



# Look out captain, I hear an ambiguous alien! A study of interpretation bias and anxiety in young children



Suzannah Stuijzand<sup>a,b</sup>, Bhisudev Chakrabarti<sup>a</sup>, Shirley Reynolds<sup>a</sup>, Helen F. Dodd<sup>a,\*</sup>

<sup>a</sup> School of Psychology and Clinical Language Sciences, University of Reading, Harry Pitt Building, Earley Gate, RG6 6AL, UK

<sup>b</sup> Institute of Higher Education and Research in Healthcare (IUFHS), Lausanne University Hospital, University of Lausanne, Lausanne, Switzerland

## ARTICLE INFO

### Keywords:

Interpretation bias  
Anxiety  
Young children  
FEMG  
Ambiguity  
Development

## ABSTRACT

There is convincing evidence that anxious children and adolescents are biased to interpret ambiguity in a negative way (Stuijzand, Creswell, Field, Pearcey, & Dodd, 2017). However, little research examines interpretation bias in children under eight years. This is due to existing measures of interpretation bias being inappropriate for young children. Consequently, we aimed to develop a new interpretation bias task for young children using tones. Children learnt to associate high tones with a 'happy alien' and low tones with an 'angry alien'. They were then asked to classify tones from the middle of the frequency range (ambiguous tones) as 'happy' or 'angry'. Corrugator muscle activity was recorded alongside behavioural responses.

A community sample of 110 children aged 4–8 years, split into high and low anxious groups, completed the task. High anxious children were more likely to interpret the ambiguous tones as negative but this effect was small and only apparent after controlling for developmental factors. Corrugator activity aligned with behavioural responses for trained but not ambiguous tones. This is the first study to assess interpretation bias in young children using behavioural and physiological measures. Results indicate the task is developmentally appropriate and has potential utility for future research.

## 1. Introduction

Children as young as three show symptoms of anxiety (Egger & Angold, 2006), and anxiety disorders have been shown to be common in early and middle childhood (Cartwright-Hatton, McNicol, & Doubleday, 2006). Anxiety in children negatively affects school attendance and social competence (Settipani & Kendall, 2013; Velting & Albano, 2001). Left untreated, anxiety disorders are associated with depression and suicidal ideation in later life (c.f. Kendall, Safford, Flannery-Schroeder, & Webb, 2004). Understanding how anxiety develops and the mechanisms by which it is maintained may help us develop effective preventions and assist children onto a healthy trajectory.

Cognitive biases are implicated as having a predisposing (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007; Eysenck, 1992; Eysenck, 1997; Williams, Watts, MacLeod, & Mathews, 1988; Williams, Watts, MacLeod, & Mathews, 1997), causal (Beck & Clark, 1997) and/or maintaining role in anxiety (Bar-Haim et al., 2007; Eysenck, 1992, 1997; Mogg & Bradley, 1998; Williams et al., 1988, 1997).

Interpretation bias is a cognitive bias where there is an increased likelihood of interpreting ambiguity as negative/threatening. A recent

meta-analysis found a robust association between interpretation bias and anxiety in children and adolescents (Stuijzand, Creswell, Field, Pearcey, & Dodd, 2017). Stuijzand et al. (2017) also found that age moderated this association, with a stronger association found between negative interpretation bias and anxiety in older children/adolescents. However, because current methods of measuring interpretation bias rely on written tasks that require relatively advanced cognitive development and/or literacy often with verbal responses, very few studies included participants under 8 years. Thus, it is not clear whether an anxiety related interpretation bias exists in young children or if the association emerges as children develop, as suggested by the moderation effect reported by Stuijzand et al. (2017).

Field and Lester (2010) describe three models of the role of development in the relationship between cognitive biases and anxiety. In the first, biases are present early, are unaffected by development, but the extent to which they are manifest is influenced by individual factors such as anxiety (integral model). In the second, all children have biases early on but only those with certain individual factors (influenced by social, emotional and cognitive development) retain the biases as they age (moderation model). In the final model, biases emerge as a function of

\* Corresponding author.

E-mail addresses: [Suzannah.ravenscroft@chuv.ch](mailto:Suzannah.ravenscroft@chuv.ch) (S. Stuijzand), [b.chakrabarti@reading.ac.uk](mailto:b.chakrabarti@reading.ac.uk) (B. Chakrabarti), [s.a.reynolds@reading.ac.uk](mailto:s.a.reynolds@reading.ac.uk) (S. Reynolds), [h.f.dodd@reading.ac.uk](mailto:h.f.dodd@reading.ac.uk) (H.F. Dodd).

<https://doi.org/10.1016/j.brat.2019.103450>

Received 18 January 2019; Received in revised form 28 July 2019; Accepted 31 July 2019

Available online 31 July 2019

0005-7967/ © 2019 Elsevier Ltd. All rights reserved.

children's emotional, social and cognitive development. Anxiety may be a cause or a consequence of the emergence of these biases (acquisition model). Field and Lester (2010) suggested that an acquisition model was best supported by the available evidence. The moderation effect reported by Stuijzand et al. (2017) is consistent with an acquisition model, but more research with young children is required in order to confirm this.

Age is of course, only a proxy for development, and showing age effects does not give any insight into the underlying developmental processes. Effortful control is a developmental factor that may be important in the development of interpretation bias. Effortful control has been defined as the ability to inhibit a dominant response to perform a subdominant response, to detect errors, and to engage in planning (Rothbart and Rueda, 2005a). Effortful control may also be needed to regulate oneself in response to threat stimuli. Sufficient abilities to control attention and use inhibition may also be required to process the threat level of stimuli. Effortful control develops considerably before age 8 (Rothbart and Rueda, 2005b) with some components (focusing and inhibition) showing relative stability from around age 7 with little difference to adult proficiency (Eisenberg, Smith, Sadovsky, & Spinard, 2004; Mezzacappa, 2004). Kindt and van den Hout (2001) proposed that the development and use of effortful control helps most children to inhibit their instinctive bias towards threat and that children who are unable to inhibit this bias are at increased risk of developing anxiety disorders. Thus, effortful control may influence the association between interpretation bias and anxiety in children. Indeed, there is evidence that effortful control may moderate the association between interpretation bias and anxiety in children (Salemink & Wiers, 2012). Given the potential influence of effortful control on the relationship between interpretation bias and anxiety in children it seems pertinent to consider effortful control as a potential moderator.

Empirical tests of Field and Lester's (2010) models of development, cognitive bias and anxiety in children have been limited because of measurement limitations. Although child anxiety can be measured by proxy (i.e. parent report) current measures of cognitive bias rely heavily on the child's verbal and cognitive skills. For example, self-report measures, e.g. the Child Ambiguous Scenario Questionnaire by Barrett, Rapee, Dadds, and Ryan (1996) requires children to imagine situations, generate and hold multiple outcomes in mind, and to select a preferred interpretation from various alternatives. In this task children hear or read an ambiguous story and have to choose or produce an interpretation of what they think may be happening. Others explicitly make use of lexical knowledge, e.g. the homophone task by Gifford, Reynolds, Bell, and Wilson (2008). Here children hear a word that could be interpreted as having a threatening or none threatening connotation and children must select a picture to match the word they heard. Some tasks to assess interpretation bias have been developed with younger children in mind, for example the Story Stem Task (Dodd, Hudson, Morris, & Wise, 2012) and the Nature Reserve Task (Field & Storksen-Coulson, 2007) but while they do not require a verbal response from the child they still rely heavily on language comprehension. For example, the Story Stem Task requires children to play out the ending to an ambiguous story which is started for them by the experimenter. In the Nature Reserve Task children are given threat-related information about a novel animal and then avoidance to that animal is subsequently assessed (Field & Storksen-Coulson, 2007). This reliance on language ability means interpretation bias can only be assessed in children who can understand the task, which limits the age range of participants and leads to significant data loss (e.g. Dodd et al., 2012 report that data was only available for 65% of their participants). As a result we do not have a clear understanding of whether an anxiety-linked interpretation bias is present in younger children.

This study aims to assess interpretation bias in 4–8 year olds using a new child-friendly task. Children are first taught to associate different tones with the emotion of an alien (happy- positive or angry-negative) and are then presented with tones (high, mid and low frequency) and asked to say if the alien is 'happy' or 'angry'.

We first aim to assess if children aged between 4 and 8 years can complete the tone task and if the task is appropriate for measuring

interpretation bias. We will conclude that the task is suitable for assessing interpretation bias in young children if the following criteria are met: 1) children are able to achieve an accuracy rate of > 60% during the training trials and can maintain this learning criteria to the trained tones throughout the task; 2) responses to the trained tones are not dependent on level of anxiety, non-verbal or verbal cognitive abilities or moderated by age or effortful control; and 3) non-verbal and verbal abilities do not predict children's responses to the ambiguous tones.

We will also assess the validity of a physiological measure of ambiguity; corrugator activity. The corrugator is the small muscle behind the eyebrow that is activated when we frown. Corrugator activity has been shown to differentiate between visual and auditory stimuli with different valences in adults (Tan et al., 2011, 2012; Dimberg, 1990; Hawk, Fischer, & Van Kleef, 2012) and children (e.g. Deschamps, Schutte, Kenemans, Matthys, & Schutter, 2012; Tottenham, Phuong, Flannery, Gabard-Durnam, & Goff, 2013). Greater corrugator activity has been found in response to negative stimuli. The corrugator has also been found to react to stimuli perceived in a negative way, reflecting the individual's subjective experience of the stimuli (Dimberg, 1990; Tan et al., 2012). Recording corrugator activity may provide a unique way to measure children's interpretation of ambiguous material that is not reliant on self-reports from young children and is free from demand characteristics. It is hypothesised that there will be more corrugator activity when children hear tones associated with an angry alien than tones associated with a happy alien.

Next, we will investigate if symptoms of anxiety are associated with responses to the ambiguous tones. It is predicted that children with high anxiety will be more likely to report the ambiguous tones as 'angry' than low anxious children. Additionally, it is expected that children in the high anxious group will show more corrugator activity when presented with the ambiguous tones, possibly indicating a negative bias.

The moderating effect of age and effortful control on the relationship between interpretation bias and anxiety will be examined. Based on the meta-analysis of Stuijzand et al. (2017) we tentatively hypothesise that the strength of the association between responses to the ambiguous tones in the task and anxiety will increase with age. Finally we will control for traits of Autism Spectrum Disorder (ASD) because these are often comorbid with anxiety in children (van Steensel, Bögels, & Perrin, 2011) and the task involves emotion judgements in a social context, which may be affected by ASD traits.

## 2. Methods

### 2.1. Participants

Participants were recruited via a range of community groups and schools as well as advertisements and leaflets in public places and local media. Parents of 351 children registered an interest in the research and completed a measure of child anxiety: the Preschool Anxiety Scale (PAS; Spence, Rapee, McDonald, & Ingram, 2001) was completed for children aged 4–6 years and the Spence Children's Anxiety Scale (SCAS; Nauta et al., 2004) was completed for children aged 7 or 8 years. Children identified by parents as having a diagnosis of Autism Spectrum disorder, Attention Deficit Hyperactivity Disorder, or as having a Learning Disability were not invited to participate because these conditions have a propensity for anxiety and/or particular cognitive processing difficulties that could influence findings ( $n = 8$ ). One hundred and fourteen children were identified as having elevated symptoms of anxiety (1SD above the existing normed mean as reported on <https://www.scaswebsite.com>) or as having low levels of anxiety (below the existing normed mean as reported on [www.scaswebsite.com](http://www.scaswebsite.com)) and agreed to take part.

One-hundred and ten children completed the task (44 females,  $M_{age} = 5.69$ ,  $SD = 1.34$ ). Of these, 97 learnt the associations necessary to complete the task. There were no differences in age ( $t(108) = -1.791$ ,  $p = .076$ ), level of anxiety ( $X^2(1) = 0.50$ ,  $p = .481$ ), or gender ( $X^2(1) = 0.23$ ,  $p = .630$ ) between those who learnt the associations

necessary to complete the task and those that did not. The 97 that learnt the associations necessary to complete the task made up the final sample (38 females,  $M_{age} = 5.08$ ,  $SD = 1.19$ , age range 4–8 years). The final sample included 62 high anxious and 35 low anxious children.

In the final sample, the majority of parents were female (95%) and reported they were the primary caregiver (97%). Ethnicity data was available for 32 children, 94% of whom were White British (remainder were White British with Arabic 3%; Australian 3%). Within the sample there were twelve sibling pairs thus 85 families were represented by the sample.

## 2.2. Measures: parent completed

### 2.2.1. Spence Child Anxiety Scale (preschool version; PAS and child version; SCAS)

Parents of children aged 4–6 years completed the PAS, a 28 item questionnaire answered on a five point Likert scale from 0 (Not true at all) to 4 (Very often true) (minimum = 0, maximum = 112). The measure has strong psychometric properties aligned with DSM-IV diagnoses and good construct validity (Spence et al., 2001). The total score of the PAS showed excellent internal consistency in this sample ( $\alpha = 0.91$ ).

Parents of children aged 7 and 8 years completed the SCAS, a parallel measure for older children. The SCAS has 38 items answered on a four point Likert scale of 0 (never) to 3 (always) (minimum = 0, maximum = 114). The SCAS has good psychometric properties (Nauta et al., 2004; Spence, 1998) and showed excellent internal consistency in this sample ( $\alpha = 0.90$ ).

### 2.2.2. Child Behaviour Questionnaire – Effortful Control Scale (CBQ-EFC; Rothbart, Ahadi, Hershey, & Fisher, 2001)

The Children's Behaviour Questionnaire is a parent report measure that assesses the temperament of children aged 3–7 years old (CBQ; Rothbart et al., 2001). The CBQ is made up of three factors, one of which is a 47 item Effortful Control Scale consisting of the subscales low intensity pleasure, inhibitory control, perceptual sensitivity and attentional focusing (Rothbart et al., 2001). In this study, the five-item attention shifting subscale was also added to the Effortful Control Scale. This subscale has shown good internal consistency when combined with attention focusing and inhibitory control to assess effortful control (Eisenberg et al., 2007). Therefore, parents answered only a 52-item scale of Effortful Control derived from the CBQ on a Likert scale of 1–7 (1 = extremely untrue, 7 = extremely true, or 8 = not applicable; minimum = 52, maximum = 364). Higher scores indicate more effortful control. In this sample, internal reliability was excellent for the total 52 items of Effortful Control Scale ( $\alpha = 0.89$ ).

### 2.2.3. The Autism Spectrum quotient: Children's version (AQ: child; Auyeung, Baron-Cohen, Wheelwright, & Allison, 2008)

The AQ: Child is a 50 item parent report measure of autistic traits, which has good psychometric properties (Auyeung et al., 2008). Parents are asked to rate each item indicating the extent to which they agree or disagree with the statements about their child using a four-point Likert scale (0 = definitely agree to 3 = definitely disagree). The higher the score the more autistic-like traits (minimum = 0, maximum = 150). The full scale showed good internal consistency in this sample ( $\alpha = 0.83$ ).

## 2.3. Measures: child completed

### 2.3.1. The Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV)

The Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV) is an individually administered standardised test of cognitive development for children 2 years 6 months–7 years 7 months. The WPPSI is administered using a standardised manual and is a highly reliable and valid measure of child general intelligence (Weiss, Keith, Zhu, & Chen, 2013). The individual scales, rather than the full test, of the WPPSI have

been previously used for research purposes (for example, Bernier, Beauchamp, Bouvette-Turcot, Carlson, & Carrier, 2013). In this study only two individual scales were completed. The scales of vocabulary acquisition and non-verbal abilities were used to assess child verbal and non-verbal cognitive abilities, specifically: verbal comprehension and block design respectively. Age equivalence was used as a metric of verbal and non-verbal cognitive abilities. To facilitate comparison between participants we chose to use the WPPSI-IV with the entire sample, despite some being aged above 7 years 7 months. Eight participants were aged above 7 years 7 months and scored at the highest age equivalence. This means that their scores may be slightly underestimated.

### 2.3.2. Interpretation bias task: Ambiguous Tones Task

**2.3.2.1. Stimuli.** To create the tones, three frequencies were selected (low, middle, high), with the middle frequency being the exact mid-point of the high and low frequency. Five computer generated instruments (a guitar, a piano, a saxophone, strings and a Wurlitzer) were then used to play each of the three tones using Logic Studio 9 (2007). An independent group of 40 children (28 females) aged 4 to 8 ( $M_{age} = 6.78$ ,  $SD = 0.77$ ) rated the three sets of tones using a 5 point Self Assessment Manikin (SAM) rating scale from very happy to very angry. The high tones ( $M = 3.81$ ,  $SD = 0.40$ ) were judged to be more positive (happy) than the low tones ( $M = 2.41$ ,  $SD = 0.95$ ;  $t(39) = 6.76$ ,  $p < .001$ ,  $d = 1.75$ ) and the ambiguous tones ( $M = 2.89$ ,  $SD = 0.65$ ;  $t(39) = 6.32$ ,  $p < .001$ ,  $d = 1.45$ ). The low tones were judged to be more negative (angry) than the ambiguous tones ( $t(39) = -2.96$ ,  $p = .005$ ,  $d = 0.59$ ). No difference was found between the mean rating given for the ambiguous tones and a neutral rating on the SAM scale (3) ( $t(39) = -1.08$ ,  $p = .288$ ). The type of instrument used to play the tones was not associated with children's valence ratings, reaction time or trait anxiety score.

An image of a space captain with his thumb up (seen during the breaks in the experimental phase; see below), an alien expressing a happy face, and an alien expressing an angry face were produced. The alien expressing happiness and anger were identical except for facial expression (see Fig. 1). All images were created using Adobe Photoshop and Illustrator for the purpose of this study.

**2.3.2.2. Design and procedure of the task.** Parents remained in the room with the children throughout the task but sat out of the child's line of sight. The task was created in Eprime 2.0 (Psychology Software Tools Inc., 2012) and had three stages: introduction, learning and assessment. See Fig. 2 for a schematic representation of the three stages.

**2.3.2.2.1. Introduction phase.** Children were introduced to the task as 'A special job for the Captain of a space ship'. They were told that the Captain 'Wants to know whether the alien coming towards the spaceship is angry or happy; We can tell this by the sound they make.' Children were told that the Captain wanted them to learn which sounds the alien makes when it is happy and when it is angry. At

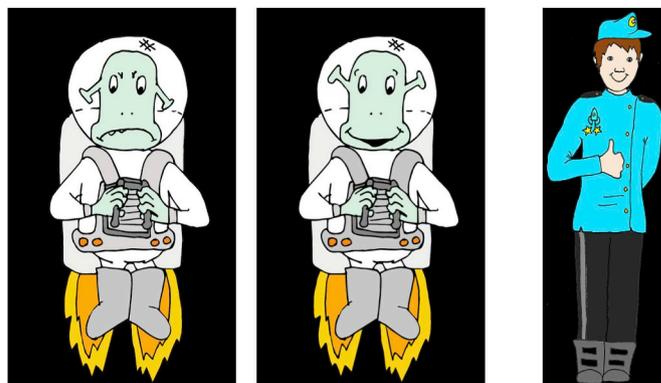


Fig. 1. Images used in the Ambiguous Tones task. From right to left: Angry Alien, Happy Alien and Captain with his Thumb up.

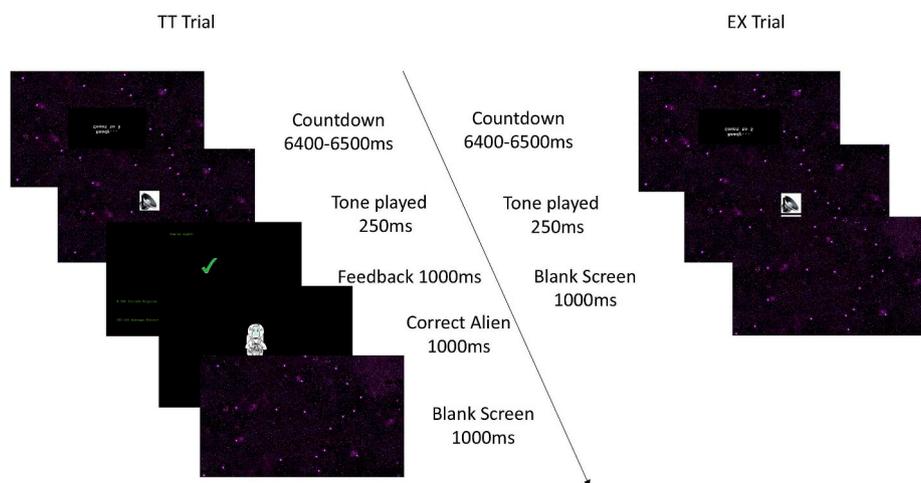


Fig. 2. Schematic Figure of the Stages of the Ambiguous Tones Task. \*only seen at the beginning of each block.

this stage the child was given headphones to wear, the picture of the ‘happy’ alien was displayed, and the high pitch tones were played. Next the picture of the ‘angry’ alien was displayed, and the low-pitched tones were played. Afterwards, children were asked if they could tell that the tones were different when the alien was happy compared to angry and they were asked how they were different. Children could hear the tones again or move on to the learning phase if it was clear they understood.

**2.3.2.2.2. Learning phase.** The learning phase included only trained tone (TT) blocks, which had 10 trials followed by a break. Each trial began with a wait of 6400 m–6500 m (jittered) during which the experimenter counted to five in the first 5000 m. After the wait period was complete a tone was played for 250 m while the child saw a space scene on the computer. The child was asked to state verbally whether they thought the tone belonged to a ‘angry’ or ‘happy’ alien. During the training block no ambiguous tones were presented. The experimenter recorded the child’s response on a mouse which then triggered on-screen feedback stating ‘You’re right’ and a green tick, or ‘Ooops’ and a red cross for 1000 m. The alien associated to the tone they had just heard was then displayed to provide further feedback. The five-high pitched and five-low pitched tones were played in each TT block in a randomised order. If the child had more than 60% of trials correct in one TT block the experimenter would ask the child if they would like to move on or do another TT block. If the child had fewer than 60% correct over one TT block the experimenter encouraged the child to do another block until they got more than 60% correct. The task was terminated if the child completed four blocks without getting more than 60% correct. Once the child achieved more than 60% correct on a block and said they were ready to move on they progressed to the assessment phase.

**2.3.2.2.3. Assessment phase.** This consisted of alternating one experimental block (EX) with one TT block. TT blocks were identical to the description given above in the Learning phase. An EX block consisted of seven or eight trials. Trials were the same as TT trials, but no feedback was given, and a blank screen was seen between the trials. Within each EX block five ambiguous tones (i.e. mid frequency) were presented with two or three trained tones (i.e. high and low frequency). The order of trials within blocks was randomised. Children could complete up to four EX blocks and four TT blocks in the assessment phase. The TT blocks were included in the assessment phase to maintain the learnt associations between the tones and the ‘happy’ and ‘angry’ aliens. Before starting each EX block children were told that this time the computer would not tell them if they were getting the answers right or not but that if they were doing a good job they would see the captain with his thumb up at the break. During breaks children always saw the captain with his thumb up for motivation. The task ended when the child had completed four experimental blocks or if the child stated that they did not wish to continue.

## 2.4. Procedure

The study had ethical approval from the University of Reading Ethics Committee (ref: 2014-025-HD). Participants attended a research session on campus during which the interpretation bias task and questionnaires outlined above were completed. During the same session participants also completed an eye-tracking task, the findings of which will be reported separately (Stuijzand, Stuijzand, Reynolds & Dodd, in preparation.). In addition, we also collected the following data, which we have no plans to publish: maternal anxiety; interpretation of ambiguous scenarios (for 7 and 8 year olds only); skin conductance response (SCR). The SCR data was not analysed because of poor data quality due to difficulty getting a good connection on children with small fingers and movement artefacts.

At the start of the session parents provided written consent and the child provided verbal assent to participate. Parents then completed questionnaires while the child completed their tasks. The order in which the children completed the tasks was based on how settled the child was but generally the children completed the eye-tracking task, the WPPSI subtasks and then the interpretation bias task. On completion of all the tasks the parent was given a debrief sheet further explaining the purpose of the study and £5 towards travel expenses. All children received a certificate, stickers and a token prize for their co-operation and time.

## 2.5. Recording physiological data

All physiological data was recorded using ADI Power Lab 8T, with an Octal Bioamp (AD Instruments, Australia) and acquired using Lab Chart 7.0 (AD Instruments, Australia).

### 2.5.1. Facial electromyography (fEMG)

Prior to starting the fEMG recording, small stickers were placed on the face of the child and the experimenter to help the child to understand where the electrodes would be placed. Once the child provided assent, the areas above the child’s left eyebrow and the middle of the forehead were cleaned with 70% alcohol prep pads (Professional Disposables, Inc., USA TD-230) and left to dry for 5 min to reduce impedance. Bipolar sensors were then placed above the eyebrow to measure corrugator activity, with the reference ground placed on the forehead using 4 mm Ag/AgCl EMG surface sensors (Discount Disposables, USA) on 5 mm collars filled with isotonic electrode gel. Children were asked to make facial expressions (angry, happy, and surprised) as a check to ensure signal from the sensors was adequate and placement was accurate.

Offline the fEMG signal was filtered using a 50Hz notch filter to remove electrical noise, and a band pass filter (500–10 Hz) was applied.

**Table 1**  
Details of Level 1 and 2 predictors included in each of the six models.

Model	Level 1 predictors	Level 2 predictors
Model 1	Valence (happy or angry sound)/ Response (happy or angry response)/Accuracy	Anxiety Autistic Quotient
Model 2	As Model 1	Model 1 + Age Age <sup>2</sup>
Model 3	As Model 1	Model 2 + Anxiety*Age Anxiety*Age <sup>2</sup>
Model 4	As Model 1	Model 2 + Effortful control
Model 5	As Model 1	Model 4 + Anxiety*Effortful Control
Model 6	As Model 1	Model 4 + Non-verbal ability Verbal ability

*Note.* Age<sup>2</sup> was included to allow for assessment of a quadratic influence of age as the influence of development may not be linear. Interactions were included to investigate moderation. Dummy variables were created for binary variables. For the variables level of anxiety, valence, response, and accuracy the reference categories were the low anxious group, positive, 'happy' and 'not accurate'. All continuous predictors were grand-mean centred (see Hox et al., 2010).

The signal was rectified, and a logarithmic transformation applied. The signal was then visually checked for signal quality and movement. The trial data was excluded if: the checks for facial expression movement were not visible in the signal; the sensor lost contact with the skin; the signal proved to be three standard deviations above or below the mean signal. As described in the procedure children counted down from five to one before stimulus onset. In fact, this countdown window was 6400–6500 ms (jittered). This was done to ensure that the countdown occurred in the first 5000 ms leaving at least an extra 1400 ms for any fEMG response to the countdown to have resolved before onset of the tone. Plots of fEMG activity following tone onset were created and visually inspected to determine the window in which the peak response to the tone would be best captured. Following this, the response window was taken as the first 300 ms after onset of the tone.

Area under the curve was calculated from stimulus onset to 300 ms post stimulus onset using the integral function in Lab Chart 7.0 (AD Instruments, Australia).

## 2.6. Statistical analysis plan

Multilevel modelling was used to analyse data from the training trials and experimental trials from the assessment phase. This was done to address the nested nature of the data: trials within children, and children within families (note some participants were siblings; Field & Wright, 2011). Also children varied in the number of trials they completed on the interpretation bias task (10–20; Hox, 2010; Hox, Moerbeek, & van de Schoot, 2010) and the anxiety groups were not of equal size. Multi-level analysis has less restrictive assumptions of variance and a balanced design (Field & Wright, 2011; Hox, 2010). The study was adequately powered for this type of analysis with more than 50 participants having valid responses to 10 or more trials involving ambiguous tones on the interpretation bias task (Bell, Morgan, Kromrey, & Ferron, 2010; Maas & Hox, 2005; Snijders, 2005). We also restricted our multi-level models to including random intercepts, thus random slopes were not included to maximise the available power for moderation analyses.

Two multilevel models were created to investigate the behavioural data: Model A investigated the influence of anxiety, age and effortful control on children's accuracy in the TT in the assessment blocks; Model B investigated whether anxiety, age and effortful control affected responses to the ambiguous tones and whether anxiety interacted with either age or effortful control to predict responses. In models A and B,

the dependent variables were binary, so logistic regressions were used and therefore model change in variance explained by each level was not meaningful and is not commented on.

Two parallel models were created to investigate corrugator activity during the task. Model C investigated the influence of anxiety, age and effortful control on corrugator activity in relation to angry and happy responses during the TT in the assessment blocks. Model D investigated the influence of anxiety, age and effortful control on corrugator activity in relation to angry and happy responses to ambiguous tones and whether anxiety interacted with either age or effortful control to predict corrugator activity to ambiguous tones. Psycho-physiological data resulted in continuous dependent variables therefore standard multi-level models were used.

Data was processed and analysed using R studio version 1.0.153 (R Core Team, 2015). R packages lme4 (Bates, Maechler, Bolker, & Walker, 2015), Pastecs (Grosjean & Ibanez, 2014) and VGAM (Yee, 2014) were used to conduct the analysis. For each model, the first step involved creating two random intercept models: 1) where the intercept was allowed to vary between participants (i.e., a two level random intercept model), 2) where the intercept was allowed to vary between both participants and families (i.e., a three level random intercept model). Using a likelihood ratio test, these models were compared against each other, as well as to a single level linear regression model. This, in combination with examining how much variance was accounted for at each level, was used to establish the appropriate model for each analysis.

Once the appropriate model had been established, predictors were added such that six models were compared (see Table 1).

Each model was compared to the random intercept model using the likelihood ratio test to assess whether adding predictors improved model fit. Should the model not show improvement the simplest model was retained i.e. the random intercept model. If a model improved model fit, subsequent models were then compared to this improved model and so on until no further models showed improvements indicating the final model had been established.

## 3. Results

### 3.1. Data preparation

Of the 97 children who learnt the associations between the frequency of the tone and the 'happy' and 'angry' aliens, 73 children had 10 or more valid trials involving the ambiguous tones. Reasons for not having enough valid trials included a technical difficulty with Eprime that led to data loss (16 participants), the remainder did not complete enough experimental blocks or did not provide responses to enough trials. There were no differences between those with 10 or more valid trials and those who had fewer than 10 valid trials on age ( $t(95) = -0.993, p = .323$ ), anxiety group ( $\chi^2(1) = 0.105, p = .746$ ), and gender ( $\chi^2(1) = 0.083, p = .773$ ) and in percentage correct that children achieved in their final practice block ( $t(95) = -4.70, p = .600$ ).

Table 2 shows the mean proportion of ambiguous tones the children reported as 'angry' (negative). The proportion of ambiguous tones reported as negative was significantly different from 50% in a one sample  $t$ -test ( $t(72) = 3.02, p = .002, d = 0.71$ ) with a medium effect. This indicates that all children had a bias towards reporting ambiguous tones as negative.

After data pre-processing and cleaning, 62 children had valid data on the fEMG measure for analysis. There were no differences in age ( $t(71) = 1.574, p = .120$ ), anxiety group ( $\chi^2(1) = 0.243, p = .622$ ), or gender ( $\chi^2(1) = 2.296, p = .130$ ) and in percentage correct that children achieved in their final practice block ( $t(71) = -0.431, p = .622$ ) between those with or without valid data. Table 2 shows the descriptive statistics of the fEMG data for both the TT and ambiguous tones from the assessment phase.

#### 3.1.1. Differences between high anxious and low anxious groups

Descriptive statistics for high and low anxiety groups can be found in Table 2. These means are reported purely for descriptive purposes.

**Table 2**  
Descriptive statistics for variables of interest by anxiety grouping.

Variables	Total (N = 73)		High Anxious group (N = 46)		Low Anxious Group (N = 27)	
	Mean	SD	Mean	SD	Mean	SD
Age	5.86	1.42	5.84	1.35	5.89	1.55
AQ total	65.33	18.54	71.09 <sup>a</sup>	14.80	56.15 <sup>a</sup>	20.39
ECS total	4.58	1.02	4.38 <sup>a</sup>	1.14	4.91 <sup>a</sup>	.68
Non-Verbal	5.65	1.28	5.61	1.31	5.71	1.25
Verbal	5.89	1.27	5.73	1.25	6.15	1.30
Prop. Ambiguous Tones Negative	56.01	19.58	58.04	20.67	52.36	17.21

<sup>a</sup> Difference between groups  $p > .05$ . The AQ total = Total score of Autistic Quotient score. ECS total = Total score of the Effortful Control Scale. Prop. Ambiguous Tones Negative = proportion of Ambiguous tones reported as Angry 'Negative'. Non-Verbal = Age equivalence on non-verbal cognitive ability task. Verbal = Age equivalence on verbal cognitive ability task.

Comparison of groups using these means is not appropriate given the clustered nature of the data and the difference in trial numbers between participants. Instead, these differences are investigated in the multi-level analyses. There was no age difference ( $t(71) = -0.155, p = .877$ ) or gender difference ( $\chi^2(1) = 0.002, p = .962$ ) between anxiety groups. However, the high anxious group had lower levels of effortful control than the low anxious group (Mann Whitney U = 870,  $p = .0045$ ) (see Table 2). Anxiety groups also differed on total autistic quotient scores, with the high anxious group having more autistic traits than the low anxious group (see Table 2) with a large effect size ( $t(68) = 3.55, p = .0007, d = 0.86$ ). There were no other significant differences between children with high anxiety and those with low anxiety symptoms. The autistic quotient score is included as a covariate in all analyses.

### 3.1.2. Did children maintain learning?

Table 3 shows that, on average, children were able to correctly identify more than 60% of the tones throughout the TT blocks in the assessment phase of the task. This was maintained on average across all experimental blocks (Table 4) and by children in both anxiety groups. Table 3 also indicates that over 70% of all children, in both the anxiety groups maintained their accuracy over the 60% threshold in each block and across all blocks. These results provide evidence towards satisfying criterion 1, of whether this task was appropriate for assessing interpretation bias in our sample.

## 3.2. Behavioural data

### 3.2.1. Multilevel analysis model A: trained tones

The first multi-level analysis (Model A) focused on the influence of anxiety and developmental proxies on accuracy in the TT blocks in the Assessment Phase. Here the dependent variable was accuracy (accurate or not accurate). Following the steps described in the data analysis plan a two-level model (trials within participants) was accepted as the appropriate analysis for the data and Model 6 was selected as the final model. The tests and decision making processes to arrive at this model can be found in Appendix 1 along with the coefficients, variance and model fit statistics for each of the models.

The results of the final model can be seen in Table 4. In the final model the only predictors influencing accuracy on the practice trials are the linear and quadratic age terms, with small effects (Chen, Cohen, & Chen, 2010). The positive coefficient indicates that, at the average of all other variables, as age increases, the likelihood of the child being accurate increases (Table 4). The negative quadratic term of age suggests that accuracy initially increases quickly with age but slows towards the older ages in the range (Table 4). There was no evidence that anxiety, non-verbal or verbal cognitive abilities or age or effortful control affected performance on the trained tones. These results contribute

towards evidence for criterion 2 of whether this task was appropriate for assessing interpretation bias in our sample.

### 3.2.2. Multilevel analysis model B: ambiguous tones

In the second multi-level analysis (Model B) the dependent variable was whether the participant reported the ambiguous tones as happy or angry per trial. This analysis was to test the hypothesis that children high in anxiety will be more likely to report the ambiguous tones as 'angry' than low anxious children and to examine the moderating effect of age and effortful control on the relationship between attention bias and anxiety. Following the steps described in the data analysis plan, a two-level model (trials within participants) was accepted as the appropriate analysis for the data and Model 6 was selected as the final model. The tests and decision-making processes to arrive at this final model can be found in Appendix 2 along with the coefficients, variance and model fit statistics for each of the models.

The final model results are shown in Table 5. In the final model, in the presence of the other predictors, anxiety was a significant predictor of interpretation bias (i.e. negative classification of the ambiguous tone ( $B = 0.42, OR = 1.53, STDError = 0.21, CI(95\%) = 1.00-2.35, p = .05$ ). The positive coefficient of the level of anxiety indicates that those with high anxiety were more likely to interpret the ambiguous tone as negative (Chen et al., 2010). Age and effortful control did not moderate the relationship between responses to ambiguous tones and anxiety. Verbal and non-verbal cognitive abilities were included in the final model but did not predict responses to the ambiguous tones; thus satisfying criterion 3 of whether this task was appropriate for assessing interpretation bias in our sample.

## 3.3. Psychophysiological data: fEMG

Descriptive statistics for corrugator activity in response to the trained and ambiguous tones can be found in Table 6.

### 3.3.1. Multilevel analysis model C: trained tones

This analysis was conducted on the fEMG data from the TT blocks in the Assessment Phase (Model C) (see Table 7). The intention of this analysis was to test the hypothesis that there would be more corrugator activity when children hear tones that they associate with an angry alien than tones they associate with a happy alien. For this analysis, participants' responses to trained tones (whether the tone was reported as 'angry' or 'happy') or the valence of the tone (whether the tone was trained to be classified as 'angry' or 'happy') could have been included as predictors of corrugator activity. As expected, response and tone valence were significantly correlated ( $r = 0.63, p < .001$ ). Given this, for Model C, two sets of models were created where valence and response were entered as level 1 predictors respectively. Only the model with response as a predictor included any significant predictors. For brevity this model is reported but the models investigating valence as a level 1 predictor can be found in Appendix 3.

Following the steps described in the data analysis plan Model 6 was selected as the final model. The tests and decision-making processes to arrive at this final model can be found in Appendix 3 along with the coefficients, variance and model fit statistics for each of the models. Only "Response" was a significant predictor of corrugator activity to the trained tones, with a small effect ( $\beta = 0.0006, B = 0.048, STDError = 0.0003, CI(95\%) = 1.00-1.001, t = 2.17, p = .03$ ). When the child reported the tone as coming from an alien who was 'angry' there was greater corrugator activity. Anxiety group and developmental proxies had no significant influence on corrugator activity to the trained tones.

### 3.3.2. Multilevel analysis model D: ambiguous tones

This analysis was conducted on the fEMG data from the ambiguous tones (Model D). The intention of this analysis was to test the hypothesis that children in the high anxious group would show more corrugator activity when presented with the ambiguous tones, possibly

**Table 3**

Average accuracy across participants per block on Trained Tones within the TT blocks of Assessment Phase and within the EX blocks of the Assessments separately, split by Anxiety.

	Total Sample			High Anxious			Low Anxious		
	N	M (SD)	% <sup>a</sup>	N	M (SD)	% <sup>a</sup>	N	M (SD)	% <sup>a</sup>
TT blocks Assessment Phase									
block 1	72	75.15 (18.93)	74.5	45	74.07 (19.42)	68.3	27	76.95 (18.30)	85.3
block 2	47	81.74 (17.73)	82	32	80.42 (14.87)	78	15	84.55 (23.04)	90
block 3	33	75.18 (16.75)	82.1	20	74.65 (17.28)	84	13	75.98 (16.57)	78.6
block 4	6	76.48 (22.38)	75	4	79.72 (26.95)	83.3	2	70 (14.14)	50
EX Blocks Assessment Phase									
block 1	73	74.20 (31.43)	74.2	46	72.46 (33.19)	71	27	77.16 (28.55)	80
block 2	71	79.11 (31.71)	78.5	45	81.11 (32.10)	74.6	26	75.64 (31.35)	85.3
block 3	46	77.17 (32.65)	73.7	31	74.73 (34.12)	70.5	15	82.22 (29.86)	79.4
block 4	32	79.17 (29.33)	77.4	19	74.56 (33.04)	74.6	13	85.90 (22.41)	82.4
All blocks	73	76.85 (24.26)	78.9	46	76.71 (24.93)	73.8	27	77.11 (23.55)	88.2

Note. TT = Trained Tones Blocks, EX = Experimental Blocks. Min = minimum average percentage accuracy, Max = maximum average percentage accuracy.

<sup>a</sup> Percentage who achieved over 60% accuracy.

**Table 4**

Final Multi-level Model predicting performance accuracy on the TT blocks from the assessment phase (Model A:6).

Fixed Effects	Estimate	Standard Error	95% CI	Z	p
Valence	.13	.12	.91–1.44	1.13	0.26
Anxiety grouping	–.39	.24	.41–1.08	–1.65	0.10
AQ	.01	.01	1.00–1.92	1.40	0.16
Age	.36	.14	1.08–1.92	2.48	0.01
Age <sup>2</sup>	–.17	.07	.73–.98	–2.22	0.03
Effortful Control	.11	.17	.80–1.57	0.66	0.51
Non-Verbal Cog	.11	.13	.87–1.44	0.87	0.38
Verbal Cog	.03	.11	.83–1.29	0.30	0.76
Random Effects					
Variance					
Intercept	.46				

Note. Non-Verbal Cog = Non-verbal cognitive abilities; Verbal Cog = verbal cognitive abilities.

R model equation: practice accuracy ~ valence + anxiety group + Autistic Quotient Total + Age + Age<sup>2</sup> + Effortful Control Scale + Age equivalence on Block Design + Age equivalence on Receptive Language + (1 | id).

**Table 5**

Final multi-level model investigating children's responses to ambiguous tones in the assessment phase (Model B:6).

Fixed Effects	Estimate (OR)	Standard Error	95% CI	Z	p
Anxiety grouping	.42 (1.53)	0.21	1.00–2.35	1.98	0.05
AQ	.005 (1.01)	0.01	.99–1.02	0.85	0.40
Age	.01 (1.01)	0.13	.78–1.32	0.11	0.91
Age <sup>2</sup>	.08 (1.09)	0.07	.95–1.25	1.19	0.24
Effortful Control	–0.01 (.99)	0.15	.73–1.36	–0.05	0.96
Non-Verbal Cog	–0.05 (.95)	0.12	.76–1.20	–0.42	0.68
Verbal Cog	0.11 (1.11)	0.10	.91–1.36	1.05	0.29
Random Effects					
Variance					
Intercept	.30				

Note. Non-Verbal Cog = Non-verbal cognitive abilities; Verbal Cog = verbal cognitive abilities.

R model equation: response ~ anxiety group + Autistic Quotient Total + Age + Age<sup>2</sup> + Effortful Control Scale + Age equivalence on Block Design + Age equivalence on Receptive Language + (1 | id).

indicating a negative bias. Following the steps described in the data analysis plan Model 5 was accepted as the final model. The results are shown in Table 8. The tests and decision-making processes to arrive at this final model can be found in Appendix 4 along with the coefficients, variance and model fit statistics for each of the models. Within this

**Table 6**

Descriptive Statistics from Corrugator activity (fEMG), assessed via Area Under the Curve (mV) by Total Sample and by Anxiety Group.

	Total Sample			High Anxious			Low Anxious		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Trained tones <sup>a</sup>	62	.0004	.002	39	.0002	.0008	23	.0008	.004
Ambiguous Tones	62	.001	.003	38	.003	.002	24	.001	.004

<sup>a</sup> Only those from the assessment phase of the task.

**Table 7**

Final multi-level model investigating children's fEMG responses to the TT trials in the Assessment phase (Model C:6).

Fixed Effects	Estimate (B)	Standard Error	95% CI	t	p
Response	.0006 (.048)	.0003	1.00–1.001	2.17	.03
Anxiety grouping	–.0006 (–.05)	.0007	.99–1.00	–.79	.43
AQ	.000007 (.02)	.00002	.99–1.00	.36	.72
Age	–.0004 (–.09)	.0004	.99–1.00	–.95	.35
Age <sup>2</sup>	.0002 (.07)	.0002	.99–1.00	.87	.39
Effortful Control	–.0003 (–.04)	.0005	.99–1.00	–.59	.56
Non-Verbal Cog	–.0003 (–.05)	.0004	.99–1.00	–.61	.54
Verbal Cog	.000006 (.12)	.0003	.99–1.00	1.77	.08
Random Effects					
Variance					
Intercept	.000005				

Note. Non-Verbal Cog = Non-verbal cognitive abilities; Verbal Cog = verbal cognitive abilities.

R model equation: Area under the curve ~ response + anxiety group + Autistic Quotient Total + Age + Age<sup>2</sup> + Effortful Control Scale + Age equivalence on Block Design + Age equivalence on Receptive Language + (1 | id).

model none of the individual predictors reached significance (see Table 8). This indicates that corrugator activity in response to the ambiguous tones was not influenced by whether the child said it was happy or angry, their level of anxiety, age or effortful control.

#### 4. Discussion

Cognitive models of anxiety propose that dysfunctional cognitions may play a predisposing, causal or maintaining role in anxiety (e.g. Bar-Haim et al., 2007; Beck & Clark, 1997; Eysenck, 1997). There is convincing evidence that anxiety is associated with an interpretation bias such that anxious individuals are more likely to interpret ambiguity in a threat-related way (Stuijzand et al., 2017). However, due to limitations with existing methods, research examining interpretation bias in young

children is scarce. Consequently, the first aim of the present study was to develop a novel task for assessing interpretation bias in children aged 4–8 years and to assess whether the task was appropriate for this age group. At the outset, we set three criteria for evaluating whether the task was appropriate, based on performance with the trained tones. The first was that children could correctly classify > 60% of the trained tones accurately and maintain this level of performance throughout the task. The results indicated that most children (75%) could do this and that this level of performance was maintained throughout the task. The second criterion was that performance on the trained tones should not be dependent on anxiety, verbal or non-verbal cognitive abilities, age or effortful control. This criterion was partially met: whether children initially learned the associations was not related to age, gender level of anxiety, verbal or non-verbal cognitive abilities or effortful control. However, older children performed better on the trained tones during the assessment phase than younger children. The third criterion was that performance on the ambiguous tones should not be related to verbal and non-verbal cognitive abilities. This was supported: multi-level analyses showed that neither verbal nor non-verbal cognitive abilities significantly predicted responses to ambiguous tones.

Our second aim was to assess whether corrugator activity provided a physiological measure of children's valence appraisal of ambiguity. Again, focusing on the trained tones, we anticipated, based on previous research that children would exhibit more corrugator activity to the trained tones they classified as belonging to the alien when angry when compared to the trained tones they classified as belonging to the alien when happy. This was supported and is in line with Dimberg (1990) and Tan et al. (2012) who also found that corrugator activity reflected participants' subjective experience of the stimuli. Corrugator response to trained tones was not predicted by age, gender, verbal or non-verbal ability, effortful control or anxiety. This provided a solid foundation for using the corrugator to explore responses to ambiguous tones.

Given that the task appeared to be appropriate for this age group, we then assessed children's behavioural and physiological responses to ambiguous (i.e. mid frequency) tones. Overall, this sample of children had a slight bias to interpret the ambiguous tones as negative (angry). It was hypothesised that children with high anxiety will be more likely to report the ambiguous tones as 'angry' than low anxious children. Additionally, it was expected that children in the high anxious group would show more corrugator activity when presented with the ambiguous tones, possibly indicating a negative bias. These hypotheses were not strongly supported. Multi-level analysis of responses to ambiguous tones indicated that children with elevated anxiety were more likely to interpret an ambiguous tone as angry than children who were not anxious, but this effect was small and only significant when age and effortful control were added to the model. Although this supports hypothesised associations between anxiety and interpretation bias, we should be cautious in interpreting our results as convincing evidence for an association in young children. In line with this, it is plausible that the addition of developmental proxies to the model simply accounted for enough variance to push the difference between anxiety groups over the significance threshold. It is therefore also plausible that had different predictors (not related to development) been added to the model they may have had the same effect. Thus, we cannot state that only in the presence of developmental variables did the differences between anxiety groups become evident.

For corrugator activity to ambiguous tones, there was no evidence that anxiety was a significant predictor. Importantly, corrugator activity did not align with the response the children made to ambiguous stimuli (angry/happy), which is inconsistent with the findings for the trained tones and previous research. For example, Tottenham et al. (2013) found the corrugator to be a physiological index of children's appraisal of ambiguous stimuli.

In summary, the results provide some indication for an anxiety-linked interpretation bias in the behavioural data but there was no evidence for this bias in the physiological data. There are a number of possible explanations for this inconsistency. First, our results may indicate that

**Table 8**

Final multi-level model investigating children's fEMG responses to the ambiguous trials in the Assessment phase (Model D:5).

Fixed Effects	Estimate (B)	Standard Error	95% CI	t	p
Response	.0009 (.04)	0.001	1.002–1.003	1.27	.34
Anxiety grouping	-.001 (-.07)	0.001	1.00–1.002	-1.32	0.20
AQ	.0001 (.01)	0.0003	1.00–1.001	0.20	0.85
Age	-.0005 (-.08)	0.0004	.99–1.00	-1.20	0.24
Age <sup>2</sup>	.0003 (.06)	0.0003	.99–1.00	.85	0.40
Effortful Control	-0.0005 (-.04)	0.0007	.99–1.00	-0.74	0.46
<hr/>					
Random Effects	Variance				
Intercept	.000006				

Note. Non-Verbal Cog = Non-verbal cognitive abilities; Verbal Cog = verbal cognitive abilities.

R model equation: Area under the curve ~ response + anxiety group + Autistic Quotient Total + Age + Age<sup>2</sup> + Effortful Control Scale + (1 | id).

interpretation bias is present at a behavioural level but is not reflected in internal reactions. This could occur because there is no actual threat within the context of the task. The children are old enough to understand that the aliens are make-believe and so an angry alien poses no objective threat. It seems feasible that this could lead to them categorising the aliens as angry without necessarily feeling threatened by them. This inconsistency in behavioural and physiological data contrasts with previous research that did find a correspondence between subjective and physiological responses (Tottenham et al., 2013). However, in their study Tottenham et al. (2013) used faces and had a greater age range of participants (6–17yrs). It is possible therefore that the difference in findings across studies is due to our use of auditory stimuli or the age range of participants. While previous work has shown no differences in corrugator activity to standardised visual and auditory stimuli in adults (Zhou, Qu, Jiao, & Helander, 2014), we do not know whether corrugator activity differs across stimuli of different modalities in children. It is possible that the corrugator activity to the ambiguous tones may reflect confusion rather than the child's valence appraisal of the ambiguous tones. However, if this was the case we would expect greater corrugator activity across the ambiguous tones in comparison to at least the trained tones reported as positive, and possibly the trained tones reported as negative. A post hoc check of the means across tone type confirmed that this was not case.<sup>1</sup> While we have no evidence that corrugator activity to the ambiguous tones reflects confusion, our analysis cannot fully rule it out. Furthermore, the greater familiarity of faces over tones may also explain why Tottenham et al. (2013) found correspondence between responses to ambiguous stimuli and corrugator activity, when we did not. It seems possible that the visual representation of the alien when happy and angry may have provided a visual bolster for the trained tones, resulting in the greater corrugator response than for the ambiguous tones, which were never paired with a visual emotional stimulus.

The role of development on the relationship between anxiety and interpretation bias in young children is still unclear. We found that development confounds behavioural responses to ambiguous tones, which suggests that development does play a role in how the relationship between interpretation bias and anxiety unfolds. However, whilst adding the developmental proxies significantly improved model fit, the main effect and interaction terms of age and effortful control were not significant predictors. Field and Lester (2010) concluded that the role of development best fitted an acquisition model for interpretation bias, i.e. that cognitive biases emerge alongside social,

<sup>1</sup> ambiguous ( $M = 0.00065$ ), happy ( $M = 0.00016$ ) and angry ( $M = 0.00071$ ).

emotional and cognitive development. Anxiety may therefore be a cause or a consequence of the biases. The recent meta-analysis of Stuijzand et al. (2017) supported a moderation model but concluded that more work with young children is required. The present research does not support the moderation or the acquisition model. We showed that children aged 4–8 years interpreted the ambiguous tones as negative; there were small differences between children with high and low anxiety and these were only significant after developmental proxies were accounted for. Ambiguous tones were interpreted as negative (angry) under 60% of the time. This was statistically different from 50% (i.e. chance) and thus taken to indicate a bias. This might suggest that in young children an interpretation bias is emerging. However, this could also support a moderation model. Having a clearer idea of how development influences the relationship between interpretation bias and anxiety in young children would help separate these models. Future research may benefit from considering development over a longer time frame and/or including other developmental factors, not assessed here.

This study has three limitations relating to the sample recruited. First, the children allocated to the elevated and low anxiety groups could have been more distinct. This may have highlighted larger group differences in behavioural and psycho-physiological data. In our sample high anxiety was defined as 1SD above the reported normed mean and low anxiety was defined as below the normed mean. This difference in criteria between the groups resulted from difficulty in recruiting children for the low anxiety group when the cut-off was initially set at 1SD below the mean. This is likely due to the positive skew of anxiety symptoms in community populations. Furthermore, this dichotomisation of a continuous variable (parent-reported anxiety) to create anxiety groups might obscure effects seen with continuous measures. Any distinction may be further amplified if samples included children with a diagnosis of an anxiety disorder, where it is reasonable to hypothesise that interpretation bias may be more negative. Second, the recruitment into the study may have attracted parents who were interested in anxiety or because they were concerned about anxiety in their child meaning that the sample may not be representative of the general population. Finally, while the original sample size was reasonable, fewer participants could be included in the final analysis and this may have meant that power was too low to adequately test moderation by age and effortful control.

A further limitation is that we were not able to evaluate the test-retest reliability of the task to evaluate stability. This will be important for future research because stability over time is required if the task is to be used within investigations that aim to link cognitive biases to other lower levels of analysis such as brain functioning or genetics. Another important focus for future research will be using a longitudinal design to better understand how developmental processes are implicated in the relationship between interpretation bias and anxiety in young children. This may help identify the role of potentially important processes that change across early childhood, such as effortful control. Such a design would also help clarify when and how interpretation biases and anxiety concerns emerge, in what order the relationships emerge, and provide partial support for causal mechanisms in anxiety.

In this study we have presented a developmentally sensitive paradigm to assess interpretation bias in children aged 4–8. We have also assessed the validity of a physiological measure of interpretation bias in young children, fEMG by comparing it to a behavioural measure. Our results indicate that the new paradigm can be used to assess interpretation bias in young children. The behavioural data indicated that high anxious children are more likely to interpret ambiguous tones as negative but this effect is only significant once developmental proxies have been statistically accounted for. The physiological measure, corrugator activity, reflected children's subjective experience of the tones but did not distinguish or indicate ambiguous tones. Therefore, we did not find that the corrugator was a valid physiological marker of children's appraisal of valence-ambiguous stimuli. Results highlight the importance of considering development in both design and analysis of

investigations of interpretation bias and anxiety in children as well as encouraging more consideration of theories of anxiety that consider a role for development.

### Conflicts of interest

The authors declare no conflict of interest.

### Open science statement

In the interests of research transparency, the authors would like to confirm that we have reported all measures, conditions, data exclusions relating to this study and the project it was contained within, and we have reported how we determined our sample sizes for this study. The authors are happy to share the task for use in future research including replication studies. Please contact the corresponding author.

### Acknowledgements

The research was funded by a SEDTC ESRC studentship (ES/J500148/1) awarded to Suzannah Stuijzand and an ESRC Future Research Leaders Grant (ES/L010119/1) awarded to Helen Dodd. Data was collected with the assistance of postgraduate student Laura Weight and undergraduate student Victoria Milner.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.brat.2019.103450>.

### References

- Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The autism spectrum quotient: Children's version (AQ-Child). *Journal of Autism and Developmental Disorders*, 38, 1230–1240. <https://doi.org/10.1007/s10803-007-0504-z>.
- Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, 133(1), 1–24. <https://doi.org/10.1037/0033-2909.133.1.1>.
- Barrett, P. M., Rapee, R. M., Dadds, M. M., & Ryan, S. M. (1996). Family enhancement of cognitive style in anxious and aggressive children. *Journal of Abnormal Child Psychology*, 24(2), 187–203. Retrieved from [http://www.safelylit.org/citations/index.php?fuseaction=citations.viewdetails&citationids0005B0005D=citjournalarticle\\_207895\\_38](http://www.safelylit.org/citations/index.php?fuseaction=citations.viewdetails&citationids0005B0005D=citjournalarticle_207895_38).
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: Automatic and strategic processes. *Behaviour Research and Therapy*, 35(1), 49–58. [https://doi.org/10.1016/S0005-7967\(96\)00069-1](https://doi.org/10.1016/S0005-7967(96)00069-1).
- Bell, B., Morgan, G., Kromrey, J., & Ferron, J. (2010). The impact of small cluster size on multilevel models: A Monte Carlo examination of two-level models with binary and continuous predictors. *JSM Proceedings, Section on Survey Research Methods*, 4057–4067. Retrieved from [http://www.amstat.org/sections/srms/Proceedings/y2010/Files/308112\\_60089.pdf](http://www.amstat.org/sections/srms/Proceedings/y2010/Files/308112_60089.pdf).
- Bernier, A., Beauchamp, M. H., Bouvette-Turcot, A. A., Carlson, S. M., & Carrier, J. (2013). Sleep and cognition in preschool years: Specific links to executive functioning. *Child Development*, 84(5), 1542–1553. <https://doi.org/10.1111/cdev.12063>.
- Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics - Simulation and Computation*, 39(4), 860–864. <https://doi.org/10.1080/03610911003650383>.
- Dimberg, U. (1990). Facial electromyographic reactions activity to auditory stimuli. *Biological Psychology*, 31, 137–147.
- Dodd, H. F., Hudson, J. L., Morris, T. M., & Wise, C. K. (2012). Interpretation bias in preschool children at risk for anxiety: A prospective study. *Journal of Abnormal Psychology*, 121(1), 28–38. <https://doi.org/10.1037/a0024589>.
- Eisenberg, N., Michalik, N., Spinrad, T. L., Hofer, C., Kupfer, A., Liew, J., et al. (2007). The relations of effortful control and impulsivity to children's sympathy: A longitudinal study. *Cognitive Development*, 22(4), 1–23.
- Eisenberg, N., Smith, C. L., Sadovsky, A., & Spinrad, T. (2004). Effortful control. *Handbook of self-regulation: Research, theory, and applications* (pp. 259–282). .
- Eysenck, M. W. (1992). *Anxiety: The cognitive perspective*. Hove, UK: Lawrence Erlbaum Associates, Inc.
- Eysenck, M. W. (1997). *Anxiety: The cognitive perspective*. Hove, UK: Psychology Press.
- Field, A. P., & Lester, K. J. (2010). Is there room for “development” in developmental

- models of information processing biases to threat in children and adolescents? *Clinical Child and Family Psychology Review*, 13(4), 315–332. <https://doi.org/10.1007/s10567-010-0078-8>.
- Field, A. P., & Storksen-Coulson, H. (2007). The interaction of pathways to fear in childhood anxiety: A preliminary study. *Behaviour Research and Therapy*, 45(12), 3051–3059. <https://doi.org/10.1016/j.brat.2007.09.001>.
- Field, A. P., & Wright, D. B. (2011). A primer on using multilevel models in clinical and experimental psychopathology research. *Journal of Experimental Psychopathology*, 2(2), 271–293. <https://doi.org/10.1017/CBO9781107415324.004>.
- Gifford, S., Reynolds, S., Bell, S., & Wilson, C. (2008). Threat interpretation bias in anxious children and their mothers. *Cognition & Emotion*, 22(3), 497–508. <https://doi.org/10.1080/02699930801886649>.
- Grosjean, P., & Ibanez, F. (2014). *pastecs: Package for analysis of space-time ecological series*.
- Hox, J. (2010). *Multilevel analysis: Techniques and applications*. Taylor & Francis.
- Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). Taylor & Francis Retrieved from <https://books.google.co.uk/books?id=yElZBwAAQBAJ>.
- Kindt, M., & van den Hout, M. (2001). Selective attention and anxiety: A perspective on developmental issues and the causal status. *Journal of Psychopathology and Behavioral Assessment*, 23(3), 193–202. <https://doi.org/10.1023/a:1010921405496>.
- Logic Studio 9 [Computer Software] (2007). Apple. Retrieved from <https://logic-studio.soft112.com>.
- Maas, C. J., & Hox, J. (2005). Sufficient sample sizes for multilevel modeling. *Journal of Research Methods for the Behavioral and Social Sciences*, 1, 86–92. <https://doi.org/10.1027/1614-1881.1.3.86>.
- Mezzacappa, E. (2004). Alerting, orienting, and executive attention: Developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. *Child Development*, 75(5), 1373–1386. <https://doi.org/10.1111/j.1467-8624.2004.00746.x>.
- Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy*, 36(9), 809–848. [https://doi.org/10.1016/S0005-7967\(98\)00063-1](https://doi.org/10.1016/S0005-7967(98)00063-1).
- Nauta, M. H., Scholing, A., Rapee, R. M., Abbott, M., Spence, S. H., & Waters, A. (2004). A parent-report measure of children's anxiety: Psychometric properties and comparison with child-report in a clinic and normal sample. *Behaviour Research and Therapy*, 42, 813–839. [https://doi.org/10.1016/S0005-7967\(03\)00200-6](https://doi.org/10.1016/S0005-7967(03)00200-6).
- Psychology Software Tools Inc (2012). Eprime 2.0. Pittsburgh, PA. Retrieved from <http://www.pstnet.com>.
- R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The children's behavior questionnaire. *Child Development*, 72(5), 1394–1408. Retrieved from <https://doi.org/10.1111/1467-8624.00355>.
- Rothbart, M. K., & Rueda, M. R. (2005a). *The development of effortful control*. American Psychological Association <https://doi.org/10.1037/11108-009>.
- Rothbart, M. K., & Rueda, M. R. (2005b). The development of effortful control. In U. Mayr, E. Awh, S. K. Keele, & M. I. Posner (Eds.). *Developing individuality in the human brain: A tribute to michael I. Posner* (pp. 167–188). (1st ed.). Washington D.C: American Psychological Association. <https://doi.org/10.1038/cddis.2011.1>.
- Salemink, E., & Wiers, R. W. (2012). Adolescent threat-related interpretive bias and its modification: The moderating role of regulatory control. *Behaviour Research and Therapy*, 50(1), 40–46. <https://doi.org/10.1016/j.brat.2011.10.006>.
- Snijders, T. a. B. (2005). Power and sample size in multilevel linear models. *Encyclopedia of Statistics in Behavioral Science*, 3, 1570–1573. <https://doi.org/10.1002/0470013192.bsa492>.
- Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behaviour Research and Therapy*, 36, 545–566. [https://doi.org/10.1016/S0005-7967\(98\)00034-5](https://doi.org/10.1016/S0005-7967(98)00034-5).
- Spence, S. H., Rapee, R., McDonald, C., & Ingram, M. (2001). The structure of anxiety symptoms among preschoolers. *Behaviour Research and Therapy*, 39(11), 1293–1316. [https://doi.org/10.1016/S0005-7967\(00\)00098-X](https://doi.org/10.1016/S0005-7967(00)00098-X).
- van Steensel, F. J. A., Bögels, S. M., & Perrin, S. (2011). Anxiety disorders in children and adolescents with autistic spectrum disorders: A meta-analysis. *Clinical Child and Family Psychology Review*, 14(3), 302–317. <https://doi.org/10.1007/s10567-011-0097-0>.
- Stuijzand, S., Creswell, C., Field, A. P., Pearcey, S., & Dodd, H. (2017). Research review: Is anxiety associated with negative interpretations of ambiguity in children and adolescents? A systematic review and meta-analysis. *Journal of Child Psychology and Psychiatry*. <https://doi.org/10.1111/jcpp.12822>.
- Stuijzand, S., Stuijzand, B. G., Reynolds S. A., & Dodd, H. F. (in preparation). Developmental influences on anxiety related attention biases in four to eight-year-olds: An eye-tracking study.
- Tan, J.-W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., et al. (2011). *Facial electromyography (fEMG) activities in response to affective visual stimulation*, Vols. 45–49. 2011 <https://doi.org/10.1109/WACI.2011.5953144>.
- Tan, J.-W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., et al. (2012). Repeatability of facial electromyography (EMG) activity over corrugator supercilii and zygomaticus major on differentiating various emotions. *Journal of Ambient Intelligence and Humanized Computing*, 3, 3–10. <https://doi.org/10.1007/s12652-011-0084-9>.
- Tottenham, N., Phuong, J., Flannery, J., Gabard-Durnam, L., & Goff, B. (2013). A negativity bias for ambiguous facial-expression valence during childhood: Converging evidence from behavior and facial corrugator muscle responses. *Emotion*, 13(1), 92–103. <https://doi.org/10.1037/a0029431>.
- Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013). WISC-IV and clinical validation of the four- and five-factor interpretative approaches. *Journal of Psychoeducational Assessment*, 31, 114–131. <https://doi.org/10.1177/0734282913478032>.
- Williams, J. M., Watts, F. N., MacLeod, C., & Mathews, A. (1988). *Cognitive psychology and emotional disorders*. New York: John Wiley & Sons, Ltd.
- Williams, J., Watts, F. N., MacLeod, C., & Mathews, A. (1997). *Cognitive psychology and emotional disorders* (2nd ed.). New York: John Wiley & Sons, Ltd.
- Yee, T. W. (2014). The VGAM package for categorical data analysis. *Journal of Statistical Software*, 32(10), 1–34.
- Zhou, F., Qu, X., Jiao, J., & Helander, M. G. (2014). Emotion prediction from physiological signals: A comparison study between visual and auditory elicitors. *Interacting with Computers*, 26, 285–302. <https://doi.org/10.1093/iwc/iwt039>.