

Children can't ignore what they hear: Incongruent emotional information leads to an auditory dominance in children

Paddy Ross^{1*}, Beth Atkins¹, Laura Allison¹, Holly Simpson¹, Catherine Duffell¹, Matthew Williams^{1,2}, & Olga Ermolina¹

1. Department of Psychology, Durham University, Durham, UK

2. Department of Psychology, Newcastle University, Newcastle, UK

***Corresponding Author:**

Dr Paddy Ross

Department of Psychology

Durham University

Science Site

South Road

Durham

DH1 3LE

01913343279

paddy.ross@durham.ac.uk

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abstract:

Effective emotion recognition is imperative to successfully navigating social situations. Research suggests differing developmental trajectories for the recognition of bodily and vocal emotion, but emotions are usually studied in isolation and rarely considered as multimodal stimuli in the literature. When presented with basic multimodal sensory stimuli, the Colavita effect suggests that adults have a visual dominance, whereas more recent research finds that an auditory sensory dominance may be present in children under 8 years of age. However, it is not currently known whether this phenomenon holds for more complex multimodal social stimuli. Here we presented children and adults with multimodal social stimuli consisting of emotional bodies and voices, asking them to recognise the emotion in one modality while ignoring the other. We found that adults can perform this task with no detrimental effects to performance, regardless of whether the ignored emotion was congruent or not. However, children find it extremely challenging to recognise bodily emotion while trying to ignore incongruent vocal emotional information. In several instances they perform below chance level, indicating that the auditory modality actively informs their choice of bodily emotion. This is therefore the first evidence, to our knowledge, of an auditory dominance in children when presented with emotionally meaningful stimuli.

Keywords: Auditory Dominance; development; emotion recognition; body; voice

Introduction

The ability to recognise emotions is a critical part of healthy development, and studies have shown different developmental trajectories for emotion recognition ability from faces (Ambadar, Schooler et al. 2005, Gao and Maurer 2010, Montiroso, Peverelli et al. 2010, Kuhn 2014), bodies (Boone and Cunningham 1998, Lagerlof and Djerf 2009, Ross, Polson et al. 2012) and vocal prosody (Sauter, Panattoni et al. 2013, Grosbras, Ross et al. 2018, Morningstar, Ly et al. 2018).

In these studies, stimuli are presented in one modality and recognition rates are recorded. However, humans exist in a multi-sensory world, with different emotion modalities seldom presented in isolation (Lewkowicz and Ghazanfar 2009). Making sense of multi-sensory information is crucial for streamlined cognitive processing, resulting in optimised accuracy and cognitive speed (Calvert and Thesen 2004, Elliot, Wing et al. 2010, Burr and Gori 2012). However, if presented with several (sometimes incongruent) different types of sensory information, how does one know to what to pay attention? The concept of sensory dominance is central to the study of multisensory processing, and there is evidence to suggest developmental differences in human's preferences for information from one modality over another.

An early demonstration of sensory dominance is described by Colavita (1974) in which participants were simultaneously presented with lights and tones. It was found that adult subjects would be unaware that a tone had been presented, and that participant's responses were dominated by the visual stimuli. This *Colavita Effect* was found by a recent meta-analysis to be very robust in adults across methodologies (Hirst, Cragg et al. 2018).

However, there appears to be the reverse effect in children. Sloutsky and Napolitano (2003) found that when presented with auditory and visual stimuli pairs, 80% of the auditory stimuli changes were detected by children compared to only 30% in the visual modality. Interestingly, when presented unimodally, children were quite able to discriminate between visual stimuli. This suggests that instead of the visual stimuli simply being more difficult to discriminate, rather, the presence of the sound appears to overshadow the visual information. Nava and Pavani (2013) found that a reverse Colavita Effect (or auditory dominance) persists in 6-7 year-olds and that a transition towards visual dominance starts at school age. Across three different experiments, they observed an auditory dominance in children up to 6-7 years of age and a clear visual dominance in children 9-12.

These experiments demonstrate sensory dominance using simple visual and auditory stimuli, but can these effects be extended to the processing of more complex and meaningful stimuli? Conducting research with adults, Koppen, Alsius et al. (2008) showed evidence of the Colavita effect using speech and semantic stimuli. Hecht and Reiner (2009) also found evidence of a visual dominance in adults when compared to both auditory and haptic stimuli. However, what evidence is there of a reverse Colavita effect using socially meaningful stimuli? Perhaps the most well-known incongruent audio-visual phenomena using more complex stimuli is the McGurk effect (McGurk and MacDonald 1976), in which a seen mouth movement changes the perceived sound. Using this effect, Massaro (1984) found that there was less of an influence of the visual aspect of the effect for children

relative to adults. More recently, Hirst, Stacey et al. (2018) found reduced susceptibility to the effect in younger children, with older children and adults more easily influenced by the visual aspect.

Using more semantically relevant stimuli, namely pictures of objects and their relative sounds (cow/mooing, bell/ringing etc.), Wille and Ebersbach (2016) also found evidence of auditory dominance in children. Using a similar paradigm to Nava and Pavani (2013), they found 6-year-olds exhibited an auditory dominance, while results were mixed in 9-year-olds. Perhaps then, a transition from auditory to visual dominance occurs at some stage between 7 and 9 years of age. It could indeed be the case that as the auditory modality is the first to develop (Birnholtz and Benacerraf 1983), it is possible that auditory dominance in children is the default, and visual dominance only occurs as a result of learning through experience (Napolitano and Sloutsky 2004, Napolitano 2006, Wille and Ebersbach 2016).

What then of emotionally meaningful stimuli? As previously mentioned, we know that there are different developmental trajectories for recognition of emotion from bodies, voices and faces (Boone and Cunningham 1998, Tonks, Williams et al. 2007, Lagerlof and Djerf 2009, Gao and Maurer 2010, Ross, Polson et al. 2012, Chronaki, Hadwin et al. 2015, Grosbras, Ross et al. 2018), mirrored by neuroimaging evidence (Bonte, Frost et al. 2013, Ross, de Gelder et al. 2014, Ross, de Gelder et al. 2019). In relation to multisensory processing, we know that when presented with emotional bodies, faces and voices, people are susceptible to being influenced by simultaneously presented incongruent information (Van den Stock, Righart et al. 2007). Van den Stock, Righart et al. (2007) found that when adult observers make judgements about the emotional content of a voice, recognition is biased toward a simultaneously perceived whole body expression, even if it is incongruent. One could argue that this is evidence of the Colavita Effect in socially meaningful stimuli; i.e. a visual dominance effect in which the content of the visual stimulus is disproportionately affecting the auditory stimulus. If this were the case, would one expect to see the reverse effect in children? In other words, would children's emotion recognition ability be adversely affected by incongruent auditory stimuli (an auditory dominance effect)?

Here we set out to address this question for the first time across two experiments. In Experiment 1 we mirrored the work of Van den Stock, Righart et al. (2007) by simultaneously presenting congruent and incongruent fearful and happy emotional bodies and voices to children and adults in an emotion recognition task. Experiment 2 was a partial replication of Experiment 1, but using two different, less orthogonal emotions, namely sadness and anger.

We hypothesised across both experiments that when asked to name the emotion in the voice and ignore incongruent emotion in the body, the Colavita Effect (visual dominance) would cause performance in adults to decrease. However, crucially, when asked to name the emotion in the body, but ignore incongruent emotion in the voice, the reverse Colavita Effect (auditory dominance) would cause performance in the children to decrease. This would provide the first evidence of auditory dominance in children using emotionally meaningful stimuli.

Experiment 1 Methods

Participants

There were 63 participants in total, 33 males and 30 females, aged 6-49. Participants were divided into 3 age groups for analysis based on the developmental work of Nava and Pavani (2013): younger children aged 7 and under ($n=32$, $M=7.26$, $SD=0.42$), older children aged 8-11 ($n=20$, $M=9.04$, $SD=0.93$) and adults aged 18 and over ($n=11$, $M=29.95$, $SD=11.97$).

All adult participants were undergraduates from Durham University and the children were recruited from an after-school club in Durham. Adults completed the study for course credits while the children received a sticker for participation. All participants had normal or corrected-to-normal vision, volunteered to take part, and the study was approved by the Psychology Department Ethics Advisory Sub-Committee at Durham University.

Stimuli

Body stimuli were taken from the Bodily Expressive Action Stimuli Test (de Gelder and Van den Stock 2011). This validated database consists of whole-body expressions from 46 actors expressing 4 emotions. The actor's face is blurred out, and the static images are de-saturated. 10 happy and 10 fearful body images were randomly selected from this set.

For the voice stimuli 10 happy and 10 fearful vocalisations were randomly selected from the Montreal Affective Voice (MAV) stimuli database (Belin, Fillion-Bilodeau et al. 2008). These stimuli are vocalisations and so free from speech or language (e.g. gasps and screams for fear, laughter for happiness). All auditory stimuli were presented using 'over ear' headphones.

Design and Procedure

Subjects took part in four different experimental blocks as illustrated in Figure 1. In block 1 (Body Only) they were presented with the 20 body stimuli (10 happy and 10 fearful) and asked 'How is this person feeling?' in a forced choice paradigm. Block 2 (Voice Only) was a similar design but with the 20 auditory stimuli.

In block 3 (Ignore Body) participants were presented with auditory and visual stimuli simultaneously. This gave the combinations of 5 Happy Body/Happy Voice, 5 Happy Body/Fearful Voice, 5 Fearful Body/Happy Voice and 5 Fearful Body/Fearful Voice stimuli. In block 3 participants were asked to ignore what they saw and asked to determine how the person was feeling based on their voice. Block 4 (Ignore Voice) presented the same stimuli, but asked participants to this time ignore what they heard, and determine how the person was feeling based on what they saw. This gave 80 trials in total across all 4 blocks. Blocks were randomised for all participants and testing took approximately 20 minutes.

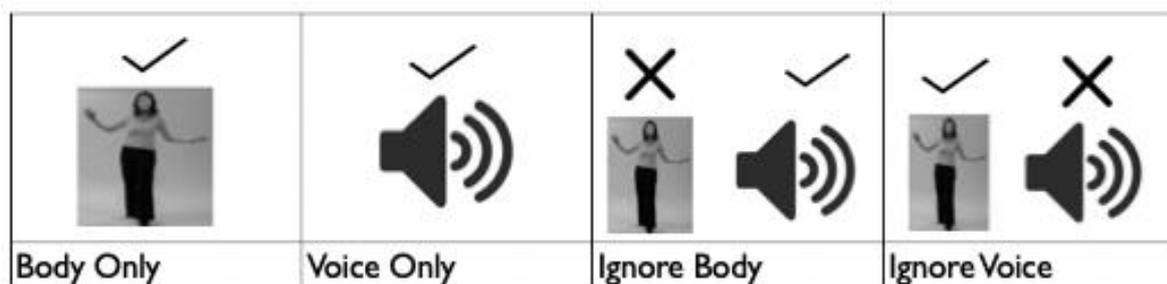


Figure 1. Block design of study detailing the 4 experimental blocks.

Participants were presented with stimuli on a 15-inch monitor using MATLAB (Mathworks, Natick, MA) and Psychtoolbox (Brainard, 1997). Participants sat 60 cm away from the screen, and visual stimuli were 4 cm × 10 cm, giving a visual angle of approximately 3.8° × 9.5°. A fixation cross was presented for 2000ms, followed by the stimuli for 1500ms. This was as long as the longest auditory utterance. ‘How is this person feeling?’ appeared on the screen until the participants chose either ‘happy’ (h key) or ‘fearful’ (f key). The relevant keys were covered with stickers denoting the emotion written in full as well as an emoticon as an added aid for younger children. This procedure was the same across all experimental blocks and is illustrated in Figure 2.

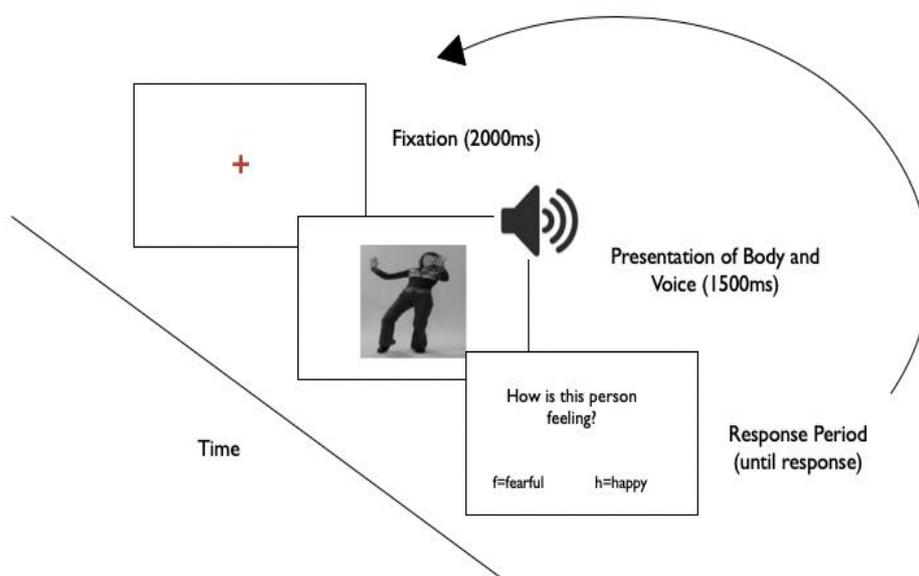


Figure 2. A typical trial sequence used throughout the study.

Experiment 1 Results

Percentage correct emotion identification scores were calculated for each emotion/modality combination for each participant and averaged across all participants to give an overall percent correct response measure. In order to explore any differences across modalities, results from the Body Alone and Voice Alone blocks were combined and mean percentage accuracies were entered into a 2x2x3 ANOVA (Emotion x Modality x Age Group).

For blocks 3 and 4 mean percentage accuracies were entered into a 4x3 ANOVA (Emotion-Modality Combination x Age Group).

Body and Voice Alone

We found no significant main effect of Age Group ($F(2,60)=.297$, $p=.744$, $\eta^2p=.010$) with young children ($M=88.7$, $SE=1.8$), older children ($M=86.88$, $SE=2.2$) and adults ($M=89.3$, $SE=3.0$) all obtaining very similar recognition accuracies. We found a main effect of Modality ($F(1,60)=40.26$, $p<.001$, $\eta^2p=.420$) driven by higher recognition accuracies for vocal ($M=96.4$, $SE=1.4$) compared to body stimuli ($M=80$, $SE=2.6$). There was also a main effect of emotion ($F(1,60)=6.36$, $p<.05$, $\eta^2p=.096$) driven by higher recognition rates for fearful ($M=91.0$, $SE=1.30$) compared to happy ($M=85.6$, $SE=2.1$) stimuli (see Figure 3).

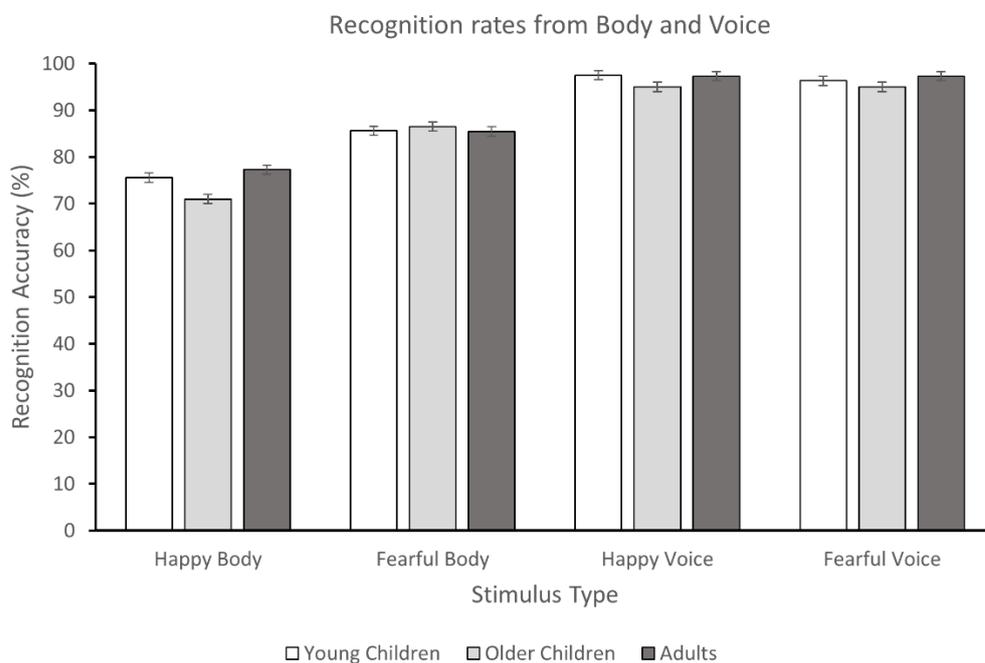


Figure 3. Average recognition rates when subjects viewed the Happy and Fearful stimuli in isolation, presented individually for emotion and modality. Error bars represent SEM.

We also observed a Modality x Emotion interaction ($F(1,60)=8.57$, $p<.005$, $\eta^2p=.125$). Post-hoc simple effect analyses revealed that this was driven by a significant difference ($t(62)=3.14$, $p<.005$) between happy ($M=74.4$, $SE=3.4$) and fearful ($M=85.9$, $SE=1.8$) recognition in the body modality, but no significant difference between the two emotions in the vocal modality ($t(62)=.814$, $p=.419$).

Crucially, Age Group did not significantly interact with any other factor, indicating similar performance across age in recognising emotions from the body and voice alone.

Combined Modalities (Identify Voice, Ignore Body)

The results of identifying the emotion from the voice while ignoring the body stimuli are presented in Figure 4a. A 4x3 ANOVA revealed no main effect of emotion-modality

combination ($F(3,180)=1.16$, $p=.327$, $\eta^2p=.019$), no main effect of age group ($F(2,60)=.403$, $p=.67$, $\eta^2p=.013$) and no age group x emotion-modality interaction ($F(6,180)=40.26$, $p=.281$, $\eta^2p=.04$).

Combined Modalities (Identify Body, Ignore Voice)

The results of identifying the emotion from the voice while ignoring the body stimuli are presented in Figure 4b. We found a main effect of emotion-modality combination ($F(3,180)=16.35$, $p<.001$, $\eta^2p=.21$) driven by Happy Body/Fearful Voice ($M=54.9$, $SE=5.0$) being recognised significantly worse ($p<.001$) than both Happy Body/Happy Voice ($M=79.8$, $SE=3.7$) and Fearful Body/Fearful Voice ($M=90.5$, $SE=1.8$). Fearful Body/Happy Voice ($M=65.94$, $SE=5.2$) was recognised significantly worse than Fearful Body/Fearful Voice ($p<.001$). We also mirrored the Body Alone findings, with Happy Body/Happy Voice being significantly worse recognised than Fearful Body/Fearful Voice ($p<.05$).

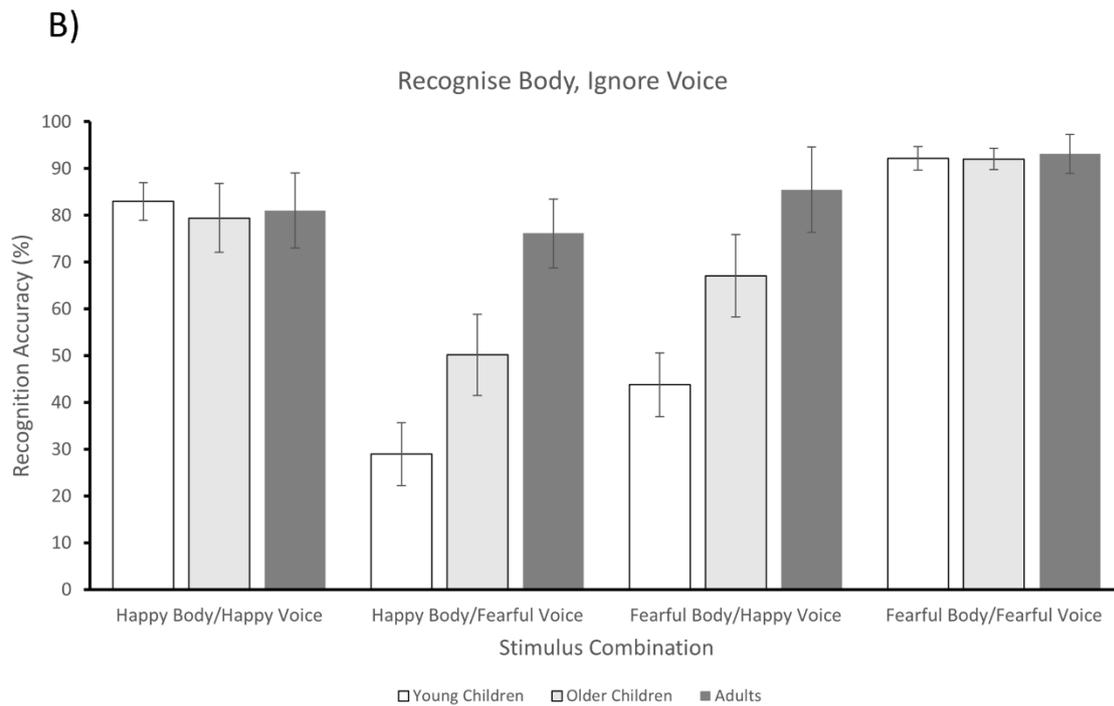
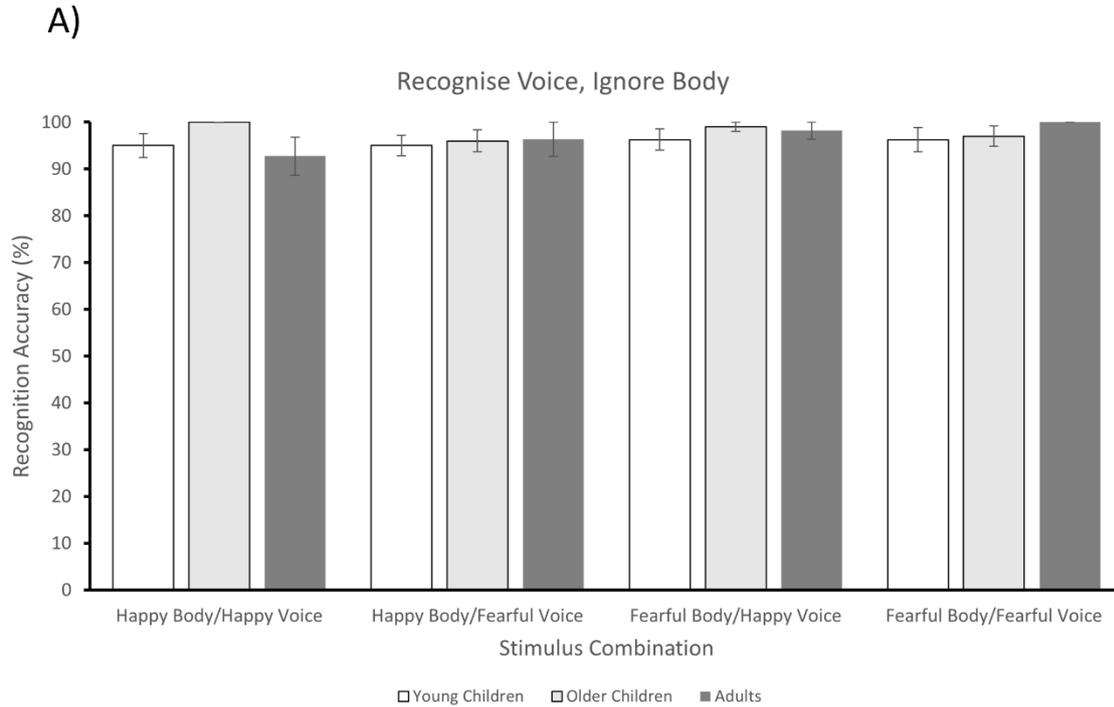


Figure 4. Average recognition rates for all participants for Happy and Fearful emotions across both combined stimulus conditions. (A: Recognise the body, Ignore the Voice; B: Recognise the Body, Ignore the Voice). It should be noted that identical stimuli were shown across these two blocks, only the instructions changed. Error bars represent SEM.

We found a main effect of Age Group ($F(2,60)=3.1, p<.05, \eta^2p=.095$) driven by significantly worse performance ($p<.05$) by the younger children ($M=65.5, SE=3.2$) compared to the adults ($M=80.9, SE=5.4$).

Crucially we also observed an age group x emotion-modality interaction ($F(6,180)=2.95$, $p<.005$, $\eta^2p=.089$). Post-hoc simple effect analyses revealed a main effect of age in both the Happy Body/Fearful Voice ($F(2,62)=5.16$, $p<.01$) and Fearful Body/Happy Voice ($F(2,62)=4.20$, $p<.05$) combinations. Post-hoc Bonferroni corrected t-tests revealed that in both cases, the main effect was being driven by significantly worse performance by the younger children compared with the adults (Happy Body/Fearful Voice: Young children ($M=37.5$, $SE=6.7$), Adults ($M=78.18$, $SE=7.36$), $p<.01$; Fearful Body/Happy Voice: Young children ($M=45.6$, $SE=6.8$), Adults ($M=78.18$, $SE=9.13$), $p<.05$). No significant differences were found between age groups in either of the congruent emotion conditions (Happy Body/Happy Voice: ($F(2,62)=.274$, $p=.762$); Fearful Body/Fearful Voice: ($F(2,62)=.021$, $p=.979$).

Cross-trial comparisons

In order to compare our bimodal incongruent and congruent conditions with unimodal baseline, we collapsed results across emotion to explore a 3 (Age: Young children, Older children, Adults) x 3 (Trial type: Congruent, Incongruent, Baseline) repeated measure ANOVA for each reporting condition (Recognise the emotion from the Vocalisation or Body). When reporting the emotion in the voice, we found no main effect of trial type ($F(2,120)=0.08$, $p=.93$, $\eta^2p=.001$) and no interaction between trial type and age group ($F(4,120)=0.99$, $p=.41$, $\eta^2p=.032$).

When reporting the emotion in the body we found a main effect of trial type ($F(2,120)=20.64$, $p<.001$, $\eta^2p=.256$). Post-hoc Bonferroni corrected t-tests found that participants scored significantly less in the Incongruent group ($M=53.5$, $SD=35.25$) compared to both the Congruent ($M=85.71$, $SD=16.43$, $p<.001$) and Baseline ($M=80.16$, $SD=16.06$, $p<.001$). There was no difference between Congruent and Baseline ($p=.74$). We also found a significant interaction between trial type and age group ($F(4,120)=4.7$, $p<.005$, $\eta^2p=.135$). Follow up analyses found this to be driven by young children and older children scoring significantly less in the incongruent condition compared to the congruent (YC: $p<.01$; OC: $p<.05$) and baseline (YC: $p<.01$; OC: $p<.05$). Adults showed no significant difference in this comparison and no group showed a significant difference between congruent and baseline.

Experiment 1 Summary

We found no differences between age groups when they were presented with body and vocal stimuli in isolation. When we combined both modalities and asked participants to ignore the body stimuli and identify the emotion from the voice, again we found no significant difference between the age groups. However, when presented with the same stimuli, but asked to ignore the voice and base their judgement on the body stimuli, we found that young children performed significantly worse than adults in the two incongruent conditions.

Experiment 2 Methods

Participants

In experiment 2 there were 76 participants in total, 30 males and 46 females, aged 6-36 years. Participants were divided into 3 age groups for analysis: younger children aged 7 and under ($n=16$, $M=7.4$, $SD=0.4$), older children aged 8-9 ($n=24$, $M=9.2$, $SD=0.5$) and adults aged 18 and over ($n=36$, $M=19.4$, $SD=2.96$).

As in Experiment 1, all adult participants were undergraduates from Durham University, but for this experiment the children were recruited from two Primary Schools in the North-east of England. Adults completed the study for course credits while the children received a sticker for participation. All participants had normal or corrected-to-normal vision, volunteered to take part, and the study was approved by the Psychology Department Ethics Advisory Sub-Committee at Durham University.

Stimuli, Design and Procedure

The design and procedure of Experiment 2 was identical to Experiment 1. The only difference being that rather than happy and fearful stimuli being chosen, 10 sad and 10 angry body images were taken from the BEAST stimuli set and 10 sad and 10 angry vocalisations were randomly selected from the MAV stimuli set (e.g. cries for sadness, 'grrrs' for anger). All auditory stimuli were presented using 'over ear' headphones. For experiment 2, 'C' was the keyboard response for sadness, and 'M' for anger. The relevant keys were covered with stickers denoting the emotion written in full as well as an emoticon as an added aid for younger children.

Experiment 2 Results

As with Experiment 1, percentage correct emotion identification scores were calculated for each emotion/modality combination for each participant and averaged across all participants to give an overall percent correct response measure. In order to explore any differences across modalities, results from the Body Alone and Voice Alone blocks were combined and mean percentage accuracies were entered into a 2x2x3 ANOVA (Emotion x Modality x Age Group). For blocks 3 and 4 mean percentage accuracies were entered into a 4x3 ANOVA (Emotion-Modality Combination x Age Group).

Body and Voice Alone

We found a significant main effect of Age Group ($F(2,73)=5.8$, $p<.005$, $\eta^2p=.137$) with adults ($M=98.9$, $SE=1.1$), scoring significantly higher than young children ($M=92.2$, $SE=1.7$), and older children ($M=93.6$, $SE=1.4$). We found no main effect of Modality ($F(1,73)=3.4$, $p=.07$, $\eta^2p=.044$). There was a main effect of emotion ($F(1,73)=5.9$, $p<.05$, $\eta^2p=.096$) driven by higher recognition rates for sad ($M=95.7$, $SE=.73$) compared to angry ($M=93.7$, $SE=1.1$) stimuli (see Figure 5).

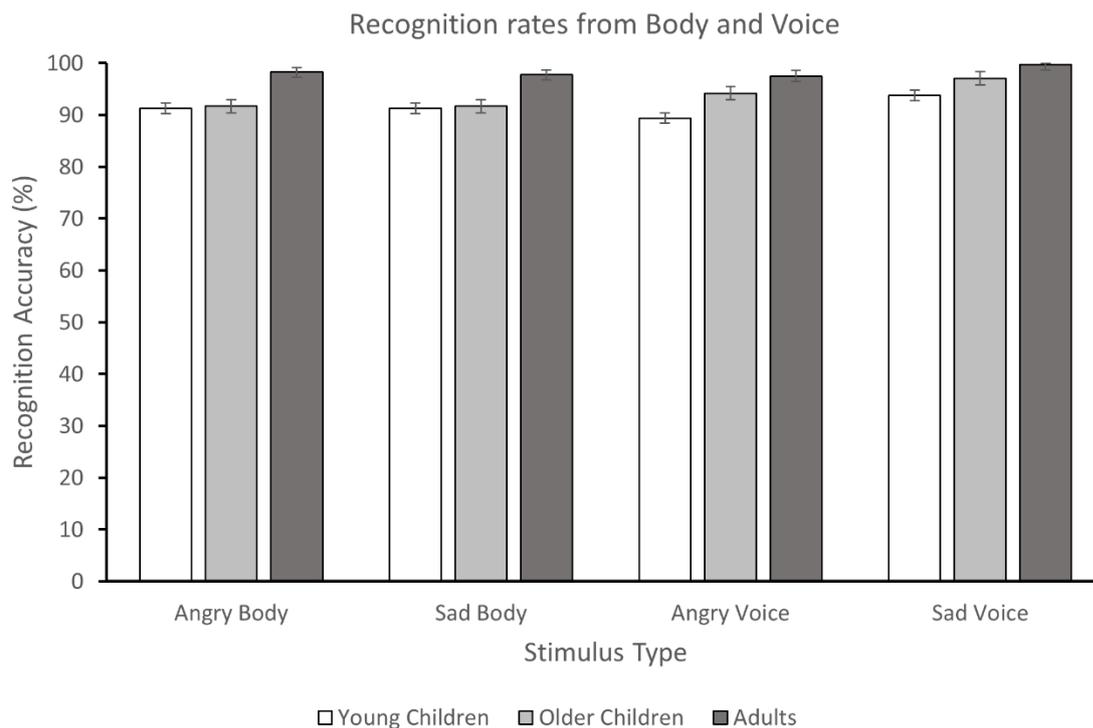


Figure 5. Average recognition rates when subjects viewed the Angry and Sad stimuli in isolation, presented individually for emotion and modality. Error bars represent SEM.

We also observed a Modality x Emotion interaction ($F(1,73)=9.34$, $p<.005$, $\eta^2p=.113$). Post-hoc simple effect analyses revealed that this was driven by a significant difference ($t(75)=3.8$, $p<.001$) between sad ($M=98.29$, $SE=.76$) and angry ($M=94.7$, $SE=1.3$) recognition in the voice modality, but no significant difference between the two emotions in the body modality ($t(75)=.231$, $p=.818$; Sad Bodies ($M=94.5$, $SE=1.1$) and Angry Bodies ($M=94.7$, $SE=1.3$)).

Again, Age Group did not significantly interact with any other factor.

Combined Modalities (Identify Voice, Ignore Body)

The results of identifying the emotion from the voice while ignoring the body stimuli are presented in Figure 6a. A 4x3 ANOVA revealed no main effect of emotion-modality combination ($F(3,219)=2.11$, $p=.1$, $\eta^2p=.028$), a main effect of age group ($F(2,73)=5.9$, $p<.004$, $\eta^2p=.139$) driven by the adults ($M=99.0$, $SE=1.3$) scoring higher on average than both the younger ($M=92.2$, $SE=1.9$) and older children ($M=93.5$, $SE=1.8$). We found no age group x emotion-modality interaction ($F(6,219)=1.54$, $p=.167$, $\eta^2p=.04$).

Combined Modalities (Identify Body, Ignore Voice)

The results of identifying the emotion from the voice while ignoring the body stimuli are presented in Figure 6b. We found a main effect of emotion-modality combination ($F(3,219)=157.8$, $p<.001$, $\eta^2p=.68$) driven by both Angry Body/Sad Voice ($M=45.9$, $SE=3.5$)

and Sad Body/Angry Voice (M=46.6, SE= 3.4) being recognised significantly worse ($p<.001$) than both Angry Body/Angry Voice (M=96.2, SE=.97) and Sad Body/Sad Voice (M=96.7, SE=1.1).

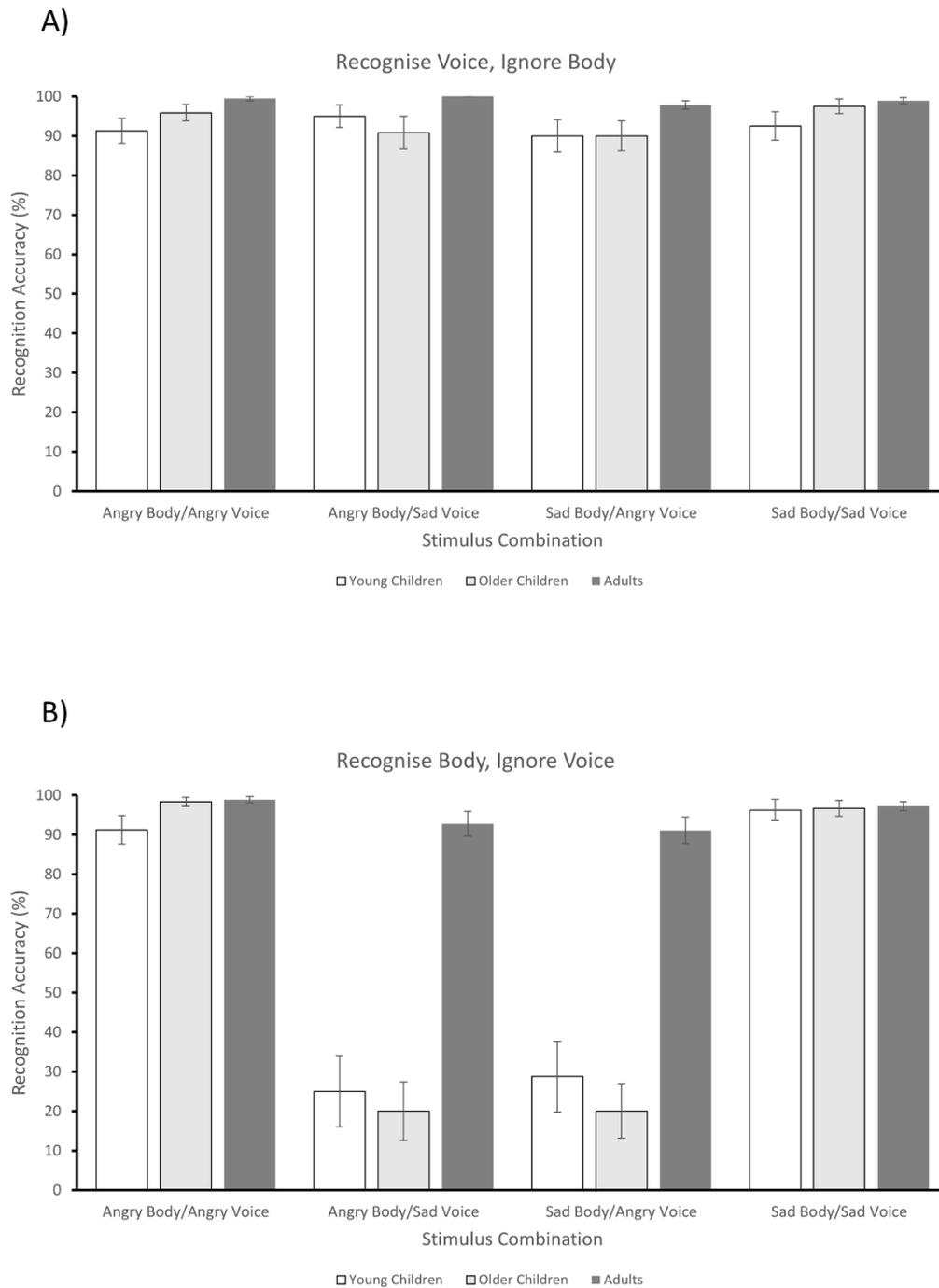


Figure 6. Average recognition rates for all participants for Angry and Sad emotions across both combined stimulus conditions. (A: Recognise the body, Ignore the Voice; B: Recognise the Body, Ignore the Voice). It should be noted that identical stimuli were shown across these two blocks, only the instructions changed. Error bars represent SEM.

We found a main effect of Age Group ($F(2,73)=66.9$, $p<.001$, $\eta^2p=.647$) driven by significantly worse performance ($p<.001$) by both the younger children (M=60.3, SE=3.4) and older children (M=58.8, SE=2.7) compared to the adults (M=95, SE=2.2).

Crucially, as in Experiment 1, we again observed an age group x emotion-modality interaction ($F(6,219)=40.0$, $p<.001$, $\eta^2p=.523$). Post-hoc simple effect analyses revealed a main effect of age in both the Angry Body/Sad Voice ($F(2,75)=5.16$, $p<.01$) and Sad Body/Angry Voice ($F(2,75)=4.20$, $p<.05$) combinations. Post-hoc Bonferroni corrected t-tests revealed that in both cases, the main effect was being driven by significantly worse performance by the younger and older children compared with the adults (Happy Body/Fearful Voice: Young children ($M=25$, $SE=9.0$), Older children ($M=20$, $SE=7.4$) & Adults ($M=92.8$, $SE=3.1$), both $p<.001$; Fearful Body/Happy Voice: Young children ($M=28.8$, $SE=8.9$), Older children ($M=20$, $SE=6.9$) and Adults ($M=91.1$, $SE=3.3$), both $p<.001$). No significant differences were found between age groups in the Angry Body/Angry Voice: ($F(2,75)=.075$, $p=.928$). However, we did find a significant difference between age groups in the Sad Body/Sad Voice condition ($F(2,75)=5.4$, $p<.01$) driven by younger children ($M=91.3$, $SE=3.6$) performing significantly worse than older children ($M=98.3$, $SE=1.2$; $p<.05$) and adults ($M=98.9$, $SE=.78$; $p<.01$).

Furthermore, we found that younger and older children were both scoring significantly below chance levels (50%) in both the incongruent conditions (Angry Body/Sad Voice: YC – $t(15)=2.8$, $p<.05$, OC – $t(15)=4.0$, $p<.001$; Sad Body/Angry Voice: YC – $t(15)=2.4$, $p<.05$, OC – $t(15)=4.3$, $p<.001$).

Cross-trial comparisons

Again, as in Experiment 1, in order to compare our bimodal incongruent and congruent conditions with unimodal baseline, we collapsed results across emotion to explore a 3 (Age: Young children, Older children, Adults) x 3 (Trial type: Congruent, Incongruent, Baseline) repeated measure ANOVA for each reporting condition (Recognise the emotion from the Vocalisation or Body).

When reporting the emotion in the voice, we found no main effect of trial type ($F(2,146)=1.89$, $p=.16$, $\eta^2p=.025$) and no interaction between trial type and age group ($F(4,146)=2.19$, $p=.073$, $\eta^2p=.057$).

When reporting the emotion in the body we found a main effect of trial type ($F(2,146)=168.5$, $p<.001$, $\eta^2p=.698$). Post-hoc Bonferroni corrected t-tests found that participants scored significantly less in the Incongruent group ($M=55.5$, $SD=44.5$) compared to both the Congruent ($M=96.97$, $SD=6.93$, $p<.001$) and Baseline ($M=94.6$, $SD=8.97$, $p<.001$). There was no difference between Congruent and Baseline ($p>.99$). We also found a significant interaction between trial type and age group ($F(4,146)=41.9$, $p<.001$, $\eta^2p=.535$). Follow up analyses found this to be driven by young children and older children scoring significantly less in the incongruent condition compared to the congruent (YC: $p<.001$; OC: $p<.001$) and baseline (YC: $p<.001$; OC: $p<.001$). Adults showed no significant difference in this comparison and adults and young children showed no significant difference between congruent and baseline. Older children, however, scored significantly higher in the congruent condition ($M=97.5$, $SD=6.7$) compared to the baseline ($M=91.7$, $SD=11.5$, $p<.05$).

Experiment 2 Summary

We found that when adults were presented with body and vocal stimuli in isolation, they scored higher on average than both groups of children. However, all groups scored over 90%. When we combined both modalities and asked participants to ignore the body stimuli and identify the emotion from the voice, again we found we found that adults outperformed children, but with all groups again scoring over 90%. However, when presented with the same stimuli, but asked to ignore the voice and base their judgement on the body stimuli, we found that younger and older children both performed significantly worse than adults in the two incongruent conditions. Here children were also scoring significantly below chance level, indicating that they were not merely guessing, but in most trials they were selecting the emotion that was in the vocal modality rather than the body, despite being instructed to ignore it.

Discussion

The aim of this study was to explore the possibility of a Colavita effect in adults, and crucially, a reverse Colavita effect (auditory dominance) in children using socially meaningful stimuli. We hypothesised that when asked to name the emotion in the voice and ignore incongruent emotion in the body adult performance would decrease. Furthermore we predicted that when instructed to state the emotion in a body (visual) stimulus while being instructed to ignore an incongruent emotion in a vocal (auditory) stimulus, children's performance would decrease. Our results supported the latter half of this hypothesis.

We did not find support for a visual dominance in adults. This appears to largely be due to a ceiling effect across our experiments. This also illustrates an interesting difference with using social/emotional stimuli compared to simple audio/visual stimuli. The question here is not whether a subject saw or heard a certain type of stimuli (as it was in Colavita, 1974; Navi and Pavani, 2013), but rather we asked subjects to make an emotional judgement regarding a stimuli. For adults, it could be that any visual dominance simply doesn't affect their ability to follow the rule and determine the emotion in the voice. Especially as they *know* that two stimuli will be present (in Colavita, 1974 in comparison, participants were told the bimodal stimuli were accidental). Whereas in the children, the reverse Colavita effect appears to greatly influence their ability to accurately categorise emotional bodies.

We found that using Happy/Fearful body and voice combinations, young children performed significantly worse compared with adults when they were asked to ignore an incongruent auditory emotion. They showed no performance issues when presented with exactly the same stimuli but asked to instead ignore the visual stimuli. A follow-up experiment using Angry/Sad body and voice combinations confirmed this finding, but also found that the older children performed just as poorly in the task as the younger children in the incongruent 'ignore the voice' trials. In this experiment children also performed significantly below chance levels.

The results complement those of Nava and Pavani (2013) and Wille and Ebersbach (2016) by showing further evidence of auditory dominance in children. Using Happy/Fearful stimuli we found that children younger than 8 years struggled with the task of labelling the emotion in the body and trying to ignore the emotion in the voice. We found an even more prominent effect using Angry/Sad stimuli, with both groups of children (under 8-year-olds and 8/9-year-olds) scoring below chance when instructed to ignore emotional auditory stimuli and label the body emotion.

Alternative Explanations

There are, however, several alternative explanations for these results, which must first be addressed in turn. Firstly, could it be that the younger children were just poorer at recognising the body emotion, leading to the auditory dominance? Previous work has shown that there is a steep developmental trajectory for emotion recognition from the body until approximately 8 years of age, by which stage children show adult-like recognition rates (Ross, Polson et al. 2012). Could this pattern of ability explain the poor performance of

young children in the incongruent trials? Given that when viewing bodies in isolation we found no significant differences between age-groups, we do not believe this to be the case. We did observe a small age difference in the Angry/Sad body identification, but this was likely caused by a ceiling effect in the adults as all age groups were over 90% accurate.

Might then the results be explained by a lack of inhibition ability in the children? As our instructions here were to ignore the voice and tell us what the emotion was in the body, one could argue that a lack of ability to follow the rule and ignore stimuli was the cause of the effect.

Evidence suggests that there is little in the way of improvement in the ability to inhibit rules throughout childhood (Davidson, Amso et al. 2006, Richardson, Anderson et al. 2018). Rather, it is the ability to switch between rules that shows a developmental trajectory. Here, if an inability to follow the rule was the issue for the younger children, then they should have arguably shown the same inability when asked to ignore the body stimuli and determine the emotion from the voice. Here, however, they showed adult-like performance and demonstrated no problems in following the rule.

Finally, as the younger and older children scored significantly below chance-level in the incongruent stimuli in Experiment 2, we can rule out that the children simply guessed the answer due to the confusing incongruent information presented to them. Therefore, we have confidence that the effect we present here is evidence of auditory dominance in children. But questions remain as to why this phenomenon would be advantageous to children, and what this can tell us about their perception of socially meaningful stimuli.

Potential Mechanisms

What processes then could account for this effect? One argument could be that participants are taking both the vocalisations and bodily movements into account and then making an emotion recognition decision based on some sort of 'weighted average' or amalgamation of the two sensory modalities. Under this account, the vocalisations are affecting the encoding of the bodily stimuli in some way, and lower accuracy is stemming from some disruption caused by the auditory stimuli. There is some evidence to suggest that visual tasks which contain concurrent auditory information show a stronger interference effect in children (Broadbent, Osborne et al. 2018). Indeed, complementing these findings, Robinson, Hawthorn et al. (2018) show that while visual distractibility doesn't differ across age, auditory interference was more pronounced early in development.

Alternatively, perhaps vocalisations have no effect (or disruption) on the body stimuli, but rather, children simply rely on the auditory stimuli when making a decision by some means of response inhibition. Contrary to the previous evidence regarding *rule* inhibition presented above, there is some evidence of a protracted development of *response* inhibition (Robinson and Sloutsky 2019). This developmental trajectory of response inhibition is shown to be linear across age (Schel and Crone 2013). These studies introduce an interesting disconnect, however, between lab based studies and natural day-to-day interactions. Namely, it is possible that in normal circumstances in which the body and voice are likely to

be congruent, where there are no rules to follow or forced choices to make, the former mechanism (disruption of encoding) is occurring. However, in the lab, it is entirely possible that by adding these additional components into emotion recognition, we tip the mechanism into the territory of response inhibition in children. We don't believe that these two mechanisms are mutually exclusive, but we should be aware of the potential disconnect between lab based and real-world conditions. Indeed, Robinson and Sloutsky (2019) found evidence for both mechanisms (termed 'Overshadowing' and 'Response competition') using subtle tweaks to their experimental design.

Perhaps then some combination of these two mechanisms explains the auditory dominance we find in the current study. One thing, however, that could be driving both mechanisms is attention. Or more specifically, a developmental difference in allocation of attentional resources.

A difference in attentional resource allocation?

One explanation for this auditory dominance could relate to attentional resources. We know that attention is limited in young children and improves across childhood (Reynolds and Romano 2016). When presented in isolation, young children were on par with adults in their ability to recognise and process the stimuli. However, when auditory and visual sensory information are presented simultaneously, children could be limited in what they can attend to due to restricted attentional resources (Napolitano and Sloutsky 2004). It is possible then that auditory dominance may emerge as auditory information is prioritised by children's limited resources (Robinson and Sloutsky 2004, Robinson and Sloutsky 2019).

This may be in line with brain imaging data detailing differences in brain maturation of the visual and auditory systems. We know that both the body-selective areas of the visual cortex (Ross, de Gelder et al. 2014, Ross, de Gelder et al. 2019) and voice areas of the temporal cortex (Bonte, Frost et al. 2013, Bonte, Ley et al. 2016) are maturing throughout childhood and adolescence. The evidence from these studies suggests that the auditory system shows an adult-like structural and selective pattern of activity earlier than the body-selective regions. Perhaps, then, younger children resort to reporting the voice in the incongruent trials in an attempt to use the best-developed neural resource, to increase ease of processing.

This may be a logical way to increase emotion recognition accuracy, as the emotions we perceive are rarely incongruent (Shackman and Pollak 2005). In real-world emotion perception, this auditory dominance may increase recognition accuracy, given that prosody is likely to accurately reflect bodily emotion also. Indeed, in Experiment 2 our older children showed significant increases in emotion recognition accuracy from the body when congruent emotional voices are included. In fact, accuracy also increases (albeit non-significantly) in both younger and older children across both experiments when given congruent vocal information on top of the body information. As a result, an evolutionary explanation of an auditory dominance in young children reconciles well with the present results. For younger children, it is imperative that threats are identified as early as possible, leading to an auditory dominance and emphasis on salient emotional information that can

be acted upon whilst a threat remains out of sight. However, for older children and adults who are taller with more developed survival skills, they are no longer completely dependent on this modality, explaining why auditory dominance is diminished as participants age.

The present research has several implications for improving education practices. Firstly, it is apparent that children under 8 years of age may respond most effectively to learning auditory information, as this is their dominant sensory modality. As visual information in learning is obviously necessary, irrelevant auditory information should be kept to a minimum. This is supported by research into language learning, which finds that the implementation of auditory-based teaching and learning styles results in more rapid memorisation of information, and increased participation in group discussions (Kayalar and Kayalar 2017). Additionally, it is important to consider the switch from auditory to visual dominance around 8 years. Teaching styles should adapt to this whilst reflecting that not all children may exhibit this switch at the same time, as 41% of children show a bimodal learning preference (Kharb, Samanta et al. 2013). Future research may also wish to apply the present findings directly to an educational setting, by investigating how sensory dominance (auditory or visual) affects learning at a range of ages.

One might also imagine that these results will be of interest to those working with populations for whom emotion recognition is somewhat atypical. For instance, Brewer, Biotti et al. (2017) showed that those individuals with autism spectrum disorder who show a reduced attention to faces could have a greater reliance on bodily cues for emotion recognition. In these individuals, given the differing levels of attention required to recognise emotion, would one expect to find evidence of auditory dominance?

It would also be useful to probe whether it is an auditory dominance for voices or for all auditory emotional information. Using emotional music, Paquette, Peretz et al. (2013) show that recognition of the 'basic emotions' (happiness, sadness, anger and fear) from music appears consistent across listeners. The same research group also showed evidence of a 'shared neural code' for aural emotional processing across different timbres, suggesting a universal acoustic code for auditory emotions that crosses aural modalities (Paquette, Takerkart et al. 2018, Sachs, Habibi et al. 2018). Therefore, by replacing the voices in the current study with emotional music stimuli, it should be possible to determine whether the auditory dominance we observe here with emotional stimuli is specifically related to the human voice, or whether it is found in other aural stimuli.

There is also the curious difference between the performance of the older children between Experiment 1 and Experiment 2. Could this be a case of familiarity with certain emotions being expressed in bodies compared to voices? Using children and adolescents, Ross et al., (2012) found no age x emotion interaction when recognising emotions from dynamic bodies, however, informally, sad and angry bodies appear to be slightly less well recognised than happy and fearful bodies. Lagerlof & Djef (2009) found the opposite results, with their youngest participants showing the highest scores for sad and angry bodies. It is hard, however, to pick apart these findings as they use different stimuli. In the current study, we have 2 groups who both used the same stimuli set and score higher, on average, when viewing sad/anger in bodies than happy/fear. This could also of course be a product of task, in which both tasks contained a binary choice. In terms of vocal emotions, Grosbras et al.

(2018) showed that across childhood and adolescence, anger was recognised significantly worse than the other 3 emotions (sadness, fear and anger). Furthermore, Happiness was significantly better recognised than all other emotions in pre-adolescents (similar to the age group we use here). This again doesn't square with the difference we find here across experiments. Perhaps our combination of emotions is having some effect here, in that one can be Sad *and* Angry, but not Happy and Fearful. So, hearing an emotional voice that could feasibly be linked to the emotional body may actually increase the likelihood of the auditory dominance. As mentioned above, if auditory dominance could be a way of increasing real-world emotion recognition accuracy when prosody is more likely to be congruent with the visual, then the effect of a more 'feasible' incongruent emotion (e.g. sadness and anger) could be to increase the auditory dominance effect. If we consider our first 'Overshadowing' mechanism, perhaps something akin to congruent processing is taking place. The Sad/Angry combination seems congruent, so more weight is put on the vocalisation in order to identify the whole stimuli. However, on the other hand it is worth remembering that here we asked participants to actively *ignore* the vocal stimuli in these trials.

To address this possibility and to further tease apart these mechanisms, the task itself could also be adapted. Here we asked subjects to ignore one modality and determine the emotion based on the other. However, we could simply ask which emotion is being portrayed rather than actively encouraging participants to ignore a modality. This could arguably let us eliminate the possibility of some sort of inhibition development being the cause of the effect and create more naturalistic conditions. Additionally, all four emotions could be contained in the same experiment, rather than splitting them as we have here. This would help to explain the difference we see in our older children across the two experiments as well as allowing us to examine the effect of emotion orthogonality on auditory dominance.

It should be noted that these interpretations may also be somewhat limited to the stimuli in this particular set. Indeed, one could imagine some differences if dynamic stimuli were used instead of static. The auditory stimuli already contain a temporal element, and this could present a disconnect between stimuli modalities. Furthermore, by not being dynamic the static bodies are arguably less ecologically valid and are perhaps less 'intense' than the vocal stimuli. Perhaps adults are able to overcome this disparity and respond correctly, but children may be more susceptible to differences in salience. This is not a trivial issue to overcome, but one method would be to create stimuli in which the vocal and body emotion are recorded simultaneously from the same actor.

There are also the ceiling effects in adults to consider. By perhaps adding noise to the stimuli or narrowing the temporal disparity in the stimuli as described above, we may see the Colavita effect influence adult's emotion recognition. Another useful addition here would be reaction time data. If the adult's recognition scores don't support the hypothesis due to a ceiling effect, it may be that the adults are slower in identifying the auditory emotion in incongruent trials (ignore body) compared to the bodies when ignoring voices. If this is the case then adding in a set response time window may also prove a way of reducing any ceiling effects in adults.

A further point that should be addressed in future work if one is going to consider attentional resources as an explanation for auditory dominance is that of formally

measuring physical attention. Here we closely monitored all subjects, but to avoid any possibility of subjects simply not viewing the visual stimulus during incongruent trials, one could implement catch trials or perhaps use eye tracking to make sure stimuli are being attended to throughout.

Conclusion

This is the first evidence to our knowledge of children showing an auditory dominance in relation to emotionally meaningful stimuli. Indeed, it seems that children find it very hard to ignore emotional auditory information even when directly instructed to do so. When presented with body and voice stimuli in isolation, children perform on par with adults. When presented with incongruent body/voice emotional stimuli and asked to ignore the body and give the emotion in the voice, children also perform on par with adults. However, when presented with the same stimuli and asked to ignore the voice and identify the emotion in the body, we find that children younger than 8 years of age are largely unable to do this. These results have fascinating implications for education, socio-emotional disorders and the study of emotion recognition development in a multi-modal environment.

References

- Ambadar, Z., J. W. Schooler and J. F. Cohn (2005). Deciphering the Enigmatic Face The Importance of Facial Dynamics in Interpreting Subtle Facial Expressions.
- Belin, P., S. Fillion-Bilodeau and F. Gosselin (2008). "The Montreal Affective Voices: a validated set of nonverbal affect bursts for research on auditory affective processing." Behav Res Methods **40**(2): 531-539.
- Birnholz, J. C. and B. R. Benacerraf (1983). "The development of human fetal hearing." Science **222**(4623): 516-518.
- Bonte, M., M. A. Frost, S. Rutten, A. Ley, E. Formisano and R. Goebel (2013). "Development from childhood to adulthood increases morphological and functional inter-individual variability in the right superior temporal cortex." Neuroimage **83**: 739-750.
- Bonte, M., A. Ley, W. Scharke and E. Formisano (2016). "Developmental refinement of cortical systems for speech and voice processing." Neuroimage **128**: 373-384.
- Boone, R. T. and J. G. Cunningham (1998). "Children's decoding of emotion in expressive body movement: the development of cue attunement." Developmental Psychology **34**(5): 1007-1016.
- Brewer, R., F. Biotti, G. Bird and R. Cook (2017). "Typical integration of emotion cues from bodies and faces in Autism Spectrum Disorder." Cognition **165**: 82-87.
- Broadbent, H., T. Osborne, M. Rea, A. Peng, D. Mareschal and N. Z. Kirkham (2018). "Incidental category learning and cognitive load in a multisensory environment across childhood." Developmental psychology **54**(6): 1020.
- Burr, D. and M. Gori (2012). Multisensory integration develops late in humans. The neural bases of multisensory processes, CRC Press/Taylor & Francis.
- Calvert, G. A. and T. Thesen (2004). "Multisensory integration: methodological approaches and emerging principles in the human brain." Journal of Physiology-Paris **98**(1-3): 191-205.
- Chronaki, G., J. A. Hadwin, M. Garner, P. Maurage and E. J. Sonuga-Barke (2015). "The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood." British Journal of Developmental Psychology **33**(2): 218-236.
- Colavita, F. B. (1974). "Human sensory dominance." Perception **3**(3): 409-412.
- Davidson, M. C., D. Amso, L. C. Anderson and A. Diamond (2006). "Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching." Neuropsychologia **44**(11): 2037-2078.
- de Gelder, B. and J. Van den Stock (2011). "The Bodily Expressive Action Stimulus Test (BEAST). Construction and Validation of a Stimulus Basis for Measuring Perception of Whole Body Expression of Emotions." Frontiers in Psychology **2**: 181.
- Elliot, M., A. Wing and A. Welchman (2010). "Multisensory cues improve sensorimotor synchronization." European Journal of Neuroscience **31**: 1828-1835.
- Gao, X. and D. Maurer (2010). "A happy story: Developmental changes in children's sensitivity to facial expressions of varying intensities." Journal of Experimental Child Psychology **107**(2): 67-86.
- Grosbras, M.-H., P. D. Ross and P. Belin (2018). "Categorical emotion recognition from voice improves during childhood and adolescence." Scientific Reports **8**(1): 14791.
- Hecht, D. and M. Reiner (2009). "Sensory dominance in combinations of audio, visual and haptic stimuli." Experimental brain research **193**(2): 307-314.
- Hirst, R. J., L. Cragg and H. A. Allen (2018). "Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect." Neuroscience & Biobehavioral Reviews.
- Hirst, R. J., J. E. Stacey, L. Cragg, P. C. Stacey and H. A. Allen (2018). "The threshold for the McGurk effect in audio-visual noise decreases with development."

- Kayalar, F. and F. Kayalar (2017). "The effects of Auditory Learning Strategy on Learning Skills of Language Learners (Students' Views)." IOSR Journal Of Humanities And Social Science (IOSR-JHSS) **22**(10): 4.
- Kharb, P., P. P. Samanta, M. Jindal and V. Singh (2013). "The learning styles and the preferred teaching-learning strategies of first year medical students." Journal of Clinical and Diagnostic Research **7**(6): 1089-1092.
- Koppen, C., A. Alsius and C. Spence (2008). "Semantic congruency and the Colavita visual dominance effect." Experimental brain research **184**(4): 533-546.
- Kuhn, L. K. (2014). Emotion Recognition in the Human Face and Voice.
- Lagerlof, I. and M. Djerf (2009). "Children's understanding of emotion in dance." European Journal of Developmental Psychology **6**(4): 409-431.
- Lewkowicz, D. J. and A. A. Ghazanfar (2009). "The emergence of multisensory systems through perceptual narrowing." Trends in cognitive sciences **13**(11): 470-478.
- Massaro, D. W. (1984). "Children's perception of visual and auditory speech." Child development: 1777-1788.
- McGurk, H. and J. MacDonald (1976). "Hearing lips and seeing voices." Nature **264**(5588): 746-748.
- Montirosso, R., M. Peverelli, E. Frigerio, M. Crespi and R. Borgatti (2010). "The Development of Dynamic Facial Expression Recognition at Different Intensities in 4- to 18-Year-Olds." Social Development **19**(1): 71-92.
- Morningstar, M., V. Y. Ly, L. Feldman and M. A. Dirks (2018). "Mid-Adolescents' and Adults' Recognition of Vocal Cues of Emotion and Social Intent: Differences by Expression and Speaker Age." Journal of Nonverbal Behavior **42**(2): 237-251.
- Napolitano, A. C. (2006). Modality dominance in young children: the underlying mechanisms and broader implications, The Ohio State University.
- Napolitano, A. C. and V. M. Sloutsky (2004). "Is a picture worth a thousand words? The flexible nature of modality dominance in young children." Child Development **75**(6): 1850-1870.
- Nava, E. and F. Pavani (2013). "Changes in Sensory Dominance During Childhood: Converging Evidence From the Colavita Effect and the Sound-Induced Flash Illusion." Child Development **84**(2): 604-616.
- Paquette, S., I. Peretz and P. Belin (2013). "The "Musical Emotional Bursts": a validated set of musical affect bursts to investigate auditory affective processing." Frontiers in psychology **4**: 509.
- Paquette, S., S. Takerkart, S. Saget, I. Peretz and P. Belin (2018). "Cross-classification of musical and vocal emotions in the auditory cortex." Ann. NY Acad. Sci **1423**: 329-337.
- Reynolds, G. D. and A. C. Romano (2016). "The Development of Attention Systems and Working Memory in Infancy." Front Syst Neurosci **10**: 15.
- Richardson, C., M. Anderson, C. L. Reid and A. M. Fox (2018). "Development of inhibition and switching: A longitudinal study of the maturation of interference suppression and reversal processes during childhood." Developmental Cognitive Neuroscience **34**: 92-100.
- Robinson, C. W., A. M. Hawthorn and A. N. Rahman (2018). "Developmental differences in filtering auditory and visual distractors during visual selective attention." Frontiers in psychology **9**: 2564.
- Robinson, C. W. and V. M. Sloutsky (2004). "Auditory dominance and its change in the course of development." Child Dev **75**(5): 1387-1401.
- Robinson, C. W. and V. M. Sloutsky (2019). "Two mechanisms underlying auditory dominance: Overshadowing and response competition." J Exp Child Psychol **178**: 317-340.
- Ross, P., B. de Gelder, F. Crabbe and M.-H. Grosbras (2014). "Body-Selective Areas in the Visual Cortex are less active in Children than in Adults." Frontiers in Human Neuroscience **8**.

Ross, P., B. de Gelder, F. Crabbe and M.-H. Grosbras (2019). "Emotion modulation of the body-selective areas in the developing brain." Developmental Cognitive Neuroscience **38**: 100660.

Ross, P. D., L. Polson and M. H. Grosbras (2012). "Developmental changes in emotion recognition from full-light and point-light displays of body movement." PLoS One **7**(9): e44815.

Sachs, M. E., A. Habibi, A. Damasio and J. T. Kaplan (2018). "Decoding the neural signatures of emotions expressed through sound." Neuroimage **174**: 1-10.

Sauter, D. A., C. Panattoni and F. Happé (2013). "Children's recognition of emotions from vocal cues." British Journal of Developmental Psychology **31**(1): 97-113.

Schel, M. A. and E. A. Crone (2013). "Development of response inhibition in the context of relevant versus irrelevant emotions." Frontiers in psychology **4**: 383.

Shackman, J. E. and S. D. Pollak (2005). "Experiential influences on multimodal perception of emotion." Child Development **76**(5): 1116-1126.

Sloutsky, V. M. and A. C. Napolitano (2003). "Is a Picture Worth a Thousand Words? Preference for Auditory Modality in Young Children." Child Development **74**(3): 822-833.

Tonks, J., W. H. Williams, I. Frampton, P. Yates and A. Slater (2007). "Assessing emotion recognition in 9-15-years olds: preliminary analysis of abilities in reading emotion from faces, voices and eyes." Brain Injury: [BI] **21**(6): 623-629.

Van den Stock, J., R. Righart and B. de Gelder (2007). "Body Expressions Influence Recognition of Emotions in the Face and Voice." Emotion **7**(3): 487-494.

Wille, C. and M. Ebersbach (2016). "Semantic congruency and the (reversed) Colavita effect in children and adults." Journal of Experimental Child Psychology **141**: 23-33.