Cognition 175 (2018) 114–121

Contents lists available at ScienceDirect

Cognition

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

Original Articles

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

Giles Hamilton-Fletcher[a,b,*], Katarzyna Pisanski[a,c], David Reby[a], Michał Stefańczyk[c], Jamie Ward[a,b], Agnieszka Sorokowska[c,d]

[a] School of Psychology, University of Sussex, Brighton, UK
[b] Sackler Centre for Consciousness Science, University of Sussex, Brighton, UK
[c] Institute of Psychology, University of Wrocław, Wrocław, Poland
[d] Smell & Taste Clinic, Department of Otorhinolaryngology, TU Dresden, Dresden, Germany

A R T I C L E I N F O

Keywords:
Blindness
Cross-modal correspondences
Pitch
Sound
Touch

A B S T R A C T

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.

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 (Roskies, 1999; Treisman, 1998). 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 (Albertazzi, Malfatti, Canal, & Micciolo, 2015; Chiou & Rich, 2012; Spence, 2011), 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

* Corresponding author at: School of Psychology, University of Sussex, Brighton, UK.
E-mail address: gh71@sussex.ac.uk (G. Hamilton-Fletcher).

https://doi.org/10.1016/j.cognition.2018.02.023
Received 9 October 2017; Received in revised form 20 February 2018; Accepted 22 February 2018

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 et al., 2016; Walker et al., 2010), 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 min. 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 33.9 y ± 12), along with twenty-seven late blind individuals (17 female, mean age 48.3 y ± 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).