CHILD DEVELOPMENT

Child Development, March/April 2013, Volume 84, Number 2, Pages 604–616

Changes in Sensory Dominance During Childhood: Converging Evidence From the Colavita Effect and the Sound-Induced Flash Illusion

Elena Nava
University of Hamburg

Francesco Pavani
University of Trento

In human adults, visual dominance emerges in several multisensory tasks. In children, auditory dominance has been reported up to 4 years of age. To establish when sensory dominance changes during development, 41 children (6–7, 9–10, and 11–12 years) were tested on the Colavita task (Experiment 1) and 32 children (6–7, 9–10, and 11–12 years) were tested on the sound-induced flash illusion (Experiment 2). In both experiments, an auditory dominance emerged in 6- to 7-year-old children compared to older children. Adult-like visual dominance started to emerge from 9 to 10 years of age, and consolidated in 11- to 12-year-old children. These findings show that auditory dominance persists up to 6 years, but switches to visual dominance during the first school years.

The notion that our everyday experiences are substantially multisensory has made research on sensory interactions a central topic of cognitive neuroscience. The behavioral consequences and the neural correlates of multisensory interactions have been extensively investigated in human adults (for reviews, see Calvert, Spence, & Stein, 2004; Macaluso & Driver, 2005; Spence & Driver, 2004). However, the developmental aspects of multisensory processing have been primarily investigated in animal models adopting a neurophysiological perspective (e.g., Stein & Meredith, 1993; Wallace, Carriere, Perrault, Vaughan, & Stein, 2006) or in human infants adopting behavioral paradigms specifically tuned for this population (for a review, see Lewkowicz & Lickliter, 1994). Thus, several questions remain unanswered, particularly in the age range covering later childhood and adolescence. Among these, the developmental trajectory for sensory dominance and multisensory interactions still remains to be characterized (though see Ernst, 2008; Gori, Dal Viva, Sandini, & Burr, 2008; Nardini, Jones, Bedford, & Braddick, 2008).

The concept of sensory dominance is central in the study of multisensory processing. It refers to the phenomenon by which one sensory modality prevails or plays a relatively dominant role over the others, when two or more sensory systems are stimulated concurrently. Classic examples of sensory dominance are the phenomena of visual capture of the perceived sound location (the ventriloquist effect; e.g., Howard & Templeton, 1966) or visual capture of the location of own body parts (the rubber hand illusion; e.g., Botvinick & Cohen, 1998). Sensory dominance has been traditionally framed within the theoretical context of the "modality appropriateness hypothesis" (Welch & Warren, 1986), which postulates that the modality that dominates is the one that is more appropriate or reliable with respect to a given task. In the examples reported above, for instance, vision typically dominates over audition and proprioception in localization tasks, because under everyday conditions vision has a higher spatial resolution with respect to these other sensory systems (see also Alais & Burr, 2004; Ernst and Banks, 2002, for a Bayesian account of multisensory integration).

We thank the director, teachers, and pupils of "Istituto Comprensivo di Tarcento" (Tarcento, Italy) and the "Istituto Comprensivo Rovereto Est" (Rovereto, Italy) for their collaboration in the present study. We also thank Gabriella Scala and Serena Siega for their help in data collection. This research was supported by a PRIN grant (Italy) to F.P. and a fellowship from the Rovereto City Council (Comune di Rovereto, Italy). We are very grateful to Professor Charles Spence for the useful comments made to improve our manuscript. It certainly has now. We also thank the two anonymous reviewers for the constructive critics and comments to the manuscript. This study has been realized thanks to the support from the Provincia Autonoma di Trento (Italy) and the Fondazione Cassa di Risparmio di Trento e Rovereto (Italy).

Correspondence concerning this article should be addressed to Elena Nava, Department of Biopsychology and Neuropsychology, University of Hamburg, Von-Melle-Park 11, 20146, Hamburg, Germany. Electronic mail may be sent to elena.nava@uni-hamburg.de.

[Correction added on 10/11/2012, after first online publication 9/24/2012: The first and second authors' affiliations have been corrected.]

0009-3920/2013/8402-0017
DOI: 10.1111/j.1467-8624.2012.01856.x

Changes in sensory dominance are not just linked to the reliability of the stimuli, but also appear to occur in relation to development. While visual dominance has typically been documented in human adults, in infants, and children up to 4 years of age, auditory dominance tends to prevail. Studies in human infants (Lewkowicz, 1988a, 1988b) have suggested that there may be an asynchronous development for the sensory systems, leading infants to prefer the auditory over the visual modality when processing multisensory events. In his studies on 6- and 10-month-old infants, Lewkowicz (1988a, 1988b) presented participants with audiovisual compounds differing in temporal characteristics (i.e., rate or duration of stimuli presentation) of either the visual or auditory component. Results showed that infants (particularly those aged 6 months) detected temporal changes in the auditory, but not in the visual modality, indicating auditory dominance in infants. Lewkowicz (1988a) suggested that this auditory dominance in early development might be a vestige of the ontogenetically asynchronous development of the sensory systems. The auditory system starts being responsive to external input much before birth (for a review on fetal sensory abilities, see Lecanuet & Schaal, 1996), whereas the visual system is the least stimulated sense, as it only receives very low light intensities in utero throughout gestation, suggesting that the visual receptors will only start being fully stimulated after birth. The term *responsiveness* should here be meant as the ability of the fetus to react to a sensory input by showing, as in the case of auditory stimulation, a motor response, or a change in heart rate (Lecanuet & Schaal, 1996).

Further behavioral studies suggested that auditory dominance could persist up to 4 years of age (Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003). These investigations aimed at drawing a developmental trajectory for sensory processing in infants aged 8, 12, and 16 months (Robinson & Sloutsky, 2004) and in 4-year-old children (Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2007; Sloutsky & Napolitano, 2003). Similar to Lewkowicz (1988a, 1988b), sensory dominance was tested by training infants and children on audiovisual compound stimuli that varied in either the auditory or visual component. Overall, results showed auditory dominance for both infants and children, although this dominance appeared to be less strong in 4-year-old children, where it was sensitive to several manipulations on the stimuli, such as changes in stimulus familiarity (Napolitano & Sloutsky, 2004). Control studies on attentional factors that may explain auditory dominance in children led these authors to conclude that children may be automatically attracted to auditory rather than visual stimuli, because of the higher alerting nature of sounds (Posner, Nissen, & Klein, 1976).

These developmental studies have clearly shown that sensory dominance is present as a phenomenon during childhood too, and that children up to 4 years of age exhibit an auditory dominance. However, as research on sensory dominance has rarely extended beyond 4 years of age, it remains unclear the developmental pattern for sensory dominance beyond this age. What research has clearly shown, though, is that adults are dominated by vision under several circumstances (e.g., Howard & Templeton, 1966; McGurk & MacDonald, 1976; Pavani, Spence, & Driver, 2000; Spence, Pavani, Maravita, & Holmes, 2008), suggesting that a change in sensory dominance must take place at some point during development. Sensory dominance in adults has been particularly investigated in multisensory conflicting situations (e.g., ventriloquist effect, rubber hand illusion), in which, although, sensory dominance could have been triggered by the particular context and task.

In 1974, Frank Colavita reported a particularly striking case of visual dominance. He found that by simply presenting an auditory and a visual stimulus, either separately or concurrently, and by asking participants to report whether a visual, an auditory, or an audiovisual stimulus was perceived, most of the participants failed to report the auditory event in the critical audiovisual (bimodal) trials. In other words, visual dominance in adults can be strong enough to override (or overshadow) awareness for the auditory event when the two stimuli are simultaneously presented (for a review of the research on the Colavita effect, see Spence, 2009). It is worth noting that in his first experiment, Colavita (1974) reported visual dominance under some confounding conditions (e.g., he did not inform participants about the presence of bimodal stimuli, and these were rarely presented). However, a series of experiments conducted by Koppen and Spence (2007a) showed that even by manipulating probability of occurrence of stimulation type (Experiment 2), response demand (Experiment 3), or attention (Experiment 4), the Colavita effect still proved robust.

The fact that one sensory modality can strongly influence or even overshadow another sense under some circumstances (e.g., as in the Colavita effect) suggests that the brain does not always equally rely on multiple sources of information. From a developmental perspective, a different sensory dominance in children as compared to adults may suggest different processes underpinning multisensory interactions. The few studies that have addressed multisensory development in school-aged children suggest that some abilities (e.g., optimal multisensory integration) may appear late in childhood, possibly as the result of recalibration of the sensory systems throughout experience (Ernst, 2008; Gori et al., 2008; Nardini et al., 2008). These studies have shown that integration of visuo haptic (Gori et al., 2008) and visuoproprioceptive (Nardini et al., 2008) information is still immature at age 8 years. In addition to these studies that have examined visuohaptic and visuoproprioceptive integration in children up to 10 years of age, other studies in school-aged children have investigated audiovisual interactions adopting the McGurk effect (Massaro, 1984; Schorr, Fox, Wassenhove, & Knudsen, 2005; Tremblay et al., 2007). In this well-known audiovisual illusion, adults typically misperceive spoken syllables when matched with incongruent lip movements, showing a strong influence of visual input. Children aged 4–10 years (Massaro et al., 1986), 5–14 years (Schorr et al., 2005), and 5–9 years (Tremblay et al., 2007) show a diminished McGurk illusion as compared to adults, thus proving more influenced by the auditory than the visual component of the stimulation. In sum, these studies suggest that the mature ability to integrate audiovisual percepts emerges late in life. Moreover, they indicate that auditory dominance may likely persist beyond 4 years of age.

In this study, we investigated the development of multisensory interactions in three school-aged groups of children using the Colavita paradigm, with the aim of directly assessing whether auditory dominance persists beyond 4 years of age and to examine when visual dominance starts to emerge. There are several reasons why we chose this paradigm over others to assess sensory dominance. First of all, it is extremely simple, which makes it easily adaptable for testing children. Second, stimuli in this experiment can be presented from the same spatial position and simultaneously, avoiding spatial and temporal confounds. Third, this paradigm does not make use of linguistic material, therefore avoiding potential confounds arising from different linguistic skills in this developmental population. Finally, and most importantly, this paradigm has undergone a series of manipulations that proved its robustness and consistency. Colavita himself (1974), for example, manipulated stimuli intensity so that they were perceived as equally intense by participants (Experiment 1) or even doubled the intensity of the auditory stimulus (Experiment 2). Both cases proved ineffective in modulating or reversing the effect. More recent investigations on this effect examined the role of stimulus type (i.e., simple vs. complex stimuli, for example, an auditory tone vs. the barking of a dog; Koppen, Alsius, & Spence, 2008), relative position of the stimuli (Koppen & Spence, 2007b), stimulus probability (Koppen & Spence, 2007a), response demands (Koppen & Spence, 2007a), and attention (Koppen & Spence, 2007a; Sinnett, Spence, & Soto-Faraco, 2007). Interestingly, although all these factors proved effective in changing the size of the effect, none could extinguish the visual dominance, thus proving the robustness and consistency of the Colavita task.

Given these premises, we hypothesized that if relative dominance of audition and vision changes during early school years, a correspondent change in the direction or magnitude of the Colavita effect (i.e., visual dominance) as a function of age could be observed.

Experiment 1a

Method

Participants. Fourteen children aged 6–7 years (7 female subjects, mean age = 6.8 years, *SD* = .3), 14 children aged 9–10 years (8 female subjects, mean age = 9.5 years, *SD* = .3), and 13 children aged 11–12 years (9 female subjects, mean age = 11.7 years, *SD* = .3) were recruited from the local state school (Istituto Comprensivo, Tarcento, Italy) to take part in the study. All children were Caucasian. All children performed the task after their parents had given informed consent. The study was approved by the Research Ethics Board of the University of Trento. All participants had normal or corrected-to-normal vision and normal hearing by self-report.

Stimuli and procedure. The visual stimulus consisted of a uniform yellow disk subtending approximately 2° of visual angle and presented at the center of a black-background computer monitor. The auditory stimulus consisted of a 4.0 kHz tone, presented at 65 dB from two loudspeakers positioned on opposite sides of the computer screen. Stimuli were generated and presented using a Dell Latitude D820 laptop. Stimulus programming, presentation, and response collection was carried out using E-Prime (http://www.pstnet.com/).

The experimental session was divided into four blocks, each comprising 100 trials. Out of the 100 trials, there were 40 trials in which the visual stimulus was presented alone (unimodal visual), 40 trials in which the auditory stimulus was presented alone (unimodal auditory), and 20 trials in which the auditory and the visual stimuli were presented simultaneously (bimodal stimulus). Stimulation type was randomized within each block. Each stimulus lasted for 50 ms, followed by 2000 ms for the manual response, and an additional 500 ms interstimulus interval (ISI). The next stimulus was thus always presented 2500 ms after the preceding stimulation, even if a response was not entered. Participants sat approximately 60 cm from the computer monitor, and were explicitly informed that they would be presented with visual stimuli, auditory stimuli, or audiovisual stimuli. They were instructed to discriminate as fast as possible between the three stimulation types and enter the corresponding response on the computer keyboard using two clearly marked keys. Half of the participants used *c* for unimodal visual stimuli, *m* for unimodal auditory stimuli, and both keys for the bimodal stimulus; for the other half of the participants, the response mapping for auditory and visual stimuli was swapped. No feedback was provided, and after each block, a short break was allowed if needed. The experiment took approximately 30 min to complete.

This experimental procedure is the standard one adopted in the literature on the Colavita effect (e.g., see Koppen & Spence, 2007a), and we only varied intervals between stimuli presentation (i.e., longer compared to experiments on adults) to make them more suitable for children. This way, any difference found in children could be indirectly compared with results found in adults.

Results

Misses. The percentage of trials without a response (i.e., misses) was entered into a mixed analysis of variance (ANOVA) with age group (6–7, 10–11, or 11–12 years old) and target stimulus (unimodal visual, unimodal auditory, and bimodal) as variables. This analysis revealed no significant main effect or interaction (all Fs < 1.9; see Table 1).

Error data. Error data (in percentage) for trials in which participants made a response were entered into a mixed ANOVA with age groups and target stimulus as variables. All children made more errors in bimodal (22%, SE = 3%) than unimodal (visual: 8%, SE = 1%; auditory: 8%, SE = 1%) trials. This effect was particularly pronounced for children aged 6–7 years (see Figure 1, top plot; see also Table 1), resulting in a significant interaction between age group and target stimulus, $F(4, 76) = 3.06$, $p = .02$, partial eta squared = .138. The main effect of age group, $F(2, 38) = 5.40$, $p = .009$, partial eta squared = .225, was also significant, but subsidiary to the higher order two-way interaction described above.

To test for the presence of the classic Colavita effect (i.e., predominant visual responses in bimodal trials) in the different age groups, we analyzed whether visual-only responses exceeded auditory-only responses in the erroneous bimodal trials. To this end, the percentage of errors in bimodal trials was entered into a mixed ANOVA with age group and response (auditory only or visual only) as variables. This analysis revealed a significant interaction between age group and response, $F(2, 38) = 3.43$, $p = .04$, partial eta squared = .151. As illustrated in Figure 1 (bottom plots), a significant Colavita effect emerged for the 9- to 10-year-olds, with more visual- than auditory-only responses in 8 out of 14 children, $t(13) = 2.79$, $p = .02$; Cohen's $d = 0.74$. This effect was even more pronounced for the 11- to 12-year-olds, in which 12 out of 13 children gave more visual than auditory responses, $t(12) = 3.82$, $p = .002$; Cohen's $d = 1.07$. By contrast, children aged 6–7 years old showed no preference for the two responses, and only 6 out of 14 made more visual- than auditory-only errors, $t(13) = 1.37$, $p = .2$. In fact, two children in this age group almost exclusively gave auditory-only responses in bimodal trials (specifically, 95% and 99% of responses were auditory only).

Response time (RT) data. Response time data for those trials in which participants responded correctly were filtered to exclude RTs faster than 150 ms or slower than 2000 ms (< 1% of the trials were excluded on the basis of this criterion). RTs were then entered into a mixed ANOVA with target modality (auditory or visual) and target type (unimodal or bimodal) as within-participant variable, as well as age group as a between-participant variable. The two children in the 6- to 7-year-old age group who failed to respond in most bimodal trials were excluded from the analysis, due to insufficient RT data for the bimodal condition.

This analysis revealed a significant main effect of target type, $F(1, 36) = 35.2$, $p < .001$, partial eta squared = .494, caused by faster responses in unimodal (825 ms, SE = 41 ms) than bimodal (911 ms,

Table 1

Summary Table Showing Percentage of Mean Response for Misses, Errors, and Reaction Times, Separately for Group and Experiment (Experiments 1a and 1b).

	Experiment 1a			Experiment 1b
	6–7 years old	9–10 years old	11–12 years old	6–7 years old
Misses (%)				
Unimodal auditory	4.0% (1.1)	2.6% (1.1)	1.2% (1.2)	2.3% (1.1)
Unimodal visual	5.4% (1.0)	2.5% (1.0)	1.1% (1.1)	3.2% (1.2)
Bimodal	5.2% (1.9)	3.4% (1.9)	3.0% (1.9)	1.7% (1.0)
Error rates (%)				
Unimodal auditory	12.3% (2.5)	7.3% (2.5)	4.1% (2.6)	8.1% (2.4)
Unimodal visual	9.8% (2.1)	7.3% (2.1)	6.3% (2.1)	4.2% (1.1)
Bimodal	35.8% (5.7)	12.4% (5.7)	16.3% (5.9)	4.7% (1.7)
Bimodal error rates (%)				
Auditory-only responses	25% (5.4)	4% (5.4)	5% (5.6)	3.9% (1.3)
Visual-only response	11 (2.3)	9% (2.3)	11% (2.4)	0.8% (0.5)
RTs of correct responses (ms)				
Unimodal				
Auditory stimulus	934 (58)	908 (54)	757 (56)	1,066 (58)
Visual stimulus	895 (48)	808 (45)	646 (47)	1,009 (65)
Bimodal				
Auditory stimulus	1,006 (65)	953 (60)	798 (62)	1,095 (61)
Visual stimulus	985 (68)	942 (63)	800 (66)	1,060 (78)

Note. Standard errors are shown in parentheses.

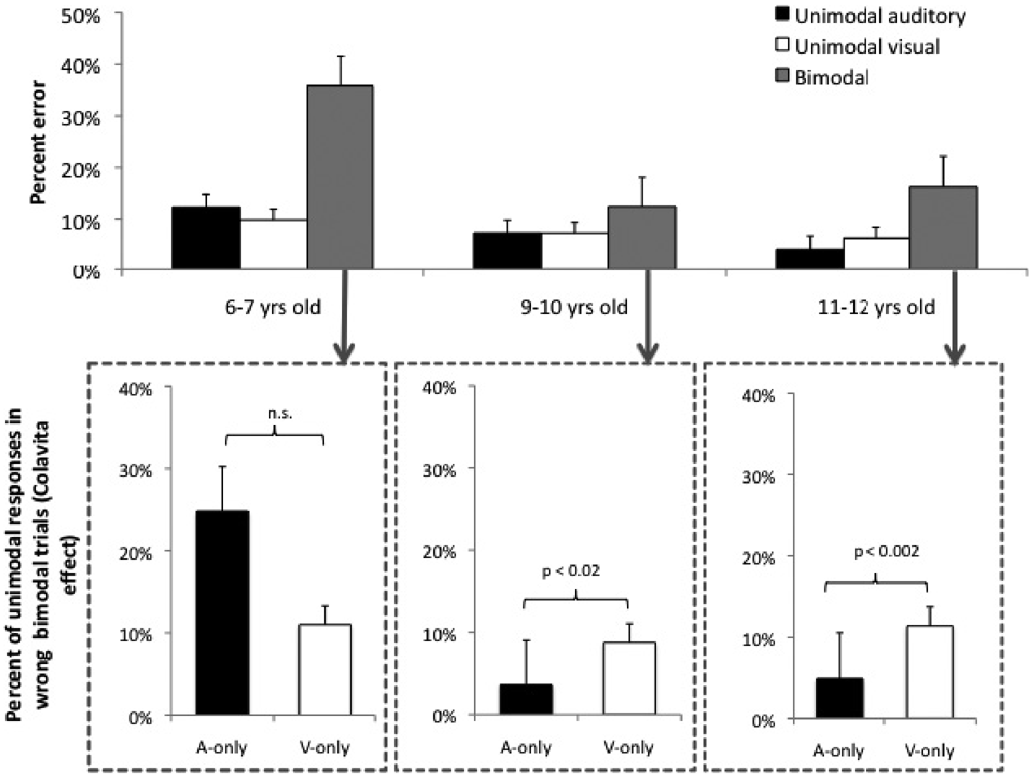

Figure 1. Top plots show percentage of errors for Experiment 1a, separately for type of stimulation (visual, auditory or audiovisual) and age group. Bottom plots show visual and auditory responses to erroneous audiovisual trials, separately for age group.
Note. Note that the difference between visual and auditory-responses was significant only for 9- to 10-year-olds and 11- to 12-year-olds.

$SE = 52$ ms) trials. There was also a significant main effect of target modality, $F(1, 36) = 27.6$, $p < .001$, partial eta squared = .435, attributable to faster responses to visual (846 ms, $SE = 45$ ms) than auditory (889 ms, $SE = 48$ ms) stimulation. However, this RT difference in responding to the different modalities emerged selectively for unimodal trials, resulting in a significant interaction between target type and target modality, $F(1, 36) = 21.8$, $p < .001$, partial eta squared = .378. Response times also decreased progressively across the three age groups (6–7 years: 955 ms, $SE = 58$ ms; 9–10 years: 898 ms, $SE = 54$ ms; 11–12 years: 750 ms, $SE = 56$ ms), resulting in a main effect of group, $F(2, 36) = 3.5$, $p = .04$, partial eta squared = .163. Most interesting, the three-way interaction between target type, target modality, and group was also marginally significant, $F(2, 36) = 3.1$, $p = .057$, partial eta squared = .148. Children aged 6–7 years old were equally fast in responding to visual and auditory unimodal stimulation (RT difference 39 ms), unlike older children that systematically responded faster to visual than auditory unimodal stimulation (9–10 years old: RT difference = 100 ms, $p = .001$ on Newman–Keuls post hoc test; 11–12 years old: RT difference = 111 ms, $p = .001$ on Newman–Keuls post hoc test; see Figure 2 and Table 1).

Discussion

Two main results emerged from Experiment 1a. First, a clear Colavita effect (i.e., a higher proportion of visual-only than auditory-only responses in response with bimodal trials) was clearly present in children aged 9–10 and 11–12 years, whereas no systematic preference for visual-only than auditory-only responses emerged in children aged 6–7 years old. Notably, two of the children in the youngest age group reported almost exclusively the auditory stimulus during bimodal trials, despite responding correctly to most visual stimuli when presented alone. The second finding of Experiment 1a is that children of 9–10 and 11–12 years of age were systematically faster in responding to visual than auditory targets when these were presented unimodally. By contrast, no RT difference in responding to the two sensory modalities emerged in children aged 6–7 years. Taken together, these findings indicate that visual dominance can emerge reliably from 9 to 10 years of age, but is still not fully developed in children aged 6–7 years.

Compared to the older groups, the youngest children showed the highest proportion of bimodal errors. This does not constitute a problem for the RT result (which was driven by performance in unimodal trials only), but raises the possibility that the different proportion of visual and auditory-responses between groups in the bimodal trials in which participants failed to make one of the two responses (i.e., the Colavita effect) could result from a difference in task-difficulty to some extent. In other words, we may have failed to reveal a Colavita effect in children aged 6–7 years because the bimodal task was far too difficult for this age group, resulting in randomly distributed visual-only and auditory-only responses. To examine this possibility, we recruited a new group of 6- to 7-

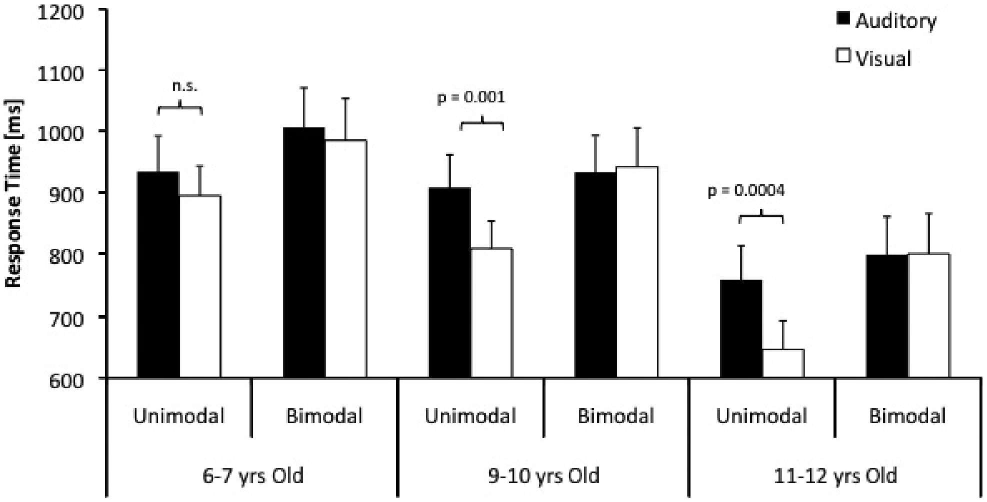

Figure 2. Response time in Experiment 1a as a function of target type and age group.
Note. Error bars indicate the standard error of the mean.

year-old children, and tested them in a simplified version of the task.

Experiment 1b

Method

Participants. Ten new participants aged 6–7 years ($n = 10$; 3 female subjects, mean age = 6.9 years, $SD = .3$) were recruited from the local state school (Istituto Comprensivo Rovereto Est, Rovereto, Italy) to take part in Experiment 1b. All children were Caucasian. All children performed the task after their parents had given informed consent. All participants had normal or corrected-to-normal vision and normal hearing by self-report.

Stimuli and procedure. These were identical to Experiment 1a with the following exceptions. The ISI between successive target presentations was increased to 6000 ms, to allow plenty of time for the response before the onset of the subsequent trial. Furthermore, the proportion of bimodal trials was increased from 20% to 33%, to limit any response bias toward the more frequent unimodal trials that might have been present in Experiment 1a.

Results

Misses. Participants failed to respond on 2.4% of the trials. The percentage of misses did not vary as a function of target stimulus, $F(2, 18) = 1.36$, $p = .3$ (see Table 1).

Error data. Error data (in percentage) for trials in which participants made a response were entered into a repeated measures ANOVA with target stimulus (unimodal visual, unimodal auditory, and bimodal) as the sole factor. This analysis revealed no statistical difference in the percentage of errors as a function of stimulation conditions, $F(2, 18) = 2.02$, $p = .2$ (see Table 1), indicating that the simplification we introduced in the paradigm proved effective in reducing the overall percentage of errors and specifically those occurring in bimodal trials (Experiment 1a: 35.8%; Experiment 1b: 4.7%). Most important, there was still a difference between the percentage of visual-only and auditory-only responses in erroneously responded bimodal trials. Children aged 6–7 years old produced more auditory-only (3.9%, $SE = 1.3\%$) than visual-only (.8%, $SE = 0.5\%$) responses, $t(9) = 3.20$, $p = .01$, two-tailed; Cohen's $d = 1.22$, resulting in a *reversed* Colavita effect.

RT data. RT data for correctly responded trials were filtered as in Experiment 1a (5% of the trials were excluded on the basis of this criterion). RTs were then entered into a repeated measures ANOVA with target modality (auditory or visual) and target type (unimodal or bimodal) as within-participant variables. This analysis revealed no significant main effect or interaction (all Fs < 2.8), confirming the pattern of results for 6- to 7-year-old children found in Experiment 1a.

Discussion

Experiment 1b shows that when the task is made easier (i.e., with longer ISI between stimuli and balanced proportions of bimodal and unimodal trials), children aged 6–7 years old exhibit a reversed Colavita effect, that is, auditory dominance. Speed of response was not affected by the simplification of the task, leading to comparable response times to auditory and visual stimuli, as previously seen in Experiment 1.

Taken together, the results of Experiment 1a and b show that auditory dominance persists beyond 4 years of age, but starts to change toward a visual dominance in early school years. This conclusion is consistent with previous evidence that showed less strong McGurk effect in primary school children (Massaro, 1984; Schorr et al., 2005; Tremblay et al., 2007). Unlike these previous findings, however, the present results are not contaminated by potential developmental changes in multisensory processing of verbal materials. Another important difference between the McGurk task and the Colavita paradigm adopted, herein, is that the interaction between visual and auditory events in the McGurk effect results in a single combined percept, whereas the interaction between the bimodal events in the Colavita task produces a competition for awareness that eventually leads to complete overshadowing of one of the two sensory modalities.

To validate and extend this novel observation of developmental changes in late childhood for multisensory perception, we tested three new groups of participants with comparable ages with the groups tested in Experiment 1 in another well-characterized audiovisual paradigm: the sound-induced flash illusion (Andersen, Tiippana & Sams, 2004; Shams, Kamitani, & Shimojo, 2002). In this task, participants judge the number of sequentially presented flashes, while trying to ignore the concurrent distracting beeps. When the number of beeps exceeds the number of flashes, participants typically report more flashes than actually presented, resulting in so-called visual fissions. By contrast, when the number of beeps is lower than the number of flashes, participants report fewer flashes than actually presented, perceiving the so-called visual fusions. As such, this paradigm is usually taken to show an overall dominance of the auditory modality when processing the temporal features of the stimuli. We adopted this audiovisual paradigm under the hypothesis that the changes in sensory dominance as a function of age that we documented in Experiments 1a and 1b should also produce a differential influence of distracting auditory stimuli in the flash-beep paradigm between younger children (aged 6–7 years old) and older children (aged 9–10 and 11–12 years old). In particular, we expected more sound-induced illusory fissions and fusions in young compared with older children.

It is worth noting that this paradigm, similar to the Colavita effect, has undergone some manipulations to test its robustness. For example, Andersen et al. (2004) manipulated the intensity of the auditory stimuli (Experiment 2, Block 1) by presenting 10 dB auditory stimuli (compared to the 80 dB stimuli used in Shams et al.'s, 2002, experiment). Interestingly, the decreased saliency of the auditory stimulus did not eliminate the illusion (although it diminished the size). Recently, Rosenthal, Shimojo, and Shams (2009) tested the possibility that the illusion could be reduced by providing participants with feedback. Even in this case, the illusion proved resistant to this type of manipulation. In addition, Setti and Chan (2011) found that the illusion persists even when stimuli are manipulated in their complexity (i.e., faces and buildings). Finally, we adopted this paradigm because, just as for the Colavita task, the sound-induced flash illusion does not entail any multisensory processing of verbal materials.

Experiment 2

Method

Participants. Ten children aged 6–7 years (3 females, mean age = 6.8 years, *SD* = .2), 12 children aged 9–10 years (2 females, mean age = 10.1 years, *SD* = .2), and 10 children aged 11–12 years (5 females, mean age = 12.2 years, *SD* = .1) were recruited from the local state school (Istituto Comprensivo Rovereto Est, Rovereto, Italy) to take part in the study. All children were Caucasian. All children performed the task after their parents had given informed consent. Eight adult participants (two females, mean age = 31 years, *SD* = 5) were also recruited from the student population from the University of Trento (Italy) to validate our version of the flash-beep paradigm. All participants had normal or corrected-to-normal vision and normal hearing by self-report.

Stimuli and procedure. Visual and auditory stimuli were the same we adopted for the Colavita task, for comparison with Experiment 1. Visual stimuli were either presented alone (visual-only conditions: 1, 2, or 3 flashes) or combined with 1, 2, or 3 auditory beeps (visual-auditory conditions). In the bimodal trials, the number of visual and auditory stimuli was either congruent (e.g., two visual stimuli paired with two auditory stimuli) or incongruent (i.e., two visual stimuli paired with one auditory stimulus). Each flash lasted 17 ms, each beep lasted 7 ms. When multiple beeps were presented, they were spaced 57 ms from one another, whereas multiple flashes were spaced 50 ms from one another. In bimodal conditions, beeps preceded the flashes by 23 ms. The duration of the stimuli and the time interval between auditory and visual stimuli in the bimodal conditions were taken from Shams et al. (2002).

Participants sat approximately 60 cm from the monitor and were asked to judge the number of sequentially presented flashes by tapping their response on the corresponding number of the computer keyboard. The experiment consisted of a short practice session and two blocks of 120 trials each (20 trials per condition), with all conditions randomly interleaved. Participants were asked to fixate a small white cross, presented at the center of the monitor throughout the experiment. Visual stimuli appeared at the center of a black-background computer monitor. The experiment took approximately 30 min to complete.

Results

V-only and V–A congruent conditions. We first assessed the ability to discriminate the number of presented flashes across groups when visual stimuli were presented alone (V-only conditions) and when they were delivered together with a matching number of beeps (V–A congruent conditions). To this end, we measured discrimination errors as the difference between the reported number of flashes in each condition and the actual number of presented flashes. Positive and negative values indicate overestimation and underestimation of the number of presented flashes, respectively.

Discrimination errors were entered into a mixed ANOVA with number of presented flashes (1, 2, or 3) and condition (V-only or V–A congruent) as within-participant variables, and group (6–7 years old, 9–10 years old, 11–12 years old, and adults) as between-participant variable. This analysis revealed a significant main effect of number of presented flashes, $F(2, 72) = 154.65$, $p < .001$, partial eta squared = .811 (see Table 2). Participants overestimated the number of flashes when a single visual stimulus was presented ($M = .17$, $SE = .03$, $p < .001$ in a t test against zero), they were overall accurate when two visual stimuli were presented ($M = .04$, $SE = .03$, $p = .2$ in a t test against zero), and underestimated the number of flashes when three visual stimuli were presented ($M = -.44$, $SE = .04$, $p < .001$ in a t test against zero). Underestimation of flash numerosity decreased with age (6–7 years: $M = -.65$, $SE = .07$; 9–10 years: $M = -.46$, $SE = .06$; 11–12 years: $M = -.39$, $SE = .07$; adults: $M = -.20$, $SE = .08$), resulting in a significant interaction between group and number of presented flashes, $F(6, 72) = 6.82$, $p = .001$, partial eta squared = .362.

The three-way interaction between group, number of presented flashes, and condition also reached significance, $F(2, 72) = 5.94$, $p < .001$. This interaction was caused by improved performance in V–A congruent as compared to V-only conditions in children aged 6–7 and 9–10 years old, particularly when two and three visual stimuli were presented. The main effect of condition as well as the two-way interactions involving this variable were also significant (all Fs > 4.2), but subsidiary to the three-way interaction discussed above.

V–A incongruent conditions (illusory trials). As anticipated above, the presentation of V–A incongruent trials could give rise to two different types of illusions: fissions (more perceived flashes than actually presented) or fusions (less perceived flashes than actually presented). Fissions could emerge from three incongruent conditions: one flash with two beeps, one flash with three beeps, or two flashes with three beeps. Fusions could emerge from the remaining three incongruent conditions: two flashes with one beep, three flashes with one beep, or three flashes with two beeps. We measured illusory errors as the difference between the reported number of flashes in each of the fission or fusion conditions and the number of reported trials in the V–A congruent condition. For example, illusory errors for the fission condition with one flash and two beeps were measured with respect to the V–A congruent condition with one flash and one beep. The rationale for choosing V–A congruent conditions as the baseline was twofold: First, they consist of bimodal stimulation similar to the illusory trials; second, and most important, they represented a performance baseline that was comparable across groups. As before, positive and negative values indicate overestimation and underestimation of the number of presented flashes, respectively.

Table 2

Average Error (Calculated as the Difference Between the Reported Number of Flashes and the Actual Number of Presented Flashes) for the Different Experimental Conditions, Separately for Group

	Experiment 2			
	6–7 years old	9–10 years old	11–12 years old	Adults
V-only				
1V	0.27 (.07)	0.28 (.07)	0.14 (.07)	0.04 (.03)
2V	−0.30 (.10)	0.03 (.09)	0.03 (.10)	0.02 (.11)
3V	−1.09 (.11)	−0.63 (.10)	−0.48 (.11)	−0.25 (.13)
V–A congruent				
IV–1A	0.15 (.06)	0.21 (.05)	0.12 (.06)	0.04 (.07)
2V–2A	0.18 (.05)	0.15 (.05)	0.05 (.05)	0.17 (.06)
3V–3A	−0.22 (.06)	−0.30 (.06)	−0.29 (.06)	−0.14 (.07)
V–A incongruent (fissions)				
IV–2A	1.09 (.10)	0.88 (.09)	0.72 (.10)	0.36 (.11)
1V–3A	1.70 (.12)	1.18 (.11)	0.82 (.12)	0.28 (.14)
2V–3A	0.71 (.09)	0.61 (.08)	0.56 (.09)	0.53 (.10)
V–A incongruent (fusions)				
1V–2A	0.85 (.11)	−0.41 (.10)	−0.41 (.11)	0.18 (.12)
1V–3A	−1.75 (.14)	−0.97 (.13)	−0.96 (.14)	−0.50 (.16)
2V–3A	−0.83 (.07)	−0.66 (.07)	−0.83 (.07)	−0.52 (.08)

Note. Standard errors are shown in parentheses. Positive (+) and negative (−) values indicate overestimation or underestimation of the number of presented flashes, respectively. Standard errors are shown in parentheses.

Although statistically nonsignificant, it should be noted that adults performed numerically better (see Table 2) compared to children, although not in all conditions. The performance variability found between groups as a function of condition suggests that the effects of redundant information did not always enhance performance. However, for the aim of the experiment, it was important to find a baseline in which children could be as accurate as possible.

Illusory errors for fission trials were entered into a mixed ANOVA with trial type (1V–2A, 1V–3A, or 2V–3A) as within-participant variable and group as between-participant variable. This analysis showed a significant value of the intercept, $F(1, 36) = 321.76$, $p < .001$, indicating an overall tendency to overestimate the number of flashes in these illusory trials ($M = .81$, $SE = .04$). Most importantly, there was a significant main effect of group, $F(3, 36) = 12.89$, $p = .001$, partial eta squared = .518, caused by the overall fission error decreasing with age

(6–7 years old: *M* = 1.17, *SE* = .09; 9–10 years old: *M* = .89, *SE* = .08; 11–12 years old: *M* = .70, *SE* = .09; adults: *M* = .39, *SE* = .10) following a linear trend. As shown in Figure 3a, this trend was further modulated by the specific audiovisual pairing, being strongest with one flash and three beeps (1V–3A) and weakest with two flashes and three beeps (2V–3A). This resulted in a significant interaction between trial type and group, $F(6, 72) = 13.15$, $p < .001$, partial eta squared = .523 (see Table 2).

Illusory errors for fusion trials were entered into a similar mixed ANOVA with trial type (2V–1A, 3V–1A, or 3V–2A) as the within-participant variable and group as the between-participant variable. This analysis showed a significant value of the intercept, $F(1, 36) = 254.55$, $p < .001$, indicating an overall tendency to overestimate the number of flashes in these illusory trials (*M* = −.63, *SE* = .06). This analysis also revealed a significant main effect of group, $F(3, 36) = 10.25$, $p < .001$, partial eta squared = .461. Again, the overall fusion error decreased with age (6–7 years old: *M* = −1.14, *SE* = .09; 9–10 years old: *M* = −.68, *SE* = .08; 11–12 years old: *M* = .73, *SE* = .09; adults: *M* = .40, *SE* = .10) following a linear trend. Similar to what was observed for fission trials, this trend was modulated by the specific audiovisual pairing, being strongest with three flashes and one beep (3V–1A) and weakest with two flashes and one beep (2V–1A; see Figure 3b). This resulted in a significant interaction between trial type and group, $F(6, 72) = 7.61$, $p < .001$, partial eta squared = .388 (see Table 2).

Discussion

The results of Experiment 2 showed that all participants misperceived the correct number of presented flashes when an incongruent number of auditory beeps were concurrently presented. We observed both illusory fissions (i.e., overestimation of the number of presented flashes) and illusory fusions (i.e., underestimation of the number of presented flashes). Most important, the sound-induced flash illusion was strongest in children aged 6–7 years (both in terms of fissions and fusions) and decreased progressively with age. This finding indicates higher relevance of the auditory input in younger children and it is in agreement with the switch from auditory to visual dominance in children of this age range we reported in Experiment 1.

When asked to identify the number of unimodally presented flashes, children aged 6–7 years old and, to some extent, children aged 9–10 years old performed worse than the other groups. This finding is not particularly surprising, given our choice to adopt exactly the same flash rate that had been adopted in the original paradigm with adults (Shams et al., 2002; see also Tremblay et al., 2007), and indicates that children aged 6–7 or 9–10 years old may be overall less accurate when enumerating a sequence of unimodally presented flashes. However, it should be emphasized that performance between groups was comparable when congruent A–V trials were considered (i.e., all participants were equally good at discriminating the number of presented flashes when they were accompanied by an equal number of beeps). As we chose performance on these A–V congruent trials as baseline, the difference between groups in the amount of sound-induced flash illusion was not confounded by the different enumeration or discrimination abilities that emerged for unimodally presented flashes. In sum, the different magnitude of the illusion can be considered as the result of developmental changes for multisensory dominance.

A sound-induced flash illusion similar to the one we have adopted has been previously tested in children by Tremblay et al. (2007). In their study, children and adolescents were grouped in three age ranges (5–9, 10–14, and 15–19 years), which were considerably wider than the ones adopted herein. In addition, although audiovisual stimulation conditions comprised all combination of one or two flashes with one or two beeps, the results were not presented in a disaggregated fashion for unimodal, A–V congruent and A–V incongruent conditions (unlike here). The results of Tremblay et al. replicated the original illusion effect (Shams, Kamitani, & Shimojo, 2000), but failed to show any main effect or interaction involving the group factor. The results of our Experiment 2 considerably extend this preliminary observation by showing comparable performance between groups when A–V congruent pairings were used and a progressively reduced magnitude of the sound-induced flash illusion, both in terms of fissions and fusions, as a function of age when A–V incongruent pairings were considered. Our results also suggest that the reason why no group effect emerged in the study by Tremblay et al. could reflect the wide age range adopted for each group in that study.

General Discussion

This study investigated sensory dominance and multisensory interactions in school-age children to

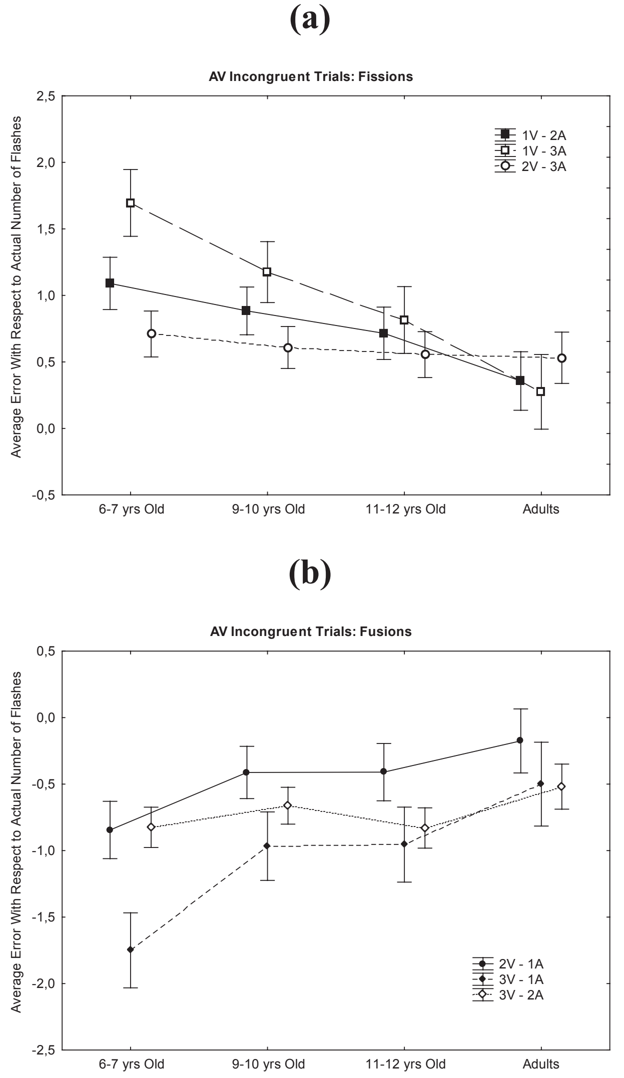

Figure 3. Average error with respect to the actual number of presented flashes in (a) fission and (b) fusion AV incongruent trials, as a function of group and stimulation condition.
Note. Positive values on *y*-axis show overestimation (a) and underestimation (b) of the number of presented flashes, respectively. Error bars show the standard error of the mean.

examine whether auditory dominance extends beyond the age of 4, and to establish when the transition toward visual dominance typically observed in adults starts to emerge. In Experiment 1, we adopted the Colavita effect (Colavita, 1974), a simple paradigm that systematically reveals visual dominance in adults. In Experiment 2, we used the sound-induced flash illusion (Shams et al., 2000, 2002), a paradigm that assesses the degree of auditory dominance by measuring to what extent enumeration of visual targets is affected by task-irrelevant auditory stimuli.

Overall, we found that auditory dominance persists until 6 years of age and that the transition toward visual dominance starts at school age. In particular, Experiment 1 showed that children aged 6–7 do not exhibit a Colavita effect (visual dominance), unlike 9- to 10-year-old and 11- to 12-year-old children, suggesting that sensory dominance undergoes a developmental change in late childhood. Interestingly, although no clear sensory dominance emerged in 6- to 7-year-old children in Experiment 1a, a clear reversal of the Colavita effect (i.e., auditory dominance) occurred in Experiment 1b. This reversal of the effect is most striking when considered with respect to the context of the adult literature, in which no reversal of the Colavita effect was ever documented (Spence, 2009). Experiment 2 (sound-induced flash illusion) validated and extended the main results of the Colavita task. We found that children aged 6–7 years proved more sensitive to the task-irrelevant auditory component during audiovisual stimulation, compared to older children and to the group of adults. This experiment supported the overall hypothesis that multisensory interactions undergo developmental changes. As we observed auditory dominance in children up to 6–7 years of age and a clear visual dominance in children aged 9–10 and 11–12 years of age in Experiment 1, it is possible that such developmental shifts in sensory dominance may underpin and modulate multisensory interactions throughout childhood.

Processing of multisensory input is typically described as advantageous, as it can enhance the ability to detect and categorize objects and events in the environment. Recent studies, however, have suggested that optimal multisensory integration occurs only relatively late in childhood (i.e., around 8–9 years of age; Ernst, 2008; Gori et al., 2008; Nardini et al., 2008). Children under 8 years of age can show behavioral responses that are dominated by individual sensory modalities, rather than the optimally weighted combination of the sensory inputs. Ernst (2008) suggested that such late development of multisensory integration could depend on the continuous calibration of the sensory systems required during development. As the sense organs grow at different time rates, children have to constantly update and recalibrate the sensory signals. For this to occur, the comparison of *individual* multisensory inputs may be more functional than their optimal integration, resulting in the lack of integration observed in childhood. Within this perspective, our effort to characterize the developmental trajectory of sensory dominance has relevant theoretical implications because changes in sensory dominance through childhood are likely to have an impact on the way sensory inputs are compared with one another and recalibrated. Furthermore, our findings clearly show that the hierarchical development of the sensory systems proposed by Lewkowicz (1988a, 1988b) to account for sensory dominance in the very first months of life extends well into childhood.

From a clinical point of view, tracing how multisensory interactions typically develop may contribute to the understanding of those cases of atypical development in which multisensory integration failures are observed. For example, autism is characterized by altered sensory processing, and recent studies (Foss-Feig et al., 2010) have shown that children with this syndrome exhibit impaired multisensory integration abilities. If multisensory integration appears early in life, then identifying behavioral patterns that differ from the typical trajectory could contribute to the development of strategies of early diagnostic and intervention. In addition, the importance of hierarchical sensory development in determining mature multisensory processing is highly relevant in those populations who initially lack one sensory modality, but start to receive inputs from the missing sense later in life. This condition is now increasingly common in people with congenital or early profound bilateral deafness, who later start to receive auditory inputs through a bionic ear (i.e., cochlear implant), or in blind individuals with inborn cataract(s), who can later reacquire visual input by having their cataracts surgically removed. For all these cases of sensory reafferentation, the typical developmental trajectory for sensory dominance and multisensory interactions could shed light on what might be expected after the missing sense has been restored.

In conclusion, our novel findings have shown that children up to 6–7 years of age show auditory dominance in different multisensory contexts in which concurrent audiovisual events are presented. This pattern of sensory dominance decreases as a function of age, suggesting a gradual change in multisensory perception during development and the consolidation of adult-like processing of multisensory inputs starting from late childhood. Further research should address the reasons why such developmental changes occur. In particular, it could be questioned whether these changes are genetically defined or are triggered by particular sensory experience. For instance, a slower maturation of the visual system could be functional for the complete achievement of linguistic skills, for which the auditory system plays a crucial role (Benasich, Thomas, Choudhury, & Leppänen, 2002). Within this evolutionary perspective, the persistence of auditory dominance up to 6- to 7-year-old children could favor language maturation. On the other hand, visual dominance could be fostered by cultural demands, in that environments of most Western societies are dominated by visual material (i.e., television, internet, etc.). Observing whether the pattern of sensory dominance we observed is also present cross-culturally could provide more evidence in support to these hypotheses.

References

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*, 257–262. doi:10.1016/j.cub.2004.01.029

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, *21*, 301–308. doi:10.1016/j.cogbrainres.2004.06.004

Benasich, A. A., Thomas, J. J., Choudhury, N., & Leppänen, P. H. T. (2002). The importance of rapid auditory processing abilities to early language development: Evidence from converging methodologies. *Developmental Psychobiology*, *40*, 278–292. doi:10.1002/dev.10032

Botvinick, M., & Cohen, J. (1998). Rubber hands "feel" touch that eyes see. *Nature*, *391*, 756. doi:10.1038/35784

Calvert, G. A., Spence, C., & Stein, B. E. (Eds.). (2004). *The handbook of multisensory processes*. Cambridge, MA: MIT Press.

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

Ernst, M. O. (2008). Multisensory integration: A late bloomer. *Current Biology*, *18*, R518–R519. doi:10.1016/j.cub.2008.05.002

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433. doi:10.1038/415429a

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., et al. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, *203*, 381–389. doi:10.1007/s00221-010-2240-4

Gori, M., Dal Viva, M., Sandini, G., & Burr, D. (2008). Young children do not integrate visual and haptic form information. *Current Biology*, *18*, 694–698. doi:10.1016/j.cub.2008.04.036

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. London: Wiley.

Koppen, C., Alsius, A., & Spence, C. (2008). Semantic congruency and the Colavita visual dominance effect. *Experimental Brain Research*, *184*, 533–546. doi:10.1007/s00221-007-1120-z

Koppen, C., & Spence, C. (2007a). Seeing the light: Exploring the Colavita visual dominance effect. *Experimental Brain Research*, *180*, 737–754. doi:10.1007/s00221-007-0894-3

Koppen, C., & Spence, C. (2007b). Spatial coincidence modulates the Colavita visual dominance effect. *Neuroscience Letters*, *417*, 107–111. doi:10.1016/j.neulet.2006.10.069

Lecanuet, J.-P., & Schaal, B. (1996). Fetal sensory competencies. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *68*, 1–23. doi:10.1016/0301-2115(96)02509-2

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

Lewkowicz, D. J. (1988b). Sensory dominance in infants: Ten-month-old infants' response to auditory-visual compounds. *Developmental Psychology*, *24*, 172–182. doi:10.1037/0012-1649.24.2.172

Lewkowicz, D., & Lickliter, R. (Eds.). (1994). *The development of intersensory perception*: *Comparative perspectives*. Hillsdale, NJ: Erlbaum.

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends in Neurosciences*, *28*, 264–271. doi:10.1016/j.tins.2005.03.008

Massaro, D. W. (1984). Children's perception of visual and auditory speech. *Child Development*, *55*, 1777–1788. doi:10.2307/1129925

Massaro, D. W., Thompson, L. A., Barron, B., & Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology*, *41*, 93–113. doi:10.1016/0022-0965(86)90053-6

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748. doi:10.1038/264746a0

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development*, *75*, 1850–1870. doi:10.1111/j.1467-8624.2004.00821.x

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology*, *18*, 689–693. doi:10.1016/j.cub.2008.04.021

Pavani, F., Spence, C., & Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological Science*, *11*, 353–359. doi:10.1111/1467-9280.00270

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, *83*, 157–171. doi:10.1037/0033-295X.83.2.157

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

Robinson, C. W., & Sloutsky, V. M. (2007). Proceedings of the 29th annual Cognitive Science Society. In D. S. McNamara & J. G. Trafton (Eds.), *Crossmodal interference: Overshadowing or response competition?* (pp. 605–610). Austin, TX: Cognitive Science Society.

Rosenthal, O., Shimojo, S., & Shams, L. (2009). Sound-induced flash illusion is resistant to feedback training. *Brain Topography*, *21*, 185–192. doi:10.1007/s10548-009-0090-9

Schorr, E. A., Fox, N. A., Van Wassenhove, V., & Knudsen, E. (2005). Auditory-visual fusion in speech perception. *Proceedings of the National Academy of Sciences*, *102*, 18748–18750. doi:10.1073/pnas.0508862102

Setti, A., & Chan, J. S. (2011). Familiarity of objects affects susceptibility to the sound-induced flash illusion. *Neuroscience Letters*, *492*, 19–22. doi:10.1016/j.neulet.2011.01.042

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*, 788. doi:10.1038/35048669

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*, 147–152. doi:10.1016/S0926-6410(02)00069-1

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: The Colavita effect revisited. *Perception & Psychophysics*, *69*, 673–686. doi:10.3758/BF03193770

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

Spence, C. (2009). Explaining the Colavita visual dominance effect. *Progress in Brain Research*, *176*, 245–258. doi:10.1016/S0079-6123(09)17615-X

Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention*. Oxford, England: Oxford University Press.

Spence, C., Pavani, F., Maravita, A., & Holmes, N.. (2008). Haptic rendering: Foundations, algorithms, and applications. In M. Lin & M. Otaduy (Eds.), *Multisensory interactions* (pp. 21–50). Wellesley, MA: AK Peters.

Stein, B., & Meredith, A. (1993). *The merging of the senses*, Cambridge, MA: MIT Press.

Tremblay, C., Champoux, F., Voss, P., Bacon, B. A., Lepore, F., & Théoret, H. (2007). Speech and non-speech audio-visual illusions: A developmental study. *PLoS ONE*, *2*, e742. doi:10.1371/journal.pone.0000742

Wallace, M. T., Carriere, B. N., Perrault, T. J., Vaughan, W., & Stein, B. E. (2006). The development of cortical multisensory integration. *Journal of Neuroscience*, *26*, 11844–11849. doi:10.1523/JNEUROSCI.3295-06.2006

Welch, R. B., & Warren, D. H.. (1986). Intersensory interactions. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 1–36). New York: Wiley.