study. *NeuroImage: Clinical*, *23*, 101917.

Chronaki, G., Broyd, S., Garner, M., Hadwin, J. A., Thompson, M. J., & Sonuga-Barke, E. J. (2012). Isolating N400 as neural marker of vocal anger processing in 6–11-year old children. *Developmental Cognitive Neuroscience*, *2*(2), 268-276.

Chronaki, G., Garner, M., Hadwin, J. A., Thompson, M. J., Chin, C. Y., & Sonuga-Barke, E. J. (2015). Emotion-recognition abilities and behaviour problem dimensions in preschoolers: evidence for a specific role for childhood hyperactivity. *Child Neuropsychol*, *21*(1), 25-40. https://doi.org/10.1080/09297049.2013.863273

Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., & Sonuga-Barke, E. J. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *Br J Dev Psychol*, *33*(2), 218-236.

Chronaki, G., Soltesz, F., Benikos, N., & Sonuga-Barke, E. J. (2017). An electrophysiological investigation of reinforcement effects in attention deficit/hyperactivity disorder: Dissociating cue sensitivity from down-stream effects on target engagement and performance. *Developmental cognitive neuroscience*, *28*, 12-20.

Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The development of cross-cultural recognition of vocal emotion during childhood and adolescence. *Sci Rep*, *8*(1), 8659.

Collin, L., Bindra, J., Raju, M., Gillberg, C., & Minnis, H. (2013). Facial emotion recognition in child psychiatry: a systematic review. *Res Dev Disabil*, *34*(5), 1505-1520. https://doi.org/10.1016/j.ridd.2013.01.008

Compas, B. E., Jaser, S. S., Bettis, A. H., Watson, K. H., Gruhn, M. A., Dunbar, J. P., ... & Thigpen, J. C. (2017). Coping, emotion regulation, and psychopathology in childhood and adolescence: A meta-analysis and narrative review. *Psychological bulletin*, *143*(9), 939.

Cooper, S., Hobson, C. W., & van Goozen, S. H. (2020). Facial emotion recognition in children with externalising behaviours: A systematic review. *Clin Child Psychol Psychiatry*, *25*(4), 1068-1085. https://doi.org/10.1177/1359104520945390

Copas, J., & Shi, J. Q. (2000). Meta-analysis, funnel plots and sensitivity analysis. *Biostatistics*, *1*(3), 247-262. https://doi.org/10.1093/biostatistics/1.3.247

Corbett, B., & Glidden, H. (2000). Processing affective stimuli in children with attention-deficit hyperactivity disorder. *Child Neuropsychol*, *6*(2), 144-155. https://doi.org/10.1076/chin.6.2.144.7056.

Corwin, J. (1994). On measuring discrimination and response bias: Unequal numbers of targets and distractors and two classes of distractors. *Neuropsychology*, *8*(1), 110.

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

Croker, S., & Maratos, F. A. (2011). Visual processing speeds in children. *Child Development Research*, *2011*(1), 450178.

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC medicine*, *11*, 1-8.

D'Agostino, A., Covanti, S., Rossi Monti, M., & Starcevic, V. (2017). Reconsidering emotion dysregulation. *Psychiatric Quarterly*, *88*, 807-825.

Dalton, P., & Lavie, N. (2004). Auditory attentional capture: effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(1), 180.

Danesh, A. A., Lang, D., Kaf, W., Andreassen, W. D., Scott, J., & Eshraghi, A. A. (2015). Tinnitus and hyperacusis in autism spectrum disorders with emphasis on high functioning individuals diagnosed with Asperger's syndrome. *International journal of pediatric otorhinolaryngology*, *79*(10), 1683-1688.

Davis, S. K., Morningstar, M., Dirks, M. A., & Qualter, P. (2020). Ability emotional intelligence: What about recognition of emotion in voices? *Personality and Individual Differences*, *160*, 109938.

De Pascalis, V., Arwari, B., D'Antuono, L., & Cacace, I. (2009). Impulsivity and semantic/emotional processing: An examination of the N400 wave. *Clinical Neurophysiology*, *120*(1), 85-92.

Della Longa, L., Nosarti, C., & Farroni, T. (2022). Emotion recognition in preterm and full-term school-age children. *International Journal of Environmental Research and Public Health*, *19*(11), 6507.

Deng, N., Sun, Y., Chen, X., & Li, W. (2022). How does self name influence the neural processing of emotional prosody? An ERP study. *Psych J*, *11*(1), 30-42. https://doi.org/10.1002/pchj.499

Deng, N., Sun, Y., Chen, X., & Li, W. (2022). How does self name influence the neural processing of emotional prosody? An ERP study. *Psych J*, *11*(1), 30-42. https://doi.org/10.1002/pchj.499

Dien, J., & Santuzzi, A. M. (2005). Application of repeated measures ANOVA to high-density ERP datasets: A review and tutorial. *Event-related potentials: A methods handbook*, 57-82.

Drechsler, R., Brem, S., Brandeis, D., Grünblatt, E., Berger, G., & Walitza, S. (2020). ADHD: Current concepts and treatments in children and adolescents. *Neuropediatrics*, *51*(05), 315-335.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455-463. https://doi.org/10.1111/j.0006-341x.2000.00455.x

Egan, G. J., Brown, R. T., Goonan, L., Goonan, B. T., & Celano, M. (1998). The development of decoding of emotions in children with externalizing behavioral disturbances and their normally developing peers. *Arch Clin Neuropsychol*, *13*(4), 383-396. https://www.ncbi.nlm.nih.gov/pubmed/14590616

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629-634. https://doi.org/10.1136/bmj.315.7109.629

Ekman, P. (1993). Facial Expression and Emotion. *American Psychologist*, *48*(4), 376-379.

Elfenbein, H. A., & Ambady, N. (2002). Is there an in-group advantage in emotion recognition?

Engelien, A., Schulz, M., Ross, B., Arolt, V., & Pantev, C. (2000). A combined functional in vivo measure for primary and secondary auditory cortices. *Hearing research*, *148*(1-2), 153-160.

Ernst, M. (2014). The triadic model perspective for the study of adolescent motivated behavior. *Brain Cogn*, *89*, 104-111. https://doi.org/10.1016/j.bandc.2014.01.006

Ernst, M., Nelson, E. E., Jazbec, S., McClure, E. B., Monk, C. S., Leibenluft, E., ... & Pine, D. S. (2005). Amygdala and nucleus accumbens in responses to receipt and omission of gains in adults and adolescents. *Neuroimage*, *25*(4), 1279-1291.

Faraone, S. V., & Biederman, J. (2005). What is the prevalence of adult ADHD? Results of a population screen of 966 adults. *Journal of attention disorders*, *9*(2), 384-391.

Faraone, S. V., Rostain, A. L., Blader, J., Busch, B., Childress, A. C., Connor, D. F., & Newcorn, J. H. (2019). Practitioner Review: Emotional dysregulation in attention-deficit/hyperactivity disorder–implications for clinical recognition and intervention. *Journal of Child Psychology and Psychiatry*, *60*(2), 133-150.

Fayyad, J., Sampson, N. A., Hwang, I., Adamowski, T., Aguilar-Gaxiola, S., Al-Hamzawi, A., ... & Kessler, R. C. (2017). The descriptive epidemiology of DSM-IV adult ADHD in the world health organization world mental health surveys. *ADHD Attention Deficit and Hyperactivity Disorders*, *9*, 47-65.

Feng, W., Störmer, V. S., Martinez, A., McDonald, J. J., & Hillyard, S. A. (2014). Sounds activate visual cortex and improve visual discrimination. *Journal of Neuroscience*, *34*(29), 9817-9824.

Feng, W., Störmer, V. S., Martinez, A., McDonald, J. J., & Hillyard, S. A. (2014). Sounds activate visual cortex and improve visual discrimination. *Journal of Neuroscience*, *34*(29), 9817-9824.

Filippi, P., Ocklenburg, S., Bowling, D. L., Heege, L., Gunturkun, O., Newen, A., & de Boer, B. (2017). More than words (and faces): evidence for a Stroop effect of prosody in emotion word processing. *Cogn Emot*, *31*(5), 879-891.

Fischer, M., Barkley, R. A., Smallish, L., & Fletcher, K. (2002). Young adult follow-up of hyperactive children: self-reported psychiatric disorders, comorbidity, and the role of childhood conduct problems and teen CD. *Journal of abnormal child psychology*, *30*, 463-475.

Flom, R., & Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Dev Psychol*, *43*(1), 238-252.

Fox, N. A., & Calkins, S. D. (2003). The development of self-control of emotion: Intrinsic and extrinsic influences. *Motivation and emotion*, *27*, 7-26.

Foxe, J. J., & Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans: A framework for defining" early" visual processing. *Experimental brain research*, *142*, 139-150.

Friedman, S. R. (2000). *Emotion recognition, social skills and adult attention-deficit/hyperactivity disorder.* , Wayne State University].

Gädeke, J. C., Föcker, J., & Röder, B. (2013). Is the processing of affective prosody influenced by spatial attention? An ERP study. *BMC neuroscience*, *14*, 1-15.

Galván, A. (2013). The teenage brain: Sensitivity to rewards. *Current directions in psychological science*, *22*(2), 88-93.

Givre, S. J., Schroeder, C. E., & Arezzo, J. C. (1994). Contribution of extrastriate area V4 to the surface-recorded flash VEP in the awake macaque. *Vision research*, *34*(4), 415-428.

Globerson, E., Amir, N., Golan, O., Kishon-Rabin, L., & Lavidor, M. (2013). Psychoacoustic abilities as predictors of vocal emotion recognition. *Attention, Perception, & Psychophysics*, *75*, 1799-1810.

Gobl, C., & Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech communication*, *40*(1-2), 189-212.

Goldberg, D., & Williams, P. (2000). General health questionnaire (GHQ). *Swindon, Wiltshire, UK: nferNelson*.

Goodman, A. & Goodman, R. (2009) 'Strengths and Difficulties Questionnaire as a Dimensional Measure of Child Mental Health', *Journal of the American Academy of Child and Adolescent Psychiatry*, 48(4), pp. 400-403

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry*, *38*(5), 581-586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x
Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry*, *38*(5), 581-586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x
Gootjes, L., Coppens, L. C., Zwaan, R. A., Franken, I. H., & Van Strien, J. W. (2011). Effects of recent word exposure on emotion-word Stroop interference: An ERP study. *International Journal of Psychophysiology*, *79*(3), 356-363.

Grabemann, M., Mette, C., Zimmermann, M., Heinrich, V., Uekermann, J., Wiltfang, J., & Kis, B. (2013). No clear effects of acute tryptophan depletion on processing affective prosody in male adults with ADHD. *Acta psychiatrica scandinavica*, *128*(2), 142-148.

Grandjean, D. (2020). Brain Networks of Emotional Prosody Processing. *Emotion Review*, *13*(1), 34-43. https://doi.org/10.1177/1754073919898522.

Gray, S. M., & Heatherington, L. (2003). The importance of social context in the facilitation of emotional expression in men. *Journal of Social and Clinical Psychology, 22*(3), 294. https://doi.org/10.1521/jscp.22.3.294.22891

Greenbaum, R. L., Stevens, S. A., Nash, K., Koren, G., & Rovet, J. (2009). Social cognitive and emotion processing abilities of children with fetal alcohol spectrum disorders: a comparison with attention deficit hyperactivity disorder. *Alcohol Clin Exp Res*, *33*(10), 1656-1670. https://doi.org/10.1111/j.1530-0277.2009.01003.x

Grosbras, M. H., Ross, P. D., & Belin, P. (2018). Categorical emotion recognition from voice improves during childhood and adolescence. *Sci Rep*, *8*(1), 14791. https://doi.org/10.1038/s41598-018-32868-3

Gross, J. J., & John, O. P. (2012). Emotion regulation questionnaire. *Journal of personality and social psychology*.

Grossmann, T. (2010). The development of emotion perception in face and voice during infancy. *Restorative neurology and neuroscience*, *28*(2), 219-236.

Groves, N. B., Kofler, M. J., Wells, E. L., Day, T. N., & Chan, E. S. M. (2020). An Examination of Relations Among Working Memory, ADHD Symptoms, and Emotion Regulation. *J Abnorm Child Psychol*, *48*(4), 525-537. https://doi.org/10.1007/s10802-019-00612-8

Gullone, E., & Taffe, J. (2012). The emotion regulation questionnaire for children and adolescents (ERQ–CA): a psychometric evaluation. *Psychological assessment*, *24*(2), 409.

Gumora, G., & Arsenio, W. F. (2002). Emotionality, Emotion Regulation, and School Performance in Middle School Children. *Journal of School Psychology*, *40*(5), 395-413.

Gustafsson, B. M., Danielsson, H., Granlund, M., Gustafsson, P. A., & Proczkowska, M. (2018). Hyperactivity precedes conduct problems in preschool children: a longitudinal study. *BJPsych open*, *4*(4), 186-191.

Hall, C. W., Peterson, A. D., Webster, R. E., Bolen, L. M., & Brown, M. B. (1999). Perception of nonverbal social cues by regular education, ADHD, and ADHD/LD students. *Psychology in the Schools*, *36*(6), 505-514.

Hanford, R. B. (2000). *Working memory and nonverbal affect recognition: Differentiation among ADHD subtypes in adults*, Emory University].

Hansen, H., & Hansen, D. (1988). Finding the Face in the Crowd: An Anger Superiority Effect. *Attitudes and Social Cognition*, *54*(6).

Hart, S. J., Green, S. R., Casp, M., & Belger, A. (2010). Emotional priming effects during Stroop task performance. *Neuroimage*, *49*(3), 2662-2670.

Harter, M. R., Aine, C., & Schroeder, C. (1982). Hemispheric differences in the neural processing of stimulus location and type: effects of selective attention on visual evoked potentials. *Neuropsychologia*, *20*(4), 421-438.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Academic Press.

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of nonverbal behavior*, *21*, 241-257.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*(1373), 1257-1270.

Holz, N., Larrouy-Maestri, P., & Poeppel, D. (2021). The paradoxical role of emotional intensity in the perception of vocal affect. *Scientific reports*, *11*(1), 9663.

Hughes, R. W. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal*, *3*(1), 30-41.

Iredale, J. M., Rushby, J. A., McDonald, S., Dimoska-Di Marco, A., & Swift, J. (2013). Emotion in voice matters: neural correlates of emotional prosody perception. *Int J Psychophysiol*, *89*(3), 483-490. https://doi.org/10.1016/j.ijpsycho.2013.06.025.

Irwin, D. A., Knott, J. R., McAdam, D. W., & Rebert, C. S. (1966). Motivational determinants of the “contingent negative variation”. *Electroencephalography and Clinical Neurophysiology*, *21*(6), 538-543.

Ito, M., & Gilbert, C. D. (1999). Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron*, *22*(3), 593-604.

Iyengar, S., & Greenhouse, J. (2009). Sensitivity analysis and diagnostics. *Handbook of research synthesis and meta-analysis*, 417-433.

Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levänen, S., ... & Belliveau, J. W. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences*, *101*(17), 6809-6814.

Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage*, *58*(2), 665-674. https://doi.org/10.1016/j.neuroimage.2011.06.035

Johnstone, S. J., Barry, R. J., & Clarke, A. R. (2013). Ten years on: a follow-up review of ERP research in attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, *124*(4), 644-657.

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of Emotion. In M. Lewis & J. Haviland (Eds.), *The Handbook of Emotion* (Vol. 2, pp. 220-235).

Joutsiniemi, S. L., Ilvonen, T., Sinkkonen, J., Huotilainen, M., Tervaniemi, M., Lehtokoski, A., ... & Näätänen, R. (1998). The mismatch negativity for duration decrement of auditory stimuli in healthy subjects. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *108*(2), 154-159.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code?. *Psychological bulletin*, *129*(5), 770.

Jusyte, A., Gulewitsch, M. D., & Schonenberg, M. (2017). Recognition of peer emotions in children with ADHD: Evidence from an animated facial expressions task. *Psychiatry Research*, *258*, 351-357. https://doi.org/10.1016/j.psychres.2017.08.066

Kan, Y., Mimura, M., Kamijima, K., & Kawamura, M. (2004). Recognition of emotion from moving facial and prosodic stimuli in depressed patients. *Journal of Neurology, Neurosurgery & Psychiatry*, *75*(12), 1667-1671.

Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain research*, *1148*, 138-148.

Kao, C., & Zhang, Y. (2023). Detecting Emotional Prosody in Real Words: Electrophysiological Evidence From a Modified Multifeature Oddball Paradigm. *J Speech Lang Hear Res*, *66*(8), 2988-2998. https://doi.org/10.1044/2023_JSLHR-22-00652

Kao, C., & Zhang, Y. (2023). Detecting Emotional Prosody in Real Words: Electrophysiological Evidence From a Modified Multifeature Oddball Paradigm. *J Speech Lang Hear Res*, *66*(8), 2988-2998. https://doi.org/10.1044/2023_JSLHR-22-00652

Kappenman, E. S., Farrens, J. L., Luck, S. J., & Proudfit, G. H. (2014). 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*, 1368.

Karalunas, S. L., Weigard, A., & Alperin, B. (2020). Emotion-cognition interactions in ADHD: Increased early attention capture and weakened attentional control in emotional contexts. *Biological psychiatry. Cognitive neuroscience and neuroimaging*, *5*(5), 520.

Kattner, F., & Ellermeier, W. (2018). Emotional prosody of task-irrelevant speech interferes with the retention of serial order. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(8), 1303.

Kattner, F., Fischer, M., Caling, A. L., Cremona, S., Ihle, A., Hodgson, T., & Föcker, J. (2024). The disruptive effects of changing-state sound and emotional prosody on verbal short-term memory in blind, visually impaired, and sighted listeners. *Journal of Cognitive Psychology*, *36*(1), 28-41.

Kelly, S. P., Gomez-Ramirez, M., & Foxe, J. J. (2008). Spatial attention modulates initial afferent activity in human primary visual cortex. *Cerebral cortex*, *18*(11), 2629-2636.

Khalfa, S., Dubal, S., Veuillet, E., Perez-Diaz, F., Jouvent, R., & Collet, L. (2002). Psychometric normalization of a hyperacusis questionnaire. *Orl*, *64*(6), 436-442.

Kis, B., Guberina, N., Kraemer, M., Niklewski, F., Dziobek, I., Wiltfang, J., & Abdel-Hamid, M. (2017). Perception of emotional prosody in adults with attention deficit hyperactivity disorder. *Acta Psychiatr Scand*, *135*(6), 506-514. https://doi.org/10.1111/acps.12719.

Kjærvik, S. L., & Bushman, B. J. (2021). The link between narcissism and aggression: A meta-analytic review. *Psychological Bulletin*, *147*(5), 477.

Klasen, M., Chen, Y. H., & Mathiak, K. (2012). Multisensory emotions: perception, combination and underlying neural processes. *Reviews in the Neurosciences*, *23*(4), 381-392. https://doi.org/10.1515/revneuro-2012-0040

Kochel, A., Schongassner, F., Feierl-Gsodam, S., & Schienle, A. (2015). Processing of affective prosody in boys suffering from attention deficit hyperactivity disorder: A near-infrared spectroscopy study. *Soc Neurosci*, *10*(6), 583-591. https://doi.org/10.1080/17470919.2015.1017111

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.

Kutas, M., & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological psychology*, *11*(2), 99-116.

Lange, K. (2010). Can a regular context induce temporal orienting to a target sound?. *International Journal of Psychophysiology*, *78*(3), 231-238.

Laukka, P., Elfenbein, H. A., Thingujam, N. S., Rockstuhl, T., Iraki, F. K., Chui, W., & Althoff, J. (2016). The expression and recognition of emotions in the voice across five nations: A lens model analysis based on acoustic features. *J Pers Soc Psychol*, *111*(5), 686-705.

Lavigne-Cerván, R., Sánchez-Muñoz de León, M., Juárez-Ruiz de Mier, R., Romero-González, M., Gamboa-Ternero, S., Rodríguez-Infante, G., & Romero-Pérez, J. F. (2022). Proposal for an Integrative Cognitive-Emotional Conception of ADHD. *International Journal of Environmental Research and Public Health*, *19*(22), 15421.

Lazzaro, I., Gordon, E., Whitmont, S., Meares, R., & Clarke, S. (2001). The modulation of late component event related potentials by pre-stimulus EEG theta activity in ADHD. *International Journal of Neuroscience*, *107*(3-4), 247-264.

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child development*, *71*(1), 107-118.

Leppanen, J. M., & Hietanen, J. K. (2001). Emotion recognition and social adjustment in school-aged girls and boys. *Scand J Psychol*, *42*(5), 429-435.

Li, X., Li, J., Hu, B., Zhu, J., Zhang, X., Wei, L., ... & Zhang, L. (2018). Attentional bias in MDD: ERP components analysis and classification using a dot-probe task. *Computer methods and programs in biomedicine*, *164*, 169-179.

Lima, C. F., Anikin, A., Monteiro, A. C., Scott, S. K., & Castro, S. L. (2019). Automaticity in the recognition of nonverbal emotional vocalizations. *Emotion*, *19*(2), 219.

Lin, Y., Fan, X., Chen, Y., Zhang, H., Chen, F., Zhang, H., ... & Zhang, Y. (2022). Neurocognitive dynamics of prosodic salience over semantics during explicit and implicit processing of basic emotions in spoken words. *Brain Sciences*, *12*(12), 1706.

Liu, P., & Pell, M. D. (2012). Recognizing vocal emotions in Mandarin Chinese: A validated database of Chinese vocal emotional stimuli. *Behavior Research Methods*, *44*(4), 1042-1051.

López-Martín, S., Albert, J., Fernández-Jaén, A., & Carretié, L. (2015). Emotional response inhibition in children with attention-deficit/hyperactivity disorder: neural and behavioural data. *Psychological medicine*, *45*(10), 2057-2071.

Loytomaki, J., Ohtonen, P., Laakso, M. L., & Huttunen, K. (2020). The role of linguistic and cognitive factors in emotion recognition difficulties in children with ASD, ADHD or DLD. *Int J Lang Commun Disord*, *55*(2), 231-242. https://doi.org/10.1111/1460-6984.12514

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. *Electroencephalography and clinical neurophysiology*, *75*(6), 528-542.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in cognitive sciences*, *4*(11), 432-440.

Ludlow, A. K., Garrood, A., Lawrence, K., & Gutierrez, R. (2014). Emotion Recognition from Dynamic Emotional Displays in Children with Adhd. *Journal of Social and Clinical Psychology*, *33*(5), 413-427. https://doi.org/DOI 10.1521/jscp.2014.33.5.413

Lundin, A., Hallgren, M., Theobald, H., Hellgren, C., & Torgén, M. (2016). Validity of the 12-item version of the General Health Questionnaire in detecting depression in the general population. *Public health*, *136*, 66-74.

Mahajan, Y., & McArthur, G. (2012). Maturation of visual evoked potentials across adolescence. *Brain and Development*, *34*(8), 655-666.

Makowska, Z., Merecz, D., Moscicka, A., & Kolasa, W. (2002). THE VALIDITY OF GENERAL HEALTH QUESTION-NAIRES, GHQ-12 AND GHQ-28, IN MENTAL HEALTH STUDIES OF WORKING PEOPLE. *International Journal of Occupational Medicine & Environmental Health*, *15*(4).

Manassis, K., Tannock, R., Young, A., & Francis-John, S. (2007). Cognition in anxious children with attention deficit hyperactivity disorder: a comparison with clinical and normal children. *Behav Brain Funct*, *3*, 4. https://doi.org/10.1186/1744-9081-3-4

Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human perception and performance*, *17*(4), 1057.

Mangun, G. R., Buonocore, M. H., Girelli, M., & Jha, A. P. (1998). ERP and fMRI measures of visual spatial selective attention. *Human brain mapping*, *6*(5-6), 383-389.

Marco, R., Miranda, A., Schlotz, W., Melia, A., Mulligan, A., Müller, U., ... & Sonuga-Barke, E. J. (2009). Delay and reward choice in ADHD: an experimental test of the role of delay aversion. *Neuropsychology*, *23*(3), 367.

Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *35*(3), 204-214.

Matthews, P. M., & Jezzard, P. (2004). Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery & Psychiatry*, *75*(1), 6-12.

Mauchand, M., & Zhang, S. (2023). Disentangling emotional signals in the brain: an ALE meta-analysis of vocal affect perception. *Cognitive, Affective, & Behavioral Neuroscience*, *23*(1), 17-29.

Maurage, P., Joassin, F., Philippot, P., & Campanella, S. (2007). A validated battery of vocal emotional expressions. *Neuropsychological Trends*, *2*(1), 63-74.

Maxim, L. A., & Nowicki, S. J. (1998). Developmental associations between non-verbal ability and social competence. Unpublished Poster presented at the meetings of the Southeastern Psychological Association meetings, Mobile, AL.

Mayer, A. R., Harrington, D. L., Stephen, J., Adair, J. C., & Lee, R. R. (2007). An event-related fMRI study of exogenous facilitation and inhibition of return in the auditory modality. *Journal of cognitive neuroscience*, *19*(3), 455-467.

Mazza, V., Turatto, M., Rossi, M., & Umiltà, C. (2007). How automatic are audiovisual links in exogenous spatial attention? *Neuropsychologia*, *45*(3), 514-522.

McClanahan, P. (1996). *Social competence correlates of individual differences in nonverbal behavior* (Unpublished master's thesis). Emory University, Atlanta, GA.

McClure, E. B., & Nowicki, J. (2001). Associations between Social Anxiety and Nonverbal Processing Skill in Preadolescent Boys and Girls. *Journal of Nonverbal behavior*, *25*(1), 3-19.

McDonald, J. J., & Ward, L. M. (2000). Involuntary listening aids seeing: Evidence from human electrophysiology. *Psychological Science*, *11*(2), 167-171.

McDonald, J. J., Teder-Salejarvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature*, *407*(6806), 906-908.

McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2005). Neural basis of auditory-induced shifts in visual time-order perception. *Nature neuroscience*, *8*(9), 1197-1202.

McDonald, J. J., Whitman, J. C., Störmer, V. S., Hillyard, S. A., & Mangun, G. R. (2013). Involuntary cross-modal spatial attention influences visual perception. *Cognitive electrophysiology of attention: signals of the mind. Academic, Oxford*, 82-94.

McKay, E., Cornish, K., & Kirk, H. (2023). Impairments in emotion recognition and positive emotion regulation predict social difficulties in adolescent with ADHD. *Clinical Child Psychology and Psychiatry*, *28*(3), 895-908.

McLaughlin, K. A., Hatzenbuehler, M. L., Mennin, D. S., & Nolen-Hoeksema, S. (2011). Emotion dysregulation and adolescent psychopathology: A prospective study. *Behaviour research and therapy*, *49*(9), 544-554.

McLennan, J. D. (2016). Understanding attention deficit hyperactivity disorder as a continuum. *Canadian Family Physician*, *62*(12), 979-982.

Melcher, T., Born, C., & Gruber, O. (2011). How negative affect influences neural control processes underlying the resolution of cognitive interference: an event-related fMRI study. *Neuroscience research*, *70*(4), 415-427.

Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: a review. *Psychological bulletin*, *112*(2), 179.

Miller, M., Hanford, R. B., Fassbender, C., Duke, M., & Schweitzer, J. B. (2011). Affect recognition in adults with ADHD. *J Atten Disord*, *15*(6), 452-460. https://doi.org/10.1177/1087054710368636

Mitchell, J. (1995). *Nonverbal processing ability and social competence in preschool children* (Unpublished Honors thesis). Emory University, Atlanta, GA.

Moehler, E., Brunner, R., & Sharp, C. (2022). Emotional Dysregulation in Children and Adolescents. *Frontiers in psychiatry*, *13*, 883753.

Monk, C. S., Nelson, E. E., McClure, E. B., Mogg, K., Bradley, B. P., Leibenluft, E., . . . Pine, D. S. (2006). Ventrolateral prefrontal cortex activation and attentional bias in response to angry faces in adolescents with generalized anxiety disorder. *Am J Psychiatry*, *163*(6), 1091-1097. https://doi.org/10.1176/ajp.2006.163.6.1091

Morningstar, M., Dirks, M. A., & Huang, S. (2017). Vocal cues underlying youth and adult portrayals of socio-emotional expressions. *Journal of Nonverbal Behavior*, *41*, 155-183.

Morningstar, M., Ly, V. Y., Feldman, L., & Dirks, M. A. (2018). Mid-Adolescents' and Adults' Recognition of Vocal Cues of Emotion and Social Intent: Differences by Expression and Speaker Age. *Journal of Nonverbal Behavior*, *42*(2), 237-251.

Morningstar, M., Mattson, W. I., & Nelson, E. E. (2022). Longitudinal change in neural response to vocal emotion in adolescence. *Soc Cogn Affect Neurosci*, *17*(10), 890-903.

Morningstar, M., Nelson, E. E., & Dirks, M. A. (2018). Maturation of vocal emotion recognition: Insights from the developmental and neuroimaging literature. *Neuroscience and Biobehavioral Reviews*, *90*, 221-230.

Morningstar, M., Nowland, R., Dirks, M. A., & Qualter, P. (2020). Loneliness and the recognition of vocal socioemotional expressions in adolescence. *Cognition and emotion*, *34*(5), 970-976.

Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Dev*, *72*(3), 834-843.

Motter, B. C. (1994). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience*, *14*(4), 2178-2189.

Mueller, S. C., Cromheeke, S., Siugzdaite, R., & Boehler, C. N. (2017). Evidence for the triadic model of adolescent brain development: Cognitive load and task-relevance of emotion differentially affect adolescents and adults. *Developmental Cognitive Neuroscience*, *26*, 91-100.

Mumme, D. L., Fernald, A., & Herrera, C. (1996). Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child development*, *67*(6), 3219-3237.

Murphy, S., Spence, C., & Dalton, P. (2017). Auditory perceptual load: A review. *Hearing Research*, *352*, 40-48.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, *24*(4), 375-425.

Nazari, M. A., Berquin, P., Missonnier, P., Aarabi, A., Debatisse, D., De Broca, A., & Wallois, F. (2010). Visual sensory processing deficit in the occipital region in children with attention-deficit/hyperactivity disorder as revealed by event-related potentials during cued continuous performance test. *Neurophysiologie Clinique/Clinical Neurophysiology*, *40*(3), 137-149.

Nelson, N. L., & Russell, J. A. (2011). Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. *J Exp Child Psychol*, *110*(1), 52-61.

Neves, L., Martins, M., Correia, A. I., Castro, S. L., & Lima, C. F. (2021). Associations between vocal emotion recognition and socio-emotional adjustment in children. *R Soc Open Sci*, *8*(11), 211412.

Nigg, J. T., Goldsmith, H. H., & Sachek, J. (2004). Temperament and attention deficit hyperactivity disorder: The development of a multiple pathway model. *Journal of Clinical Child and Adolescent Psychology*, *33*(1), 42-53.

Nigg, J. T., Stavro, G., Ettenhofer, M., Hambrick, D. Z., Miller, T., & Henderson, J. M. (2005). Executive functions and ADHD in adults: Evidence for selective effects on ADHD symptom domains. *Journal of Abnormal Psychology*, *114*(4), 706-717. https://doi.org/10.1037/0021-843x.114.3.706

Noordermeer, S. D. S., Luman, M., Buitelaar, J. K., Hartman, C. A., Hoekstra, P. J., Franke, B., Faraone, S. V., Heslenfeld, D. J., & Oosterlaan, J. (2020). Neurocognitive Deficits in Attention-Deficit/Hyperactivity Disorder With and Without Comorbid Oppositional Defiant Disorder. *J Atten Disord*, *24*(9), 1317-1329. https://doi.org/10.1177/1087054715606216

Norvilitis, J. M., Casey, R. J., Brooklier, K. M., & Bonello, P. J. (2000). Emotion appraisal in children with attention-deficit/hyperactivity disorder and their parents. *Journal of Attention Disorders*, *4*(1), 15-26.

Novak, K. D., & Foti, D. (2015). Teasing apart the anticipatory and consummatory processing of monetary incentives: An event-related potential study of reward dynamics. *Psychophysiology*, *52*(11), 1470-1482.

Nowicki S, Mitchell J. (1998). Accuracy in identifying affect in child and adult faces and voices and social competence in preschool children. Genet. Soc. Gen. Psychol. Monogr. 124, 39–60.

Nowicki, S., & Duke, M. P. (1992). The Association of Children's Nonverbal Decoding Abilities With Their Popularity, Locus of Control, and Academic Achievement. *The Journal of* Genetic Psychology, 153(4), 385-393.

Nowicki, S., & Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. *Journal of Nonverbal behavior*, *18*, 9-35.

Nowicki, S., Bliwise, N., & Joinson, C. (2019). The association of children's locus of control orientation and emotion recognition abilities at 8 years of age and teachers' ratings of their personal and social difficulties at 10 years. *Journal of Nonverbal Behavior*, 43, 381-396.

Oades, R. D., Dmittmann-Balcar, A., & Zerbin, D. (1997). Development and topography of auditory event-related potentials (ERPs): Mismatch and processing negativity in individuals 8–22 years of age. *Psychophysiology*, *34*(6), 677-693.

Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol Rev*, *108*(3), 483-522. https://doi.org/10.1037/0033-295x.108.3.483.

Orwin, R. G., & Vevea, J. L. (2009). Evaluating coding decisions. *The handbook of research synthesis and meta-analysis, 2*, 177-203.

Orwin, R. G., & Vevea, J. L. (2009). Evaluating coding decisions. *The handbook of research synthesis and meta-analysis, , 2*, 177-203.

Ozdemir, H., & Rezaki, M. (2007). General health questionnaire-12 for the detection of depression. *Turk Psikiyatri Dergisi, 18*(1), 13-21. Retrieved from <Go to ISI>://WOS:000245259500003

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Rev Esp Cardiol (Engl Ed)*, *74*(9), 790-799. https://doi.org/10.1016/j.rec.2021.07.010

Pang, E. W., & Taylor, M. J. (2000). Tracking the development of the N1 from age 3 to adulthood: an examination of speech and non-speech stimuli. *Clinical Neurophysiology*, *111*(3), 388-397.

Papp, S., Tombor, L., Kakuszi, B., Balogh, L., Réthelyi, J. M., Bitter, I., & Czobor, P. (2020). Impaired early information processing in adult ADHD: a high-density ERP study. *BMC psychiatry*, *20*, 1-10.

Park, S., & Chang, H. (2022). Developmental trajectory of inattention and its association with depressive symptoms in adolescence: Peer relationships as a mediator. *Frontiers in psychology*, *12*, 736840.

Parker, M. M., Hunnicutt Hollenbaugh, K. M., & Kelly, C. T. (2021). Exploring the impact of child-centered play therapy for children exhibiting behavioral problems: A meta-analysis. *International Journal of Play Therapy*, *30*(4), 259.

Paulin, J., Andersson, L., & Nordin, S. (2016). Characteristics of hyperacusis in the general population. *Noise and Health*, *18*(83), 178-184.

Paulmann, S., & Kotz, S. A. (2008). An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo-and lexical-sentence context. *Brain and Language*, *105*(1), 59-69.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, *51*, 195-203.

Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. A. (2009). Recognizing Emotions in a Foreign Language. *Journal of Nonverbal Behavior*, *33*(2), 107-120.

Perchet, C., Revol, O., Fourneret, P., Mauguiere, F., & Garcia-Larrea, L. (2001). Attention shifts and anticipatory mechanisms in hyperactive children: an ERP study using the Posner paradigm. *Biol Psychiatry*, *50*(1), 44-57. https://doi.org/10.1016/s0006-3223(00)01119-7.

Peschard, V., Gilboa-Schechtman, E., & Philippot, P. (2017). Selective attention to emotional prosody in social anxiety: a dichotic listening study. *Cognition and Emotion*, *31*(8), 1749-1756.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature reviews neuroscience*, *9*(2), 148-158.

Pfabigan, D. M., Lamplmayr-Kragl, E., Pintzinger, N. M., Sailer, U., & Tran, U. S. (2014). Sex differences in event-related potentials and attentional biases to emotional facial stimuli. *Frontiers in psychology*, *5*, 1477.

Philippot, P., & Feldman, R. S. (1990). Age and social competence in preschoolers' decoding of facial expression. *Br J Soc Psychol*, *29 (Pt 1)*, 43-54.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., ... & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology*, *37*(2), 127-152.

Pierce, J. (2020) [Computer software]. Pavlovia. Retrieved March 2020 from https://www.psychopy.org/

Plećević, V., Djoković, S., & Kovačević, T. (2021). Recognition of emotions and affective attitudes in children with attention-deficit hyperactivity disorder. 78(10). *Vojnosanitetski pregled*, *78*(10).

Polanczyk, G. V., Willcutt, E. G., Salum, G. A., Kieling, C., & Rohde, L. A. (2014). ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. *International journal of epidemiology*, *43*(2), 434-442.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology*, *118*(10), 2128-2148.

Pollak, S. D., & Tolley-Schell, S. A. (2003). Selective attention to facial emotion in physically abused children. *Journal of abnormal psychology*, *112*(3), 323.

Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical neurophysiology*, *111*(2), 220-236.

Posner, J., Nagel, B. J., Maia, T. V., Mechling, A., Oh, M., Wang, Z., & Peterson, B. S. (2011). Abnormal amygdalar activation and connectivity in adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *50*(8), 828-837.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cereb Cortex*, *14*(6), 619-633. https://doi.org/10.1093/cercor/bhh023

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological psychology*, *92*(3), 492-512.

Raiguel, S. E., Lagae, L., Gulyàs, B., & Orban, G. A. (1989). Response latencies of visual cells in macaque areas V1, V2 and V5. *Brain research*, *493*(1), 155-159.

Raij, T., McEvoy, L., Mäkelä, J. P., & Hari, R. (1997). Human auditory cortex is activated by omissions of auditory stimuli. *Brain research*, *745*(1-2), 134-143.

Ralli, M., Romani, M., Zodda, A., Russo, F. Y., Altissimi, G., Orlando, M. P., ... & Turchetta, R. (2020). Hyperacusis in children with attention deficit hyperactivity disorder: a preliminary study. *International Journal of Environmental Research and Public Health*, *17*(9), 3045.

Rapport, L. J., Friedman, S. R., Tzelepis, A., & Van Voorhis, A. (2002). Experienced emotion and affect recognition in adult attention-deficit hyperactivity disorder. *Neuropsychology*, *16*(1), 102-110. https://doi.org/10.1037//0894-4105.16.1.102

Rauss, K., Schwartz, S., & Pourtois, G. (2011). Top-down effects on early visual processing in humans: A predictive coding framework. *Neuroscience & Biobehavioral Reviews*, *35*(5), 1237-1253.

Ray, D. C., Schottelkorb, A., & Tsai, M. H. (2007). Play therapy with children exhibiting symptoms of attention deficit hyperactivity disorder. *International Journal of Play Therapy*, *16*(2), 95.

Riley, R. D., Higgins, J. P., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ*, *342*, d549. https://doi.org/10.1136/bmj.d549

Rodriguez, P. D., & Baylis, G. C. (2007). Activation of brain attention systems in individuals with symptoms of ADHD. *Behavioural Neurology*, *18*(2), 115-130.

Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery & Psychiatry*, *57*(12), 1518-1524.

Ross, E. D., & Monnot, M. (2008). Neurology of affective prosody and its functional–anatomic organization in right hemisphere. *Brain and language*, *104*(1), 51-74.

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *Neuroimage*, *28*(4), 848-858.

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. https://doi.org/10.1016/j.neuropsychologia.2007.12.013

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, *107*(6), 2408-2412.

Sauter, D. A., Panattoni, C., & Happe, F. (2013). Children's recognition of emotions from vocal cues. *Br J Dev Psychol*, *31*(Pt 1), 97-113.

Schaefer, A., & Gray, J. R. (2007). A role for the human amygdala in higher cognition. *Reviews in the Neurosciences*, *18*(5), 355-364.

Scherer, K. R. (1995). Expression of Emotion in Voice and Music. *Journal of Voice*, *9*(3), 235-248.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-cultural psychology*, *32*(1), 76-92.

Schirmer, A., & Adolphs, R. (2017). Emotion Perception from Face, Voice, and Touch: Comparisons and Convergence. *Trends Cogn Sci*, *21*(3), 216-228.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn Sci*, *10*(1), 24-30. https://doi.org/10.1016/j.tics.2005.11.009.

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2005). On the role of attention for the processing of emotions in speech: Sex differences revisited. *Cognitive Brain Research*, *24*(3), 442-452.

Schlittmeier, S. J., Feil, A., Liebl, A., & Hellbrück, J. (2015). The impact of road traffic noise on cognitive performance in attention-based tasks depends on noise level even within moderate-level ranges. *Noise and health*, *17*(76), 148-157.

Schönwiesner, M., Krumbholz, K., Rübsamen, R., Fink, G. R., & von Cramon, D. Y. (2007). Hemispheric asymmetry for auditory processing in the human auditory brain stem, thalamus, and cortex. *Cerebral Cortex*, *17*(2), 492-499.

Schultz, D., Izard, C. E., Ackerman, B. P., & Youngstrom, E. A. (2001). Emotion knowledge in economically disadvantaged children: Self-regulatory antecedents and relations to social difficulties and withdrawal. *Development and psychopathology, 13*(1), 53-67.

Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Curr Opin Neurobiol*, *14*(2), 139-147. https://doi.org/10.1016/j.conb.2004.03.017

Schweitzer, J. B., & Sulzer-Azaroff, B. (1995). Self-control in boys with attention deficit hyperactivity disorder: Effects of added stimulation and time. *Journal of Child Psychology and Psychiatry*, *36*(4), 671-686.

Sells, R. C., Liversedge, S. P., & Chronaki, G. (2024). Vocal emotion recognition in attention-deficit hyperactivity disorder: a meta-analysis. *Cognition and Emotion*, *38*(1), 23-43.

Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology*, *23*(6), 695-703.

Seymour, K. E., Chronis-Tuscano, A., Iwamoto, D. K., Kurdziel, G., & MacPherson, L. (2014). Emotion regulation mediates the association between ADHD and depressive symptoms in a community sample of youth. *Journal of abnormal child psychology*, *42*, 611-621.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *J Am Acad Child Adolesc Psychiatry*, *39*(1), 28-38. https://doi.org/10.1097/00004583-200001000-00014.

Shaw, P., Gilliam, M., Liverpool, M., Weddle, C., Malek, M., Sharp, W., ... & Giedd, J. (2011). Cortical development in typically developing children with symptoms of hyperactivity and impulsivity: support for a dimensional view of attention deficit hyperactivity disorder. *American Journal of Psychiatry*, *168*(2), 143-151.

Shaw, P., Stringaris, A., Nigg, J., & Leibenluft, E. (2014). Emotion dysregulation in attention deficit hyperactivity disorder. *Am J Psychiatry*, *171*(3), 276-293. https://doi.org/10.1176/appi.ajp.2013.13070966

Silverman, M. H., Jedd, K., & Luciana, M. (2015). Neural networks involved in adolescent reward processing: an activation likelihood estimation meta-analysis of functional neuroimaging studies. *NeuroImage*, *122*, 427-439.

Sinzig, J., Morsch, D., & Lehmkuhl, G. (2008). Do hyperactivity, impulsivity and inattention have an impact on the ability of facial affect recognition in children with autism and ADHD?. *European child & adolescent psychiatry*, *17*, 63-72.

Sjöwall, D., Bohlin, G., Rydell, A. M., & Thorell, L. B. (2017). Neuropsychological deficits in preschool as predictors of ADHD symptoms and academic achievement in late adolescence. *Child Neuropsychology*, *23*(1), 111-128.

Sjowall, D., Roth, L., Lindqvist, S., & Thorell, L. B. (2013). Multiple deficits in ADHD: executive dysfunction, delay aversion, reaction time variability, and emotional deficits. *J Child Psychol Psychiatry*, *54*(6), 619-627. https://doi.org/10.1111/jcpp.12006

Slotnick, S. D. (2020). The experimental parameters that affect attentional modulation of the ERP C1 component. *The Cognitive Neuroscience of Attention*, 41-50.

Sokolova, E., Oerlemans, A. M., Rommelse, N. N., Groot, P., Hartman, C. A., Glennon, J. C., ... & Buitelaar, J. K. (2017). A causal and mediation analysis of the comorbidity between attention deficit hyperactivity disorder (ADHD) and autism spectrum disorder (ASD). *Journal of autism and developmental disorders*, *47*, 1595-1604.

Sonuga-Barke, E. J. (1994). On dysfunction and function in psychological theories of childhood disorder. *Child Psychology & Psychiatry & Allied Disciplines*.

Sonuga-Barke, E. J. (2005). Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biol Psychiatry*, *57*(11), 1231-1238. https://doi.org/10.1016/j.biopsych.2004.09.008

Sonuga-Barke, E. J. S., & Fairchild, G. (2012). Neuroeconomics of attention-deficit/hyperactivity disorder: Differential influences of medial, dorsal, and ventral prefrontal brain networks on suboptimal decision making? *Biological Psychiatry*, *72*(2), 126–133. https://doi.org/10.1016/j.biopsych.2012.04.004

Sonuga-Barke, E. J., Becker, S. P., Bölte, S., Castellanos, F. X., Franke, B., Newcorn, J. H., ... & Simonoff, E. (2023). Annual Research Review: Perspectives on progress in ADHD science–from characterization to cause. *Journal of child psychology and psychiatry*, *64*(4), 506-532.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & psychophysics*, *59*(1), 1-22.

Spence, C., & McDonald, J. (2004). The Cross-Modal Consequences of the Exogenous Spatial Orienting of Attention.

Steinberg, E. A., & Drabick, D. A. (2015). A developmental psychopathology perspective on ADHD and comorbid conditions: The role of emotion regulation. *Child Psychiatry & Human Development*, *46*, 951-966.

Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors.

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rucker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. (2011).

Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, *343*, d4002. https://doi.org/10.1136/bmj.d4002

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.

Stifter, C. A., & Fox, N. A. (1986). Preschool children's ability to identify and label emotions. *Journal of Nonverbal Behavior*, *10*(4), 255-266.

Stolarova, M., Keil, A., & Moratti, S. (2006). Modulation of the C1 visual event-related component by conditioned stimuli: evidence for sensory plasticity in early affective perception. *Cerebral cortex*, *16*(6), 876-887.

Störmer, V. S., McDonald, J. J., & Hillyard, S. A. (2009). Cross-modal cueing of attention alters appearance and early cortical processing of visual stimuli. *Proceedings of the National Academy of Sciences*, *106*(52), 22456-22461.

Stuss, D. T., Alexander, M. P., Floden, D., Binns, M. A., Levine, B., McIntosh, A. R., ... & Hevenor, S. J. (2002). Fractionation and localization of distinct frontal lobe processes: Evidence from focal lesions in humans. *Principles of frontal lobe function*, 392-407.

Suslow, T., Lemster, A., Koelkebeck, K., & Kersting, A. (2023). Interpersonal problems and recognition of facial emotions in healthy individuals. *Front Psychiatry*, *14*, 1139051.

Szuromi, B., Czobor, P., Komlósi, S., & Bitter, I. (2011). P300 deficits in adults with attention deficit hyperactivity disorder: a meta-analysis. *Psychological medicine*, *41*(7), 1529-1538.

Tait, R. J., French, D. J., & Hulse, G. K. (2003). Validity and psychometric properties of the General Health Questionnaire-12 in young Australian adolescents. *Australian & New Zealand Journal of Psychiatry*, *37*(3), 374-381.

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. . *Journal of Developmental and Life-Course Criminology*, *2*(1), 85-112.

Taylor, E., Chadwick, O., Heptinstall, E., & Danckaerts, M. (1996). Hyperactivity and conduct problems as risk factors for adolescent development. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(9), 1213-1226.

Techentin, C. (2009). *Integrating Semantic and Emotional Information in Boys with and Without ADHD During an Auditory Stroop-like Task*, University of New Brunswick].

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. *Biol Psychol*, *79*(2), 216-222. https://doi.org/10.1016/j.biopsycho.2008.05.004

Thomas, R., Sanders, S., Doust, J., Beller, E., & Glasziou, P. (2015). Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. *Pediatrics*, *135*(4), e994-e1001.

Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. *The development of emotion regulation*.

Trentacosta, C. J., & Fine, S. E. (2010). Emotion Knowledge, Social Competence, and Behavior Problems in Childhood and Adolescence: A Meta-analytic Review. *Social Development*, *19*(1), 1-29.

Tripp, G., & Alsop, B. (2001). Sensitivity to reward delay in children with attention deficit hyperactivity disorder (ADHD). *Journal of Child Psychology and Psychiatry*, *42*(5), 691-698.

Uekermann, J., Kraemer, M., Abdel-Hamid, M., Schimmelmann, B. G., Hebebrand, J., Daum, I., Wiltfang, J., & Kis, B. (2010). Social cognition in attention-deficit hyperactivity disorder (ADHD). *Neurosci Biobehav Rev*, *34*(5), 734-743. https://doi.org/10.1016/j.neubiorev.2009.10.009

Van Dessel, J., Sonuga-Barke, E., Moerkerke, M., Van der Oord, S., Lemiere, J., Morsink, S., & Danckaerts, M. (2020). The amygdala in adolescents with attention-deficit/hyperactivity disorder: structural and functional correlates of delay aversion. *The World Journal of Biological Psychiatry*.

Van Rooijen, R., Ploeger, A., & Kret, M. E. (2017). The dot-probe task to measure emotional attention: A suitable measure in comparative studies?. *Psychonomic bulletin & review*, *24*, 1686-1717.

Vidal-Ribas, P., Brotman, M. A., Salum, G. A., Kaiser, A., Meffert, L., Pine, D. S., ... & Stringaris, A. (2018). Deficits in emotion recognition are associated with depressive symptoms in youth with disruptive mood dysregulation disorder. *Depression and anxiety*, *35*(12), 1207-1217.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of statistical software*, *36*(3), 1-48.

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, *37*(2), 190-203.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn Sci*, *9*(12), 585-594. https://doi.org/10.1016/j.tics.2005.10.011.

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia*, *45*(1), 174-194.

Waddington, F., Hartman, C., de Bruijn, Y., Lappenschaar, M., Oerlemans, A., Buitelaar, J., & Rommelse, N. (2018). Visual and auditory emotion recognition problems as familial cross-disorder phenomenon in ASD and ADHD. *European Neuropsychopharmacology,*, *28*(9), 994-1005.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *nature*, *203*(4943), 380-384.

Wang, Y., Tang, Z., Zhang, X., & Yang, L. (2022). Auditory and cross-modal attentional bias toward positive natural sounds: Behavioral and ERP evidence. *Front Hum Neurosci*, *16*, 949655. https://doi.org/10.3389/fnhum.2022.949655.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence--Second Edition (WASI-II)* [Database record]. APA PsycTests.https://doi.org/10.1037/t15171-000

Westfall, J. (2015). PANGEA: Power analysis for general ANOVA designs. *Unpublished manuscript. Available at http://jakewestfall. org/publications/pangea. pdf*, *4*.

Wild-Wall, N., & Falkenstein, M. (2010). Age-dependent impairment of auditory processing under spatially focused and divided attention: an electrophysiological study. *Biological psychology*, *83*(1), 27-36.

Williams, L. M., Hermens, D. F., Palmer, D., Kohn, M., Clarke, S., Keage, H., ... & Gordon, E. (2008). Misinterpreting emotional expressions in attention-deficit/hyperactivity disorder: evidence for a neural marker and stimulant effects. *Biological psychiatry*, *63*(10), 917-926.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences*, *90*(18), 8722-8726.

Woldorff, M. G., Tempelmann, C., Fell, J., Tegeler, C., Gaschler-Markefski, B., Hinrichs, H., ... & Scheich, H. (1999). Lateralized auditory spatial perception and the contralaterality of cortical processing as studied with functional magnetic resonance imaging and magnetoencephalography. *Human brain mapping*, *7*(1), 49-66.

Young, A. W., Fruhholz, S., & Schweinberger, S. R. (2020). Face and Voice Perception: Understanding Commonalities and Differences. *Trends Cogn Sci*, *24*(5), 398-410.

Zimmer, U., Wendt, M., & Pacharra, M. (2022). Enhancing allocation of visual attention with emotional cues presented in two sensory modalities. *Behavioral and Brain Functions*, *18*(1), 10.

Zlotnik, S., Attias, J., Pratt, H., & Engel-Yeger, B. (2018). Neurophysiological manifestations of auditory hypersensitivity correlate with daily life experiences. *Neuroscience and Medicine*, *9*(01), 29.

Zuberer, A., Schwarz, L., Kreifelts, B., Wildgruber, D., Erb, M., Fallgatter, A., Scheffler, K., & Ethofer, T. (2020). Neural Basis of Impaired Emotion Recognition in Adult Attention-Deficit/Hyperactivity Disorder. *Biol Psychiatry Cogn Neurosci Neuroimaging*. https://doi.org/10.1016/j.bpsc.2020.11.013

Zupan, B. (2015). Recognition of high and low intensity facial and vocal expressions of emotion by children and adults. *Journal of Social Sciences and Humanities*, *1*(4), 332-344.

Zupan, B., & Eskritt, M. (2024). Facial and vocal emotion recognition in adolescence: A systematic review. *Adolescent Research Review*, *9*(2), 253-277.