A Signal Detection Approach to the Perception of Affective Prosody in Anxious Individuals: A Developmental Study

By

MEGAN HUMPHREY

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Masters of Science in Psychology

Victoria University
2009
Abstract
The present study used a Signal Detection approach to the study of prosody perception in children and adults who self-reported high levels of anxiety. Seventy-one children aged eight and nine years, and 85 adults listened to filtered speech and were required to discriminate angry, fearful and happy tones of voice. Anxiety levels were not associated with perception of affective prosody in adults. Levels of anxiety were related to children’s criterion but not sensitivity to prosody. Highly anxious children were significantly more liberal in reporting fearful prosody compared to low anxious children. Analyses of total responses suggest that this criterion is reflective of an interpretation bias as opposed to a response bias. Given that the interpretation bias was observed in children and not adults, it is possible that the bias may mark a vulnerability to develop further anxiety. This is consistent with previous experimental findings in other modalities as well as integrative models of anxiety development that identify such cognitive biases as predisposing factors. Furthermore, regardless of anxiety level, children were comparable to adults in their accuracy for fearful prosody, yet were significantly poorer than adults in their accuracy for angry and happy prosody. This suggests that fear may be one of the first emotions children learn to identify.
Acknowledgements

I would firstly like to thank Dr Gina Grimshaw for supervising this thesis. Thank you Gina for your commitment, support and feedback throughout the year. Your assistance in helping me on my journey to completing this thesis is greatly appreciated.

Thank you to Desiree Cheer for providing her voice for the recording of the stimuli. I would like to acknowledge Hazel Godfrey. Thank your Hazel for your assistance with the acoustic analysis of sound files. And to my fellow students who provided support and feedback throughout the year.

Considerable thanks also to the principals and teachers of the three schools involved for kindly allowing me to come into the school and conduct this research. Thanks also to the parents/caregivers, children and university students for making this research possible through their participation.
Table of Contents

Abstract 2
Acknowledgements 3
Table of contents 4

Introduction 7
Anxiety and the processing of emotional expressions 7
The processing of vocal expressions of emotion: Affective prosody and anxiety 8
Signal Detection Theory 12
Why use the developmental approach? 16
Emotion recognition impairment in anxious children 20
Prosody perception in anxious children 21
Signal Detection Theory paradigms with children 21
The present study 22

Method 24
Participants 24
Materials and stimuli 24
Procedure 28
Design 30
Statistical analysis 31

Results 32
Participant characteristics 32
Sensitivity 33
Criterion 35
False Alarm analysis for specific emotional comparisons 38

Discussion 42
Anxiety-related findings – Criterion 43
A Comparison with child prosody literature 43
How does this fit with other SDT findings? 44
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence of a predisposing factor</td>
<td>45</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>48</td>
</tr>
<tr>
<td>Role of depression</td>
<td>51</td>
</tr>
<tr>
<td>Developmental findings</td>
<td>52</td>
</tr>
<tr>
<td>General discussion</td>
<td>56</td>
</tr>
</tbody>
</table>

References 57

Appendix A: Letter to the Principal 69
Appendix B: Information Sheet and Consent to Parents and Caregivers 72
Appendix C: Information Sheet and Consent to IPRP Students 75
Appendix D: Debrief form for IPRP students 77
Appendix E: Debrief form for Parents and Caregivers 79
List of Tables and Figures

Table 1. Acoustic Parameters for Angry, Fearful, Happy and Neutral Prosodic Sentences………………………………………………………………………………………………28

Table 2. Participant Demographics, Anxiety and Depression Scores (Standard Deviations) anxious adults and children……………………………………………………………… …………32

Table 3. Mean sensitivity scores (d’) and (standard deviations) for low, high, and combined anxious children and adults……………………………………………………………… …………34

Table 4. Mean criterion (c) and (standard deviations) for low, high, and combined anxious adults and children…………………………………………………………………………………… 36

Table 5. False Alarms and standard deviations as a function of target and distracter for low, high and combined anxious adults……………………………………………………………… …………39

Table 6. False Alarms and standard deviations as a function of target and distracter for low, high and combined anxious children……………………………………………………………………………….41

Figure 1. A comparison of children’s and adults’ ability to discriminate different emotional prosodies. …………………………………………………………………………………………………35

Figure 2. Response biases to different emotional prosodies in low and high anxious adults and children. …………………………………………………………………………………………………37
A Signal Detection Approach to the Perception of Affective Prosody in Anxious Individuals: a Developmental Study

Communicating and understanding others’ emotional states is an essential skill in establishing and maintaining social and interpersonal relationships across the lifespan. In recent times the processing of emotion has received increasing attention in the exploration of emotional disorders. Research suggests that individuals who suffer from both clinical and sub clinical levels of anxiety process emotional stimuli in a quantifiably different way to individuals with low levels of anxiety. The majority of research has explored the recognition of facial expressions of emotions; much less is known about the relationship between anxiety and the recognition of vocal emotional expression. The present study used a developmental approach to explore anxiety-related differences in the interpretation of emotional tone of voice (affective prosody).

Anxiety and the processing of emotional expressions

Both enhanced and diminished processing of facial emotional expressions in highly anxious individuals have been reported. It appears that the methodology and more importantly the analyses used, play an important role in the effects observed. The majority of studies that have used basic emotion recognition or identification tasks report that anxious individuals are better than non-anxious individuals at recognising negative (usually fearful) facial expressions. When adults are presented with a series of images of emotional facial expressions and asked to choose the emotion depicted, high-trait anxiety adults are more accurate at recognizing fearful facial expressions than low trait anxiety adults (Richards, French, Calder, Webb, Fox & Young, 2002; Surcinelli, Codispoti, Montebarocci, Rossi & Baldaro, 2006). There is some evidence of enhanced recognition of negative emotions other than fear. For example, Joormann and Gotlib (2006) found that adults with social phobia showed enhanced identification of angry (but not fearful) faces when compared to controls.
Additionally, there is some evidence that individuals high in anxiety are not only superior at detecting negative emotions such as fear, but also **poorer** at detecting positive emotions such as happiness (Silvia, Allan, Beauchamp, Maschauer, & Workman, 2006).

The findings in this area are mixed, however. There is some evidence that anxiety is associated with poorer identification of negative emotions. On an emotion recognition task, adults with a diagnosis of panic disorder were significantly poorer than controls at accurately recognising facial expressions of sadness and anger (Kessler, Roth, Wietersheim, Deighton & Traue, 2007). Other studies have reported no significant differences between high and low anxious adults in their ability to accurately identify emotional facial expressions depicting different emotional categories (Cooper, Rowe & Penton-Voak, 2008; Philippot & Douilliez, 2005; Schofield, Coles & Gibb, 2007).

**The processing of vocal expressions of emotion: Affective prosody and anxiety**

Taken together, the literature to date suggests that the processing of emotional facial expressions is susceptible to individual differences in anxiety. Yet much less is known about the recognition of emotions in other modalities including vocal expression, and the lack of studies utilising auditory stimuli to examine cognitive biases in anxiety has been noted in the literature (Heinrichs & Hofmann, 2001). The question of whether the processing of emotional tone of voice is also disrupted in anxious individuals is important for several reasons. Firstly, the emotional tone of a speaker’s voice is a critical piece of information that we use in everyday life to establish meaning of speech. Children are known to use prosodic information from infancy onward to help them to interpret linguistic material and emotional tone of voice continues to be an integral component of language understanding throughout the life span. Similar to the recognition of facial expressions of emotions, understanding emotional tone of voice is a critical skill that aids daily social interactions. Disruptions in this skill likely lead to disruptions in interpersonal functioning. Secondly, neuropsychological research suggests that
Anxiety-related interpretation bias

the underlying neural substrates involved in the recognition of facial and vocal expressions of emotions are shared (Zupan, Neumann, Babbage, & Willer, 2009). Therefore one could predict that a population that demonstrates atypical recognition of facial expressions may also show atypical recognition of emotional prosody.

Prosody is a non-lexical component of speech and consists of linguistic prosody and affective prosody. Linguistic prosody helps to produce meaning by differentiating lexical components through stress. For example, the sentences “are you coming tomorrow?” and “are you coming tomorrow?” can denote very different meanings when the word that is stressed is changed. Emotional or affective prosody refers to the variation of pitch, rhythm, intensity, and speech rate used to convey emotion through speech. For example, vocal expressions of discrete emotions such as, happiness or fearfulness correspond with a distinct idiosyncratic pattern of the prosodic components that are characteristic of vocal expressions of that emotion (Banse & Scherer, 1996; Pell, 2001). The production of these prosodic components involves both voluntary and involuntary processes; involuntary processes are related to physiological changes associated with emotional state, and voluntary processes reflect the manipulation of the voice to communicate emotional information.

The past five years have seen rapid development in the study of affective prosody, and recently the role of affective prosodic interpretation in clinical disorders has been examined. The ability to interpret a speaker’s emotional state through their tone of voice has been found to be affected in a variety of mental health disorders, such as schizophrenia (Bozikas et al. 2006) and remitted bipolar disorder (Bozikas et al. 2007). Only two studies to date have explored prosody perception in anxious adults. In the first study, Quadflieg, Wendt, Mohr, Miltner, and Straube (2007) examined prosody perception in 15 participants with generalised social phobia, and 15 healthy controls. Participants listened to 384 pseudo-words without semantic meaning, spoken in a happy, sad, fearful, angry, disgusted, or neutral tone.
of voice, taken from the “Magdeburger Prosodie-Korpus”. Participants indicated on a questionnaire which of the six basic emotional categories the word expressed. Consistent with the recognition of facial expressions, highly anxious adults correctly identified more fearful and sad emotional prosody. They were also significantly poorer at identifying happy prosody compared to controls. The groups did not differ in their ability to identify neutral or disgust prosody. Participants also completed a self-report measure of depression and interestingly, the effect of anxiety on the perception of sad and fearful prosody remained when depression was controlled for, but the happy effect was eliminated. This suggests that depression was more strongly associated with the poorer accuracy for happy voices, but anxiety was associated with the enhanced performance for fearful and sad prosodies. Additionally, depression was associated with a poor ability to detect anger prosody.

More recently Freeman, Hart, Kimbrell and Ross (2009) examined prosody perception in adults with combat-related post traumatic stress disorder (PTSD). Eleven male adults with a diagnosis of PTSD were administered the comprehension component of the Aprosdia Battery, which required them to identify six emotional prosodies; happy, sad, disinterested, neutral, surprised and angry. Adults with PTSD performed significantly worse on the emotion identification task than the twelve controls. Unfortunately the analysis did not examine differences between the different discrete emotions, but rather reported a global deficit. It is important to note however, that although Freeman et al. excluded participants with any neurological disease including head injury including loss of consciousness, stroke, history of alcoholism and any neurodegenerative diseases, comparisons between the population examined by Freeman et al. and other anxiety disordered adults should be made with caution. PTSD has distinctively different symptoms and a markedly different aetiology than other anxiety disorders, particularly chronic life-long combat-related PTSD (as assessed by Freeman et al. 2009) which is often associated with other cognitive impairments (Uddo,
Vasterling, Brailey, & Sutker, 1993; Vasterling et al., 2002). The preliminary findings by Quadflieg and colleagues (2007) and Freeman and colleagues (2009) suggest that clinical levels of anxiety (in this case social phobia and PTSD) are associated with an altered capacity to interpret emotional prosody. However it is not yet known if sub-clinical levels of anxiety are also associated with changes in this critical skill.

The anxiety-related prosodic findings found in socially phobic patients (Quadflieg et al., 2007) is consistent with the general picture emerging from the cloudy research on the influence of anxiety on recognition of facial expressions of emotion. That is, anxiety appears to be associated with an enhanced ability to identify or recognise fear in others, and perhaps other negative emotions such as anger or sadness. Additionally, anxiety may be associated with a decreased ability to identify or recognise positive emotions such as happiness, however it is possible that this effect is moderated by co-morbid depression. As noted earlier, discrepant findings have been reported in the prosodic (e.g. Freeman et al., 2009) and facial expression literature (Kessler at al., 2009). One possible explanation for the inconsistency in findings of anxiety’s influence on emotional processing is that the reported studies have not systematically examined whether enhanced recognition of negative emotional expressions (that is, either facial or vocal) primarily reflects a difference in the sensitivity to discriminate discrete emotions in others or a response bias, i.e. a tendency to report particular emotions.

Most of the studies described above that report anxiety-related differences in emotion perception utilise emotion recognition or identification tasks. For example participants are required to listen to a sentence or word and choose from a list of possible emotions which best describes what the actor is feeling. With the given paradigm however, the interpretation of results can be difficult. Consider the following example; an individual is presented with 30 randomly-assigned sentences; 10 happy, 10 angry and 10 fearful, and they are asked to indicate which emotion is expressed when presented. If they respond “happy” to all 30 trials,
they would score 100% accuracy for the identification of happy tone of voice, and 0% accuracy for the identification of angry and fearful tones. This score does not in fact reflect the person’s ability to discriminate happy tone of voice, yet in the given paradigm, this can be the only conclusion that the researcher could draw. This is often the case when looking at accuracy scores or percentage of correctly identified stimuli. Yet when looking at the pattern of responding made by this participant it is clear that his/her score in fact reflects a liberal response criterion for happy prosody rather than an enhanced ability to detect happy prosody.

One way to get around this problem is to present the participant with a target emotion, e.g. ‘happy’ amongst other distracter emotions with the task to correctly detect the happy tone of voice. That is, for every sentence that is presented, the participant needs to make a decision, (i.e. Is this happy? or not happy?). This methodology allows the responses to be analyzed with signal detection theory, to derive a score of sensitivity, which reflects the person’s ability to detect the happy voice, and a measure of response criterion, which reflects the person’s likelihood to make a ‘happy’ response under conditions of uncertainty.

*Signal Detection Theory*

The current study used a signal detection approach (SDT: Green & Swets, 1966; Macmillan & Creelman, 2005) to distinguish between an individual’s ability to detect a given affective prosody (e.g. fear), and their tendency to label a voice of any given prosody as fearful. Specifically, this research determines whether anxious individuals are better at detecting fear, or anger, in another’s voice, or are simply more likely to report that all voices sound threatening regardless of the actual tone of voice. Similarly, for positive stimuli, are they really poor at detecting happy prosody, or simply less likely to report it?

Interestingly, SDT as a theoretical framework and the corresponding experimental paradigm have seldom been applied to the study of anxiety disorders, and never been applied to the perception of affective prosody, even though SDT is the generally preferred method
Anxiety-related interpretation bias

when perceptual thresholds are sought. According to SDT (Macmillan & Creelman, 2005) one’s task of detecting a stimulus involves two independent processes. Firstly, performance depends on sensitivity ($d’$): the individual’s ability to discriminate the target stimulus from the absence of it or from other stimuli. Secondly, an individual’s performance also depends on their response criterion ($c$): their general tendency to respond with “yes” or “no” when asked if the target stimulus is present. Describing their accuracy in terms of $d’$ and $c$ will better reflect how individuals are processing these emotional stimuli. The exact nature of the mechanisms underlying the anxiety-related effects reported in previous studies is still largely unknown. Processing differences associated with anxiety may stem from atypical perceptual sensitivity to emotional stimuli or from a tendency to interpret emotional stimuli as negative.

Some studies described earlier in the facial expression literature claim that the observed anxiety-related effects for negative emotional stimuli are due to sensitivity and not criterion/response bias (e.g. Surcinelli et al., 2006; Joorman & Gotlib, 2006; Richards et al., 2002). From the reported method and analyses, these studies used paradigms that did allow for the control of response criterion (e.g. two alternative forced choice), yet they did not allow for the independent measurement of response bias and sensitivity. In fact when the method and analyses of these studies are explored further, results suggest that differences may be in bias, rather than sensitivity. For example, Richards et al. report that high-trait anxiety adults demonstrate an enhanced sensitivity for identifying fearful faces. However, their results suggest that anxious adults were more likely to categorise all faces as fearful, indicating a response bias, rather than enhanced sensitivity. This ambiguity in interpretation highlights the need for signal detection methodology to allow for extraction of measures of response criterion and sensitivity.

Across a variety of anxiety subtypes (spider phobia, trait anxiety, and social anxiety) and stimulus modalities (lexical and facial expressions), the majority of studies that have used
Anxiety-related interpretation bias

A signal detection methodology report anxiety-related differences in response criterion, and less often in sensitivity. Interestingly, in contrast to the non-SDT studies that report anxiety is associated with an *enhanced* ability to recognise or identify negative emotions, the SDT studies that have found differences in sensitivity report the opposite pattern; high levels of anxiety are associated with *poorer* ability to discriminate fear in others (Garner, Baldwin, Bradley & Mogg, 2009; Frenkel, Lamy, Algom & Bar-Haim, 2008).

Of these studies, some report anxiety related differences in both sensitivity and criterion. For example, anxious individuals displayed poorer perceptual sensitivity for both, mildly threatening faces and moderately happy faces as well as displaying a significantly less conservative criterion when judging mild and moderate degrees of fearful facial expression (Frenkel et al., 2008). That is, the non-anxious participants displayed a tendency to judge mildly and moderately fearful facial expressions as ‘less fearful’ whereas the anxious group did not show this trend. The authors suggest that the observed normative tendency to under-report fear in facial expressions is likely to be an effective strategy in most daily situations, as it allows one to ignore irrelevant or non-imminent threat. Additionally, differences in response criterion to positive emotional faces were also observed, such that anxious individuals displayed a conservative criterion, and non-anxious individuals displayed a liberal criterion to moderately happy faces (Frenkel et al., 2008). Together, the findings of Frenkel et al. suggest that compared to non-anxious individuals, anxious adults demonstrate two distinct biases. First, they lack the normative bias to interpret mild and moderate degrees of fearful stimuli as ‘less threatening’, and they lack the bias to interpret most positive stimuli as ‘more positive’. The authors describe this as a ‘double setback’ for anxious individuals.

Conflicting sensitivity results (but similar criterion results) were reported in a study that applied the SDT paradigm to systematically measure the influence of anxiety on the perception of positive and negative emotional social behaviours (Veljaca & Rapee, 1998).
Low and high socially anxious adults were asked to give a short speech in front of an audience who were trained to provide positive and negative cues of evaluation while the participant was speaking. When participants were required to detect these behaviours, highly anxious adults displayed a higher sensitivity to negative audience behaviours (e.g. frowns) than positive audience behaviours (e.g. smiling), whereas low anxious adults showed the opposite pattern; that is they were more sensitive to positive feedback. Consistent with the findings by Frenkel et al. (2008), highly anxious adults also displayed a more liberal response criterion to report negative audience behaviours than the low anxious adults.

Other studies report anxiety-related differences solely in response criterion. Anxious individuals demonstrate a more liberal criterion to reporting negative emotional expressions than low anxious individuals when judging different social cues of emotional expressions, from video clips displaying tone of voice, facial and body expressions (Winton, Clark, & Edelmann, 1995). Similarly, anxious adults demonstrated a liberal criterion compared to non anxious adults in classifying words as threatening (Manguno-Mire, Constans & Geer, 2005). Additionally, some studies have reported anxiety-related criterion differences specific to threatening stimuli. For example, in an experiment in which highly anxious spider-fearful and non-fearful participants were asked to decide whether a picture of a spider, beetle, or butterfly was presented, highly anxious participants were not better at detecting spiders than the healthy controls, they were however, more liberal in reporting that they had seen a spider or beetle (Becker & Rinck, 2004). Others authors have failed to find any differences in either sensitivity or criterion (e.g. Sawchuk, Meunier, Lohr & Westendorf, 2002).

Although some inconsistencies are observed, the emerging picture from the literature to date suggests that anxious adults demonstrate poorer sensitivity and a more liberal criterion to negative emotional stimuli. The literature supporting a liberal criterion or response bias is more robust than the literature supporting differences in sensitivity. This pattern has primarily
been observed in the facial modality. To my knowledge, the present study is the first to systematically examine both response criterion and sensitivity for affective prosody perception in anxious individuals. It is also the first study to examine these effects in both adults and children using the same task.

*Why use the developmental approach?*

Few studies have explored anxiety-related differences in information processing across both adults and children. One way of bridging the gap between child and adult anxiety literature is to include both populations in the same study completing the same task. Therefore, the second aspect of this study was to take a developmental approach to exploring the relationship between anxiety and prosody perception. This novel approach of studying the developmental trajectory of anxiety-related differences in sensitivity and criterion of prosody perception is important. First, this approach allows for a direct comparison of adult and child performance. Furthermore, establishing the presence of differences in either response criterion or in the ability to discriminate emotional prosody in anxious children may encourage further research that aims to establish whether such anxiety-related differences mark a vulnerability for the development of adulthood anxiety.

Models of adulthood anxiety (e.g. Clark & Wells, 1995; Rapee & Heimburg, 1997) suggest that abnormalities in the processing of emotional stimuli are implicated in the maintenance of anxiety. However, exploring these patterns across children and adults may help to inform about not only the maintenance but also the development of anxiety. Thus it seems surprising that no one has ever studied children’s and adult’s affective prosody processing in the same study, using the same task. Some studies have attempted to make comparisons of anxiety-related effects on the recognition of facial expressions of emotions across studies; i.e. comparing one adult study (e.g. Richards et al., 2002) to subsequent child studies (e.g. Richards et al., 2007). However, these studies have altered the task requirements
Anxiety-related interpretation bias

for each age group, making comparisons across studies problematic. It is important to note that the idiosyncrasy of prosodic stimuli means that developmental comparisons are only meaningful if the same stimuli are used for adults and children. It is problematic to make comparisons in the ability to interpret, detect or process emotional prosody across studies because performance in any given study is very dependant on the stimuli used. Given that most studies exploring prosody perception use their own unique stimuli, what is more conclusive, and the process that was adopted for the current study, was not to attempt to compare absolute levels of performance with those in other studies, but rather to compare groups and conditions within the study.

What is known from the research on the development of prosody is that there is a window of development between the ages of four and twelve when a gradual and steady development in affective prosody perception occurs. Evidence suggests that by about age eleven children are able to perceive most emotional categories in voice as well as adults. The first study to ever systematically examine the ability of children to perceive affective prosody was Gates (1927; cited in Dimitrovosky, 1964) who assessed children from grade three to grade eight in their ability to perceive emotional meaning from a recording of the alphabet recited to express each of nine different emotional categories. Gates reported that the ability to identify the emotional expression was positively related to age. Unfortunately, however, only a very brief summary of this report was published (as reported by Dimitrovosky, 1964) therefore, overall accuracy levels, and accuracy for different emotions in the sample are not known.

The next study to follow on from Gates’ work was Dimitrovosky (1964), who assessed prosody perception in 224 children aged five to twelve years. Children listened to semantically-neutral sentences depicting the following emotions; anger, happiness, love and sadness. The children were asked to respond by pointing to one of four stick figure drawings
which represented the categories of emotional meaning. Before using the drawings as responses in the emotion identification task, children were trained to associate each of the four drawings with the appropriate emotional word. The overall finding was a gradual, steady progressive increase in ability to identify emotional prosody with age. Five year olds averaged 33% accuracy which gradually increased through to 12 year olds who average 65% accuracy.

In Dimitrovsky’s (1964) sample, children demonstrated the best recognition for sad tones of voice, followed by angry, happy and loving, respectively. However this finding must be considered in conjunction with the observation that the children in this sample also significantly favoured negative emotions in their responding, i.e. responding “sad” and “angry” significantly more often than “loving” or “happy” (Dimitrovsky, 1964). Furthermore, while there was a decrease in the number of errors made with age, the pattern of a negative response bias remained stable, suggesting it is particularly characteristic of childhood. This finding highlights again the need again for SDT analysis in research in this area to assess and control for such tendencies.

Several years later the developmental trend in prosody perception was further explored and supported in a cross culture study. Matsumoto and Kishimoto (1983) presented young American and Japanese children (aged four to nine years) with content free speech depicting four emotion types; happiness, anger, surprise and sadness. The authors found that between the ages of four and nine there is a gradual and generally steady increase in children’s ability to identify discrete emotions in speech, such that accuracy tends to rise from about 30% to about 80%, slightly higher accuracy than reported in the earlier study by Dimitrovovsky (1964). The authors report that the trend of increasing emotion perception ability with age continues through each of the successive ages, yet there are differences in accuracy between different emotions. It is not until age 7 that children can accurately identify
Anxiety-related interpretation bias

all four emotion types at a level higher than that of chance. Matsumoto & Kishimoto reported that the increase is apparent across cultures, yet at different rates. For example, by age six, American children could identify all four emotions (i.e. happiness, anger, surprise and sadness) above chance level, yet at the same age, Japanese children could accurately identify all emotions with the exception of anger, which did not emerge until age seven.

Further studies have reported that children’s ability to perceive happy, sad and mad prosodies gradually increases from the age of six to ten, with eleven year olds achieving perfect performance on an emotion recognition task (Cohen, Prather, Town, & Hynd, 1990). Similar results were found with children aged five and half years to eight and a half years (Doherty, Fitzsimons, Asenbauer and Staunton, 1999).

The literature clearly suggests children can accurately identify emotional prosody. However, it may be that although young children have the ability, the task of inferring others emotional state through tone of voice is susceptible to distraction in younger children. Morton & Trehub, (2001) tested the ability of children between age four and ten years of age to correctly label happy and sad prosody when presented with conflicting semantic content. When the prosodic and semantic emotional cues conflicted, children eight years of age and younger judged how the speaker was feeling based on the semantic content (i.e. what the speaker said) even when instructed to listen to the tone of voice, whereas children between nine and ten years of age were divided in their focus between semantic content and prosody, and adults judged the speaker’s feelings by her tone of voice. Further experiments found that when the content was obscured by use of a foreign language and by low pass filtering the English spoken utterances, children between the ages of four and ten years were able to accurately identify happy and sad emotional prosodies. This suggests that the initial observation that children under 8 years of age rely predominantly on semantic cues to form judgement of the speakers emotional state reflects a tendency to accord greater weighting to
Anxiety-related interpretation bias

semantic cues rather than an inability to identify the prosodic cues, and may reflect developmental changes in attention rather than prosodic perception. These studies did not examine each emotion separately but rather the trends reflect a general increase in the ability to process emotional tone of voice.

The conflicting rates of accuracy between the studies are to be expected when using idiosyncratic stimuli. Some use natural speech, pseudowords or filtered speech, all of which introduce unique processing demands to the experimental task. Additionally studies suggest that there is large variability in prosody perception ability within the age groups. For example in Dimitrovosky’s (1964) sample, the highest performing five-year-olds were more accurate than the lowest performing twelve-year-olds (Dimitrovosky, 1964). It is clear from the literature that the age range of seven to eleven years is a window of developmental change in prosodic processing, and therefore a good candidate age range in which to capture developmental effects related to anxiety. It is reasonable to hypothesise that anxiety may have its strongest effects while the ability to perceive emotional prosody is developing.

Emotion recognition impairment in anxious children

The adult literature suggests that anxious adults show differences in the way that they process emotional stimuli when compared to non-anxious adults, yet corresponding studies with anxious children are few. The research that is available however, tends to suggest that childhood anxiety is related to a decreased ability to interpret emotions in others. As in the adult literature, the majority of these studies have examined facial expressions of emotion. For example, children with social phobia (Simonian, Beidel, Turner, & Berkes, 2001) and children and adolescents with various anxiety disorders demonstrated a general (non-emotion specific) decrease in ability to accurately identify facial expressions of emotions (Easter, et al., 2005) when compared to healthy controls. Consistent with the adult literature, some
studies report no differences in emotion perception abilities in anxious youth (e.g. McClure, Pope, Hoberman, Pine & Leibenluft, 2003).

Prosody perception in anxious children

To date there has been one study to explore anxiety-related differences in affective prosody perception in children. McClure and Nowicki (2001) found that children aged between eight and ten years, who reported more anxious avoidant behaviours and more distress in social situations, had poorer accuracy at recognising vocal expressions of emotion. Unfortunately McClure and Nowicki did not differentiate between positive and negative affective prosodies, and did not use SDT. Thus it is difficult to make conclusions about what may have driven the reported global decrease in prosody perception for this group of anxious children.

SDT paradigms with children

Just as the literature utilizing signal detection analysis on emotion perception in anxious adults is sparse, even less research has been done with a child population. Findings to date seem to be consistent with the adult findings. Few studies have reported differences in sensitivity, with more studies reporting criterion differences. The few sensitivity findings in child samples when a strict SDT paradigm is used have found that anxiety is not associated with enhanced sensitivity, but rather a decreased sensitivity to discriminate emotional expressions in others. For example, highly anxious children were less able to discriminate angry from happy facial expressions than low anxious children (Richards et al., 2007) and children with social phobia were less sensitive than healthy controls when recognising facial expressions of various emotions (Simonian et al., 2001). In a less conclusive study, anxious children were significantly more likely to report seeing an emotion (both positive and negative) in an emotionally neutral face than were the controls (Melfsen & Florin, 2002).
Consistent with the adult findings, there have been several studies that suggest childhood anxiety is associated with a liberal criterion to report negative emotional stimuli, including prosody, social and physical cues, yet they have not systematically extracted measures of sensitivity or bias. For example in the study described earlier exploring prosody perception in anxious children, socially anxious children more frequently mislabelled fearful voices as sad (McClure & Nowicki, 2001). In a peer interaction task, there was no anxiety-related difference in ability to accurately identify hostile intent in peer interactions, but anxious children displayed a tendency to misinterpret non-hostile situations as hostile (Bell-Dolan, 1995). Additionally, anxious adolescents when faced with anxiety provoking tasks did not display any differences to non-anxious adolescents in objective physiological arousal, yet they reported more subjective physiological arousal than their non-anxious counterparts (Anderson & Hope, 2009).

Finally the role of depression in emotional communication has been noted across studies, and most studies control for depression either statistically or through participant selection. Those that do not are problematic because anxiety and depression are highly associated, and the influence of depression on emotional processing, particularly for anger has been noted across studies (Luck & Dowrick, 2004; Quadflieg et al., 2007). Kessler et al. (2007) found that emotion recognition differences between anxious and non-anxious groups disappeared when depression was accounted for and Eley et al. (2008) found that depression was related to threat interpretation in eight-year-old twins but anxiety was not.

The present study

The present study had two clear goals. The first objective was to systematically assess the distinct roles of both perceptual sensitivity and response criterion in the processing of affective prosody in anxious and non-anxious individuals. This was done using SDT. Based on previous SDT studies in anxious populations, it was expected that anxiety related
differences would be seen in the form of a liberal criterion for negative prosody, particularly fear. It was hypothesized that differences in sensitivity would not be observed, or that highly anxious individuals would demonstrate poorer sensitivity across prosodies.

Secondly, this study aimed to explore whether any differences in sensitivity and criterion vary across the developmental trajectory. Such differences may help to inform of the role that the hypothesised bias plays in the development and/or maintenance of anxiety. Given that no previous studies have compared child and adult populations across the same stimuli and task paradigm, there was little evidence on which to base a conclusive hypothesis regarding any developmental trend to be observed.

The age group of children aged eight and nine was chosen because it fits within the age group identified as a developmental window for prosody perception, in which children are able to correctly identify most emotional prosodic categories whilst still performing at a level different to adults (Dimitrovosky, 1964; Matsumoto & Kishimoto, 1983). Pilot testing ensured that children of this age range are able to effectively use the keyboard to indicate their response, thus allowing the same task to be used for adults and children.

In the present study participants listened to sentences that were low-pass filtered. This process was used for two reasons; it removed all semantic meaning, and it allowed for emotional tone of voice to still be heard, yet made the task more difficult. It is at such a level of uncertainty that differences in the processing of emotional stimuli have been reported (Frenkel et al., 2008) and studies that do not use such methods are often confounded with ceiling effects (see Quadflieg et al., 2007; Ellis et al., 1997). Also, it is theoretically important to remove semantic content. There is always an interaction between semantic content and prosody; removing the semantic content allows examination of the effect of prosody in isolation. Additionally this process was intended to minimise the tendency of younger children to become distracted by the semantic content of utterances, thus affecting
Anxiety-related interpretation bias

performance in prosodic evaluation (e.g. Morton & Trehub, 2001; Morton, Trehub, & Zelazo, 2003).

Method

Participants

Adults

Eighty-five adult participants (23 males and 62 females) were recruited through the Introductory Psychology Research Participation Program (IPRP). Participants were Psychology undergraduates from Victoria University of Wellington. The average age of adult participants was 20 years (SD = 4.67, range = 16-43 years). All participants reported being native English speakers and having no visual problems that were not corrected by glasses or contacts and no known hearing problems. All participants reported not currently being treated for depression (either through counselling or SSRI medications). They received one hour credit toward compulsory four hours research participation as a course requirement.

Children

Seventy-one children (36 eight-year-olds and 35 nine-year-olds, 35 males and 36 females) were recruited through schools in the Wellington and Manawatu regions of New Zealand. Schools were contacted via letters sent to the principals (see appendix A) and followed up via phone calls from the researcher. Information letters and consent forms were sent home to the parents/caregivers (see appendix B) of all children aged eight and nine years who were enrolled at the participating schools. All children who returned signed consent forms were recruited for participation in the present study. The study was approved by the Victoria University of Wellington’s Human Ethics Committee.

Materials and stimuli

Anxiety in the adult sample was measured with the Beck Anxiety Inventory (BAI; Beck & Steer, 1990). Depression in the adult sample was measured with the Beck Depression
Anxiety-related interpretation bias

Inventory – Second edition (BDI-II; Beck, Steer, & Brown, 1996). The BAI is a 21 item self-report measure; each item is descriptive of subjective, somatic, or panic-related symptoms of anxiety in accordance with the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Health Disorders Fourth Edition (DSM-IV; 1994). The respondent is required to rate their experience of each symptom over the past week on a four-point scale ranging from zero to three. The item scores are summed to obtain a total score, ranging from zero to 63. The BAI has been demonstrated to hold high internal consistency, item-total correlations ranging from .30 to .71 and good validity (Beck, Epstein, Brown & Steer, 1988). Likewise, the BDI-II is a 21-item self-report measure intended to assess the severity of depressive symptoms in accordance with the DSM-IV (American Psychiatric Association; APA, 1994). It similarly asks participants to rate their experience of each symptom on a four-point scale ranging from zero to three, with a total score of 63. However, the time period for the BDI-II is over the past two weeks.

Anxiety and depression in children were measured with the Beck Youth inventories of Emotional and Social Impairment- second edition (age 7-14; Beck, Beck, Jolly & Steer, 2005); the Anxiety Inventory, (BAI-Y) was used to measure anxiety and the Depression Inventory, (BDI-Y) was used to measure depression. The BAI-Y includes 20 items that reflects children's specific worries, fears and physiological symptoms associated with anxiety consistent with the anxiety disorders criteria of the DSM–IV (APA, 1994). The BDI-Y includes 20 items related to a child's or adolescent’s negative thoughts, feelings and somatic symptoms and correspond to the depression criteria of the DSM–IV (APA, 1994). A recent review concluded that although the youth inventories are not without problems, the depression and anxiety scales are a useful tool for mental health professions for accurately measuring disturbances in these states (Bose-Deakins & Floyd, 2004). Both youth inventories have a total score of 60.
The auditory stimuli for the emotion recognition task were 60 sentences spoken in a New Zealand accent by an adult female actress. Each of the 60 sentences was semantically neutral and was spoken in four different tones of voice; neutral, happy, angry and fearful (producing a total of 240 tokens). All the sentences of a particular emotional prosody were recorded in a single block, to maximise the actress’s emotional consistency across sentences. For the neutral prosody the actress was instructed to put as little emotion as possible into her voice; for the fearful prosody, the actress was told to aim for a panic, high arousal tone of voice; for angry prosody the actress aimed for cold bitter anger (low arousal); and for the happy prosody the actress was told to aim for a joyful, high arousal tone of voice. The digital stimuli were recorded at a professional sound studio with a Neumann U87 microphone, using the software Protools version 7, controlled by a Macintosh G5 computer. The recordings were made in one channel (mono) at 24 bits and 44100KHz. The editing software Audacity version 1.2.6 was used to duplicate the mono recordings to make stereo files, convert the files to 16 bits, add 40 ms of silence to the beginning of each sentence, and to equate the sentences for peak amplitude.

The sentences were low pass filtered to remove all perceivable semantic content. This was done using Praat version 4.6.06 system for doing phonetics by computer (Boersma, 2001). Filtering was done using a Pass Hann Band of 0-550 Hz. After the low pass filtering, the sentences were re-amplified. In line with signal detection theory, the stimulus must include ‘noise’, making the stimulus more difficult to perceive. In research on facial expressions of emotions, this is typically achieved by morphing the faces to display various degrees of any given emotion. With prosodic stimuli this can be done by low pass filtering. This procedure makes the speech unintelligible, but preserves the emotional tone of voice and eliminates concurrent semantic and phonological processing. This procedure yields stimuli
that sound like the participant is listening to someone speaking on the other side of wall. Prior to filtering the sentences were piloted to ensure that they reflected the target emotions.

Extraction of the acoustic parameters was done using Praat. The filtered sentences were analysed to calculate: Fundamental frequency or F₀ (with an auto-correlation method; floor 75hz, ceiling 500hz); Standard deviation of F₀ (Variability); Duration and Number of words. Fundamental frequency is the physical characteristics of the acoustic stimulus that most closely corresponds to perceived pitch. The average parameters for each emotional category are represented in Table 1. The duration included the 40ms of silence at the beginning of each sentence and given that the same sentences were used for each emotional category, the average number of words was 7.1. Speech Rate (words per minute) was calculated as such; (Number of words/ms) x 60000.

Table 1

<table>
<thead>
<tr>
<th>Prosody</th>
<th>Mean F₀ (Hz)</th>
<th>SD F₀ (Hz)</th>
<th>Speech rate (wpm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angry</td>
<td>198.87</td>
<td>43.03</td>
<td>213.13</td>
</tr>
<tr>
<td>Fearful</td>
<td>347.98</td>
<td>57.42</td>
<td>225.50</td>
</tr>
<tr>
<td>Happy</td>
<td>273.32</td>
<td>72.75</td>
<td>184.30</td>
</tr>
<tr>
<td>Neutral</td>
<td>180.5</td>
<td>41.1</td>
<td>152.87</td>
</tr>
</tbody>
</table>

The 60 sentences were divided into two lists (i.e. List A and B). Each participant completed 3 blocks, each one a mixture of sentences spoken in the target emotional voice and an equal mixture of the three distracter emotions. That is, each participant completed a happy block, a fearful block and an angry block. Neutral sentences did not appear as targets, they were only used as distracters. Within each block 60 sentences were presented (30 target emotion and 30 distracter emotions), for a total of 180 trials per participant. The distracter
sentences were made up of a random selection from the list that was not the target list for that block, that is, targets and distracters were filtered versions of different sentences. For example, for participant one, block one was a “Happy block” therefore happy tone of voice was the target, and 30 sentences from Happy list A were presented. The distracter sentences for this block included 10 neutral, 10 angry and 10 fearful sentences from list B. The assignment of lists to emotion conditions and block order were counterbalanced across participants. The experiment was programmed and presented using a Dell PC running Psychology Software Tools’ E-Prime Suite version 1.0 (Schneider, Eschman, & Zuccolotto, 2002).

Procedure

Adults

Each participant came into the lab for 30 minutes and was given an information and consent form to sign (see appendix C) before completing the listening task and questionnaires. Participants were told that they were going to participate in a computerized experiment (approximately 20 minutes), and fill out some questionnaires (approximately 10 minutes).

The task was completed in a large room with computers at individual cubicles, desks and chairs. Participants completed the experiment individually on their own computer, with up to 6 participants completing the experiment at any one time in the room. They were told to read the instructions on the screen. Firstly participants completed 12 practice trials. They were told to pay close attention, because the sentences that they were to hear were muffled and would sound as if they were listening through a wall. Instructions were presented on the computer monitor with black font on a white background. The experimenter read the following instructions verbatim from the screen:
“You will be listening to sentences that sound muffled. Although you will not be able to hear what the person is saying, I want you to judge the emotional tone of voice”.

“Let's practice. The tone of voice you will hear will either be happy, fearful, or angry. Click on the emotion that you think best describes the actor’s tone of voice”.

Participants were told to start the practice trials, in which they were required to identify the emotions. The practice trials consisted of 3 sentences spoken in each of three different prosodies. These sentences were not used in the experiment. After the practice trials the participants were asked if they had any questions before moving onto the experiment which consisted of 2 blocks of 60 sentences. They were told that the blocks would be slightly different to the practice trials in that they would simply have to choose “Happy or not happy”; “Angry or not angry”; and “Fearful or not fearful”.

They were instructed to choose the answer that best suited the tone of voice. The task was split into 3 blocks, 60 trials per block. Participants listened to the sentences through earphones while sitting at a computer, and followed the instructions on the computer to make a forced choice response on the computer key pad after each sentence, i.e. Press 1 = Target emotion, 2 = Not target emotion. The computer recorded responses. The instruction screen was as follows:

“In this first block of trials, you will be listening for a [Target Emotion] voice. If the voice is [Target Emotion], press the 1 key. If it is a different emotion, press the 2 key”

Participants were not given any feedback on their responses. Following the experiment, each participant was given the questionnaires to complete; the questionnaire had the subject’s participant number on the top, and they were instructed to not put any identifying information (e.g. name) on the questionnaire. The experimenter read through the instructions on the top of the questionnaire aloud, and told the participants to ensure they had
Anxiety-related interpretation bias

read it carefully before they began. Order of questionnaires was counterbalanced, such that half participants completed the BAI first, and half completed the BDI first. After the experiment was finished, the participants were verbally debriefed about the conditions and aims of the study, given a debriefing form (see appendix D) and asked if they had any questions.

Children

Procedure for the child sample was similar to that of adults in all aspects, except that data collection took place at the child’s school in a separate room provided for testing by the school. Testing for children took place over two separate days. On the first day all children completed the questionnaires. This was done by the experimenter individually with each child. The experimenter read through each question with the child and the child marked their response on the record form. However, for children who reported they could read through the questions themselves, the experimenter read the first question aloud, and the children read through the remaining questions themselves. The children were instructed to ask the experimenter if they had any questions or if there were words that they did not understand.

On the subsequent day the children completed the computerised listening task. At any given time, there was up to three children completing the task at once, each on a separate lap top computer that was provided by the experimenter. Each lap top computer had good quality Panasonic stereo headphones (model RP-HT160; XBS) with circumaural cushions to maximise sound quality and minimise distraction for the children. Each child was given a small gift as a thank you for their participation and a debrief form (see appendix E) to give to their parents/caregivers at the end of the second session.

Design

The study was a mixed model 3 x 2 x 2 design. The independent (within-subject) variable was emotion (happy, angry, fearful), the between subject variables were anxiety level
(high/low) and age (adult/child). The dependant variables were sensitivity measured by $d'$ and response bias measured by $c$ (see below).

**Statistical analysis**

Overall accuracy rates for each individual were calculated by adding their correct responses across each of the three types of emotional prosody. Because the listening task was a signal detection task, a signal detection analysis was conducted on the participants’ response. Hit rates (HR) were calculated for each individual by dividing their hits (e.g. a happy response to a happy trial) by the number of target present trials (i.e. hits/30). A false alarm is when a participant gives a target (e.g. happy) response to non target (or distracter) sentence. False alarm rates (FA) were calculated for each individual by dividing the false alarms by the number of each given emotional prosody sentence (i.e. false alarms/30). Additionally separate false alarm rates were computed for each distracter emotion, e.g. on a happy block, a fearful FA rate was calculated when a participant called a fearful distracter happy and an angry FA rate was calculated when a participant called an angry distracter sentence happy. This allowed for the examination of how well the participant could discriminate the target emotion from each distracter.

$Z$ scores were calculated for each hit rate and false alarm rate. These scores were used to calculate sensitivity ($d'$) and criterion ($c$) scores for each individual with the following equations described by Macmillan and Creelman (2005):

$$d' = z(HR) - z(FAR) \quad c = - \frac{1}{2} [z(HR) + z(FAR)]$$

Sensitivity ($d'$) is a measure of the participant’s ability to differentiate between a given emotional prosody and distracters. A sensitivity score of zero reflects an inability to
discriminate a given prosody from the distracters. That is for a participant who is just as likely to say “happy” on a happy or not happy trial their hits and false alarms will be equal, giving them a sensitivity score of zero. As a participant’s sensitivity increases, they make more hits and fewer false alarms because they are becoming more accurate at differentiating a given emotional prosody from the distracters (i.e. noise), that is, increased perceptual sensitivity results in larger sensitivity values.

The response bias score or criterion (c) reflects the degree to which ‘happy’ responses or “not happy” responses are preferred. A positive value of criterion reflects a tendency to say ‘not a target’ (conservative bias), whereas a negative value of criterion reflects a tendency to say ‘target’ (liberal bias). d’ and c are independent except when d’ approaches the ceiling or floor. Measures of sensitivity and criterion were calculated for each of the three target emotions.

Results

Participant characteristics

Both adult and child participants were separated into two groups; high anxiety and low anxiety by means of a median split based on their overall score on the BAI or BAI-Y. Demographics of the participant groups is reported in Table 2.

Table 2

**Participant Demographics, Anxiety and Depression Scores (Standard Deviations)**

<table>
<thead>
<tr>
<th></th>
<th>Adults</th>
<th>Children</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low anxiety</td>
<td>High anxiety</td>
</tr>
<tr>
<td>Mean Age</td>
<td>20.72 (5.99)</td>
<td>18.47 (2.40)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n of males</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>n of females</td>
<td>30</td>
<td>32</td>
</tr>
</tbody>
</table>
Anxiety-related interpretation bias

| Anxiety (BAI and BAI-Y) | 5.89 (3.98) | 14.30 (8.39) | 9.46 (3.70) | 21.38 (5.03) |
| Depression (BDI and BDI-Y) | 5.23 (2.96) | 15.90 (4.79) | 7.05 (3.27) | 15.74 (6.95) |
| Correlation BAI*BDI | 0.38** | 0.54* | 0.64** | 0.42* |

Note the BAI and BDI have maximum scores of 63, the BAI-Y and BDI-Y have maximum scores of 60.
* p<0.05
** p<0.001

Dependent variables were analysed in a 2 (anxiety: high, low) x 2 (age: adult, child) x 3 (emotional prosody: happy, fearful, angry) mixed ANOVA with anxiety and age as between-subjects variables and emotional prosody as a within-subject variable. Depression scores as measured by the BDI were included as a covariate for all analyses. All effects of emotion were further explored with post-hoc t-tests using the Bonferroni corrected significance level of $p = 0.016$ to compare specific emotions while controlling for multiple comparisons. Sex was included in the initial analysis. No sex differences were observed, therefore it was excluded for subsequent analysis.

**Sensitivity**

Table 3 presents the mean sensitivity (d’) scores of each emotion for both adult and child participant groups. The ability to discriminate between a target emotional prosody and distracters (d’) was analysed in a 2 (anxiety: high, low) x 2 (age: adult, child) x 3 (emotional prosody: happy, fearful, angry) mixed ANOVA with anxiety and age as between-subjects variables and emotional prosody as a within-subjects variable.

There was a main effect of age, $F(2, 302) = 29.30, p < 0.001$ and a main effect of emotion, $F(2, 302) = 23.50, p < 0.01$. These were qualified by an age x emotion interaction, $F(2, 302) = 44.316, p < 0.01$. Follow up independent samples t-tests showed that adults were significantly better than children in their ability to detect angry and happy prosodies, $t(154) = 7.05, p < 0.001$, and $t(154) = 10.17, p < 0.001$ (respectively). Yet the two groups did not significantly differ in their ability to detect fearful prosody, $t(154) = -1.56, p = 0.12$. In fact
children displayed non-significantly better discrimination than adults for fearful voices. (Refer to figure 1). There were no effects of anxiety.

Table 3

Mean sensitivity scores (d’) and (standard deviations) for low, high and combined anxious adults and children

<table>
<thead>
<tr>
<th>Emotional prosody</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant group</td>
<td>Angry</td>
<td>Fearful</td>
<td>Happy</td>
</tr>
<tr>
<td>Adults</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low anxiety</td>
<td>1.32 (0.82)</td>
<td>1.66 (0.61)</td>
<td>1.68 (0.78)</td>
</tr>
<tr>
<td>High anxiety</td>
<td>1.27 (0.65)</td>
<td>1.49 (0.69)</td>
<td>1.82 (0.76)</td>
</tr>
<tr>
<td>Combined</td>
<td>1.30 (0.74)</td>
<td>1.60 (0.65)</td>
<td>1.75 (0.77)</td>
</tr>
<tr>
<td>Children</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low anxiety</td>
<td>0.50 (0.72)</td>
<td>1.90 (1.25)</td>
<td>0.71 (0.56)</td>
</tr>
<tr>
<td>High anxiety</td>
<td>0.49 (0.62)</td>
<td>1.78 (1.59)</td>
<td>0.55 (0.55)</td>
</tr>
<tr>
<td>Combined</td>
<td>0.49 (0.67)</td>
<td>1.84 (1.41)</td>
<td>0.63 (0.56)</td>
</tr>
</tbody>
</table>

Given the main effects and interaction involving age, sensitivity was analyzed in adults and children separately. When adults and children were examined separately, adults displayed a main effect of emotion $F(2, 164) = 4.32, p < 0.05$. Adults were significantly better at detecting happy and fearful tones of voice than they were at detecting angry tone of voice, $t(82) = -4.64, p < 0.001$ and, $t(82) = -3.17, p < 0.001$, respectively.
Children also showed a main effect of emotion, \( F(2, 136) = 18.71, p < 0.001 \). But in contrast to adults, post hoc t-tests revealed that children were significantly better at detecting fearful tone of voice than they were at detecting angry, \( t(70) = -8.89, p < 0.001 \), and happy tone of voice, \( t(70) = 8.20, p < 0.001 \). Children were equally poor at discriminating angry and happy tones of voice, \( t(70) = -0.56, p = 0.123 \).

![Figure 1. A comparison of children’s and adult’s ability to discriminate different emotional prosodies. Error bars represent standard error values. Note, higher d’ values reflect better ability, and lower values reflect poorer ability. * p < 0.001](image)

**Criterion**

Table 4 presents the mean criterion (c) scores of each emotion for both adult and child participant groups. Overall, adults were very conservative on happy trials, less conservative on angry trials and slightly liberal on fearful trials. Children showed little to no bias on happy trials, and were quite liberal on fearful and angry trials.
Table 4

Mean criterion (c) and (standard deviations) for low, high and combined anxious adults and children

<table>
<thead>
<tr>
<th>Participant group</th>
<th>Emotional prosody</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Angry</td>
</tr>
<tr>
<td>Adults</td>
<td></td>
</tr>
<tr>
<td>Low anxiety</td>
<td>0.18 (0.39)</td>
</tr>
<tr>
<td>High anxiety</td>
<td>0.10 (0.40)</td>
</tr>
<tr>
<td>Combined</td>
<td>0.15 (0.39)</td>
</tr>
<tr>
<td>Children</td>
<td></td>
</tr>
<tr>
<td>Low anxiety</td>
<td>-0.14 (0.29)</td>
</tr>
<tr>
<td>High anxiety</td>
<td>-0.08 (0.31)</td>
</tr>
<tr>
<td>Combined</td>
<td>-0.11 (0.30)</td>
</tr>
</tbody>
</table>

*Note.* Positive values reflect a conservative criterion and negative values reflect a liberal criterion.

Participants’ response biases were analysed in a 2 (anxiety: high, low) x 2 (age: adult, child) x 3 (emotional prosody: happy, fearful, angry) mixed ANOVA with anxiety and age as between-subjects variables and emotional prosody as a within-subject variable. There was an interaction between emotion and BDI, $F(2, 302) = 3.12, p < 0.05$ in the omnibus ANOVA. There was no main effect of depression, $F(2, 302) = 0.31, p = 0.58$. This was followed up with correlational analyses; the BDI correlated with angry bias, $r = -0.17, p < 0.05$, indicating that the higher a participant’s depression score was, the more likely they were to call a
sentence angry during the angry block. Given that children and adults varied in their bias (see below) children and adults were then separated. When this was done, the effect approached significance for adults, \( r = -0.21, p = 0.056 \) but disappeared for the children, \( r = -0.05, p = 0.70 \).

There was a main effect of emotion, \( F(2, 302) = 9.50, p < 0.001 \) and a main effect of age, \( F(2, 302) = 38.33, p < 0.001 \). These were qualified by a significant emotion x age x anxiety interaction, \( F(2, 302) = 3.220, p < 0.04 \). This interaction was explored by examining the effects of emotion and anxiety for adults and children separately.

When adults and children were analysed separately, adults showed a main effect of emotion, \( F(2, 164) = 9.155, p < 0.01 \) such that they were more conservative during the angry block and the happy block than the fearful block, \( t(84) = 3.04, p < 0.01 \) and \( t(84) = -5.19, p < \)
Anxiety-related interpretation bias

0.001, respectively. There was no main effect, $F(2, 164) = 0.28, p = 0.60$ or interaction with anxiety, $F(2, 164) = 0.10, p = 0.91$.

In contrast, the children showed no main effect of emotion, $F(2, 136) = 1.37, p = 0.26$, but the interaction between anxiety and emotion approached significance $F(2, 136) = 2.69, p = 0.07$. Although both groups demonstrated a liberal response bias overall, independent t-tests showed that highly anxious children were significantly more liberal in their responding, that is more likely to respond “yes” when fearful was a target than low anxious children, $t(69) = 2.11, p < 0.05$. The two groups did not differ significantly in their responding to the other two emotion types, i.e. angry and happy.

*False Alarm analysis for specific emotional comparisons*

When a participant makes a false alarm, they inaccurately identify a distracter as a target. The initial analysis analyzed participants’ abilities to discriminate each prosodic target from all other distracters combined. A supplementary analysis was conducted to determine whether there were specific distracter emotions that were most commonly misidentified as a target and whether these errors differed as a function of age or anxiety.

First, the number false alarms were computed for each distracter emotion. Recall all together there were three target emotions (angry, fearful, happy) and there were four distracter emotions (angry, fearful, happy and neutral). On each block there was one target emotion and three distracters. Therefore on a happy block, a fearful FA rate was calculated when a participant called a fearful distracter happy, an angry FA rate was calculated when a participant called an angry distracter happy, and a neutral FA rate was calculated when a participant called a neutral distracter happy. This allowed for the examination of how well the participant could discriminate the target emotion from each individual distracter emotion (see tables 5 and 6). Given the large differences observed in hit and false alarm rates, this analysis was done separately for adults and children.
Table 5

*False Alarms and (standard deviations) as a function of target and distracter for low, high and combined anxious adults*

<table>
<thead>
<tr>
<th>Participant group</th>
<th>Distracter emotion</th>
<th>Target emotion</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Angry</td>
<td>Fearful</td>
</tr>
<tr>
<td><strong>Low Anxiety</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angry</td>
<td>-</td>
<td>0.82 (0.72)</td>
<td>2.68 (1.90)</td>
</tr>
<tr>
<td>Fearful</td>
<td>1.38 (1.62)</td>
<td>-</td>
<td>1.72 (2.16)</td>
</tr>
<tr>
<td>Happy</td>
<td>2.11 (1.56)</td>
<td>1.75 (1.44)</td>
<td>-</td>
</tr>
<tr>
<td>Neutral</td>
<td>3.63 (3.14)</td>
<td>0.75 (0.80)</td>
<td>1.42 (1.73)</td>
</tr>
<tr>
<td><strong>High Anxiety</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angry</td>
<td>-</td>
<td>0.83 (0.62)</td>
<td>2.38 (1.50)</td>
</tr>
<tr>
<td>Fearful</td>
<td>1.34 (1.22)</td>
<td>-</td>
<td>1.73 (1.70)</td>
</tr>
<tr>
<td>Happy</td>
<td>2.39 (1.87)</td>
<td>1.98 (1.71)</td>
<td>-</td>
</tr>
<tr>
<td>Neutral</td>
<td>4.13 (2.87)</td>
<td>0.67 (0.51)</td>
<td>0.94 (0.83)</td>
</tr>
<tr>
<td><strong>Combined (all adults)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angry</td>
<td>-</td>
<td>0.82 (0.67)</td>
<td>2.53 (1.72)</td>
</tr>
<tr>
<td>Fearful</td>
<td>1.36 (1.43)</td>
<td>-</td>
<td>1.72 (1.94)</td>
</tr>
<tr>
<td>Happy</td>
<td>2.25 (1.71)</td>
<td>1.86 (1.57)</td>
<td>-</td>
</tr>
<tr>
<td>Neutral</td>
<td>3.87 (3.00)</td>
<td>0.71 (0.67)</td>
<td>1.19 (1.38)</td>
</tr>
</tbody>
</table>

*Note* False Alarms out of a total of ten.
Although there are differences in the false alarm rates of the emotions, all emotions have reasonably low rates of false alarms. The exception is that neutral is often misinterpreted as angry. A 2 (anxiety; high, low) x 3 (distracter) ANOVA was calculated for each target emotion separately. Whenever effects of distracter emotion were observed, they were followed up with post-hoc t-tests with a Bonferroni correction. Given the large differences in performance between adults and children in their ability to discriminate the different emotions, the analyses of False Alarms were conducted for each age group separately.

For adults, main effects of distracter emotion were observed for all target prosodies, $F(2, 164) = 5.32, p < 0.05$; $F(2, 164) = 22.14, p < 0.001$; $F(2, 164) = 4.58, p < 0.05$, for angry, fearful and happy respectively. For angry targets adults made fewer false alarms when fearful was the distracter than when happy or neutral were, $t(84) = -5.05, p < 0.001$ and $t(84) = -6.36, p < 0.001$, respectively. Additionally, adults made fewer false alarms when happy was the distracter compared to when neutral was the distracter, $t(84) = -4.28, p < 0.001$. For fearful targets few false alarms were made, however more false alarms were made when happy was the distracter than when angry or neutral were the distracters, $t(84) = -6.30, p < 0.001$, and $t(84) = 7.01, p < 0.001$, respectively. Also, they made more false alarms when angry was distracter than when neutral was, $t(84) = 2.5, p < 0.05$. For happy targets adults made more false alarms when angry was the distracter than when fearful or neutral were the distracters, $t(84) = 3.02, p < 0.005$ and $t(84) = 7.34, p < 0.001$, respectively, indicating that they were more likely to confuse angry with happy prosody. They also made more false alarms when fearful was the distracter than when neutral was, $t(84) = 2.02, p < 0.05$. Importantly, there were no effects of anxiety on any target.
Table 6.

False Alarms and (standard deviations) as a function of target and distracter for low, high and combined anxious children

<table>
<thead>
<tr>
<th>Participant group</th>
<th>Distracter emotion</th>
<th>Target emotion</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Angry</td>
<td>Fearful</td>
<td>Happy</td>
</tr>
<tr>
<td>Low anxiety</td>
<td>Angry</td>
<td>2.41 (1.95)</td>
<td>3.91 (2.20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fearful</td>
<td>3.16 (2.86)</td>
<td>- 4.39 (2.72)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Happy</td>
<td>4.23 (2.36)</td>
<td>3.20 (2.12)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Neutral</td>
<td>6.42 (2.45)</td>
<td>1.80 (1.68)</td>
<td>2.70 (2.03)</td>
</tr>
<tr>
<td>High anxiety</td>
<td>Angry</td>
<td>2.56 (2.47)</td>
<td>4.49 (2.35)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fearful</td>
<td>2.71 (2.41)</td>
<td>- 3.93 (2.78)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Happy</td>
<td>3.93 (2.04)</td>
<td>4.25 (2.37)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Neutral</td>
<td>6.81 (2.24)</td>
<td>2.66 (2.83)</td>
<td>3.94 (2.41)</td>
</tr>
<tr>
<td>Combined (all children)</td>
<td>Angry</td>
<td>2.48 (2.20)</td>
<td>4.18 (2.72)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fearful</td>
<td>2.94 (2.64)</td>
<td>- 4.17 (2.74)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Happy</td>
<td>4.01 (2.20)</td>
<td>1.51 (1.29)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Neutral</td>
<td>6.61 (2.34)</td>
<td>2.21 (2.33)</td>
<td>3.30 (2.29)</td>
</tr>
</tbody>
</table>

Note: False Alarms out of a total of ten.
Overall children were good at discriminating fear from all emotions, poor at discriminating happy from all emotions (even fearful), and they could discriminate angry from fear, but not from happy or neutral. Similar to adults, children displayed main effects of distracter emotion for angry and fearful target prosodies, $F(2, 136) = 10.09, p < 0.001$, $F(2, 136) = 3.05, p < 0.05$, respectively. Yet, unlike the adults, children did not show a main effect when happy was the target prosody, $F(2, 136) = 0.77, p < 0.46$. This is likely a reflection of the children’s poor ability to discriminate happy tone of voice from any distracters. For angry targets, children made significantly more false alarms when neutral was the distracter compared to when happy, $t(70) = -5.60, p < 0.001$, or fearful, $t(70) = -7.24, p < 0.001$ were. Additionally, they made more false alarms when happy was the distracter than when fearful was, $t(70) = -3.39, p < 0.005$. For fearful targets, children made significantly more false alarms when happy was the distracter compared to when angry or neutral was the distracter, $t(70) = -3.98, p < 0.001$, $t(70) = 4.52, p < 0.001$ (respectively). Children did not differ in their false alarm rates when angry and neutral were the distracters, $t(70) = 1.61, p = 0.11$ Again, there were no effects of anxiety

Discussion

This study examined two aspects of affective prosody perception. Firstly, the processing of happy, angry and fearful affective prosody was examined in a sample of individuals who self-reported high levels of anxiety using a signal detection paradigm (Macmillan & Creelman, 2005). This methodology was used to systematically measure the relationship between anxiety, criterion and sensitivity in the processing of discrete emotions. Signal detection theory has been applied to exploring anxiety related differences in the perception of emotional expressions in other modalities yet to my knowledge this is the first study to apply it to the perception of emotional prosody in anxious individuals. Secondly these differences were examined across the developmental trajectory, using the same stimuli
Anxiety-related interpretation bias

and same task in a sample of children aged eight and nine years old and an adult sample. This novel approach was used to make direct comparisons between child and adult performance and to encourage future research to explore other anxiety related differences across the developmental trajectory in an attempt to examine the roles of such differences in the development and/or maintenance of adulthood anxiety.

**Anxiety-related findings**

Firstly, as expected, anxiety related differences were observed in individuals’ criterion and not in their sensitivity to emotional prosody. Specifically, high levels of anxiety in childhood were associated with a more liberal criterion to report a fearful tone of voice. Secondly, and importantly, the relationship between anxiety and criterion was observed only for children and not for adults. That is, highly anxious children were more likely to label a voice fearful than their low anxious counterparts, whereas highly anxious adults were no more likely to label a voice fearful than low anxious adults. This difference suggests that the mechanisms that support anxiety-related effects on interpretation of emotional tone of voice are different in children and adults.

**A Comparison with child prosody literature**

The perception of emotional prosody in anxious children has rarely been examined. The little research available reports that sub-clinical levels of anxiety are associated with poorer ability to interpret others’ vocal emotional communication (McClure & Nowicki, 2001). Unfortunately McClure and Nowicki did not differentiate between positive and negative affective prosodies, or discrete emotions, and again did not use SDT; thus it is difficult to make conclusions about what may have been driving the poorer accuracy for the children. In light of the present criterion finding, poor performance in anxious children observed by McClure and Nowicki may be reflective of a liberal criterion for fear, i.e. that they tend to call more voices fearful, regardless of whether they were fearful or not, and thus
reducing their overall accuracy across a range of emotions. The present study observed no effects of anxiety on criterion for detecting anger or happiness. However it is possible that differences in anger or happiness were not observed because children’s sensitivity levels for angry and happy prosody approached the floor.

_How does this fit with other SDT findings?_

When we step away from the prosody literature, this present criterion finding complements several earlier observations that anxious children and adolescents make more threatening, unpleasant and negative interpretations of social situations (Barrett, Rapee, Dadds, & Ryan, 1996; Chorpita, Albano, & Barlow, 1996; Higa & Daleiden, 2008; In-Albon, Dubi, Rapee, Schneider, in press; Micco & Ehrenreichm, 2008; Muris, Merckelbach, & Damsma, 2000; Vassilopoulos & Banerjee, 2008; Waters, Craske, Bergman, & Treanor, 2008); produce more threatening than neutral responses when presented with threat-neutral homophones (Hadwin, Frost, French, & Richards, 1997;& Taghavi, Moradi, Neshat-Doost, Yule, & Dalgleish, 2000); misinterpret non-hostile peer interaction as hostile (Bell-Dolan, 1995); and report more subjective physiological arousal in the absence of increased objective physiological arousal (Anderson & Hope, 2009) than their non-or-low anxious counterparts. This is the first study to report the presence of an interpretation bias for fearful prosody.

It is important to note that bias can exist on two levels; on the immediate interpretation of stimuli - the assignment of meaning to a stimulus (interpretation bias) and at the level of the motor output - making the response (response bias). Few studies are designed to distinguish between the two. However, in the present study if the observed anxiety related differences in criterion were a reflection of output response bias, one could expect to see differences in response bias between the high and low anxious groups across all emotional prosodies, which would reflect a participant’s tendency to favour one particular response (e.g. more liberal responding by anxious participants across all emotions). Differences in criterion,
Anxiety-related interpretation bias

however, were only observed for fearful voices, providing clear evidence that anxiety in
children is related to an interpretation bias or the assignment of fearful meaning to voices.
High anxious children were more likely than low anxious children to label all distracter
emotions as fearful.

In the literature anxious adults too have displayed negative interpretation biases with
words and sentences (e.g., Calvo & Castillo, 1997, 2001; Calvo, Eysenck, & Castillo, 1997;
Calvo, Eysenck, & Estevez, 1994; Eysenck, Mogg, May, Richards, & Mathews, 1991;
MacLeod & Cohen, 1993; Taghavi et al, 2000); story completion tasks (Huppert, Pasupuleti,
Foa & Mathews, 2007); social situations (Butler & Mathews, 1983; Kessler et al., 2007); and
faces (Yoon & Zinbarg, 2008). However no bias in the present adult sample was observed
and this was notably not due to ceiling effects, as the adults were not on the ceiling for any of
the emotions. It may be that anxious adults show an interpretation bias to some emotional
stimuli, however not to emotional prosody. The mechanisms underlying why the bias is not
shown for affective prosody are unknown, it may be that there is a particular characteristic of
vocal expressions of emotions that excludes this modality from interpretation biases in
adulthood. Given that there is so little research published exploring prosody perception in
anxious individuals, and none with a SDT approach, further investigation is warranted to
explore this developmental finding.

Evidence of a predisposing factor

The interesting dissociation that children who self-report higher levels of anxiety
demonstrate a tendency to assign negative (fearful) meaning to emotional tone of voice, but
adults who self-report high levels of anxiety do not demonstrate the same cognitive bias,
suggests that the observed interpretation bias may be associated with the development of
anxiety. That is, perhaps the tendency to interpret other’s tone of voice as fearful serves as a
marker of vulnerability to the further development of anxiety. It is therefore possible that in
adulthood this interpretation bias is attenuated such that adults acquire more knowledge about the emotional and social communication of others and thus are better able to make accurate interpretations of vocal expressions of emotions. A longitudinal study would be advisable to determine if interpretive bias in childhood reflects a vulnerability to clinical anxiety in adulthood. Evidence for this developmental attenuation comes from findings that non-anxious adults judge faces that are mild and moderately fearful as ‘less fearful’ (conservative bias) whereas the anxious adults do not show this tendency (Frenkel et al., 2008). Frenkel et al. suggest that the tendency in non-anxious adults to under-report fear in other’s faces is likely to be an effective time saving strategy in most everyday interactions, as it allows one to ignore mild, non imminent signals of threat. It is possible that this tendency develops in late childhood or early adulthood. This would explain why the children in the current sample did not demonstrate this conservative bias. In fact even low anxious children in this study demonstrate a more liberal criterion than low and high anxious adults. Clearly this is simply one possible explanation and warrants further study to explore its potential validity.

Further experimental support for the causal role of the observed interpretation bias in the development of anxiety comes from findings that inducing a negative bias to homophones increases anxiety reactivity to a subsequent stressful situation (Wilson, MacLeod, Mathews, & Rutherford, 2006). If the manipulation of interpretation bias in previously non anxious adults can cause a subsequent increase in anxiety for these individuals, it is likely that children who demonstrate a similar bias are thus vulnerable to develop further anxiety.

The present findings were observed in children with sub-clinical levels of anxiety therefore the relevance to clinical anxiety disorders is uncertain. Various cognitive models of adulthood anxiety (e.g. Clark & Wells, 1995 & Rapee & Heimburg, 1997) suggest that cognitive processes such as attentional biases are strongly involved in the maintenance of anxiety once established. However given that the anxiety effects were observed only in
children and not in adults, it is possible interpretation biases such as interpreting emotional tone of voice as fearful is involved in the development as well as the maintenance of anxiety.

Interpretation biases in anxious children are well established (Vasey & Macleod, 2001) and some aspects of biased information processing and cognitive processes have been encapsulated in earlier models of anxiety (Daleiden & Vasey, 1997; Vasey, Daleiden, Williams, & Brown, 1995) as well as contemporary, more integrative models (Vasey & Dadds, 2001). Vasey and Dadds suggest that cognitive biases and distortions mark just one of many factors that predispose to anxiety disorders. Such integrative models suggest that a predisposing factor such as an interpretative bias in isolation is not sufficient to develop further anxiety; however it is likely that most anxious children will begin their pathway with one of the predisposing factors, and the presence of one factor may increase the likelihood that other predisposing factors will occur (Vasey & Dadds, 2001). For example, a child who has a temperament high in behavioural inhibition traits (first risk factor; Lonigan & Phillips, 2001) will often by nature of their temperament, behave in a shy and avoidant manner. This behaviour will often elicit protective parenting behaviours (second risk factor; Dadds & Roth, 2001) which inadvertently decreases the child’s sense of control and autonomy, and limits their exposure to fearful situations (third risk factor; Dadds & Roth, 2001; Menzies & Harris, 2001) thus limiting their chances to master such contexts. Although there is a large evidence base for such risk factors, to my knowledge this is the first study to provide empirical support for the presence of an interpretation bias to emotional prosody in anxious children. Notably, the paradigm used in the present study does alloy for causality to inferred; it is possible that the interpretation bias observed is only a symptom of childhood anxiety. Hopefully this novel finding, however will encourage further research to explore how the observed interpretation bias may interact with various other individual, interpersonal and environmental factors to either produce and or be associated with childhood anxiety.
Research suggesting that individuals can be trained out of negative interpretation biases presents promising ideas for intervention given the current finding. Using a basic behaviourist approach, two studies demonstrated this training effect by presenting high anxious (yet sub-clinical) adults with various social scenarios. Participants were reinforced for resolving such situations in a positive or non-negative manner (e.g. the training phase). Following training, participants could not only demonstrate their learned skill in a further interpretation task, but also predicted that they would be less socially anxious in the future than the control group (Mathews, Ridgeway, Cook, & Yiend, 2007; Murphy, Hirsch, Mathews, Smith, & Clark, 2007). This promising research suggests that methods such as reinforcing more positive interpretations of prosody may also decrease the observed interpretation bias in anxious children with the expectation that such a change will decrease their level of anxiety. This research is consistent with the literature supporting the use of cognitive behavioural treatments that target information processing in therapy for childhood anxiety (Chorpita & Southam-Gerow, 2006). However such interpretation training is likely to be most effective if family members of the children are also involved, as research suggests the negative interpretation bias is highly influenced by parents’ interpretations (Barrett et al., 1996). It is also not clear that such training alone would be sufficient to decrease a child’s level of anxiety, as various other factors are known to be involved in the development of childhood anxiety, such as temperament (Kagan, 1999; Kagan & Snidman, 1999), social referencing (De Rosnay, Cooper, Tsigaras, & Murray, 2006) and parenting behaviours (Dadds & Roth, 2001) to name a few (for a comprehensive list of predisposing factors see Vasey & Dadds, 2001).

**Sensitivity**

Various methodological differences in the studies to date that have explored prosody perception in anxious adult (Quadflieg et al., 2007 and Freeman et al., 2009) and children
MClure & Nowicki, 2001) make comparisons difficult. Moreover none of these previous studies used SDT analysis. However of notable difference is that the two reported studies in adult samples observed anxiety related differences in prosody perception – although in differing directions, and the present study failed to find any anxiety-related differences in prosody perception measured by d’ for adults. In the first adult study, Quadflieg et al. examined prosody perception in 15 subjects with generalised social phobia and 15 controls. Anxious adults were significantly more superior at identifying fearful and sad prosody compared to non-anxious adults. More recently Freeman et al. examined prosody perception using the Aprosodia Battery in 11 male adults with combat related post traumatic stress disorder (PTSD). Adults with PTSD performed significantly poorer on the emotion identification task than controls (Freeman et al. 2009). Unfortunately only a global deficit in emotion recognition was reported and differences between the different discrete emotions were not analysed. Therefore the two previous studies with clinically anxious adults provided conflicting results, suggesting that clinical levels of anxiety (in this case social phobia and PTSD) may be associated with an altered capacity to interpret emotional prosody, however the nature of the difference, whether it is an enhanced ability or a deficit remains unclear.

There could be several reasons why no anxiety related differences were observed in the adult sample in this study when two previous studies have reported (albeit conflicting) anxiety-related differences. The present research population included participants with sub-clinical levels of anxiety. Previous reports of impaired identification of affective prosody in anxious adults was reported in a sample of 11 male veterans with an average age of 57 years and a clinical diagnosis of chronic and life long PTSD (Freeman et al., 2009). The observation of enhanced accuracy for negative affective prosody perception in anxious adults was reported in a study with 15 participants with a clinical diagnosis of social phobia (Quadflieg et al., 2007). The present study sample included 85 adults, male and female with
a mean age of 20 years. Therefore clear differences in population demographics including sex, age (Laukka & Juslin, 2007; Paulmann, Pell & Kotz, 2008) and level of anxiety is evident between the three studies, likely influencing three different outcomes. The present study found no anxiety-related effects in an adults sample with sub-clinical levels of anxiety. Clearly further research that carefully considers sample characteristics, subtypes of anxiety and experimental task is warranted.

Enhanced or increased sensitivity to negative emotional stimuli and poorer sensitivity to positive emotions in other modalities has also been noted (Veljaca & Rapee, 1998), however the majority of the SDT research fails to find differences in sensitivity (Sawchuk et al., 2002; Winton et al., 1995;) or suggests that anxiety is associated with poorer sensitivity to negative stimuli (Garner et al., 2000; Frenkel et al., 2008). Careful exploration of the methodologies of these studies reveals a clear distinction between those studies that report increased sensitivity and those that report poorer sensitivity. In the earlier study by Veljaca and Rapee, low and high socially anxious adults were asked to give a short 5 minute speech in front of an audience and detect positive or negative audience behaviours. Highly anxious adults displayed a higher sensitivity to negative audience behaviours (e.g. frowns) than positive audience behaviours (e.g. smiling), whereas low anxious adults showed the opposite pattern; that is they were more sensitive to positive feedback.

Compare this methodology with those of the studies reporting poorer sensitivity (Frenkel et al., 2008; Richards et al., 2007; Simonian et al., 2001) that tend to present all participants with emotional stimuli, e.g. a face depicting a discrete emotion for a short period, and require them to make a speeded decision. In this situation, the participant has only one task to do, and can allocate all their attentional resources to it. These notable methodological differences between studies may account for equivocal findings. It seems apparent that anxiety-related enhanced sensitivity is only observed when the task requires participants to
allocate or direct their attention to emotional stimuli while there are competing or concurrent stimuli present. Under such conditions, anxious individuals may display an attentional bias to negative information and thus demonstrate heightened sensitivity to detect such stimuli. However under conditions where there is no competition for attention, anxious individuals are no better at detecting negative stimuli.

This notion fits with theoretical models of anxiety (Clark & Wells, 1995; Rapee & Heimburg, 1997) which suggest that anxiety is exacerbated or maintained by a preferential attention to sources of negative emotion or threat in the environment. If this is the case, superior sensitivity will only be observed under conditions where attention can be allocated in different directions and not under conditions such as the present paradigm where participants’ attentional resources were entirely allocated toward the auditory stimulus. Under such conditions it is likely that if any differences in sensitivity were to be observed, they would be a reflection of actual ability to detect the emotion. In contrast it is possible that in studies that report differences in sensitivity, the observed differences in fact reflect the tendency of anxious individuals to allocate more attention to emotional stimuli – thus increasing their ability to detect them. A large body of literature has explored attention allocation in both anxious adults and children (Brotman et al., 2007; Clarke, MacLeod, & Shirazee 2008; Fulcher, Mathews, & Hammerl, 2008; Helfinstein, White, Bar-Haim, & Fox, 2008; Koster, Crombez, Verschueren, & De Houwer, 2004; Lee, & Telch, 2008; Mogg, Mathews & Eysenck, 1992; Mogg, Holmes, Garner, & Bradley, 2008; Telzer et al., 2008) and the consensus seems to be that anxious individuals demonstrate a bias to allocate more attention to negative emotional depictions.

Role of depression

Previous observations in anxious adults have reported differences in criterion for happy faces as well as anxiety-related differences in sensitivity (Frenkel et al., 2008), neither
of which was observed in this sample. However in that study, like many studies in this area, depression was not statistically controlled. The present study along with previous research (e.g. Qualdfleig et al., 2007) demonstrates the importance of controlling for depression. Depression in the present adult sample was associated with a liberal bias when angry was the target. Depression scores correlated negatively with angry bias indicating that the higher a participant’s depression score was, the more likely they were to label a sentence as angry during the angry block. When children and adults were analysed separately, this effect only held true for adults, indicating that a high level of depression in adulthood is associated with a liberal response bias to angry voices, but this pattern is not present in childhood. It is possible that such a pattern was not found in the child sample because their level of discriminability for angry voices was so low – at chance level. Depression in adulthood has previously been associated with decreased accuracy for recognising happy tone of voice (Qualdfleig et al., 2007) yet the influence on angry tone of voice observed in this sample is a new finding.

**Developmental findings**

The second key finding of the current study was the developmental difference observed in sensitivity to different affective prosodies. Adults were able to correctly discriminate all three target emotions; angry, fearful and happy. Yet children were only able to successfully discriminate one of the three emotional prosodies well; fearful. This was regardless of anxiety level. Children showed significantly better sensitivity for fearful tone of voice than they did for angry and happy, for both of which they demonstrated very poor sensitivity. The observation that children were in fact comparable to adults at discriminating fearful tones of voice suggests that they were able to understand the task and respond appropriately. Because block order was counterbalanced across participants, it also cannot be the case that poor discrimination of angry and happy reflected either practice or fatigue.
effects. Thus children’s poor sensitivity to angry and happy tones of voice likely reflects a true lack of discriminability, not an inability to understand the task.

The age group for the present study of children (aged eight to nine years) was chosen because previous research suggests that by this age, children should be fairly accurate when identifying all prosodies, while still performing slightly poorer than adults (Cohen et al., 1990; Cohen, Branch & Hynd, 1994; Doherty et al., 1999; Matsumoto & Kishimoto, 1983). At face value, the current findings suggest that perhaps children develop their ability to accurately perceive angry and happy emotional prosodies later in life than previous research suggests. However, as described earlier, given that each study assessing affective prosody perception tends to use its own idiosyncratic stimuli, drawing comparisons across studies is problematic. Furthermore some previous studies may have overestimated children’s ability to identify emotions by excluding children from analyses who used obvious strategic response strategies (Morton & Trehub, 2001), whereas no children were excluded from the analysis in the present study. It would be inappropriate to conclude from this study that children cannot discriminate angry and happy prosody from distracters. Rather in the context of the present experiment, children were worse than adults in the discriminability of anger and happiness while performing at adult levels in the discriminability of fear. What this finding does suggest, is that when children are faced with increased perceptual ambiguity, or when they find a task more difficult, differences in their ability to discriminate various discrete emotions may emerge.

Previous studies have tended to report a gradual steady increase between the ages of four and twelve in the ability to identify emotional prosody, but few have examined such differences across discrete emotional prosodies. One of the earliest studies to systematically examine the ability of children to perceive affective prosody, Dimitrovosky (1964), reported that children demonstrated the highest recognition for sad tones of voice, followed by angry,
happy and loving, respectively (fear was not assessed). However this finding must be considered in conjunction with the findings that the children in that sample also favoured negative emotions in their responding, i.e. responding “sad” and “angry” significantly more often than “loving” or “happy”. Furthermore, while there was a decrease in the number of errors made with age, the pattern of a negative response bias remained stable, suggesting it is particularly characteristic of childhood. This bias finding highlighting again the need for signal detection analysis to control for such tendencies.

The current study found highest sensitivity for fearful tones of voice. Studies to date have assessed children’s ability to identify sad, angry, mad, happy, loving and surprise prosody (Cohen et al., 1990, Cohen et al., 1994; Dimitrovsky, 1964; Doherty et al., 1999; Matsumoto & Kishimoto, 1983; Morton & Trehub, 2001), however to my knowledge, this is only the second study to date to measure children’s sensitivity to fearful prosody and the first study to observe the superior sensitivity for fear over other prosodies. Unlike Dimitrovsky’s (1964) finding of superior accuracy for negative emotions that was in fact driven by a tendency to produce more negative responses, the current study extracted measures of sensitivity and response criterion, and observed differences between emotional prosodies in sensitivity and not bias. In the only other known study to date to examine fearful prosody perception, Fujiki, Spackman, Brinton and Illig (2008) examined the ability of children with language impairment to identify various emotional prosodies from speech. The eight-to-ten year old control group demonstrated highest accuracy for happiness, followed by anger, sadness and fear respectively. Clearly, this is in contrast to the differences across emotions in the present study. One explanation may be the stimuli and task used. Participants in Fujiki et al.’s study listened to the following passage:

“It was the first day of school. I got ready early. I wanted to see who was in my class. I walked in my class and sat down. Pat came in and sat next to me. Then the
teacher walked in the room. I knew this year would be different.”

It is likely that children were using the prosody to aid their semantic interpretation of the passage, which is different to processing the prosody in isolation as the current study required.

From an evolutionary perspective it seems theoretically likely that fear is one of the first emotions that children learn to identify. Being able to accurately interpret fear in another’s voice appears more closely linked to survival than being able to interpret anger or happiness in others. Additionally there is some evidence that fear is one emotion that is preserved in older adults’ prosody perception relative to angry, happy and sad prosodies (Ruffman, Henry, Livingstone & Phillips, 2008).

In regards to relative identification of different discreet emotions, previous research suggests that surprise is one of the first emotions that Japanese and American children aged 4 learn to recognise before the recognition of sadness, happiness and anger (Matsumoto & Kishimoto, 1983). However that study did not include fearful prosody, and so a comparison of the development of fear and surprise warrants further investigation. It may also be that surprise and fearful prosodies share some of the same prosodic components, and are more perceptually salient to children. For example, both are characterised by a high fundamental frequency (Banse & Schere, 1996). The finding that accurate perception of anger may be one of the later skills to develop is consistent with previous findings examining the perception of discrete emotions in sequential age groups that found that anger was the last emotion, after surprise, happiness and sadness that Japanese children learnt to recognise (Matsumoto & Kishimoto, 1983; for contrasting results see control group in Fujiki et al., 2008).

Additionally, given that there is some evidence of cultural differences between the relative development of different emotions, future research should include such demographics in their
sample. Ethnicity and culture were not recorded in the present study, although all participants were native speakers of English.

General discussion

In summary, anxious children are not more sensitive to negative tones of voice as some previous literature suggests, rather, elevated levels of anxiety in childhood are associated with a liberal bias to interpret voices as fearful. This observed anxiety-related pattern is not present in anxious adults. Non-clinical levels of anxiety in adulthood do not appear to be associated with sensitivity to or a response bias when processing happy, angry and fearful tones of voice. The presence of such an interpretation bias in childhood may mark a vulnerability to develop further anxiety; however this tendency to interpret others’ tone of voice in a fearful manner may be attenuated in adulthood as a normative tendency to under-report fear becomes more adaptive.

The second key finding came from the examination of developmental trends in prosody perception regardless of anxiety. Eight and nine year old children are able to detect fearful tones of voice at a level comparable to adults, but they have great difficulty, and in fact very poor accuracy at detecting angry and happy tones of voice. This finding suggests that fear may be one of the first emotional states that children learn to decode from others’ tone of voice. Both findings encourage further research into these two domains to further explore the role that interpretation biases may play in the development of anxiety, and the development of prosody perception for discrete emotions.
References


Anxiety-related interpretation bias


Dear Principal,

We are writing to invite children in your school to participate in a research project investigating how people perceive and interpret emotional tone of voice and how a person’s level of anxiety may influence this interpretation. This project has been granted approval from the Victoria University Human Ethics Committee.

What is the purpose of this research?
This research will allow us to examine how people perceive and interpret emotional tone of voice and how a person’s level of anxiety may influence this interpretation.

Who is conducting the research?
- This study is being carried out by myself, Megan Humphrey, under the supervision of Dr Gina Grimshaw. This research is currently in the process of being approved by the University ethics committee.
- The research will contribute to My Masters thesis.

What is involved if you give consent for students at your school to participate?
- All aspects of this research will take place at your school and we aim to begin data collection early in 2009.
- The children will complete two short questionnaires where he/she will respond to questions such as “I worry people might get mad at me” and “I feel sorry for myself”. We anticipate that the questionnaires will take no more than ten minutes to complete.
- Additionally, the children will complete an emotion perception task in which he/she will listen to voices and judge the emotional tone of voice, e.g. happy voice, sad voice, neutral voice. We anticipate this task will take no longer than 20 minutes.
- During the research, each child is free to withdraw at any point before the experiment has been completed.

Privacy and Confidentiality
- We will keep all consent forms and information collected for at least five years after publication.
• The resulting Masters thesis will be kept in the Victoria University Library, and submitted for marking and publication.
• Individual children will never be identified in my research project or in any other presentation or publication.
• In accordance with the requirements of some scientific journals and organisations, each child’s coded questionnaire may be shared with other competent researchers and may be used in other related studies.
• A copy of the coded data will remain in the custody of Gina Grimshaw at Victoria University.

**What happens to the information that the children provide?**
- The overall findings may be submitted for publication in a scientific journal, or presented at scientific conferences.
- The overall findings will form part of a, PhD thesis, Masters thesis, or Honours research project that will be submitted for assessment.

Any questions may be directed to the principal research for this project, Megan Humphrey (detail above) **Thank you for taking the time to consider our request for participation in this research.**

Yours sincerely,

Megan Humphrey (MSc Student)
Statement of consent

I have read the information about this research and any questions I wanted to ask have been answered to my satisfaction.

I agree to participate in this research. I understand that I can withdraw my consent at any time, prior to the end of my participation.

Name: __________________________________

Signature: ________________________________

Date: ____________________________________

Student ID: _______________________________

Copy to:
[a] participant,
[b] researcher (initial both copies below)
What is the purpose of this research?
The purpose of this study is to help us understand how children perceive and interpret vocal expressions of emotion and how a child’s level of anxiety may influence their interpretation.

Who is conducting the research?
- This study is being carried out by Megan Humphrey, a Masters student in clinical psychology, under the supervision of Dr. Gina Grimshaw. This research has been approved by the University Human Ethics Committee.

What is involved if you give consent for your child to participate?
- All aspects of this research will take place at your child’s school
- Your child will complete two short questionnaires where he/she will respond to questions such as “I worry people might get mad at me” and “I feel sorry for myself”. We anticipate that the questionnaire will take no more than ten minutes to complete.
- Additionally, your child will complete an emotion perception task in which he/she will listen to voices and judge the emotional tone of voice, e.g. happy voice, sad voice, neutral voice.. We anticipate this task will take no longer than 20 minutes.
- During the research, your child is free to withdraw at any point before the experiment has been completed.

Privacy and Confidentiality
- We will keep your consent forms and information collected for at least five years after publication.
- The resulting Masters thesis will be kept in the Victoria University Library, and submitted for marking and publication.
- Your child will never be identified in my research project or in any other presentation or publication.
- In accordance with the requirements of some scientific journals and organisations, your child’s coded questionnaire may be shared with other competent researchers and may be used in other related studies.
- A copy of the coded data will remain in the custody of Dr. Gina Grimshaw at Victoria University.

What happens to the information that you provide?
• The overall findings may be submitted for publication in a scientific journal, or presented at scientific conferences.
• The overall findings will form part of a Masters thesis that will be submitted for assessment.

Thank you for your participation
As a token of our appreciation, we will give your child a small gift as a thank you for their time and effort, at the end of the study.

If you would like to know the results of this study, they will be available approximately from December 2009 and can be sent to you via email or mail if requested. If you would like to be sent a copy of these, please leave your email address on the returned consent form.

If you have any further questions regarding this study please contact any one of us above. Thank you for considering participation in this research.

Megan Humphrey
Statement of consent

I have read the information about this research and any questions I wanted to ask have been answered to my satisfaction.

I agree for my child to participate in this research. I understand that he/she can withdraw my consent at any time, prior to the end of their participation.

Name: __________________________________

Signature: __________________________________

Date: __________________________________

Student ID: __________________________________

Copy to:
[a] participant,
[b] researcher (initial both copies below)

Please return one copy of this to the school and keep one copy for your own information. Thank you.
Information Sheet and consent form

Megan Humphrey
MSc Student
megan.humphrey@vuw.ac.nz

Dr. Gina Grimshaw
Supervisor
gina.grimshaw@vuw.ac.nz

What is the purpose of this research?
This research will allow us to examine how people perceive and interpret emotional tone of voice and how this relates to their level of anxiety.

Who is conducting the research?
• We are a team of researchers in the School of Psychology at Victoria University of Wellington. Dr. Grimshaw is supervising this project. This research has been approved by the University Human Ethics Committee.

What is involved if you agree to participate?
• If you agree to participate in this study, you will complete a short questionnaire about your mood where you will respond to questions such as “I worry about the future” and “I feel sorry for myself”. We anticipate that the questionnaire will take you no more than five minutes to complete.
• You will also complete an emotion perception task in which you will listen to voices and judge the emotional tone of voice. We anticipate this task will take no longer than 20 minutes.
• During the research you are free to withdraw, without any penalty at any point before the experiment has been completed.

Privacy and Confidentiality
• We will keep your consent forms and survey for at least five years after publication.
• You will never be identified in my research project or in any other presentation or publication.
• In accordance with the requirements of some scientific journals and organisations, your coded survey may be shared with other competent researchers.
• Your coded data may be used in other, related studies.
• A copy of the coded data will remain in the custody of Dr. Gina Grimshaw at Victoria University.

What happens to the information that you provide?
• The data you provide may be used for one or more of the following purposes:
  • The overall findings may be submitted for publication in a scientific journal, or presented at scientific conferences.
  • The overall findings will form part of a Masters thesis that will be submitted for assessment.

If you would like to know the results of this study, they will be available approximately from December 2009 and can be sent to you via email or mail if requested. (If so, please write your email address on the consent form)

If you have any further questions regarding this study please contact any one of us above.

Thank you for considering participation in this research.

Megan Humphrey
Statement of consent

I have read the information about this research and any questions I wanted to ask have been answered to my satisfaction.

I agree to participate in this research. I understand that I can withdraw my consent at any time, without penalty, prior to the end of my participation.

Name: ________________________________

Signature: ______________________________

Date: ________________________________

Student ID: ______________________________

Copy to:
  [a] participant,
  [b] researcher (initial both copies below)
Debriefing statement

Thank you for participating in this experiment. This study examined how people perceive and interpret emotional tone of voice, and the ways in which this perception and interpretation may be influenced by their anxiety level.

Previous research suggests that anxious individuals are poor at detecting positive emotions in other modalities such as facial expressions or words, and that this may in fact be a maintaining factor which causes further anxiety. The same line of research also suggests that anxious individuals are better at interpreting negative emotions such as fear, sadness and anger in facial expression. Yet little research has explored emotional interpretation through tone of voice in anxiety. The small body of research to date suggests that anxious individuals are indeed poor at detecting positive prosody such as happiness and are enhanced at detecting negative prosody such as sadness and fearfulness.

However, from the findings to date, we do not know if anxious individuals are really better at interpreting negative tone of voice, or if they are simply more likely to report a voice as being negative. I am interested in understanding the exact mechanisms behind the interpretation bias. That is, are anxious individual better at detecting fear, anger or sadness in another’s voice, or are they simply more likely to report that all voices sound threatening regardless of the actual tone of voice. Likewise with positive stimuli, are they really poor at detecting happy tone of voice, or simply less likely to report it?

The outcome of this study will be important: Anxiety affects a large proportion of individuals during their lifetime. Having an understanding of specific deficits associated with anxiety will add great value to clinician’s abilities to understand and treat anxiety. For example, clinicians will be better informed on what types of thought patterns and perceptual impairments need to be the focus of treatment.

If you have any further questions regarding this study please contact any one of us above.
Thank you for participating in this research.

**Recommended Reading**

Debriefing statement

Thank you for consenting for your child to participate in this experiment. This study looked at how children perceive and interpret emotional tone of voice, and the ways in which this perception and interpretation may be influenced by their anxiety level.

Previous research suggests that anxious individuals are poor at identifying positive emotions in others, for example happy facial expressions. Additionally, anxious individuals are better at identifying negative emotions such as fear, sadness and anger in facial expression. There is some suggestion that these differences in the perception of others emotions may add to the difficulties faced by these people and may maintain their anxiety. The small body of research on vocal expressions of emotion suggests that anxious individuals are indeed poor at identifying positive emotional tone of voice such as happiness and are enhanced at identifying negative tones of voice such as sadness and fearfulness.

However, from the findings to date, we do not know if anxious individuals are really better at interpreting negative tone of voice per se, or if they are simply more likely to report all voices as being negative. I am interested in understanding the exact mechanisms behind the interpretation bias. That is, are anxious individuals better at detecting fear, anger or sadness in another’s voice, or are they simply more likely to report that all voices sound threatening to them regardless of the actual tone of voice. Likewise, are they really poor at detecting happy tone of voice, or simply less likely to report it?

The outcome of this research will have numerous valuable practical applications. Anxiety affects a large proportion of individuals during their lifetime. Having an understanding of specific deficits associated with anxiety at different ages will add great value to psychologists’ abilities to understand and treat anxiety. For example, professionals working with anxious children will be better informed on what types of thought patterns and
perceptual impairments need to be the focus of treatment and how this may vary as children grow up.

Given that the questionnaires used to measure anxiety in this study were used purely in a research context and not as a diagnostic tool, we are unable to make any informed comments regarding any individual child’s anxiety level. For this reason we will not be providing parents with their child’s individual scores on these measures. However, if you are concerned about your child’s anxiety or depression level, contact details of appropriate services have been provided on the bottom of this page.

Thank you to both yourself and your child for your involvement in this research.
Any queries please contact any one of us above.

Yours sincerely

Megan Humphrey and Dr. Gina Grimshaw

Contact regarding any mental health issues:

Victoria University Psychology Clinic
5th Floor, Easterfield Building
Victoria University of Wellington
Kelburn Parade

Phone: 04 463 6400
Email psychclinic@vuw.ac.nz