Research Report

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
21(1) 21–25

DOI: 10.1177/0956797609354734
http://pss.sagepub.com
SAGE

Preverbal Infants' Sensitivity to Synaesthetic Cross-Modality Correspondences

Peter Walker[1], J. Gavin Bremner[1], Uschi Mason[1], Jo Spring[1], Karen Mattock[1], Alan Slater[2], and Scott P. Johnson[3]
[1]Lancaster University, [2]University of Exeter, and [3]University of California, Los Angeles

Abstract

Stimulation of one sensory modality can induce perceptual experiences in another modality that reflect synaesthetic correspondences among different dimensions of sensory experience. In visual-hearing synaesthesia, for example, higher pitched sounds induce visual images that are brighter, smaller, higher in space, and sharper than those induced by lower pitched sounds. Claims that neonatal perception is synaesthetic imply that such correspondences are an unlearned aspect of perception. To date, the youngest children in whom such correspondences have been confirmed with any certainty were 2- to 3-year-olds. We examined preferential looking to assess 3- to 4-month-old preverbal infants' sensitivity to the correspondences linking auditory pitch to visuospatial height and visual sharpness. The infants looked longer at a changing visual display when this was accompanied by a sound whose changing pitch was congruent, rather than incongruent, with these correspondences. This is the strongest indication to date that synaesthetic cross-modality correspondences are an unlearned aspect of perception.

Keywords

synaesthesia, infant perception, cross-modal perception, preferential looking

Received 3/6/09; Revision accepted 6/2/09

Synaesthesia occurs when stimulation of one sensory channel induces perceptions normally associated with a different sensory channel. Although there is much that is idiosyncratic about the experiences of individual synaesthetes, their experiences are constrained by a set of underlying cross-modality correspondences. These correspondences reflect the alignment of different dimensions of sensory experience (e.g., between visual brightness and auditory pitch) and can also be observed in nonsynaesthetes (where their influence is called *weak synaesthesia*; see Martino & Marks, 2001). Claims that neonatal perception is synaesthetic for all infants (James, 1890/1950; Maurer, Pathman, & Mondloch, 2006; Mondloch & Maurer, 2004) suggest that these cross-modality correspondences are an innate aspect of perception.

Cross-modality correspondences qualify most clearly as synaesthetic either when at least one of the dimensions involved is *metathetic* (i.e., concerned with a qualitative aspect of sensory experience) or when two *prothetic* dimensions (i.e., concerned with the amount of sensory experience) are aligned in opposition to their *more-than* interpretation (e.g., when less size is aligned with more brightness and more speed; see Walker & Smith, 1985).[1] When two prothetic dimensions are aligned according to their shared *more-than* interpretation (e.g., when more visual brightness is aligned with more auditory loudness), the cross-modality associations that arise need not be synaesthetic. Instead, they might reflect the fact that the two modalities are conveying equivalent (redundant) information about the same aspect of an object or event (as when vision and audition provide equivalent information about the strength of a stimulus). Lewkowicz and Turkewitz (1980) emphasized the latter interpretation in their seminal study demonstrating that 3- to 4.5-week-old infants respond to the equivalence of different levels of visual brightness and auditory loudness.

As a metathetic dimension, auditory pitch is involved in cross-modality correspondences with visuospatial height (Melara & O'Brien, 1987; Miller, Werner, & Wapner, 1958; Roffler & Butler, 1967), visual sharpness (Marks, 1987; O'Boyle & Tarte, 1980), visual brightness (Collier & Hubbard, 2001; Marks, 1978; Walker & Smith, 1984, 1985; Ward, Huckstep, & Tsakanikos, 2006), lightness in weight (Walker & Smith, 1984), and smallness in size (Parise & Spence, 2008; Perrott, Musicant, & Schwethelm, 1980; Walker &

Corresponding Author:
Peter Walker, Department of Psychology, Lancaster University, Lancaster LA1 4YF, United Kingdom
E-mail: p.walker@lancaster.ac.uk

Smith, 1984). Some of these studies confirmed the impact of cross-modality correspondences on perception (e.g., Miller et al., 1958; Parise & Spence, 2008; Roffler & Butler, 1967). Strong evidence for this also comes from visual-hearing synaesthesia, where higher pitched sounds induce visual images that are brighter, smaller, higher in space, sharper, and more likely to be moving than those induced by lower pitched sounds (Karwoski & Odbert, 1938; Marks, 1987; Ward et al., 2006).

If neonatal perception is synaesthetic, then infants should reveal their sensitivity to these cross-modality correspondences. Wagner, Winner, Cicchetti, and Gardner (1981) claimed to have shown that 9- to 13-month-olds appreciate the correspondence between visual height (specifically, an up and down pointing arrow) and auditory pitch (an ascending and descending tone). As noted by Wagner et al., however, arrows are conventional symbols, whose significance needs to be learned, and the infants may not have interpreted them as representing ascending and descending change. Instead, Wagner et al. proposed that the infants looked at the top of an up arrow, and at the bottom of a down arrow, because these are the locations at which the highest density of visual information is found. Wagner et al. implied, therefore, that the difference in the elevation of infants' fixation is the critical factor in correspondence with the ascending and descending pitch of a sound. Wagner et al. offered no supporting evidence for the presence and significance of such a correspondence in adults and did not monitor the elevation of the infants' fixation, and so we contend that the youngest children in whom synaesthetic cross-modality correspondences have been confirmed with any certainty were 2- to 3-year-olds (Maurer, Pathman, & Mondloch, 2006; Mondloch & Maurer, 2004).

This age range does not preclude mediation through postnatal learning. Such learning could be based on a child's exposure to language, because synaesthetic cross-modality correspondences find expression in the visible, spoken, and conceptual aspects of language (Marks, 1978). For example, across a range of languages, smaller, brighter, sharper, and faster objects tend to have names with higher pitched vowel sounds (Marks, 1978; Maurer et al., 2006). To test the perceptual innateness of synaesthetic cross-modality correspondences, while avoiding any learning based on language comprehension, we used a preferential looking procedure to assess 3- to 4-month-old preverbal infants' sensitivity to two synaesthetic correspondences involving auditory pitch.

In the first of two experiments, we examined the correspondence between auditory pitch and visuospatial height. In the second experiment, we examined the correspondence between auditory pitch and visual sharpness. The existence of these two perceptual correspondences has been confirmed with several observations. For example, when adults point to the perceived visual location of a sound whose source is hidden from view, the elevation of this location is determined by the pitch of the sound, and not by the actual location of the source of the sound (Roffler & Butler, 1967). And when an incidental sound rises and falls in pitch, it induces corresponding visual illusory movement (i.e., upward vs. downward autokinetic movement, respectively) in a stationary spot of light (Miller et al., 1958). In addition to being observed in visual-hearing synaesthesia, the correspondence between auditory pitch and visual sharpness emerges when adults judge the sharpness of auditory tones varying in pitch using a scale defined by the polar adjectives *sharp* and *blunt* (Walker & Smith, 1985). And when, in a preliminary phase of the present study, people rated pointed and curved variants of several geometric shapes (including those in Fig. 1b) on a scale with endpoints defined as *would make a high/low-pitched sound if it came to life or was struck by another object*, pointedness was again associated with higher pitch.

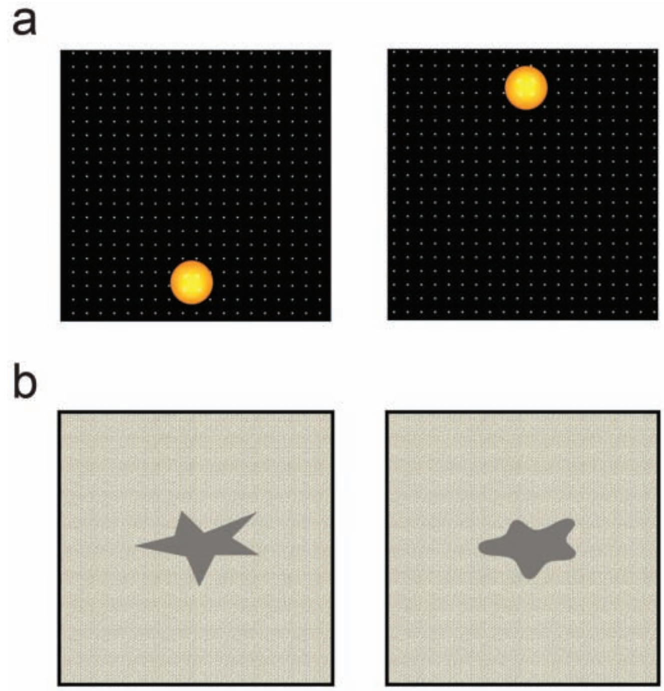

Fig. 1. Examples of animations presented as stimuli in the (a) pitch-height and (b) pitch-sharpness experiments. In (a), the moving ball is shown at the extremes of its vertical trajectory. In (b), the extreme forms of pointedness of the static morphing shape are shown. All images are reproduced to scale.

Method

Participants

Eight male and 8 female infants completed the first (pitch-height) experiment (mean age = 128 days, range: 114 to 145 days). A further 2 infants were unable to complete this experiment because of excessive restlessness. Eight male and 8 female infants completed the second (pitch-sharpness) experiment (mean age = 129 days, range: 91 to 150 days). A further

3 infants were unable to complete the experiment because of excessive restlessness.

Materials and procedure

A Barco screen was integrated into one side of a 3-m^2 cubicle created with thick, black curtains. QuickTime animations appeared within a 48 × 48 cm screen area (25.6° × 25.6°), each lasting a maximum of 60 s. Each infant sat on her or his mother's knee, viewing the animations at eye level and from a distance of 1 m. The mothers fixed their own gaze on the top of their child's head. Each infant's eye fixations were monitored and recorded on video, and an animation was removed immediately after he or she looked elsewhere for a single period of 2 s or more. The total length of time an infant looked directly at an animation was later determined by two independent judges, who were unaware of the nature of the animation being observed by the infant.

In the pitch-height experiment, the infants watched a 7-cm (4°) diameter orange ball (average luminance = 39 cd/m^2) move up and down a 32-cm vertical trajectory in front of a 20 × 20 grid of small, white dots on an otherwise dark field (see Fig. 1a). The deletion and accretion of the dots served to strengthen the impression that the ball was a solid object moving in its own depth plane. The ball moved at a constant speed of 12.8 cm/s and paused for 50 ms at each endpoint. Each animation was accompanied by the sound of a sliding whistle, whose fundamental frequency changed at a constant rate, between 300 and 1700 Hz, over a period of time coinciding with an individual phase of the animation (2.5 s). The loudness of the sound increased and then decreased between 54 and 80 dB (against an uncontrolled ambient noise level of 42 dB) within each phase of the animation, peaking when the fundamental frequency of the sound was midrange. The natural sound of a sliding whistle has a loudness profile of this type, and this profile was retained during sound editing for two reasons: first, to preserve the naturalness of the sound (which we considered especially important for infants) and, second, in the absence of individual equal-loudness contours for our participants, to ensure that changes in perceived loudness were not confounded with changes in pitch. In the *congruent* condition, the pitch of the sound rose and fell as the ball rose and fell. In the *incongruent* condition, the pitch of the sound fell and rose as the ball rose and fell. Every infant viewed three congruent animations interleaved with three incongruent animations, with 8 infants viewing a congruent animation first and 8 viewing an incongruent animation first.

In the pitch-sharpness experiment, the infants watched computerized animations of a static, uniformly gray (13 cd/m^2) geometric shape morphing constantly between two extreme forms of pointedness (see Fig. 1b). The shape appeared against a beige, hopsack background (average luminance = 28 cd/m^2) and in its pointed and curved extreme forms had a surface area of 100.3 and 99.7 cm^2, respectively. With this small difference in surface area, the overall size of the shape did not appear to change as it morphed (as judged by a set of independent adult judges in a preliminary study). The morphing was accompanied by the same sound used in the pitch-height experiment. Each phase of the animation (morphing from one extreme to the other) again lasted 2.5 s, with a brief 50-ms pause at each extreme. In the *congruent* condition, the pitch of the sound rose and fell as the shape became more and less pointed. In the *incongruent* condition, the pitch of the sound fell and then rose as the shape became more and less pointed. Other aspects of the method were identical to the pitch-height experiment.

Results

In both experiments, a high level of agreement was confirmed between the two judges in their estimates of each infant's individual viewing times across the six trials (mean Pearson $r = .99$, $p < .01$, in both cases).

Auditory pitch and visuospatial height

Twelve of the 16 infants looked longer at a congruent animation than at an incongruent animation ($p = .038$ on a binomial test, $p_{rep} = .90$). The average time infants looked at congruent and incongruent animations was 28.5 s (*SD* = 10.5) and 21.7 s (*SD* = 11.9), respectively. Because the looking times were positively skewed, they were log-transformed before analysis of variance confirmed the significance of the effect of congruity, $F(1, 15) = 9.22$, $p = .008$, $p_{rep} = .94$, $\eta_p^2 = .38$.

Auditory pitch and visual sharpness

Twelve of the 16 infants looked longer at a congruent animation than at an incongruent animation ($p = .038$ on a binomial test, $p_{rep} = .90$). The average time infants looked at congruent and incongruent animations was 20.1 s (*SD* = 7.9) and 15.9 s (*SD* = 6.3), respectively. Analysis of variance of the log-transformed looking times confirmed the significance of the effect of congruity, $F(1, 15) = 7.57$, $p = .015$, $p_{rep} = .96$, $\eta_p^2 = .34$. The percentage reduction in looking time caused by introducing incongruity was comparable across both experiments (i.e., 24% and 21% for the pitch-height and pitch-sharpness experiments, respectively).

Discussion

Our results confirm young infants' sensitivity to the correspondences linking auditory pitch with visuospatial height and visual sharpness. They provide the strongest indication to date that synaesthetic cross-modality correspondences are an unlearned aspect of perception and are consistent with claims that neonatal perception is synaesthetic.

Some might wish to argue that our infants were old enough to have learned the correspondences. This raises two questions. First, where might the correspondences originate, if not in language? Second, could there be sufficient exposure to these origins to allow the correspondences to be learned within the first 3 or 4 months of life? It is possible that relevant co-occurrences exist in the real world, and exposure to these ensures perceptual learning of the underlying dimensional correspondences. For example, when they hear people or animals vocalize (Bee, Perrill, & Owen, 2000; Davies & Halliday, 1978; Harrington, 1987) and objects strike each other (Grassi, 2005; Perrott et al., 1980), infants will encounter co-occurrences reflecting the correspondence between size and pitch, and these encounters might be sufficient to establish perceptual expectations linking size and pitch. It is difficult, however, to identify natural co-occurrences capable of supporting other correspondences, including the two correspondences examined here. With regard to sharpness and pitch, a potentially relevant co-occurrence might be mediated by the hardness of natural materials. Sharp and pointed objects tend to be formed from hard materials, for which reason the objects are also likely to make relatively high-pitched sounds when struck (Freed, 1990; Klatzky, Pai, & Krotkov, 2000; van den Doel & Pai, 1998). Future research must explore the incidence of natural co-occurrences to determine if the typical environment to which young infants are exposed could support perceptual learning of the synaesthetic cross-modality correspondences observed in adulthood. Confirmation would not, of course, mean that sensitivity to the correspondences is acquired during infancy. Indeed, we doubt very much that infants' encounters with objects within the first 3 to 4 months could support such learning, in part because any relevant co-occurrences will not be in evidence with complete consistency (e.g., objects that look sharp will sometimes make a low-pitched sound, and vice versa). Furthermore, as implied by the claim that neonatal perception is synaesthetic, acquisition of the correspondences could be phylogenetic rather than ontogenetic, with appropriate neural circuitry in place at birth (Maurer & Mondloch, 2006).

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interests with respect to their authorship and/or the publication of this article.

Note

1. The distinction between prothetic and metathetic dimensions (continua) of sensory experience is, in essence, a distinction between quantity and quality (Stevens, 1957). Prothetic dimensions are concerned with how much of a type of sensation is being experienced (e.g., how loud) and have been linked to variations in the level of excitation within a sensory channel. Metathetic dimensions deal with qualitative differences in sensory experience and have been linked to variations in the location of maximum excitation in a sensory channel.

References

Bee, M.A., Perrill, S.A., & Owen, P.C. (2000). Male green frogs lower the pitch of acoustic signals in defense of territories: A possible dishonest signal of size? *Behavioral Ecology*, *11*, 169–177.

Collier, W.G., & Hubbard, T.L. (2001). Judgements of happiness, brightness, speed and tempo change of auditory stimuli varying in pitch and tempo. *Psychomusicology*, *17*, 36–55.

Davies, N.B., & Halliday, T.R. (1978). Deep croaks and fighting assessment in toads *Bufo bufo. Nature*, *274*, 683–685.

Freed, D.J. (1990). Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events. *Journal of the Acoustical Society of America*, *87*, 311–322.

Grassi, M. (2005). Do we hear size or sound? Balls dropped on plates. *Perception & Psychophysics*, *67*, 274–284.

Harrington, F.H. (1987). Aggressive howling in wolves. *Animal Behavior*, *35*, 7–12.

James, W. (1950). *Principles of psychology*. New York: Dover. (Original work published 1890)

Karwoski, T.F., & Odbert, H.S. (1938). Color-music. *Psychological Monographs*, *50*(2, Whole No. 222).

Klatzky, R.L., Pai, D.K., & Krotkov, E.P. (2000). Perception of material from contact sounds. *Presence*, *9*, 399–410.

Lewkowicz, D.J., & Turkewitz, G. (1980). Cross-modal equivalences in early infancy: Auditory-visual intensity matching. *Developmental Psychology*, *16*, 597–607.

Marks, L.E. (1978). *The unity of the senses: Interrelations among the modalities.* New York: Academic Press.

Marks, L.E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 384–394.

Martino, G., & Marks, L.E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, *10*, 61–65.

Maurer, D., & Mondloch, C.J. (2006). The infant as synesthete? In Y. Munakata & M.H. Johnson (Eds.), *Attention and performance XXI: Processes of change in brain and cognitive development* (pp. 449–471). Oxford, England: Oxford University Press.

Maurer, D., Pathman, T., & Mondloch, C.J. (2006). The shape of boubas: Sound-shape correspondences in toddlers and adults. *Developmental Science*, *9*, 316–322.

Melara, R.D., & O'Brien, T.P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, *116*, 323–336.

Miller, A., Werner, H., & Wapner, S. (1958). Studies in physiognomic perception: V. Effect of ascending and descending gliding tones on autokinetic motion. *The Journal of Psychology*, *46*, 101–105.

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

O'Boyle, M.W., & Tarte, R.D. (1980). Implications for phonetic symbolism: The relationship between pure tones and geometric figures. *Journal of Psycholinguistic Research*, *9*, 535–544.

Parise, C., & Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neuroscience Letters*, *442*, 257–261.

Perrott, D.R., Musicant, A., & Schwethelm, B. (1980). The expanding-image effect: The concept of tonal volume revisited. *Journal of Auditory Research*, *20*, 43–55.

Roffler, S.K., & Butler, R.A. (1967). Localization of tonal stimuli in the vertical plane. *Journal of the Acoustical Society of America*, *43*, 1260–1266.

Stevens, S.S. (1957). On the psychophysical law. *Psychological Review*, *64*, 153–181.

van den Doel, K., & Pai, D.K. (1998). The sounds of physical shapes. *Presence*, *7*, 382–395.

Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). "Metaphorical" mapping in human infants. *Child Development*, *52*, 728–731.

Walker, P., & Smith, S. (1984). Stroop interference based on the synaesthetic qualities of auditory pitch. *Perception*, *13*, 75–81.

Walker, P., & Smith, S. (1985). Stroop interference based on the multimodal correlates of haptic size and auditory pitch. *Perception*, *14*, 729–736.

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*, 264–280.