Journal of Experimental Child Psychology 230 (2023) 105632

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Turn that music down! Affective musical bursts cause an auditory dominance in children recognizing bodily emotions

Paddy Ross [a,1,*], Ella Williams [a,b,1], Gemma Herbert [a], Laura Manning [a], Becca Lee [a]

[a] Department of Psychology, Durham University, Durham DH1 3LE, UK
[b] Oxford Neuroscience, University of Oxford, Oxford OX3 9DU, UK

ARTICLE INFO

Article history:
Received 10 October 2022
Revised 16 December 2022

Keywords:
Auditory dominance
Development
Emotion recognition
Body
Music

ABSTRACT

Previous work has shown that different sensory channels are prioritized across the life course, with children preferentially responding to auditory information. The aim of the current study was to investigate whether the mechanism that drives this auditory dominance in children occurs at the level of encoding (*overshadowing*) or when the information is integrated to form a response (*response competition*). Given that response competition is dependent on a modality integration attempt, a combination of stimuli that could not be integrated was used so that if children's auditory dominance persisted, this would provide evidence for the overshadowing over the response competition mechanism. Younger children (≤7 years), older children (8–11 years), and adults (18+ years) were asked to recognize the emotion (happy or fearful) in either nonvocal auditory musical emotional bursts or human visual bodily expressions of emotion in three conditions: unimodal, congruent bimodal, and incongruent bimodal. We found that children performed significantly worse at recognizing emotional bodies when they heard (and were told to ignore) musical emotional bursts. This provides the first evidence for auditory dominance in both younger and older children when presented with modally incongruent emotional stimuli. The continued presence of auditory dominance, despite the lack of modality integration, was taken as supportive evidence for the overshadowing explanation. These

* Corresponding author.
E-mail address: paddy.ross@durham.ac.uk (P. Ross).
[1] Paddy Ross and Ella Williams contributed equally to this work and are co-first authors.

https://doi.org/10.1016/j.jecp.2023.105632
0022-0965/

findings are discussed in relation to educational considerations, and future sensory dominance investigations and models are proposed.

Introduction

Humans have opportunities for rich and varied sensory experiences, with each sense providing information that is pieced together to form one's perception of the world (Ernst & Bülthoff, 2004). Although useful content may be extracted from singular modalities, research suggests that multimodal information enhances sensory events (Pascual-Leone & Hamilton, 2001) and improves cognitive processes (e.g., memory; Wammes et al., 2019). As such, multisensory processing plays a significant role in human experience and functioning. Critical to any multisensory processing discussion is sensory dominance, which describes the way in which the brain prioritizes one source of information over others when two or more sensory stimuli are presented simultaneously (Nava & Pavani, 2013). Colavita (1974) gave one of the earliest demonstrations of this modality preference. Adult participants were asked to decide whether they had perceived a light flash (visual) or brief tone (auditory) in unimodal trials (visual *or* auditory) and, critically, bimodal trials (visual *and* auditory). In these bimodal trials, participants most commonly reported the visual stimulus. This was taken as evidence for a visual dominance (VD), later coined the *Colavita effect,* in adults; the visual information appeared to override the auditory event.

For children, a collection of studies suggests that they show a reverse Colavita effect or auditory dominance (AD) (see meta-analysis in Hirst et al., 2018). Lewkowicz (1988) demonstrated that 6-month-old infants, despite being able to discriminate changes in unimodal trials, detected temporal changes in auditory but not visual aspects of compound presentations. Such detection was indicated by greater looking times only in response to auditory changes (i.e., dishabituation). Moreover, in a study by Sloutsky and Napolitano (2003), participants were taught to recognize compound stimuli, specifying locations of prizes. When presented with novel stimuli containing either trained auditory or visual components, children most commonly chose stimuli containing the auditory portion. These results further suggest that children prioritize auditory aspects of multimodal stimuli.

Considering the discussed evidence, sensory dominance appears to follow a developmental trajectory, with different modalities being prioritized at different ages (Hirst et al., 2018). However, there is a lack of consensus regarding when the transition to a VD occurs; one study discovered AD in 14-year-olds (Schorr et al., 2005), whereas another study found, across three different experiments, an AD in children up to 6 or 7 years of age and a clear VD in children aged 9 to 12 years (Nava & Pavani, 2013).

So far, most of the discussed literature has used basic artificial (as opposed to naturally occurring) bimodal stimuli. However, a more natural example of multimodal presentations may be found in the display of emotions. Ross et al. (2021) recently investigated sensory dominance in relation to emotionally meaningful stimuli. It was found that younger and older children struggled to ignore what they could hear (i.e., they demonstrated AD) when asked to recognize emotions in bodily expressions while listening to incongruent emotional human vocalizations. Emotion recognition accuracy decreased in incongruent bimodal trials when the children were instructed to respond only to the visual stimulus.

However, one question left unanswered by Ross et al. (2021) relates to the mechanism by which this AD transpires. Two possibilities labeled *response competition* and *overshadowing* were proposed (Robinson & Sloutsky, 2019). Response competition describes instances in which both auditory information and visual information are processed, however, when a response is required, children rely more heavily on the auditory stimulus. On the other hand, overshadowing suggests that auditory information overrides visual input during encoding, leading to attenuated visual encoding. Because the human vocalizations used in Ross et al.'s (2021) study could feasibly come from the figures in the visual stimuli, attempts to integrate the two modalities in the response phase and give the emotion one heard when asked to recognize the emotion one saw would be plausible. Therefore, their results could not support one mechanism over the other.

To dissociate these two potential mechanisms, instead of human vocalizations, the current study used short excerpts of music with strong emotional valences known as musical emotional bursts (MEBs). These MEBs were selected from a set of stimuli created and validated by Paquette et al. (2013) and are the "musical counterpart" of the affective voices used in Ross et al. (2021). Given that this nonvocal auditory stimulus could no longer come from the human figures, any attempt to integrate the two modalities (a central tenet to response competition) would be inhibited as it would not make sense for the music to be coming from the visual body stimuli. Therefore, it was posited that if AD persisted for this combination of stimuli, overshadowing would be a more likely explanation.

Here, we hypothesized that, as in Ross et al. (2021), young children would display an AD when hearing incongruent emotional music while recognizing emotional bodies. This effect was not expected for older children or adults.

Method

Participants

A total of 90 participants were tested (45 male). These participants were split into three age groups: younger children aged 7 years and under (n = 38; M_{age} = 5.90 years, SD = 1.03), older children aged 8 to 11 years (n = 31; M_{age} = 8.90 years, SD = 1.01), and adults aged 18 years and over (n = 21; M_{age} = 34.43 years, SD = 20.53). This categorization was informed by Nava and Pavani's (2013) results where AD was observed in children up to 7 years old and mirrors Ross et al. (2021).

Both adult and child participants had normal or corrected-to-normal vision and hearing. The adult participants were undergraduate students from Durham University who completed the study for course credits. The children either were tested during an after-school club at St Mary's RC primary school, Barnard Castle, United Kingdom, during a Sunday school at King's Church Durham, or were tested at a Junior Scientist event held at Durham University.

All adult participants gave informed consent before taking part in the study. Each child was asked whether they were happy to participate prior to testing, and a parent/guardian gave informed consent. All participants and the parents of the children were debriefed after the study. The study was approved by the psychology department advisory subcommittee at Durham University.

Stimuli

Body (visual) stimuli

The body stimuli were images from the Bodily Expressive Action Stimuli Test (BEAST; de Gelder & Van den Stock, 2011), which contains 254 whole-body expressions using 46 different actors portraying four different emotions (anger, fear, happiness, and sadness). The current study used 10 randomly selected happy images and 10 randomly selected fearful images from this database.

Musical (auditory) stimuli

The musical stimuli were taken from the Musical Emotional Bursts dataset, which contains 80 short musical excerpts that portray three emotions (fear, happiness, and sadness) as well as neutrality (Paquette et al., 2013). Each MEB was played on either the violin or the clarinet. Again, the current study used 10 randomly selected happy MEBs and 10 randomly selected fearful MEBs. These auditory stimuli were played through a set of headphones.

Design and procedure

This was a partial replication of Ross et al.'s (2021) study using four different experimental blocks, as illustrated in Fig. 1. In Block 1 (body only), participants were presented with the bodily expressions of emotion stimuli (10 happy and 10 fearful) and were asked to identify the emotions. The design used

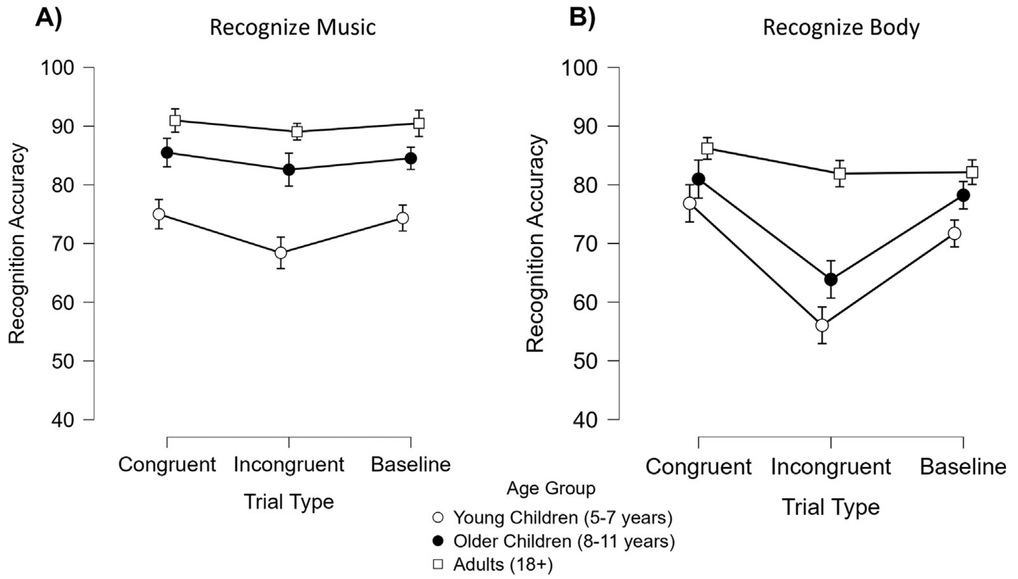

Fig. 1. Average recognition rates for trial types when participants were asked to recognize the emotional music bursts (A) and the emotion bodies (B). Error bars represent standard errors.

a forced-choice paradigm where the participants needed to select happy or fearful. Block 2 (music only) was an identical design but instead used the MEBs.

In Block 3 (ignore body), participants were presented with the bodily expressions and MEBs simultaneously. This combination consisted of 5 Happy Body/Happy Music, 5 Happy Body/Fearful Music, 5 Fearful Body/Happy Music, and 5 Fearful Body/Fearful Music stimuli. During Block 3, participants were asked to ignore what they saw and base their emotion identification on the music. Block 4 (ignore music) presented the same stimuli but asked participants to ignore the music and base their emotion identification on what they saw. There were 80 trials in total across all four blocks. Blocks and stimuli presentation were randomized for all participants, and each test took approximately 20 min.

For each trial using MATLAB (The MathWorks, Natick, MA, USA) and Psychtoolbox (Brainard, 1997), a fixation cross in the center of the 15-inch screen was displayed for 2000 ms. Participants sat 60 cm away from the screen, and visual stimuli were 4 $\times$ 10 cm, giving a visual angle of approximately 3.8 $\times$ 9.5°. Following the fixation cross, the visual stimulus was presented on the monitor and/or the musical stimulus was played through headphones for a total of 1500 ms. Participants were instructed to identify the emotion portrayed using either the F key for fearful stimuli or the H key for happy stimuli. Stickers with a happy or fearful face were placed on top of the appropriate keys as an aid for younger children. This procedure was repeated for all four experimental blocks.

Results

Cross-trial comparisons

By collapsing results across emotion, it was possible to compare bimodal congruent and incongruent trials with the unimodal baseline condition. Accordingly, a 3 (Age Group) $\times$ 3 (Trial Type) mixed analysis of variance (ANOVA) was conducted for each reporting condition (i.e., report music [Blocks 2 and 3] or report body [Blocks 1 and 4]).

An initial look at the baseline accuracy revealed that adults were significantly better at recognizing emotions from the music (M = 90.48, SD = 8.6) than from the body (M = 82.14, SD = 12.1, p < .05), whereas no difference was found in the accuracy of either child group across the modalities (older children—music: $M = 84.52$, $SD = 10.9$; body: $M = 78.23$, $SD = 18.7$, $p = .09$; younger children—music: $M = 74.34$, $SD = 14.5$; body: $M = 71.71$, $SD = 17.8$, $p = .42$).

When reporting the emotion in the music, we found no main effect of trial type, $F(2, 174) = 1.94$, $p = .15$, $\eta_p^2 = .01$, and no interaction of trial type and age group, $F(4, 174) = 0.36$, $p = .84$, $\eta_p^2 = .004$ (see Fig. 1A). We did see a main effect of age group, $F(2, 87) = 23.48$, $p < .001$, $\eta_p^2 = .35$, driven by adults and older children showing significantly higher accuracy than younger children (both *p*s < .001).

When reporting the emotion in the body, we also found a main effect of age group, $F(2, 87) = 8.45$, $p < .001$, $\eta_p^2 = .16$, driven by adults showing significantly higher overall accuracy compared with older children ($p < .05$) and younger children ($p < .001$). We found a main effect of trial type, $F(2, 1.74) = 17.93$, $p < .001$, $\eta_p^2 = .079$. Post hoc Bonferroni-corrected *t* tests found that participants scored significantly lower in the incongruent condition ($M = 64.78$, $SD = 22.8$) compared with both the congruent condition ($M = 80.44$, $SD = 18.5$, $p < .001$) and the baseline condition ($M = 76.39$, $SD = 17.35$, $p < .001$). There was no difference between the congruent and baseline conditions ($p = .31$). Crucially, we also found a significant interaction between trial type and age group, $F(4, 174) = 2.50$, $p < .05$, $\eta_p^2 = .022$. Follow-up analyses found this to be driven by younger children and older children scoring significantly lower in the incongruent condition compared with the congruent condition (younger children: $p < .001$; older children: $p < .001$) and the baseline condition (younger children: $p < .001$; older children: $p < .001$). Adults showed no significant difference in this comparison, and no group showed a significant difference between the congruent and baseline conditions (see Fig. 1B).

Discussion

The current results demonstrated auditory dominance (AD) in younger and older children. Specifically, incongruent MEBs impaired both groups of children's abilities to recognize emotion in the body. No such impairment was seen for adults. In addition, there were no significant differences between congruent and visual baseline trials in any age group when asked to recognize the emotion in the body, further indicating that the incongruent auditory information caused the impairment. In short, both groups of children demonstrated clear AD; their pronounced performance impairment was observed when asked to ignore incongruent auditory information.

These results did not fully support the hypothesis given that it was predicted that AD would occur only for younger children. Opposing this prediction, AD was also evident in older children. This finding contradicts Wagener et al.'s (2021) results where incongruent emotional music did not worsen 8- to 12-year-olds' facial emotion recognition abilities. Some possibilities for this disagreement could be, first, that task demands were different; Wagener et al. used facial emotion, whereas the current study tested bodily expressions of emotion. Second, the stimuli presentation times differed. In the current study, visual and auditory stimuli lasted for 1500 ms, whereas in Wagener et al.'s design the stimuli remained until participants responded. It could be argued, therefore, that the older children had longer to employ attentional flexibility, which counteracted the AD. It is still plausible that these older children processed the auditory component first but then had time to make a cognitive shift and encode the visual information (Robinson & Sloutsky, 2004).

Although contrasting with Wagener's et al.'s (2021) results, the current study did support Ross et al.'s (2021) findings; AD was again observed for younger and older children using emotional stimuli. Beyond this support for AD, the current study explored potential explanatory mechanisms. The current results demonstrated that AD persists when integration attempts (key features of response competition) are constricted. Of the two models being investigated, therefore, these results support the overshadowing mechanism.

However, it should be noted that although they are from different modalities, both signals are still of emotions and may well be getting integrated at some stage in the processing hierarchy (Aviezer et al., 2008). Unfortunately, due to the current design of the study and the binary response choice offered to participants, we were unable to explore whether there was any biasing effect from one modality in the judgment of emotion from another modality. This should be a key issue to address in future work given that more choice would allow some measure of information integration. It would also allow us to discriminate whether errors in the incongruent condition are due to AD or simply to the participants getting the answers wrong.

However, if we are to accept the overshadowing account here, an explanation is beneficial. One such exposition lies within an attentional resources account. It is known that, compared with adults, children have more limited attentional resources (Matusz et al., 2015), and this could mean that they prioritize one modality over another (Robinson & Sloutsky, 2019). Auditory information may be selected because the auditory system is more developed/efficient at these younger ages (Birnholz & Benacerraf, 1983; Robinson & Sloutsky, 2019).

In addition, auditory information is usually more dynamic and/or transient, so attention may be directed preferentially to auditory stimuli over more stable/permanent visual stimuli. In evolutionary terms, and in accord with the (unreported) greater recognition accuracies for fearful stimuli found in the current study, this preferential attending is especially important for threatening stimuli. Although these suggestions provide neat explanations for auditory overshadowing in children, it is hard to apply them to the VD reported elsewhere in adults (Colavita, 1974). Accordingly, it is probable that a combination of mechanisms explains sensory dominance across the life course (Barnhart et al., 2018; Robinson & Sloutsky, 2019).

An alternative explanation could involve the reliability of the individual signals. In other words, if emotion recognition is easier from one of the two modalities, then are children simply relying on the more reliable emotional signal in making their judgments? We found that older and younger children showed no difference in their ability to recognize emotions across the two modalities in the baseline condition. Therefore, here we found no evidence that performance on the baseline tasks predicted AD, so this is unlikely to be behind the effect.

Limitations and implications

It could be argued that the task was too complicated for the youngest participants. However, if this were the case, similar degrees of impairment would be expected across all experimental blocks. Instead, children showed a particular pattern of impairment when asked to respond to the body in incongruent trials. In addition, as a point of clarification, the youngest participants (3–4 years) were removed and the analysis was re-run. With this alteration, the same AD was found, so it is doubtful that the task demands confounded our results.

It has also been suggested that familiarity with auditory stimuli may influence AD. Specifically, Robinson and Sloutsky (2004) proposed that reduced familiarity with auditory stimuli intensifies AD during encoding processes because unfamiliar stimuli require more attention, thereby leaving fewer resources for visual processing. It is unlikely that participants were familiar with the MEBs prior to the experiment, and so it is acknowledged that this may have augmented the AD. To tease apart the contribution of unfamiliarity, future studies could use more familiar emotional music (e.g., famous classical pieces) to test the extent to which AD remains. To minimize effects of familiarity changes within the experiment (i.e., practice effects), block order was randomized for each participant.

Hearing abilities here were assessed informally, but no information was taken regarding the musical habits or abilities of participants. Perhaps in future work this could be taken into account given that level of familiarity with the instruments/type of auditory stimuli involved could influence performance on these tasks.

The current results have implications for educational settings. The described AD could explain why infants rely more heavily on auditory speech signals compared with visual cues when learning words (Havy et al., 2017). Furthermore, Havy and Zesiger (2017) discovered that children make cross-modal representations of words when presented with auditory depictions, but visual presentations remain unimodal, suggesting greater efficiency of auditory learning. Given that the current results imply that children preferentially attend to auditory input, AD may be capitalized on by using predominantly auditory teaching methods.

Or, in instances where visual resources are necessary for teaching, irrelevant background noise should be minimized to mitigate effects of AD. It is also important to acknowledge that our findings suggest that AD occurs during encoding. Therefore, modality considerations apply to learning processes regardless of whether an imminent decision is required (e.g., through testing). Moreover, given that emotions play a prominent role in school settings (e.g., achievement emotions; Pekrun & Stephens, 2012), it is particularly noteworthy that AD has been consistently observed using emotional stimuli.

Future directions

In the future, additional age groups (e.g., secondary school-aged children) could be tested to determine when AD/overshadowing might diminish. In addition, with reference to the suggestion that the more transient nature of auditory stimuli might act as a driving force for AD, future researchers could minimize differences between auditory and visual stimuli by using dynamic video clips rather than static images. One could also test additional modalities (e.g., haptic) to see whether auditory input overshadows other sensory information. However, more basic stimuli may be required for such an investigation because it is unclear how one might portray emotions via other senses.

Moreover, to determine whether this potential auditory overshadowing truly originates from the greater efficiency of the auditory system in children, a future study could test those who were born deaf and have since become hearing (e.g., through cochlear implants). We know that nonverbal emotion recognition and social competence suffers in those with cochlear implants (Wiefferink et al., 2012, 2013). Furthermore, being able to understand emotional sounds shows a quality of life improvement (Schorr et al., 2009). Indeed vocal emotion recognition scores are positively correlated with self-reported quality of life (Luo et al., 2018). Due to the initial auditory deprivation, one would expect the visual system to instead be more developed; accordingly, if system efficiency drives a potential overshadowing mechanism, VD would be expected. Some support for this premise has been found (Schorr et al., 2005), but the generalizability remains limited.

Finally, it would be useful to determine whether this potential auditory overshadowing leads to complete neglect of visual processing or just reduced encoding. Previous work suggests that additional congruent sensory information improves emotion perception (Paulmann & Pell, 2011). At first glance, the current results appear to support this claim; participants performed better in congruent bimodal trials in the ignore music block, albeit nonsignificantly. A "bonus" effect of congruency would depend on some degree of alternative modality processing, however, and this undermines the premise of complete auditory overshadowing. According to a complete overshadowing account, in the current study's congruent trials, children would have been responding only to the MEBs even when they were asked to ignore them. Given that we found overall better recognition accuracies for MEBs over visual stimuli, this could explain why the results for congruent trials in the ignore music block were better than the visual baseline without the need for any visual processing; the children were simply responding to the more easily recognized MEBs.

Furthermore, to further test whether AD is a result of complete or partial overshadowing, one might employ a similar method to Robinson & Sloutsky (2007), in which bimodal stimuli were first presented, followed by a test phase (e.g., discrimination task) involving unimodal visual components. By requiring a response based solely on visual components of the initially bimodal stimulus, it becomes possible to determine whether this portion was processed or completely neglected.

Conclusion

The current study provides further evidence for AD in children using emotional stimuli. Because the stimuli could not be integrated, the continued presence of AD potentially points toward the overshadowing mechanism. It has been noted, however, that overshadowing might not provide a comprehensive explanation for all sensory dominance phenomena across the lifetime. The implications of the current results extend particularly to educational practices, suggesting prescriptions for optimum teaching resources and learning environments. Future studies could further illuminate the mechanisms driving sensory dominance effects by testing individuals from different age groups across different tasks and with different sensory experiences.

Data availability

Data will be made available on request.

References

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., ... Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*(7), 724–732. https://doi.org/10.1111/j.1467-9280.2008.02148.x.

Barnhart, W. R., Rivera, S., & Robinson, C. W. (2018). Different patterns of modality dominance across development. *Acta Psychologica, 182*, 154–165. https://doi.org/10.1016/j.actpsy.2017.11.017.

Birnholz, J. C., & Benacerraf, B. R. (1983). The development of human fetal hearing. *Science, 222*(4623), 516–518. https://doi.org/10.1126/science.6623091.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics, 16*(2), 409–412. https://doi.org/10.3758/BF03203962.

de Gelder, B., & Van den Stock, J. (2011). The Bodily Expressive Action Stimulus Test (BEAST): Construction and validation of a stimulus basis for measuring perception of whole body expression of emotions Article 181. *Frontiers in Psychology, 2*. https://doi.org/10.3389/fpsyg.2011.00181.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8*(4), 162–169. https://doi.org/10.1016/j.tics.2004.02.002.

Havy, M., Foroud, A., Fais, L., & Werker, J. F. (2017). The role of auditory and visual speech in word learning at 18 months and in adulthood. *Child Development, 88*(6), 2043–2059. https://doi.org/10.1111/cdev.12715.

Havy, M., & Zesiger, P. (2017). Learning spoken words via the ears and eyes: Evidence from 30-month-old children Article 2122. *Frontiers in Psychology, 8*.

Hirst, R. J., Cragg, L., & Allen, H. A. (2018). Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. *Neuroscience & Biobehavioral Reviews, 94*, 286–301. https://doi.org/10.1016/j.neubiorev.2018.07.012.

Lewkowicz, D. J. (1988). Sensory dominance in infants: I. Six-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*(2), 155–171. https://doi.org/10.1037/0012-1649.24.2.155.

Luo, X., Kern, A., & Pulling, P. (2018). Vocal emotion recognition performance predicts the quality of life in adult cochlear implant users. *Journal of the Acoustical Society of America, 144*(5), EL429–EL435. https://doi.org/10.1121/1.5079575.

Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, R., & Scerif, G. (2015). Multi-modal distraction: Insights from children's limited attention. *Cognition, 136*, 156–165. https://doi.org/10.1016/j.cognition.2014.11.031.

Nava, E., & Pavani, F. (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. https://doi.org/10.1111/j.1467-8624.2012.01856.x.

Paquette, S., Peretz, I., & Belin, P. (2013). The "Musical Emotional Bursts": A validated set of musical affect bursts to investigate auditory affective processing Article 509. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00509.

Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. In C. Casanova & M. Ptito (Eds.). *Progress in brain research* (Vol. 134, pp. 427–445). Elsevier. https://doi.org/10.1016/S0079-6123(01)34028-1.

Paulmann, S., & Pell, M. D. (2011). Is there an advantage for recognizing multi-modal emotional stimuli? *Motivation and Emotion, 35*(2), 192–201. https://doi.org/10.1007/s11031-011-9206-0.

Pekrun, R., & Stephens, E. J. (2012). Academic emotions. In K. R. Harris, S. Graham, T. Urdan, J. M. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors* (pp. 3–31). American Psychological Association. https://doi.org/10.1037/13274-001.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development, 75*(5), 1387–1401. https://doi.org/10.1111/j.1467-8624.2004.00747.x.

Robinson, C. W., & Sloutsky, V. M. (2007). Auditory dominance: Overshadowing or response competition? In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 29). Cognitive Science Society. https://escholarship.org/uc/item/2jv5c1rk

Robinson, C. W., & Sloutsky, V. M. (2019). Two mechanisms underlying auditory dominance: Overshadowing and response competition. *Journal of Experimental Child Psychology, 178*, 317–340. https://doi.org/10.1016/j.jecp.2018.10.001.

Ross, P., Atkins, B., Allison, L., Simpson, H., Duffell, C., Williams, M., & Ermolina, O. (2021). Children cannot ignore what they hear: Incongruent emotional information leads to an auditory dominance in children. *Journal of Experimental Child Psychology, 204*. https://doi.org/10.1016/j.jecp.2020.105068 105068.

Schorr, E. A., Fox, N. A., van Wassenhove, V., & Knudsen, E. I. (2005). Auditory-visual fusion in speech perception in children with cochlear implants. *Proceedings of the National Academy of Sciences, 102*(51), 18748–18750. https://doi.org/10.1073/pnas.0508862102.

Schorr, E. A., Roth, F. P., & Fox, N. A. (2009). Quality of life for children with cochlear implants: Perceived benefits and problems and the perception of single words and emotional sounds. *Journal of Speech, Language, and Hearing Research, 52*(1), 141–152. https://doi.org/10.1044/1092-4388(2008/07-0213.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*(3), 822–833. https://doi.org/10.1111/1467-8624.00570.

Wagener, G. L., Berning, M., Costa, A. P., Steffgen, G., & Melzer, A. (2021). Effects of emotional music on facial emotion recognition in children with autism spectrum disorder (ASD). *Journal of Autism and Developmental Disorders, 51*(9), 3256–3265. https://doi.org/10.1007/s10803-020-04781-0.

Wammes, J. D., Jonker, T. R., & Fernandes, M. A. (2019). Drawing improves memory: The importance of multimodal encoding context. *Cognition, 191*. https://doi.org/10.1016/j.cognition.2019.04.024 103955.

Wiefferink, C. H., Rieffe, C., Ketelaar, L., De Raeve, L., & Frijns, J. H. M. (2013). Emotion understanding in deaf children with a cochlear implant. *Journal of Deaf Studies and Deaf Education, 18*(2), 175–186. https://doi.org/10.1093/deafed/ens042.

Wiefferink, C. H., Rieffe, C., Ketelaar, L., & Frijns, J. H. M. (2012). Predicting social functioning in children with a cochlear implant and in normal-hearing children: The role of emotion regulation. *International Journal of Pediatric Otorhinolaryngology, 76*(6), 883–889. https://doi.org/10.1016/j.ijporl.2012.02.065.