The role of visual experience in the emergence of cross-modal correspondences

Article (Accepted Version)

Hamilton-Fletcher, Giles, Pisanski, Katarzyna, Reby, David, Stefańczyk, Michal, Ward, Jamie and Sorokowska, Agnieszka (2018) The role of visual experience in the emergence of cross-modal correspondences. Cognition, 175. pp. 114-121. ISSN 0010-0277

The Role of Visual Experience in the Emergence of Cross-Modal Correspondences

Giles Hamilton-Fletcher[1,2,*], Katarzyna Pisanski[1,3], David Reby[1], Michał Stefańczyk[3], Jamie Ward[1,2], and Agnieszka Sorokowska[3,4]

[1]School of Psychology, University of Sussex, Brighton, UK
[2]Sackler Centre for Consciousness Science, University of Sussex, Brighton, UK
[3]Institute of Psychology, University of Wrocław, Wrocław, Poland.
[4]Smell & Taste Clinic, Department of Otorhinolaryngology, TU Dresden, Dresden, Germany
*e-mail: gh71@sussex.ac.uk

Cross-modal correspondences describe the widespread tendency for attributes in one sensory modality to be consistently matched to those in another modality. For example, high pitched sounds tend to be matched to spiky shapes, small sizes, and high elevations. However, the extent to which these correspondences depend on sensory experience (e.g. regularities in the perceived environment) remains controversial. Two recent studies involving blind participants have argued that visual experience is necessary for the emergence of correspondences, wherein such correspondences were present (although attenuated) in late blind individuals but absent in the early blind. Here, using a similar approach and a large sample of early and late blind participants (N=59) and sighted controls (N=63), we challenge this view. Examining five auditory-tactile correspondences, we show that only one requires visual experience to emerge (pitch-shape), two are independent of visual experience (pitch-size, pitch-weight), and two appear to emerge in response to blindness (pitch-texture, pitch-softness). These effects tended to be more pronounced in the early blind than late blind group, and the duration of vision loss among the late blind did not mediate the strength of these correspondences. Our results suggest that altered sensory input can affect cross-modal correspondences in a more complex manner than previously thought and cannot solely be explained by a reduction in visually-mediated environmental correlations. We propose roles of visual calibration, neuroplasticity and structurally-innate associations in accounting for our findings.

Keywords: Blindness; Cross-modal correspondences; Pitch; Sound; Touch.

1. Introduction

Our senses provide us with a broad array of signals from the environment. We preferentially bind these signals together into coherent objects or events (Treisman, 1998; Roskies, 1999). These multisensory signals are more likely to be bound together if they are congruent in terms of temporal, spatial, or semantic factors (Spence, 2011). However, another basis for integrating across modalities is through matching specific stimulus features (e.g. pitch, colour). This preferential matching reflects a variety of processing biases, some of which manifest from unconscious intuitions that certain sensory properties relate to or map onto those of other senses. Such unconscious biases can be easily demonstrated by asking seemingly nonsensical questions - such as whether listening to a high-pitched tone is perceptually closer to white or black? The answers to these questions reveal widespread matching preferences between seemingly separate sensory features, collectively known as 'cross-modal correspondences' (Spence, 2011).

Cross-modal correspondences affect attention, perceptual processing, multisensory integration, and aesthetics (Chiou & Rich, 2012; Spence, 2011; Albertazzi, Malfatti, Canal & Micciolo, 2015), and have even been speculated to play a role in language evolution (Ramachandran & Hubbard, 2001). Furthermore, cross-modal correspondences are also present in other species (Ratcliffe, Taylor & Reby, 2016), from chimpanzees mapping high pitch to high luminance (Ludwig, Adachi & Matsuzawa, 2011) to birds and mammals mapping low pitch to large body size (Morton, 1977).

The origins of cross-modal correspondences are controversial (Spence & Deroy, 2012), and likely vary in accordance with the specific association in question. The currently prevailing explanations relate to neurological, statistical or mediating factors (Spence, 2011). Neurological accounts suggest that common cortical representations of separate sensory features relate to their perceived similarity; for example, increased loudness is associated with increased brightness (Bond & Stevens, 1969), where it is argued that experiencing either loud or bright stimuli increases the neural activity in the primary auditory and visual cortices respectively (Goodyear & Menon, 1998; Jäncke, Shah, Posse, Grosse-Ryuken & Müller-Gärtner, 1998; Mulert et al., 2005), and as a result these sensory features become associated through an 'intensity-matching' process. A related neurological account suggests that while individual sensory features may be qualitatively distinct, they can be abstracted to a common mechanism for gauging their magnitude or polarity in the parietal cortex, such that two sensory features perceived as 'high' on a given scale become associated (Walsh, 2003). Statistical accounts suggest that regularly co-occurring pairs of stimuli become associated with one another and internalised. For example, positive relationships between pitch and elevation are a regular feature of natural soundscapes, as well as a by-product of our ears' ability to filter frequencies based on elevation (Parise, Knorre & Ernst, 2014). Finally, cross-modal correspondences can further become reinforced through mediating factors such as culture (e.g. pitch-height in musical notation), language (e.g. 'high' used to describe both sound and elevation), or emotional connotation (certain sounds and colours can be matched on emotional valence – see Palmer, Schloss, Xu & Prado-León, 2013).

The role of sensory experience in cross-modal correspondences has most often been studied from a developmental perspective. The argument is that if a young infant demonstrates a given association, then the association is likely to be non-learned (i.e. innate) (Mondloch & Maurer, 2004; Nava, Grassi & Turati, 2016; Walker et al., 2010; 2018 - although see Spence & Deroy, 2012). A complementary approach is to study adults with sensory deficits such as deafness or blindness. This allows researchers to examine whether the presence of appropriate sensory experiences, at different stages of development and for different amounts of time, is necessary for the emergence of cross-modal correspondences.

Blindness is often associated with a substantial change in both unisensory and multisensory processing. For instance, improved horizontal monoaural localisation abilities in the early blind are also correlated with activation of the visual cortex, indicating that the additional neural resources available can alter unisensory processing (Gougoux, Zatorre, Lassonde, Voss & Lepore, 2005). However, the ability to accurately localise sound sources in the vertical spatial plane is commonly impaired in early blind individuals. This finding has been ascribed to the lack of visual calibration for correlating the observed spatial location of sound sources with subtle changes in frequency-filtering by the pinna (Lewald, 2002; Zwiers, Van Opstal & Cruysberg, 2001). The lack of a calibrating visual reference frame in congenitally blind persons can also influence multisensory spatial integration between hearing and touch (Hötting, Rösler & Röder, 2004). For example, the detection of touch among congenitally blind persons exhibits less interference from task-irrelevant auditory cues than does touch detection in sighted persons, suggesting a reduction in audio-tactile integration in the congenitally blind (Hötting & Röder, 2004). This illustrates how visual processes can underlie seemingly non-visual audio-tactile interactions - either through access to 'visual regions' of the brain, or by vision calibrating the other senses.

Three studies recently examined how cross-modal correspondences are affected by blindness. Eitan, Ornoy, and Granot (2012) found that sighted participants associated increasing tonal pitch with increasing verticality in the spatial plane, whereas early-blind participants associated it with increasing proximity. The mechanisms driving this group difference are unclear, however it may be related to the increased importance of egocentric co-ordinates in spatial processing for the congenitally blind (Iachini, Ruggiero & Ruotolo, 2014; Pasqualotto, Spiller, Jansari & Proulx, 2013), and thus may represent a variation of the pitch-height correspondence, rather than a qualitatively new correspondence per se. The relationship between pitch-height and sightedness was further explored by Deroy, Fasiello, Hayward and Auvray (2016) who found that congruent and incongruent correspondences between tonal pitch and tactile-spatial elevation affected information processing on an implicit association task among sighted, but not among early/late blind participants. These findings are particularly interesting because pitch-height relationships are observed in early infancy both for visual and tactile height (Nava, Grassi & Turati, 2016; Walker et al., 2010; 2018), suggesting that pitch-height correspondences need to be maintained by visual experience in order to manifest in adult auditory-tactile interactions (Occelli, Spence & Zampini, 2009). Finally, examining audio-tactile correspondences for object features, Fryer, Freeman and Pring (2014) presented blind and sighted participants with shapes they could feel but not see. The researchers found that the tendency for (blindfolded) sighted people to match 'bouba' and 'kiki' sounds to specific haptic-shapes (round and angular respectively) was reduced in a late blind group, and absent in the early blind. This finding suggests that the formation of sound-shape correspondences may require visual experience and appears to support the statistical account for associations between hearing and touch, such that individuals with sensory impairments experience fewer statistical correlations in the environment, ultimately reducing their association with one another. However, this pattern may not reflect all correspondences, since the range of associations between hearing and touch tested to date is limited and, in some cases, the strength of these cross-modal correspondences can be quite low. Furthermore, many studies examining the influence of blindness on correspondences have had to contend with low statistical power resulting from relatively low numbers of blind participants (e.g. Deroy et al., 2016).

In the present paper, we compare sound-touch correspondences among one-hundred and twenty-two sighted (blindfolded), late-blind, and early-blind adults. Our study includes a wider range of audio-tactile correspondences than previously examined, allowing us to test whether the influence of visual experience on the strength and direction of sound-touch associations depends on the given tactile dimension. In our experiment, on each trial, participants were presented with either a low-pitched (200 Hz) or high-pitched tone (2000 Hz) and were tasked with choosing which tactile object from a pair 'best' matched the tone. Across 10 sets of tactile object pairs, we examined 5 tactile dimensions: shape, texture, softness, size, and weight. Each pair differed primarily on a single tactile dimension (e.g., 'size,' small vs large) and did not vary on the other experienced dimensions (e.g., the 'size' stimuli did not differ in weight, texture, softness, or shape). All combinations of pitch and object-pairs were presented to participants in a random order over 20 minutes. The late-blind (n=27) and early-blind (n=32) groups of participants consisted of fully-blind adults who were also homogenous in their current visual abilities with no light perception (further participant details can be found in table 1).

2. Method

2.1 Participants.

Sixty-three sighted individuals were recruited (39 female, mean age 33.9y ± 12), along with twenty-seven late blind individuals (17 female, mean age 48.3y ± 11.4), and thirty-two early blind individuals (16 female, mean age 34.38 ± 9.83). The late blind individuals had an average vision loss of 18.5 years (± 12.7), or 39% (± 26.3) of their life as fully blind. The inclusion criteria for the 'early blind' group was the presence of 'significant eyesight loss before reaching 2 years of age,' and hence before the completion of the critical period of visual development (following Wiesel, 1982 – see Pisanski, Oleszkiewicz, & Sorokowska, 2016; Sorokowska, 2016). The early and late blind groups were blind due to a wide range of aetiologies (see table 1). Mean number of years of education for participants was 16.11 ± 3.15, 13.78 ± 3.13, and 16.17 ± 3.70 for the sighted, late blind and early blind groups respectively. All late blind and early blind individuals reported no light or residual perception. Sixteen participants reported some level of hearing impairment (7 sighted, 6 late blind, 3 early blind), the extent of impairment was self-rated at an average of 4.5 ± 2.1 out of ten, and was not found to significantly affect responses (see supplemental analyses). All participants could discriminate between low and high tones and so were not omitted as outliers. Participants were recruited from associations for blind people in Poland. Participants were reimbursed with 30 US dollars for their time. The study was performed in accordance with the Declaration of Helsinki on Biomedical Studies Involving Human Subjects and was approved by the Ethical Committee of the Institute of Psychology, University of Wrocław (project no. 2013/11/B/HS6/01522).

Table 1. Demographics and aetiologies of the early and late-blind groups.

Early Blind Group (*n*=32)				**Late Blind Group (*n*=27)**			
Age	Gender	Blind (yrs)	Aetiology	Age	Gender	Blind (yrs)	Aetiology
17	Male	17	Retinopathy of prematurity (ROP)	23	Female	6	Stargart disease
18	Female	18	Retinopathy of prematurity (ROP)	26	Female	22	Retinoblastoma
19	Male	19	Retinopathy of prematurity (ROP)	29	Male	1.5	Chemical poisoning
23	Male	23	Retinoblastoma	32	Female	8	Retinal detachment
23	Female	23	Retinopathy of prematurity (ROP)	36	Female	12	Accident resulting in retinal detachment
25	Male	25	Unknown	40	Female	24	Glaucoma and retinal detachment
27	Male	27	Retinopathy of prematurity (ROP)	41	Male	30	Retinal detachment
28	Female	28	Leber congenital amaurosis	41	Female	38	Optic nerve atrophy
28	Female	28	Congenital retinal detachment	45	Female	8	Glaucoma
30	Male	30	Retinoblastoma	47	Male	17	Diabetes
31	Female	31	Unknown	47	Female	11	Accident resulting in retinal detachment
31	Male	31	Retinopathy of prematurity (ROP)	48	Female	20	Retinal detachment
32	Female	32	Congenital optic nerve atrophy	50	Male	2	Retinal detachment and glaucoma
32	Male	32	Retinopathy of prematurity (ROP)	52	Male	20	Retinitis pigmentosa
33	Male	33	Maternal toxoplasmosis	52	Female	28	Diabetes resulting in polyneuropathy
34	Male	34	Retinopathy of prematurity (ROP)	53	Female	5	Retinitis pigmentosa
34	Male	34	Congenital cataracts	54	Male	50	Meningitis
34	Female	34	Retinopathy of prematurity (ROP)	56	Male	24	Accident resulting in retinal detachment
34	Female	34	Congenital optic nerve atrophy	56	Female	28	Diabetes
37	Female	37	Leber congenital amaurosis	56	Male	3.5	Diabetes
37	Male	37	Retinopathy of prematurity (ROP)	56	Female	10	Stargardt disease
38	Male	38	Retinopathy of prematurity (ROP)	57	Female	15.5	Chronic uveitis and optic nerve atrophy
38	Male	38	Retinitis pigmentosa and optic nerve atrophy	57	Male	5	Glaucoma and retinal detachment
40	Female	40	Optic nerve damage	60	Female	40	Chronic uveitis
41	Female	41	Unknown (probably maternal medication)	61	Male	34	Chemical burn resulting in anophthalmia
42	Female	42	Retinitis pigmentosa and optic nerve atrophy	64	Female	22	Glaucoma
44	Female	44	Microphthalmia	64	Female	16	Retinal detachment
44	Male	44	Retinoblastoma				
44	Female	44	Retinopathy of prematurity (ROP)				
50	Male	50	Glaucoma				
53	Female	53	Optic nerve atrophy				
59	Female	59	Congenital optic nerve atrophy				

Note: 'Blind (yrs)' refers to the total number of years the participant has experienced blindness at the time of testing.

2.2 Apparatus.

Auditory stimuli consisted of a low-pitched (200 Hz) and high-pitched (2000 Hz) pure tone, both 1 second in duration and loudness-equalised to 40 phons by altering their amplitude in line with loudness-equalisation curves (ISO, 2003). The sounds were saved as wav files at a sampling rate of 44100 Hz.

The tactile stimuli consisted of twenty unique objects. These were divided into ten pairs. Objects within each pair varied on only one of the five tactile dimensions of interest, namely either in their shape, texture, softness, size or weight. For each tactile dimension, there were two pairs of stimuli, within which object A and object B varied on that dimension alone (see figure 2), while the other dimensions remained either constant, or were inaccessible to participants (for comparisons of each object's shape, texture, softness, size and weight please consult supplemental table 1).

The shape stimuli consisted of sets of objects that were either rounded or angular. Set 1 contained two commercially available 3D shapes, a sphere and a pyramid obtained from a local art store, while set 2 contained a disc and triangle cut from a 0.2 cm thick sheet of balsa wood. The texture stimuli included sets of objects that were either smooth or rough. Set 1 contained two wooden squares and set 2 contained two wooden spheres. In both sets, the smooth stimulus had a thin wood layer applied to the surface, whereas the rough stimulus had coarse large sand grains applied to the surface creating a rough texture. The softness stimuli consisted of objects that varied in their compliance to pressure. Set 1 consisted of two spheres, the softer stimulus was made of foam while the harder stimulus was made of hard plastic, both stimuli were given a felt texture as this would not affect the compliance of the softer stimulus. Set 2 consisted of two blocks previously used in Simner and Ludwig (2012), the soft block was given a rating of 40 Newtons (range 30-50), while the hard block was given a rating of 270 Newtons (range 240-300), both had a cloth texture that would also comply with the softer stimulus. The size stimuli consisted of objects that varied in their physical dimensions but not their overall weight. To achieve this, set 1 consisted of wooden spheres of different sizes (diameters of 5 cm and 7.5 cm), the smaller sphere was hollowed out and additional metal weights and plaster filling were placed in the centre of the sphere until the final weight matched the larger sphere. The second set of size stimuli consisted of a coin and wooden disc (diameters 2.75 cm and 8 cm respectively) covered in felt. The weight stimuli consisted of objects only varying in their physical weight. Set 1 contained a filled in rubber sphere and a hollow hard plastic sphere, both covered in felt. Set 2 contained two smooth wooden spheres, one of which was hollowed out and filled with metal and plaster to increase its weight relative to the lighter wooden sphere.

A commercially available blindfold known as the 'Mindfold®' mask (Mindfold Inc., Colorado, USA) was used to block visual stimulation for sighted participants during the task; this allows for the eyes to be open or shut without physical impedance while eliminating all light. The sound stimuli (low and high tones) were played through Sennheiser HD 201 professional headphones using a custom computer interface.

2.3 Procedure.

Participants completed the experiment in individual sessions. They were seated within the experimental room and were read the following instructions: "You will be presented with a series of bags. Each bag will contain two objects; these objects differ substantially in one way. Concentrate on the difference between these objects. After you have both objects in your hands, you will be played one of two tones. It is your task to indicate which of the two objects 'goes best' with the tone that you hear. You will do this for a number of different objects. There are no right or wrong answers, and if you are not sure, choose one anyway."

The participants were then issued headphones and had the high and low tones played to them in a random order as a demonstration, and told, "these are the two tones that you will hear on a given trial." Sighted individuals were given Mindfold® blindfolds and told to keep them on for the entire duration of the task in order to help them concentrate on the tactile feel of the objects. On each trial, participants were presented with a bag containing 1 of the 10 sets of stimuli to tactually explore. When participants were confident they had identified the main difference between the tactile objects in the bag, they were played either the low or high tone, and asked to present the object that 'goes best' with that sound to the experimenter who recorded the response.

Each participant completed a total of 20 trials, consisting of each of the 10 sets of stimuli presented twice, once with the low-pitched tone and once with the high-pitched tone, in a fully randomised order. This task was given within a larger testing battery examining olfactory cues and memory. While the individual task took on average 20 minutes to complete, the larger battery occurred over 1 hour and 30 minutes. After the battery was completed participants were debriefed as to the nature of the study.

3. Results

To first evaluate overall trends in cross-modal correspondences across groups, a single aggregate score was created for each participant on each tactile dimension. This score was created from the participants' tone-object selections on a 0-4 scale where 4 indicates complete alignment with prior expectations, 2 indicates no correspondence, and 0 indicates a reversal in correspondence. The predicted cross-modal correspondences based on prior research with sighted individuals were that high pitch would be associated with: angular shapes, smoother textures, softer compliance, smaller size and lighter weights (Eitan & Timmers, 2010; Evans & Treisman, 2009; Gallace & Spence, 2006; Marks, 1987; Walker & Smith, 1985). The data were first analysed in SPSS v.23 using a two-way mixed ANOVA with sightedness (sighted, late blind, early blind) as the between-participant factor and tactile dimension (shape, texture, softness, size, weight) as the within-participant factor. Mauchly's test indicated that the assumption of sphericity had been violated χ^2 (9) = 78.63, *P*<.001, with ε <.75, thus Greenhouse-Geisser corrections were used. Levene's test indicated that the assumption of homogeneity of variance had been met for each of the tactile dimensions.

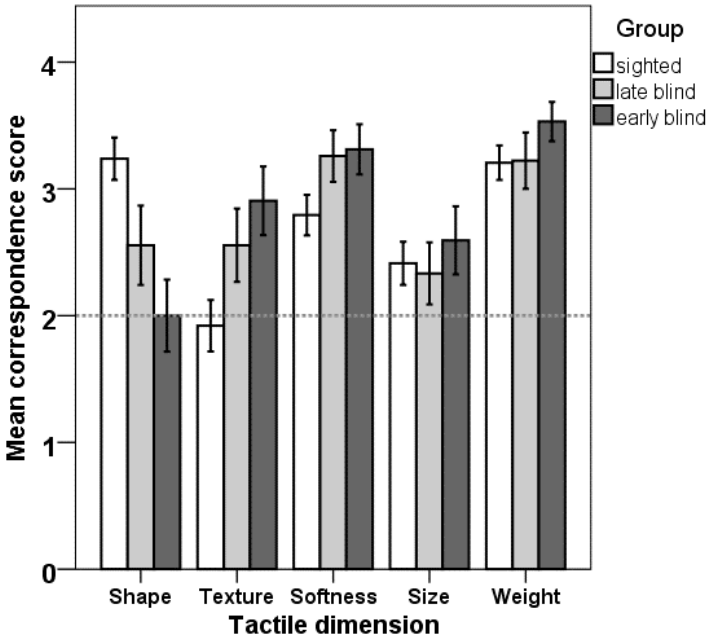

Figure 1 | Agreement in auditory-tactile correspondences among sighted, late and early blind adults. Increasing values indicate cross-modal correspondences in the expected direction (i.e., high pitch matched with angular shapes, smoother textures, softer compliance, smaller size and lighter weights). Y-axis indicates correspondences in line with prior evidence, a score of 0 indicates a complete reversal of correspondence, a score of 2 indicates no correspondence (dotted line), and a score of 4 indicates perfect agreement with prior expectations. Error bars indicate 1 standard error of the mean.

There was no main effect of sightedness in this model, *F*(2, 119)=1.20*, P*=.306, η_p^2 = 0.02, indicating no systematic variation in the strength of cross-modal correspondences for sighted, late blind and early blind participants when collapsing across tactile dimensions. Thus, contrary to prior evidence, blindness was not associated with an overall reduction in cross-modal correspondences across audio-tactile stimuli. However, there was a main effect of tactile dimension, *F*(2.98, 354.47)=8.774, *P*<.001, η_p^2 = 0.06. Bonferroni-corrected posthoc tests showed that some cross-modal correspondences were stronger in their influence than others. Collapsing across sightedness groups, 'weight' was significantly stronger than 'shape' (M_{diff}=0.722, *P*=.002), 'texture' (M_{diff}=0.859, *P*<.001), and 'size' (M_{diff}=0.873, *P*<.001), and 'softness' was significantly stronger than 'texture' (M_{diff}=0.661, *P*=.003) and 'size' (M_{diff}=0.675, *P*=.012). As such, pitch-weight and pitch-softness correspondences appear to be among the strongest cross-modal correspondences, when ignoring the effect of sightedness. There was however an interaction between sightedness and tactile dimension, *F*(5.96, 354.47)=4.08, *P*=.001, η_p^2 = 0.06, indicating that different groups varied in how strongly they matched tones with specific tactile characteristics (see figure 1), this is expanded upon in the next section.

To explore relationships among tonal pitch, sightedness and tactile dimension more closely, the data were then analysed in SPSS v.23 using a series of Generalised Linear Mixed Models (GLMMs) across groups of sighted, late blind, and early blind participants. Pitch tone (low, high) and tactile dimension (shape, texture, softness, size, weight) were included in the GLMM as fixed factors. We used a binary logistic regression link on the dependent variable of object selection (A, B). This created a 0-1 proportion scale (where 0 indicates entirely object B selections, 0.5 indicates no correspondence, and 1 indicates entirely object A selections) which allowed us to examine specific pitch-object associations without enforcing assumptions from prior evidence. Participant identity was entered as a random variable, and age as a covariate.

The omnibus model revealed a significant main effect of tone, *F*(1, 2.4)=166.6, *P*<.001, showing that overall, participants did exhibit systematic correspondences between tones and tactile dimensions. There were significant interactions between sightedness and tone (*F*(2, 2.4)=27.6, *P*<.001) and between tone and tactile dimension (*F*(4, 2.4)=74.1, *P*<.001) on object selection, justifying the need for independent models for each tactile dimension. Listener age showed no effect (*F*(1, 2.4)=0.13, *P*=.78) and was therefore not included in subsequent models.

To explore the interaction term, five independent models were constructed to examine each correspondence. In all five models there was a significant main effect of pitch on participants' choice of object (pitch-shape *F*(1, 482)=41.8, *P*<.001; pitch-texture *F*(1, 482)=22.4, *P*<.001; pitch-softness *F*(1, 482)=112.0, *P*<.001; pitch-size *F*(1, 482)=20.9, *P*<.001; pitch-weight *F*(1, 482)=143.6, *P*<.001). That is, there was evidence for a cross-modal correspondence in each case. In all five models there was no main effect of sightedness (pitch-shape $F(2, 482)=0.6$, $P=.55$; pitch-texture $F(2, 482)=1.2$, $P=.31$; pitch-softness $F(2, 482)=0.37$, $P=.95$; pitch-size $F(2, 482)=0.05$, $P=.95$; pitch-weight $F(2, 482)=0.92$, $P=.40$). The absence of a main effect of sightedness merely means that the groups were not differently biased in their overall object choices (e.g. selecting object A more often). For three of the models the interaction term was significant (pitch-shape $F(2, 482)=11.2$, $P<.001$; pitch-texture $F(2, 482)=11.2$, $P<.001$; pitch-softness $F(2, 482)=5.0$, $P=.007$). As illustrated in figure 2, visual experience has a positive influence on the emergence of pitch-shape correspondences (i.e. they are stronger in the sighted) but a negative influence on pitch-texture and pitch-softness correspondences (i.e. they are stronger in the blind). The remaining two associations (pitch-size and pitch-weight) showed no interaction ($F(2, 482)=0.92$, $P=.56$ and $F(2, 482)=2.0$, $P=.14$ respectively). That is, correspondences linking low pitch to large size and heavy weight were equally apparent in all groups and, hence, appear to be independent of visual experience. The GLMMs are logistic models with a logit link function, and hence effect sizes are reported as odds ratios (as per Hosmer, Lemeshow & Sturdivant, 2013). Odds ratios were conducted for each sightedness-pitch-touch condition and can be found in supplemental table 2.

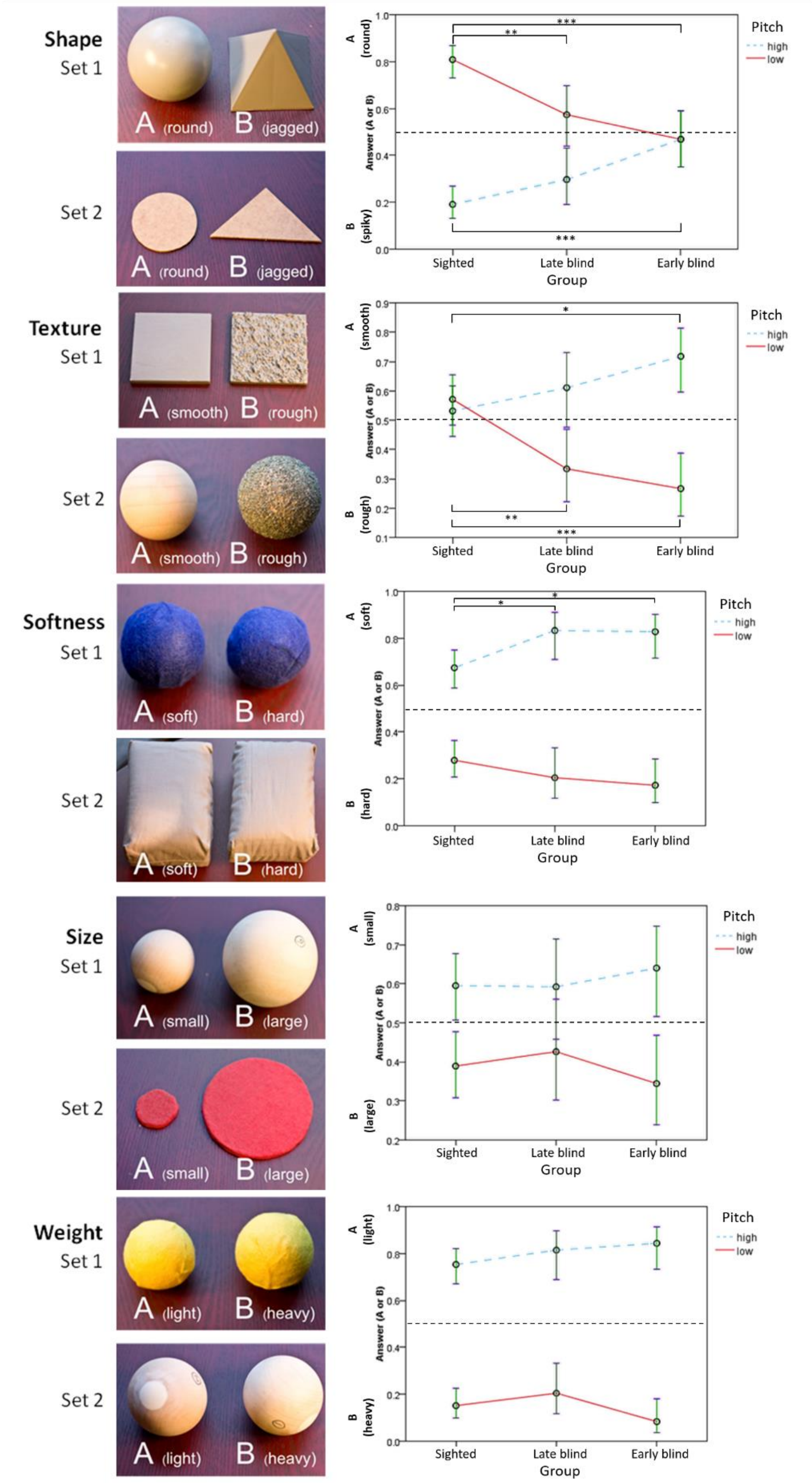

Figure 2 | Auditory-tactile correspondences. Preferential matching for object (A, B) across tone (low, high) and sightedness (sighted, late blind and early blind) for objects that only varied on the tactile dimensions of shape, texture, softness, size and weight (ordered top to bottom). Pairwise contrasts reported with Bonferroni-corrected significance levels. Error bars indicate 95% confidence intervals. Dotted midline indicates no cross-modal correspondence. Key: *=*p*<.05, **=*p*<.01, ***=*p*<.001.

4. Discussion

Previous studies examining the role of visual experience in influencing audio-tactile interactions have found that blindness is associated with a reduction in integration in terms of spatial processing (Hötting, Rösler & Röder, 2004; Hötting & Röder, 2004), pitch-height relationships (Deroy et al., 2016; Eitan, Ornoy & Granot, 2012), and sound-shape correspondences (Fryer, Freeman & Pring, 2014). Here we show that blindness does not result in a universal decrease in audio-tactile correspondences across all object features. Rather, although some correspondences are reduced or absent in blind adults (pitch-shape), other correspondences are maintained in the absence of visual experience (pitch-size, pitch-weight), and others appear to emerge with blindness (pitch-texture, pitch-softness). Below we consider each of these relationships in turn as well as their implications for currently prevailing theories regarding the experience of environmental correlations through vision as a mediating factor in the development and maintenance of audio-tactile correspondences.

Our finding that pitch-shape correspondences (low pitch as rounded, high pitch as angular) vary as a function of visual experience mirrors previous research on sound symbolism (i.e. the tendency for meaning to be conveyed in sound features), wherein matching auditory "bouba" and "kiki" to angular and jagged shapes was attenuated in the late blind and eliminated in the early blind (Fryer, Freeman & Pring, 2014). However, ours is the first demonstration that a similar pattern is observed for correspondences that utilise even more fundamental auditory characteristics such as individual frequencies (Marks, 1987). Historically, shape was considered by philosophers, such as Aristotle, to be metamodal (termed a 'common sensible', i.e. a property that is shared by multiple senses). This led others to give an affirmative answer to hypothetical questions such as: can a congenitally blind person, on recovering sight, recognise a sphere or cube by vision alone? But evidence from sight-recovered patients does not support instant cross-modal transfer (Held et al., 2011). Our study is also consistent with the idea that shape is not intrinsically metamodal, suggesting instead that visual and haptic shape processing are distinct, rather than just being two paths to 'shape processing.' If it were metamodal, then correspondences between sound and visual shape in sighted people should be equally apparent in the blind in terms of sound and haptic shape. It is possible that cross-modal correspondences between sound-shape derive specifically from statistical regularities in the visual environment or else require some more generic involvement of visual experience in the development of haptic shape perception. Whatever the mechanism, we hypothesise that it is relatively specific to shape because the same pattern was not found for other metamodal properties (size and texture).

We found that pitch-texture (smooth-rough) and pitch-softness (soft-hard) correspondences tended to be stronger in the blind than the sighted. Bonferroni-adjusted pairwise contrasts (see figure 2) showed that this result reflected increased tendencies for early and late blind persons to associate low pitch with rough textures and high pitch with softness, and an increased tendency for early blind persons to associate high pitch with smooth textures, compared to sighted controls. Eitan and Timmers (2010) found that sighted individuals associated low-pitched music with words like 'rough' and 'hard,' and high-pitched music with the words 'smooth' and 'soft' – however these verbal associations do not necessarily translate to physical correspondences or actual tactile experiences of texture and softness. Our results suggest that the experience of smooth and rough tactile textures does not actually map onto high or low pitch among the sighted. Other research on sighted persons has found that reading the nonsense words 'bouba' and 'kiki' is associated with the tactile feel of satin (smooth) and sandpaper (rough), respectively (Etzi, Spence, Zampini & Gallace, 2016); however, this seemingly contradictory finding could reflect grapheme-shape correspondences rather than the acoustic properties of those stimuli.

Considering that real-world experiences of tactile texture/softness paired with sound are unlikely to differ between the blind and sighted, what could account for this increased tendency to associate higher pitch with smoother textures and softer compliance in the two blind groups? It could be that blindness has a different level of influence on the different constituent parts that make up these multisensory couplings. One potential explanation relates to neuroplasticity of the visual cortex. Pitch-luminance is one of the strongest cross-modal correspondences in sighted persons, as seen by its influence on reaction times, its choice as the dominant visual characteristic for audiovisual matching, and its role in suppressing other correspondences such as pitch-chroma (Hamilton-Fletcher et al., 2017; Jonas, Spiller & Hibbard, 2017; Spence, 2011; Ward, Huckstep & Tsakanikos, 2006). However, with visual deprivation, regions normally responsive to luminance become responsive to tactile texture instead (Merabet et al., 2007; 2008; Stilla & Sathian, 2008). As such, in response to these changes to the functioning of the 'visual cortex' in the blind, we would expect an increase in the neural resources responsive to processing the tactile dimensions explored in the pitch-texture and pitch-softness conditions. If these changes in neural functioning are responsible for altering the correspondences expressed by the blind groups, this would be indicative of a structural explanation, rather than the currently prevailing statistical account.

Finally, there were two cross-modal correspondences that appear to be independent of visual experience, namely pitch-size and pitch-weight. Blind persons associated low pitch with heavier/larger objects to a comparable degree to sighted controls. The cross-modal correspondence between low tonal pitch (Evans & Treisman, 2009; Gallace & Spence, 2006) or vocal pitch (Pisanski et al., 2014; Ratcliffe et al., 2016) and perceived largeness is well documented. There is also evidence of a correspondence between pitch and the haptic-size of response keys (Walker & Smith, 1985), with smaller sizes congruent with higher pitches.

In related work, it has been found that sighted, late blind and early blind adults can accurately assess the relative body sizes of men and women from the formant frequencies of their voices (Pisanski, Oleszkiewicz & Sorokowska, 2016; Pisanski et al., 2017), with no differences in accuracy among blind and sighted listeners, providing further evidence that visual experience is not necessary to utilise pitch-size heuristics. Moreover, sound-size associations among blind participants assessing women's body size from the voice exhibit the same pattern of errors as those of sighted adults, suggesting that blind persons also use a similar, albeit erroneous, rule-of-thumb to estimate size (i.e., mapping not only low formant frequencies, but also low fundamental frequency/voice pitch to large size, despite the lack of a corresponding physical relationship between vocal pitch and body size within sexes; Pisanski et al., 2017). Pitch-size associations among blind persons may be acquired through pairings of other non-visual modalities (e.g., sound localisation indicating spatial elevation) or may be innate, present at birth in both human and non-human animals. The absence of an interaction between sightedness and the presence of pitch-size correspondences is troublesome for statistical accounts, as congenitally blind persons will have fewer sensory routes and exemplars of any naturally-occurring pitch-size associations, since many examples are easier to see than touch (e.g. elephants, birds). By contrast, for the pitch-weight correspondence the lack of an interaction with visual experience is expected (assuming it is based on a statistical correspondence) because the weight of an object is perceived via haptic exploration and cannot be directly experienced from the visual modality. Another related factor in assessing both size and weight is that of the perceived density of the object (Chouinard, Large, Chang & Goodale, 2009), it may be possible to examine its influence on correspondences from both the matching strategies and size/weight estimations made by participants.

Of the three audio-tactile correspondences that are influenced by visual experience, our results indicate that the late-blind group is situated between the sighted and early-blind groups, thus corresponding to differences in the degree of visual experience across groups. The late blind group's 'moderate tendency' could be homogeneous (i.e. all late-blind are more moderate), or heterogeneous (i.e. matching tendencies vary with visual experience). More detailed analyses showed that the late-blind's matching tendencies do not correlate with the proportion of life they have been blind, and so appear to be homogeneous. Another consideration is that different groups could have treated different pairs of stimuli (e.g. circles versus spheres) differently in their matching preferences; however, it was found that all stimulus sets belonging to the same tactile dimension were treated similarly within each group. These additional analyses can be found in the supplementary material.

5. Conclusion

Our results provide the first evidence that blindness is not associated with a universal decrease in audio-tactile correspondences, instead we reveal that while this does occur for some correspondences (pitch-shape), new correspondences can emerge or grow stronger with blindness (pitch-texture and pitch-softness), whereas others remain unaffected (pitch-size and pitch-weight). This evidence suggests that the currently prevailing preference to explain cross-modal correspondences as arising from experienced multisensory couplings (Deroy & Spence, 2013; Parise, 2016; Spence & Deroy, 2012) does not fare as well when attempting to explain multimodal perception in people with limited senses and reduced experience with multisensory interactions. Instead, alternative explanations from neuroplasticity, visual calibration and structurally-innate couplings may fare better at explaining this pattern of effects.

Code availability. The data that support the findings of this study are available as supplementary materials.

Acknowledgements

This work was supported by the RM Phillips Foundation (2014-2016 to J.W., G.H-F.), Polish Ministry of Science and Higher Education (2015–18 to K.P., 2013–16 to A.S.), the Foundation for Polish Science (START 2015 to K.P.), the Polish National Science Centre (OPUS 2013/11/B/HS6/01522 to A.S.), and the European Commission (H2020-MSCA-IF-2014-655859 to K.P.).

Author contributions

G.H-F., K.P. and J.W. designed the study with discussion from other co-authors, G.H-F. designed, collected and created all materials, the experiment was implemented by A.S. and conducted by M.S., analysis was conducted by D.R., K.P., and G.H-F., while the manuscript was written up by G.H-F., K.P., and J.W. with support from all co-authors.

Additional information

Supplementary information is available for this paper.

Correspondence should be addressed to G.H-F.

Competing interests

The authors have no competing interests to declare.

References

Albertazzi, L., Malfatti, M., Canal, L., & Micciolo, R. (2015). The hue of angles—was Kandinsky right?. *Art & Perception*, *3*(1), 81-92. doi: 10.1163/22134913-00002025

Bond, B., & Stevens, S. S. (1969). Cross-modality matching of brightness to loudness by 5-year-olds. *Perception & Psychophysics*, *6*(6), 337-339. doi: 10.3758/BF03212787

Chiou, R., & Rich, A. N. (2012). Cross-modality correspondence between pitch and spatial location modulates attentional orienting. *Perception*, *41*(3), 339-353. doi: 10.1068/p7161

Chouinard, P. A., Large, M. E., Chang, E. C., & Goodale, M. A. (2009). Dissociable neural mechanisms for determining the perceived heaviness of objects and the predicted weight of objects during lifting: An fMRI investigation of the size–weight illusion. *Neuroimage*, *44*(1), 200-212. doi: 10.1016/j.neuroimage.2008.08.023

Deroy, O., & Spence, C. (2013). Are we all born synaesthetic? Examining the neonatal synaesthesia hypothesis. *Neuroscience & Biobehavioral Reviews*, *37*(7), 1240-1253. doi: 10.1016/j.neubiorev.2013.04.001

Deroy, O., Fasiello, I., Hayward, V., & Auvray, M. (2016). Differentiated audio-tactile correspondences in sighted and blind individuals. *Journal of Experimental Psychology: Human Perception and Performance*, 42(8), 1204. doi: 10.1037/xhp0000152

Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition*, *114*(3), 405-422. doi: 10.1016/j.cognition.2009.10.013

Eitan, Z., Ornoy, E., & Granot, R. Y. (2012). Listening in the dark: Congenital and early blindness and cross-domain mappings in music. Psychomusicology: Music, Mind, and Brain, 22(1), 33-45. doi: 10.1037/a0028939

Etzi, R., Spence, C., Zampini, M., & Gallace, A. (2016). When sandpaper is 'kiki' and satin is 'bouba': an exploration of the associations between words, emotional states, and the tactile attributes of everyday materials. *Multisensory Research*, *29*(1-3), 133-155. doi: 10.1163/22134808-00002497

Evans, K. K., & Treisman, A. (2009). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, *10*(1), 6-6. doi: 10.1167/10.1.6

Fryer, L., Freeman, J., & Pring, L. (2014). Touching words is not enough: How visual experience influences haptic–auditory associations in the "Bouba–Kiki" effect. *Cognition*, *132*(2), 164-173. doi: 10.1016/j.cognition.2014.03.015

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophysics*, *68*(7), 1191-1203. doi: 10.3758/BF03193720

Goodyear, B. G., & Menon, R. S. (1998). Effect of luminance contrast on BOLD fMRI response in human primary visual areas. *Journal of Neurophysiology*, *79*(4), 2204–2207

Gougoux, F., Zatorre, R. J., Lassonde, M., Voss, P., & Lepore, F. (2005). A functional neuroimaging study of sound localization: visual cortex activity predicts performance in early-blind individuals. *PLoS Biol*, *3*(2), e27. doi: 10.1371/journal.pbio.0030027

Hamilton-Fletcher, G., Witzel, C., Reby, D., & Ward, J. (2017). Sound properties associated with equiluminant colours. *Multisensory Research*, *30*(3-5), 337-362. doi: 10.1163/22134808-00002567

Held, R., Ostrovsky, Y., de Gelder, B., Gandhi, T., Ganesh, S., Mathur, U., & Sinha, P. (2011). The newly sighted fail to match seen with felt. *Nature Neuroscience*, *14*(5), 551-553. doi: 10.1038/nn.2795

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons.

Hötting, K., Rösler, F., & Röder, B. (2004). Altered auditory-tactile interactions in congenitally blind humans: an event-related potential study. *Experimental brain research*, *159*(3), 370-381. doi: 10.1007/s00221-004-1965-3

Hötting, K., & Röder, B. (2004). Hearing cheats touch, but less in congenitally blind than in sighted individuals. *Psychological Science*, *15*(1), 60-64. doi: 10.1111/j.0963-7214.2004.01501010.x

Iachini, T., Ruggiero, G., & Ruotolo, F. (2014). Does blindness affect egocentric and allocentric frames of reference in small and large scale spaces?. *Behavioural Brain Research*, *273*, 73-81. doi: 10.1016/j.bbr.2014.07.032

Jäncke, L., Shah, N. J., Posse, S., Grosse-Ryuken, M., & Müller-Gärtner, H. W. (1998). Intensity coding of auditory stimuli: an fMRI study. *Neuropsychologia*, *36*(9), 875-883. doi: 10.1016/S0028-3932(98)00019-0

Lewald, J. (2002). Vertical sound localization in blind humans. *Neuropsychologia*, *40*(12), 1868-1872. doi: 10.1016/S0028-3932(02)00071-4

Jonas, C., Spiller, M. J., & Hibbard, P. (2017). Summation of visual attributes in auditory–visual crossmodal correspondences. *Psychonomic bulletin & review*, *24*(4), 1104-1112. doi: 10.3758/s13423-016-1215-2

Ludwig, V. U., Adachi, I., & Matsuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans. *Proceedings of the National Academy of Sciences*, *108*(51), 20661-20665. doi: 10.1073/pnas.1112605108

Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(3), 384. doi: 10.1037/0096-1523.13.3.384

Merabet, L.B., Hamilton, R., Schlaug, G., Swisher, J.D., Kiriakopoulos, E.T., Pitskel, N.B., Kauffman, T. & Pascual-Leone, A., 2008. Rapid and reversible recruitment of early visual cortex for touch. *PLoS One*, *3*(8), e3046. doi: 10.1371/journal.pone.0003046

Merabet, L. B., Swisher, J. D., McMains, S. A., Halko, M. A., Amedi, A., Pascual-Leone, A., & Somers, D. C. (2007). Combined activation and deactivation of visual cortex during tactile sensory processing. *Journal of Neurophysiology*, *97*(2), 1633-1641. doi: 10.1152/jn.00806.2006

Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 133-136. doi: 10.3758/CABN.4.2.133

Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *American Naturalist*, 855-869. doi: 10.1086/283219

Mulert, C., Jäger, L., Propp, S., Karch, S., Störmann, S., Pogarell, O., Möller, H-J., Hegerl, U. (2005). Sound level dependence of the primary auditory cortex: Simultaneous measurement with 61-channel EEG and fMRI. *Neuroimage*, *28*(1), 49–58. doi: 10.1016/j.neuroimage.2005.05.041

Nava, E., Grassi, M., & Turati, C. (2016). Audio-visual, visuo-tactile and audio-tactile correspondences in preschoolers. *Multisensory Research*, *29*(1-3), 93-111. doi: 10.1163/22134808-00002493

Occelli, V., Spence, C., & Zampini, M. (2009). Compatibility effects between sound frequency and tactile elevation. *Neuroreport*, *20*(8), 793-797. doi: 10.1097/WNR.0b013e32832b8069

Palmer, S. E., Schloss, K. B., Xu, Z., & Prado-León, L. R. (2013). Music–color associations are mediated by emotion. *Proceedings of the National Academy of Sciences*, *110*(22), 8836-8841. doi: 10.1073/pnas.1212562110

Parise, C. V. (2016). Crossmodal correspondences: standing issues and experimental guidelines. *Multisensory Research*, *29*(1-3), 7-28. doi: 10.1163/22134808-00002502

Parise, C. V., Knorre, K., & Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences*, *111*(16), 6104-6108. doi: 10.1073/pnas.1322705111

Pasqualotto, A., Spiller, M. J., Jansari, A. S., & Proulx, M. J. (2013). Visual experience facilitates allocentric spatial representation. *Behavioural Brain Research*, *236*, 175-179. doi: 10.1016/j.bbr.2012.08.042

Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'connor, J. J., & Feinberg, D. R. (2014). Return to Oz: Voice pitch facilitates assessments of men's body size. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1316-1331. doi: 10.1037/a0036956

Pisanski, K., Oleszkiewicz, A., & Sorokowska, A. (2016). Can blind persons accurately assess body size from the voice?. *Biology Letters*, *12*(4), 20160063. doi: 10.1098/rsbl.2016.0063

Pisanski, K., Feinberg, D., Oleszkiewicz, A., & Sorokowska, A. (2017). Voice cues are used in a similar way by blind and sighted adults when assessing women's body size. *Scientific Reports*, *7*. doi: 10.1038/s41598-017-10470-3

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia--a window into perception, thought and language. *Journal of Consciousness Studies*, *8*(12), 3-34.

Ratcliffe, V. F., Taylor, A. M., & Reby, D. (2016). Cross-modal correspondences in non-human mammal communication. *Multisensory Research*, *29*(1-3), 49-91. doi: 10.1163/22134808-00002509

Roskies, A. L. (1999). The binding problem. *Neuron*, *24*(1), 7-9. doi: 10.1016/S0896-6273(00)80817-X

Sorokowska, A. (2016). Olfactory performance in a large sample of early-blind and late-blind individuals. *Chemical senses*, *41*(8), 703-709. doi: 10.1093/chemse/bjw081

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971-995. doi: 10.3758/s13414-010-0073-7

Spence, C., & Deroy, O. (2012). Crossmodal correspondences: Innate or learned?. *i-Perception*, *3*(5), 316-318. doi: 10.1068/i0526ic

Stilla, R., & Sathian, K. (2008). Selective visuo-haptic processing of shape and texture. *Human Brain Mapping*, *29*(10), 1123-1138. doi: 10.1002/hbm.20456

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *353*(1373), 1295-1306. doi: 10.1098/rstb.1998.0284

Walker, P., & Smith, S. (1985). Stroop interference based on the multimodal correlates of haptic size and auditory pitch. *Perception*, *14*(6), 729-736. doi: 10.1068/p140729

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, *21*(1), 21-25. doi: 10.1177/0956797609354734

Walker, P., Bremner, J. G., Lunghi, M., Dolscheid, S., Dalla Barba, B., & Simion, F. (2018). Newborns are sensitive to the correspondence between auditory pitch and visuospatial elevation. *Developmental Psychobiology, 60*(2), 216:223. doi: 10.1002/dev.21603

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483-488. doi: 10.1016/j.tics.2003.09.002

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all?. *Cortex*, *42*(2), 264-280. doi: 10.1016/S0010-9452(08)70352-6

Wiesel, T. N. (1982). The postnatal development of the visual cortex and the influence of environment. *Bioscience Reports*, *2*(6), 351-377. doi: 10.1007/BF01119299

Zwiers, M. P., Van Opstal, A. J., & Cruysberg, J. R. (2001). A spatial hearing deficit in early-blind humans. *The Journal of Neuroscience*, *21*(9), 1-5.