Developmental Psychology
2009, Vol. 45, No. 1, 175–189

0012-1649/09/$12.00 DOI: 10.1037/a0014171

Synesthesia: A New Approach to Understanding the Development of Perception

Ferrinne Spector and Daphne Maurer
McMaster University

In this article, the authors introduce a new theoretical framework for understanding intersensory development. Their approach is based upon insights gained from adults who experience synesthesia, in whom sensory stimuli induce extra cross-modal or intramodal percepts. Synesthesia appears to represent one way that typical developmental mechanisms can play out by magnifying connections present in early life that are pruned and/or inhibited during development but persist in muted form in all adults. As such, the study of synesthesia provides valuable insights into the nature of intersensory development. The authors review evidence on the perceptual reality and neural basis of synesthesia, then summarize developmental models and evidence that its underlying mechanisms are universal in adults. They illustrate how evidence for consistent sensory associations in adults leads to predictions about toddlers' perception and present 3 bodies of work that have confirmed those hypotheses. They end by describing novel hypotheses about intersensory development that arise from this framework. Such intersensory associations appear to reflect intrinsic sensory cortical organization that influences the development of perception and of language and that may constrain the learning of environmentally based associations.

Keywords: synesthesia, cross-modal, perceptual development, multisensory, pruning

At birth, babies can already recognize the mother's voice (DeCasper & Fifer, 1980; Ockleford, Vince, Layton, & Reader, 1988), but they must learn to associate that voice with the correct face. Later they must learn to associate the sound "ay" with the written letter *A* and a perfumey smell with roses. Classic theoretical approaches by Piaget (1952) and Gibson (1966) have generated empirical studies on infants' ability to transfer information learned in the lab across modalities and to recognize naturally occurring cross-modal correspondences. These studies indicate that some cross-modal associations are readily grasped by young infants, whereas others emerge later; researchers have explained the differences on the basis of the integration of schemas (Piaget) or amodal versus arbitrary correspondences (Gibson). In this article, we introduce an alternative theoretical framework for understanding the development of cross-modal perception derived from recent findings about adult synesthesia.

For an adult with synesthesia, a stimulus induces not only the usual percept but also a seemingly automatic additional percept, often in a different sensory modality. For example, sound may induce color, with the specific color tied to the pitch or timbre of the sound: For a particular synesthete, a high *C* on a trumpet will induce the color vermillion, whereas the same note on a violin or a lower note on the trumpet will induce the color pink. Among the most common types of synesthesia is *colored grapheme synesthesia*, in which black letters and digits induce colored percepts, and *time–space synesthesia*, in which units of time (weekdays, months, digits) are laid out regularly in space (e.g., Simner et al., 2006). Recent behavioral and neuroimaging studies (e.g., Hubbard, Arman, Ramachandran & Boynton, 2005; Muggleton, Tsakanikos, Walsh & Ward, 2007; Simner et al., 2006) have established the perceptual reality of synesthetic percepts and have documented that it occurs in roughly 5% of the adult population.

In this article, we present evidence that adult synesthesia can inform the study of the development of perception and even of language, because it appears to represent one way in which normal developmental mechanisms can play out. In a sense, it magnifies connections present in early life that are pruned and/or inhibited during development and that persist in muted form in all adults. Thus, the cross-modal connections that can be documented in adult synesthetes can be used to derive novel hypotheses about cross-modal links likely to be present in early childhood.

We begin by summarizing the scientific evidence about the nature of synesthesia and its neural basis. We then describe the two current theories of the developmental origins of synesthesia and note the implications of each for understanding typical development. We illustrate the value of our approach by showing how three hypotheses derived from adult synesthetes have led to novel tests of perception and language in toddlers and by describing additional hypotheses that could be tested. We end by contrasting this approach to the traditional approaches of Piaget and Gibson.

Synesthesia

There are at least 54 types of synesthesia, the majority of which involve extra visual percepts of color (Day, 2007). Synesthetes

---

Ferrinne Spector and Daphne Maurer, Department of Psychology, Neuroscience, and Behavior, McMaster University, Hamilton, Ontario, Canada.

This research was funded by a grant from the National Science and Engineering Research Council of Canada, awarded to Daphne Maurer, and an Ontario Graduate Scholarship, awarded to Ferrinne Spector.

Correspondence concerning this article should be addressed to Daphne Maurer, Department of Psychology, Neuroscience, and Behavior, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1. E-mail: maurer@mcmaster.ca

report that they have had the extra percepts all their lives, and many experience more than one type of synesthesia. Some synesthetes indicate that the extra percepts are projected into specific locations in space, where they may be superimposed on real world stimuli (projectors), whereas others report that the extra percepts are in the mind's eye (associators; Dixon, Smilek, & Merikle, 2004; see also Ward & Sagiv, 2007). Although the specific mappings vary among synesthetes, there is nevertheless some consistency across individuals and between the extra cross-modal percepts of synesthetes and the cross-modal associations of nonsynesthetic adults. It is from those consistencies that we derive developmental predictions later in this article.

Prevalence and Perceptual Reality

Synesthesia was historically dismissed as a rare phenomenon arising from overly vivid imagination—mainly in women. However, recent studies of the general population suggest that it occurs in 4%–5% of adults and is as common in men as in women (Simner et al., 2006). Recent studies have also used a variety of techniques to establish its perceptual reality. One method has been to study the specificity and consistency of synesthetic percepts over time, compared to the cross-modal associations of the typical population. An early study of nine synesthetes with colored word hearing established consistency as a hallmark of synesthesia: The synesthetes chose the same color 92% of the time when given a surprise retest after 1 year, compared to 38% consistency in a control group given a retest that they had been told to expect after 1 week (Baron-Cohen, Harrison, Goldstein, & Wyke, 1993). Similarly high rates of consistency have been shown in later studies of colored hearing synesthesia (Asher, Aitken, Farooqi, Kurmani, & Baron-Cohen, 2006; de Thornley Head, 2006; Ward, Huckstep, & Tsakanikos, 2006), colored grapheme synesthesia (Dixon, Smilek, Cudahy, & Merikle, 2000; Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007; Edquist, Rich, Brinkman, & Mattingley, 2006; Mattingley, Rich, Yelland, & Bradshaw, 200l), time–space synesthesia (Smilek, Callejas, Dixon, & Merikle, 2007), and word–taste synesthesia (Ward & Simner, 2003), with no overlap between the worst synesthetic score and the best control score.

Additional evidence for the perceptual reality of synesthesia comes from studies showing that synesthetic percepts interact with the perception of real-world stimuli as would be expected if they were perceptual phenomena, interfering with typical percepts if they are incongruent and facilitating typical percepts if they are congruent. The most prominent example is Stroop interference. Much as a color word (*green*) interferes with naming an incompatible color of ink (*red*), individuals with color grapheme synesthesia have difficulty naming the color of ink of a letter or digit if it induces an incompatible synesthetic color (Dixon et al., 2000; Mattingley et al., 200l; Mattingley, Payne, & Rich, 2006; Ward et al., 2006). The incongruency causes longer reaction times and induces pupillary dilation (Paulsen & Laeng, 2006). For example, a synesthete who reports that the digit 1 induces the color green will be slower to name the color of a red 1 than a black or green 1. In fact, subjects report an inability to suppress the incompatible synesthetic color, much as we cannot prevent ourselves from automatically processing the meaning of the word when it is irrelevant to the task of naming the color of ink (Mattingley et al., 2001). Stroop-like interference is largest when the real-world and synesthetic colors are opponent colors (red–green or blue–yellow), an effect suggesting that synesthetic color arises from the same opponent color mechanisms in the primary visual cortex and/or extrastriate visual area V4 as mediate normal color vision (Nikolié, Lichti, & Singer, 2007).

Stroop-like interference also occurs for colored hearing: Synesthetes for whom auditory tones induce colors are slower to name the color of a visual patch when listening to an irrelevant tone that induces an incongruent rather than congruent color (Ward et al., 2006). Similarly, when individuals with colored graphemes are asked to name a letter formed from the arrangement of smaller letters (i.e., a large black letter *A* formed by the arrangement of smaller black *E*s), they are slower if the two letters induce incongruent synesthetic colors (Rich & Mattingley, 2003). Like the results for the Stroop effect, these results suggest that the synesthetic percepts are automatic and not easily suppressed.

Synesthetic colors act like real colors in inducing a number of other perceptual phenomena—apparent motion, grouping during binocular rivalry, figure–ground segmentation, visual search, and attentional priming—in some cases inducing effects as strong as the interactions among real colors and in others effects that are similar but less strong (Hubbard et al., 2005; Hubbard, Manohar, & Ramachandran, 2006; Kim, Blake, & Palmieri, 2006; Laeng, Svartdal, & Oelmann, 2004; Mattingley et al., 2001, 2006; Palmieri, Blake, Marois, Flanery, & Whetsell, 2002; Ramachandran & Azoulai, 2006; Smilek et al., 2007; Smilek, Dixon, Cudahy, & Merikle, 2001; Smilek, Dixon, & Merikle, 2003; Ward et al., 2006; but see Edquist et al., 2006; Sagiv, Heer, & Robertson, 2006). Like the data on consistency over time, the behavioral findings that synesthetic percepts behave like typical percepts in perceptual and attentional tasks have established their perceptual reality.

Neural Basis of Synesthesia

Recent studies using functional magnetic resonance imaging (fMRI) reveal that during synesthetic percepts there is activation in the expected brain areas (e.g., areas normally activated by the inducing stimulus). In addition, there is activation in brain areas, usually contiguous, that respond when an external stimulus induces the synesthetic percept, even though, for the synesthete, the external stimulus is not needed. An example comes from an fMRI study of a synesthete ("JIW") who tastes words: Specific words evoke highly specific tastes ("Philip" evokes the taste of "oranges not quite ripe"; Ward & Simner, 2003, p. 241). When JIW listened to words in the scanner, there was activation in the primary gustatory cortex (Brodman's area 43) that did not occur when he listened to tones (Ward & Simner, 2003). Similarly, when synesthetes with colored hearing listen to inducers in the scanner (e.g., the sound of the letter *L* or of a trumpet that induces red), there is activity in visual cortical area V4/ V8, which plays a key role in the processing of color and form in typical perception (Aleman, Rutten, Sitskoorn, Dautzenberg, & Ramsey, 2001; Gray et al., 2006; Hubbard et al., 2005; Nunn et al., 2002; Paulesu et al., 1995; Stevens, Hansen, & Blakemore, 2006; Winawer & Witthoft, 2004). In addition, in some studies, there was activation of the primary visual cortex, of a number of higher visual association areas, and of areas in the parietal cortex in the angular gyrus that bind color to shape.

Similarly, in synesthetes for whom black letters evoke colored percepts (i.e., who have colored grapheme synesthesia), viewing letters in the scanner causes activation in visual cortical area V4/V8, with some reports of additional activation in lower visual areas, including primary visual cortex V1, and in a number of higher cortical areas, including the intraparietal cortex (Hubbard et al., 2005; Rouw & Scholte, 2007; Sperling, Prvulovic, Linden, Singer, & Stirn, 2006; but see Rich et al., 2006; Weiss, Zilles, & Fink, 2005). A role for the parietal cortex was confirmed by two studies using transcranial magnetic stimulation (TMS) to temporarily deactivate specific areas of the parietal cortex in synesthetes with colored graphemes: Silencing the right parietal–occipital area and, in some subjects, the right parietal area or the left parietal–occipital area, reduced the interference between synesthetically induced and real colors (Esterman, Verstynen, Ivry, & Robertson, 2006; Muggleton et al., 2007; see Rouw & Scholte, 2007, for converging evidence of increased parietal connections in the left parietal cortex of grapheme color synesthetes). Like the fMRI results, these findings suggest a role for the areas of the parietal cortex that are involved in binding color to shape in typical perception. Overall, the neuroimaging results suggest that the brain connections mediating synesthesia are similar to those mediating typical perception.

Developmental Origins of Synesthesia

There are two predominant theories about the development of synesthesia, both of which are rooted in normal developmental processes. In the case of the cross-activation theory, synesthesia arises when the pruning of synapses is not completed between some contiguous brain areas (e.g., Maurer & Maurer, 1988; Ramachandran & Hubbard, 2001). In the case of the disinhibited feedback theory, synesthesia arises when the reentrant feedback that develops postnatally from higher cortical areas onto lower sensory cortical areas is not strong enough to inhibit effects from connections between primary sensory cortical areas (Grossenbacher & Lovelace, 2001). By either account, cross-modal effects similar to those seen in adult synesthesia are expected to occur during early childhood and to persist in muted form even in typical adults.

Cross-Activation

In adults, each sensory cortical area is specialized for the processing of information from one sensory modality: Neurons in the visual cortex respond to input from the eyes; neurons in the auditory cortex respond to input from the ears, etc. The cross-activation theory is based on evidence that, across a number of species, sensory cortical areas are initially not as specialized as they become. Instead, there are transient connections between sensory cortical areas that are pruned during childhood in an experience-dependent manner. For example, in the kitten there are transient connections between the visual, auditory, tactile, and motor cortices (Dehay, Bullier, & Kennedy, 1984; Dehay, Kennedy, & Bullier, 1988); although such ubiquitous over-wiring does not appear to occur in the infant monkey brain, there are transient connections from auditory cortex to visual area V4 (Kennedy, Batardiere, Dehay, & Barone, 1997), the color area that is active when adult synesthetes with colored hearing listen to an inducing sound. There is indirect evidence that the same phenomenon—a superabundance of connections between sensory cortical areas followed by experience-dependent pruning—occurs in humans and that the extra connections are functional during early childhood. For example, in the newborn, tactile stimulation of the wrist evokes activity over the somatosensory cortex, as it does in adults, but unlike adults, the response is enhanced if accompanied by the sound of white noise (Wolff, Matsumiya, Abrohms, van Velzer, & Lombroso, 1974). In young infants, spoken language elicits activity over the auditory cortex, as expected, but unlike adults, it evokes just as much activity over the visual cortex; with age, the activity over the visual cortex diminishes, but it does not disappear until about age 3 (Neville, 1995). Converging evidence comes from a study that used positron emission tomography as 2-month-olds watched faces: The faces elicited more activity than the control visual stimulus in the right inferior temporal gyrus, near the classic fusiform face area of adults, but unlike adults, they also elicited more activity in the left auditory cortex and left Broca's area that is later specialized exclusively for language (Tzourio-Mazoyer et al., 2002; see also Huttenlocher, 1984; Huttenlocher, 1994; Huttenlocher & Dabholkar, 1997; Huttenlocher & de Courten, 1987; Huttenlocher, de Courten, Garey, & Van der Loos, 1982). Combined with the animal data, these findings suggest that there are functional connections between sensory cortical areas during early childhood that are later pruned.

According to the cross-activation account, synesthesia occurs when some of the connections between sensory cortical areas (usually ones that are contiguous) are not pruned (Ramachandran & Hubbard, 2001). Thus, synesthetic color is evoked because connections between area V4/V8 and areas mediating the perception of sound, words, graphemes, or taste were not pruned. Evidence for such extra connections was obtained in a recent study using diffusion tensor imaging to trace white matter tracks: Adults with colored grapheme synesthesia showed evidence of greater connectivity than did controls in three brain regions, including the word form area in the inferior temporal cortex that lies contiguous to V4/V8; the strength of hyperconnectivity correlated with the strength of projecting the synesthetic colors onto the inducing black letters (Rouw & Scholte, 2007). Even in typical adults, some of those connections may not be pruned, leading to synesthetic-like cross-modal effects that do not reach conscious perception (see below).

Although the reason for the reduced pruning in synesthetes is not known, a genetic factor is suggested by its tendency to run in families (e.g., Ward & Simner, 2005). In addition, studies of sensory deprivation suggest indirectly that the pruning of connections between sensory cortical areas is shaped by experience: When the normal input is missing because the child is blind or deaf, the primary sensory cortex missing its normal input does not develop normal specialization, but instead responds to input from other sensory modalities (reviewed in Maurer, Lewis, & Mondloch, 2005). In adults blind from an early age, Braille reading (Sadato et al., 1996), as well as the untrained tactile tasks of discriminating between vibrotactile gratings (Burton, Sinclair, & McLaren, 2004) and between embossed roman letters (Burton, McLaren & Sinclair, 2006), recruit the visual cortex, including both extrastriate and primary visual cortices (for reviews, see Amedi, Merabet, Bermpohl, & Pascual-Leone, 2005; Maurer et al., 2005). The ability of the blind group to discriminate Braille and embossed roman letters is impaired when visual cortical activity is reduced by applying TMS over the medial occipital cortex or from an occipital stroke (Cohen et al., 1997, 1999; Hamilton, Keenan, Catala, & Pascual-Leone, 2000), whereas sighted individuals' ability to discriminate embossed roman letters is impaired by TMS applied only over the sensorimotor cortex. In adults with congenital blindness, the visual cortex also responds to auditory stimuli and perhaps even language (e.g., Burton, Snyder, Diamond, & Raichle, 2002; Röder, Rösler, Hennighausen, & Nacker, 1996; Röder, Stock, Bien, Neville, & Rösler, 2002; Sadato et al., 1998; Sadato, Okado, Honda, & Yonekura, 2002). Thus, it appears as if the visual cortex is recruited after early blindness for tactile and auditory perception. Similarly, in kittens whose eyes were removed at birth, neurons in the visual cortex give well-tuned responses to sound (Yaka, Yinon, & Wollberg, 1999). These data converge in indicating that the specialization of sensory cortical areas is tuned by experience from the expected sensory modality, which may be favored because its input is stronger, faster, or more coherent over space and time than input from other sensory modalities. In the absence of the expected input—and perhaps to a lesser extent when there is a genetic predisposition to synesthesia—connections from the "wrong" sensory modality remain and influence perception.

Disinhibited Feedback

The alternative explanation of the development of synesthesia is that it arises from altered feedback from higher cortical areas onto lower sensory cortices (Grossenbacher & Lovelace, 2001). In the typical adult, the feedback strengthens firing of neurons consistent with the expected stimulus (e.g., neurons tuned to horizontal and vertical orientations when a square is expected or begins to be perceived) and inhibits inconsistent firing (e.g., neurons tuned to diagonal orientations or responding to input from an unexpected modality, such as audition). What happens in synesthesia, according to this account, is that some of the inhibitory feedback is disinhibited, allowing primary sensory cortical areas to be activated by unexpected input from the wrong sense. This account rests on the presumption that connections between sensory cortices are not all eliminated by pruning; instead, some remain but are normally functionally inhibited. Evidence for this presumption has emerged in recent animal studies: Sensory cortical areas that were traditionally thought to be unimodal receive input from other sensory areas. Thus, in the adult marmoset, there is evidence that the primary auditory cortex receives inputs from a number of visual cortical areas and from the somatosensory cortex; and in the monkey, the response of neurons in the primary auditory cortex is modulated by simultaneous visual input (reviewed in Bulkin & Groh, 2006). Also in monkeys, there is evidence that neurons in the primary visual cortex with receptive fields in the periphery receive input from the primary auditory cortex (Falchier, Clavagnier, Barone, & Kennedy, 2002; see also Cappe & Barone, 2005; Rockland & Ojima, 2003).

Indirect evidence for connections between sensory cortical areas in human adults comes from studies of the visual cortex of typical sighted adults after a period of visual deprivation (blindfolding; Pascual-Leone & Hamilton, 2001). Over the course of 5 days of blindfolding, sighted individuals were taught to discriminate tactile patterns. From Day 2 onward, the visual cortex was increasingly active during these tactile tasks, and the somatosensory cortex was increasingly less active (Pascual-Leone & Hamilton, 2001). TMS over the visual cortex disrupted the tactile discrimination as much as it does in subjects blind from an early age. A similar pattern of activation over the visual cortex occurred when the blindfolded adults discriminated between auditory tones, a task unlikely to be mediated by visual imagery. A half day after the blindfold was removed—after normal visual input was restored—auditory and tactile stimulation no longer caused activation of the visual cortex, and TMS over the visual cortex no longer interfered with discrimination, even if the blindfold was temporarily restored. In a more recent study, adults were blindfolded only briefly while they performed tactile discriminations in an fMRI scanner (Merabet et al., 2007). The tactile stimuli produced significant activation in the primary visual cortex and significant deactivation in higher levels of the visual pathway, as would be expected if the activation of the primary visual cortex by touch is evident only if higher level responses are suppressed.

Collectively, these data suggest that the visual cortex of the typical adult favors visual input because it is stronger, faster, and/or more coherent and because neural responses to other inputs are normally inhibited. If the visual input is missing, the visual cortex readily responds to those other inputs. The evidence for short-term, reversible changes in cortical activation patterns after blindfolding suggests that functional connections between primary sensory cortical areas that persist into adulthood and are typically inhibited but can be disinhibited in the case of sensory deprivation and perhaps synesthesia. As would be expected, chemicals likely to modulate the level of inhibition (alcohol, caffeine) affect the reported intensity of synesthetic percepts (e.g., Ward & Simner, 2003).

Implications

The theories of synesthesia have implications for understanding the development of the typical child. By both explanatory accounts, development involves the proliferation of connections between sensory cortical areas and then the specialization of each sensory cortex for a particular sensory modality. The specialization is driven by experience, which influences which connections are pruned and shapes reentrant feedback. The inhibitory aspects of that feedback are likely to be especially slow to develop (see Burkhalter, 1993, for evidence for the visual cortex). Moreover, some of the extra, wrong connections appear to persist into adulthood as conscious percepts in synesthetes and as influences on perception in the typical adult (see next sections). These explanations imply that such cross-modal influences will be even stronger during early childhood, before pruning of many of the excess connections between sensory cortices and before the development of inhibitory reentrant feedback. They also imply that synesthesia is an exaggeration of processes common to us all (e.g., Marks, 1975, 1982; Mulvenna & Walsh, 2006; Ramachandran & Hubbard, 2001; Sagiv & Ward, 2006; Ward et al., 2006). That exaggeration makes explicit the connections in the typical adult brain that might not otherwise be suspected and the connections that are likely to have influenced the perception of the developing child.[1] These developmental processes—of experience-dependent pruning and reentrant feedback—have clear adaptive value and an unexpected side-effect: cross-modal and cross-dimensional correspondences that are not readily explained by learning.

Clues to the associations present in early development come from the consistencies in the relationship between specific inducers (e.g., the pitch of a middle C) and specific synesthetic percepts (e.g., red) among individual synesthetes and in the cross-modal associations of typical adults (e.g., common color associates to middle C). Particularly compelling are cases where there is evidence for the same connection in the percepts of synesthetes and the cross-modal associations of typical adults (e.g., is middle C red for both?). In the next section, we give three examples of hypotheses we have generated based on such consistency and have tested in typically developing toddlers. In each case, the cross-modal or cross-dimensional correspondence is not obvious: It is not amodal and not predominant in the environment. Yet the consistencies among synesthetic and nonsynesthetic adults suggest that it may arise from the intrinsic wiring of the nervous system and hence influence the perception of young children. In presenting this argument, we do not imply that adults' cross-modal perception arises only from the effects of pruning and inhibition on intrinsic connections among sensory cortical areas. Learning clearly also plays a role as the child learns face–voice associations, the color of familiar foods, the noise that trains make, etc., but the influences we describe here may also constrain that learning, facilitating the learning of some associations and interfering with the learning of others.

Using Synesthesia to Make Predictions About Toddlers' Perception

Although the specific synesthetic percepts of each synesthete are idiosyncratic, there is some consistency across individuals and with the cross-modal associations of typical adults. In this section, we illustrate three bodies of work in which we use those consistencies to derive hypotheses about the cross-modal and cross-dimensional perception of toddlers.

Pitch–Lightness

Synesthetic adults with colored hearing experience brighter percepts in response to sounds of higher pitch and darker percepts in response to sounds of lower pitch (e.g., a higher pitched C elicits a brighter red; Marks, 1974; Ward et al., 2006). Likewise, nonsynesthetic adults match tones of higher pitch to lighter colors and tones of lower pitch to darker colors (Marks, 1974; Ward et al., 2006). Moreover, the pitch of a distracting noise affects their accuracy and speed in making a two-choice luminance discrimination: They are faster and more accurate if the distracter has a higher auditory frequency when the correct answer is the lighter of the two visual stimuli (Marks, 1987). Similarly, the luminance of a distracting light affects their performance when discriminating auditory frequency: They are faster and more accurate on trials when the distracter is lighter if the correct answer is the higher auditory frequency (Marks, 1987). These consistencies between synesthetic and typical adults suggest there are natural mappings between pitch and lightness that may be present in early childhood.

Toddlers (2.5–3 years of age) demonstrate the same pitch–lightness correspondence as do adults. This was shown in a study in which toddlers observed two simultaneously bouncing balls, one light and one dark, accompanied by a lower pitched or higher pitched sound. When asked which ball was making the noise, toddlers consistently matched the lower pitched sound to the darker ball and the higher pitched sound to the lighter ball (Mondloch & Maurer, 2004). This correspondence between pitch and lightness is unlikely to arise from experience with the association in the environment, as lighter objects do not consistently make higher pitched sounds in the world (e.g., a brown mouse has a high-pitched squeak, the same as a white mouse). Thus, pitch and lightness are associated sensory dimensions that could be naturally biased by cortical connectivity between neighboring sensory cortical areas.

Toddlers, like both synesthetic and nonsynesthetic adults, also connect higher pitch to smaller balls (Mondloch & Maurer, 2004; see also Marks, 1974). This could be an additional example of a natural bias, or it might arise from experience (e.g., mice squeak but lions roar; children have higher pitched voices than adults). Alternatively, it could result from a dynamic interplay of natural biases and learning. For example, there may be an initial natural bias to associate pitch and size. This would help the developing child to understand the statistics of the environment. Learning these statistics would reinforce the strength of this association as the child gains experience in a world where smaller organisms tend to make higher pitched sounds.

Color–Letters

The second case concerns the connection between letters of the alphabet and color. Although each individual grapheme–color synesthete has a unique colored alphabet, there are some letters of the alphabet that tend to be associated with the same colors across synesthetes (e.g., ~ 40% of synesthetes say that *A* is red, perhaps because they learned early in life that "*A* is for apple" and apples are red; Day, 2004; Rich, Bradshaw, & Mattingley, 2005; Simner et al., 2005). Likewise, nonsynesthetic adults do not typically associate letters with colors, but when asked to do so, they tend to agree on the choice for some letters of the alphabet—the same ones for which synesthetes with colored graphemes show consistency (Rich et al., 2005; Simner et al., 2005). Some of the consistent letter–color associations appear to be based upon literacy: For example, English-speaking subjects commonly associate *G* with

---

[1] A red herring to understanding the development of synesthesia is a concern about whether the child was exposed early in life to the specific associations that form the synesthesia. For example, did the first alphabet book of an adult with color grapheme synesthesia who reports scarlet *C*s and yellow *D*s have that exact coloring? The only documented match is from a single synesthete whose parent had kept the refrigerator magnets from her childhood (Witthoft & Winawer, 2006). A large-scale study in Australia found no correlation between the common color–grapheme pairings in synesthetes and the books or refrigerator magnets available during their childhood (Rich, Bradshaw, & Mattingley, 2005). However, even if there were a correlation with early books or fridge magnets, this would not explain why synesthetes, but not typically developing children, retain the specific color–grapheme connection, nor how what started as an association becomes an automatically triggered extra percept.

green. However, some of the consistent color–letter associations cannot be explained easily by literacy. For example, at levels far exceeding chance, English-speaking adults, whether synesthetic or nonsynesthetic, associate *X* and *Z* with black, *O* and *I* with white, and *C* with yellow (Day, 2004; Rich et al., 2005; Simner et al., 2005). Although more cross-lingual studies are needed, an initial study of German-speaking adults suggests that at least some color–letter associations are consistent between English-speaking synesthetes and English- and German-speaking nonsynesthetic adults (Simner et al., 2005).

Because letters are stimuli that have to be learned, one might not expect preliterate children to associate letters with colors in the same manner as do adults. Although toddlers may recognize the letter *G*, they do not know that *G* is the first letter of the word *green* and hence are unlikely to associate *G* with *green* on the basis of knowledge of the written language. Consistent with this prediction, in the four cases we have tested, toddlers did not consistently associate letters to colors when there is an apparent literacy basis for their mapped colors in adults (e.g., *G* for *green*; *A* for *red*; *B* for *blue*; *Y* for *yellow*; Spector & Maurer, 2007b, 2007c, 2008a). These associations were demonstrated by typical adults and literate 7- and 9-year-old children (Spector & Maurer, 2007a, 2007b, 2008a). However, in four other cases where the adult association does not have a ready literary explanation, toddlers made the association to the same colors as are common in synesthetic percepts and typical adult cross-modal associations: They expected the *X* and *Z* to be hidden in a black box and the *I* and *O* to be hidden in a white box (Spector & Maurer, 2007b, 2007c, 2008a). Further, the consistent matching for *I*, *Z*, *X*, and *O* was based upon the shape and not the sound of the letter (Spector & Maurer, 2007d, 2008a). These results indicate that although some shape–color associations are based on learning the written language, others cannot be easily explained by learning. Rather, these associations may result from naturally biased associations between shape and color that reflect intrinsic cortical connectivity between neighboring sensory cortical areas. These connections seem to persevere into adulthood, as shown by the persistence into adulthood of the associations not readily explained by literacy (*O* and *I*, white; *X* and *Z*, black; Day, 2004; Rich et al., 2005; Simner et al., 2005). Furthermore, although intrinsic sensory cortical organization may initially bind color to shape, the development of literacy can induce additional associations, as shown by the emergence around age 7 in English-speaking children of the association of *A* to red and *G* to green, *B* to blue and *Y* to yellow—perhaps as a result of differential recruitment of higher order networks as letters take on meaning (Spector & Maurer, 2008a).

Sound–Shape

In the third case, we evaluated whether intrinsic cross-modal associations may influence the development not only of perception but also of language. The hypothesized connections were based on evidence that typical adults have biases to associate specific shapes and properties of stimuli to particular sounds (Kohler, 1929; Lindauer, 1990; Marks, 1996; Ramachandran & Hubbard, 2001). For example, sharp visual shapes go with words that produce a small, constricted movement of the tongue and mouth (e.g., *spike*, *point*). This idea is supported by an experiment in which adults were asked to make a forced choice between a rounded and a jagged shape as the referent for a nonsense word. When asked to match the nonsense words *takete* and *maluma* to the shapes, most adults answered that *takete* was the jagged shape and *maluma* was the rounded shape (Kohler, 1929; Lindauer, 1990). This effect has been replicated with modified shapes and words (e.g., *kiki* and *bouba*) in English-speaking adults and in 8- to 14-year-old children who spoke Swahili and the Bantu dialect of Kitongwe, but not English (Davis, 1961; Holland & Wertheimer, 1964; Ramachandran & Hubbard, 2001). Ramachandran and Hubbard (2001) speculate that these phenomena arise from connections between contiguous cortical areas mediating decoding of the visual percept of the nonsense shape (round or angular), the appearance of the speaker's lips (open and round or wide and narrow), and the feeling of saying the same words oneself. They argue that these connections lead to natural mappings between sound and shape that sometimes lead to synesthesia but which are present in some form in everyone.

In addition to shape correspondences to nonsense words (e.g., *kiki*), consistency in shape associations with sound can occur for real words and symbols in foreign languages. For example, Hebrew-speaking adults with no knowledge of Chinese matched Chinese characters to their corresponding Hebrew words with an accuracy above chance (Koriat & Levy, 1979). Likewise, when tested with Huambisan (a Jivaroan language from north-central Peru), naïve English-speaking adults accurately sorted words they heard in Huambisan into those naming birds and those naming fish (Berlin, 1994), perhaps because bird names tend to contain nonrounded vowels and fish names tend to contain rounded vowels. Such findings support the idea that some sound symbolism is universal across languages. Thus, we may have naturally biased associations between nonrounded vowel sounds and objects that contain angular contours (like a bird) and between rounded vowel sounds and objects that contain rounded contours (like a fish). This natural bias may have driven the evolution of the Huambisan language to designate words with nonrounded vowels to represent birds. Likewise, these associations may have influenced the evolution of our own language, such that we can easily figure out the meaning of words in an unknown language (see Day, 2004; Koriat & Levy, 1979; Nuckolls, 1999; and Tanz, 1971, for supporting cross-language evidence). These naturally biased associations may also influence the language development of an individual child and contribute to the ease with which the child learns semantic mappings. The relationship between the natural mappings and the semantics of the language will be one of mutual influence: As the child acquires the vocabulary of the language, some of the natural correspondences between shape and sound will be reinforced, and others will be altered because they are not common in the child's language (see Nuckolls, 1999; Smith & Sera, 1992).

From this perspective, we expect that language-learning children will associate nonsense words with rounded vowels with unfamiliar rounded shapes and nonsense words with unrounded vowels with unfamiliar angular shapes. A vowel is considered rounded or nonrounded based on the shape of the mouth and lips when pronouncing the phoneme represented by the vowel. For example, for the phoneme /i/, as in the word *beat*, the corners of the mouth are drawn back into a narrow elliptical shape, whereas for the phoneme /o/, as in *code*, the lips are rounded and slightly protruded (Dale, 1976). We tested for the predicted association in English-speaking toddlers to ascertain whether the influence is present early enough in development that it might indeed influence the learning of language. We played a game with English-speaking toddlers in which we contrasted four pairings of nonsense words containing rounded and nonrounded vowels and asked the child to choose which of two unfamiliar shapes, one round and one angular, the nonsense words corresponded to. The contrasting shapes were ones known to be optimal for stimulating cortical area V4v, the area active during forms of synesthesia that involve language (see above). As predicted, toddlers, like the nonsynesthetic control adults, associated the nonsense words that contained nonrounded vowels (e.g., "tee-tay," "tuh-kee-tee") with the jagged shapes and the nonsense words with rounded vowels (e.g., "go-gaa," "maa-boo-maa") with the rounded shapes (Maurer, Pathman, & Mondloch, 2006). There is also evidence to suggest that there are no natural mappings in toddlers between contrasting consonants (stop vs. approximate) and rounded versus jagged shapes (Spector & Maurer, 2008b; see also Holland & Wertheimer, 1964; Westbury, 2005).

We cannot rule out an experiential explanation for sound–shape mappings in toddlers, as toddlers have had enough experience with words and the objects that they represent to pick up statistical regularities in English semantics. It is possible that words that have nonrounded vowels tend to represent objects that are sharp and that words that have rounded vowels tend to represent objects with curved contours. However, it is also possible that this effect represents a naturally biased association between shape and the sound of the phoneme, between shape and the sight of the shape of the mouth when producing the sound, and/or between shape and the feeling (amount of oral constriction) needed to produce the same sound oneself, and that these natural associations influenced the evolution of language itself, as described above (Ramachandran & Hubbard, 2001). Within this framework, words used to represent sharper objects would develop to have phonemes that constrict the mouth, taking advantage of natural biases. Clearly we cannot differentiate natural biases from learning by studying people who have had significant experience with the statistics of language mappings, such as adults and even toddlers. Rather, we should examine the understanding of sound–shape correspondences in infants who are just beginning to learn language and hence have not developed a large enough vocabulary to notice the regularities. Similar studies across infants learning different languages would also be helpful, because languages appear to differ in the extent to which they make use of the natural correspondences (Day, 2004; Nuckolls, 1999).

The first two examples included in this section (pitch–lightness and color–letter) involve sensory associations that are present in adults and toddlers and unlikely to be based entirely on learning. Instead, they seem to be natural biases in sensory associations that could reflect cortical organization present at birth. As such, we predict that the same patterns of association will be present in infants. The origins of the vowel–shape association are less clear and need to be elucidated by studies of infants. In future studies, it would also be fruitful to test the other common forms of sound symbolism in infants and young children who are in the process of learning language.

Evidence From Adults: New Hypotheses About Development

The evidence of consistency in sensory associations between groups and across ages provides support for the idea that functional connections within and between sensory areas that are present at birth persevere to some extent into adulthood. This perseverance may be selectively exaggerated in synesthesia through the mechanisms common to us all: selective pruning and inhibition. From this perspective, consistent associations found among synesthetes or among typical adults are likely to be present in early development as influences on perception, constraints on the learning of environmentally based associations, and effects on the readiness with which new words are learned. In this section, we describe such consistent associations that have not yet been tested in children.

A Common Code for Magnitude

Some cross-modal matching seems to be the manifestation of a multisensory code for magnitude: more in one modality translates into more in another modality. For example, nonsynesthetic adults match louder tones (more sound) to larger objects (more size; Smith & Sera, 1992) and louder noises to brighter lights (more light; Marks, 1987). Similarly, synesthetic adults with colored hearing report that louder sounds induce brighter percepts (Marks, 1974).

This kind of multisensory magnitude association occurs for more complex displays as well. For example, adults' perception of note duration can be modified by visual information. This was demonstrated by an interesting study in which subjects were asked to rate the duration of a note from a marimba, a percussion instrument played by raising a mallet into the air, then rapidly lowering the hand and bouncing the mallet off a bar (Schutz & Lipscomb, 2007). Percussionists agree that the length of the gesture of mallet lowering does not affect the duration of the physical note. When given only auditory or only visual information, adults did not rate notes with the same actual duration as different when they were accompanied by long or short gestures. However, when given both auditory and visual information, adults rated the notes from a long gesture as longer in duration than those from a short gesture, despite the fact that the notes were of the same duration and the subjects had been instructed to make their ratings on the basis of auditory information only (Schutz & Lipscomb, 2007). Thus, for adults, the magnitude of visual information in the form of length of arm movement influences the magnitude of the perceived note duration.

Magnitude matching also affects the production of language. In one experiment, adults watched an animation of a dot moving left or right at different speeds and were asked to report the direction in which the dot was moving (left or right; Shintel, Nusbaum, & Okrent, 2006). Analysis of the verbal reports revealed that the mean duration of speech was shorter (i.e., words spoken faster) when describing the direction of the faster dots and longer (i.e., words spoken slower) for the slower dots. Independent listeners were accurate at guessing the speed of the dot based on the speech alone. This represents a more complex example of magnitude translation, namely, a translation involving visual speed and motor phonation. Similarly, adult speakers spontaneously increase or decrease pitch when describing motion up or down, respectively (Shintel et al., 2006). This pitch change is not related to the phonetics of the words, as demonstrated by a control condition in which participants were instructed to say "bup" or "bown."

Concurrent multisensory stimulation can also affect the perceived magnitude of a stimulus. For example, adults perceive white noise presented with light as louder than when presented alone (Odgaard, Arieh, & Marks, 2004). Similarly, adults rate near-threshold light flashes as brighter when accompanied with a burst of white noise (Stein, London, Wilkinson, & Price, 1996). In the realm of the chemical senses, adding red to a solution that smells like strawberry increases adults' perception of the intensity and pleasantness of the odor (Zellner & Kautz, 1990), just as increasing the amount of red color increases the perceived sweetness of a sucrose solution (Johnson & Clydesdale, 1982). Although some of these effects may be based on learned expectancy or modulation of attention, they occur even when the color and odor or flavor are mismatched: Red lemon smells stronger than pink lemon, which in turn smells stronger than colorless lemon (Zellner & Kautz, 1990).

The correspondence between magnitudes in different modalities could be learned from the statistics of the environment: Larger objects do tend to make louder sounds, for example. However, this cross-modal magnitude translation extends to examples not readily explained by learning (e.g., speed–phonation), and there is evidence for a form of it at birth. After habituation to a brighter or darker light, newborns respond less to an intense or soft sound, respectively, a result suggesting that they had habituated to the intensity of the light and translated it into the auditory domain (Lewkowicz & Turkewitz, 1980). However, it is possible that a cross-modal magnitude code represents another natural bias in the associations between sensory modalities without specific learning. Such a code may facilitate the learning of cross-modal correspondences that are present in the environment. From this perspective, humans are born equipped with the cortical connections necessary to make sense of correspondences they can expect to find in the world, thus leaving more energy for learning correspondences that are individually meaningful, such as Mom's voice and face. This perspective could be tested by probing connections documented in adults, but it has not yet been tested developmentally. Specifically, we predict that from birth, infants will display the same pattern of association as adults between loudness and size and lightness, spatial magnitude and note duration, and vision and tactile acuity. As soon as they learn to speak, they will also show the same connections as adults from pitch and speed of phonation to the movement of an object.

Cross-Modal Influences Not Based on Magnitude

In addition to a common intersensory code for magnitude, there are other consistent associations found in adults that provide clues about the nature of early intersensory perception. We provide examples in this section.

Influences on vision from timbre and pitch. Just as pitch and lightness are correlated sensory dimensions (see above), so are timbre (the quality of a pitch that differs when it is played, for example, by different instruments) and chroma (color saturation). Synesthetic adults with colored hearing report that the timbre of a pitch affects the saturation of the induced color, such that a note played from a piano induces a more saturated color percept than a pure tone of the same pitch, with midrange notes (e.g., C) reported as eliciting the most colorful percepts (Ward et al., 2006). Nonsynesthetic adults asked to associate pitch with color show the same relationship between timbre, pitch, and chroma, although they are less consistent in their choices from an initial to a repeat test 2–3 months later, and their color associations do not interfere with performance on behavioral tasks like the Stroop test or spatial cueing. The common patterns among synesthetes and between synesthetic and nonsynesthetic adults suggest that there are natural correspondences between timbre, pitch, and color that may be present early in development and influence the child's learning of music.

Synesthetic adults also report that higher pitched sounds induce more angular percepts than do lower pitched sounds (Marks, 1974). Likewise, nonsynesthetic adults are faster at responding to angular–rounded shapes when simultaneously presented with high-pitched–low-pitched tones, respectively (Marks, 1987). Thus, we predict that young children will show these additional cross-modal associations between pitch and shape that will influence cross-modal perception and the ease of language mapping (words with higher pitched vowels will be more easily associated to angular shapes; metaphors consistent with the mappings will be easier to learn).

Influences on the perception of visual events. One of the best examples of multisensory modification of unimodal stimuli is the illusory flash effect, in which the perception of a visual stimulus is induced by sound (Shams, Kamitani, & Shimojo, 2000). Specifically, when a single flash is presented concurrently with multiple short beeps, adults perceive the single flash as multiple flashes. Likewise, when a single flash is presented along with tactile stimulation in the form of two concurrent taps, adults report seeing two flashes (Violentyev, Shimojo, & Shams, 2005). The timing of the auditory illusory flash effect and the evoked related potentials associated with it suggest that it results from direct connections from the auditory to visual cortex, rather than feedback from the higher multimodal cortex (Mishra, Martinez, Sejnowski, & Hillyard, 2007)—the same connections that appear to be functional in early infancy and to be only partially pruned and/or inhibited during development (see above). Thus, we predict that there should be strong illusory flash effects in infants and young children that may make it difficult for them to accurately disambiguate a multimodal environment into separate unimodal events—an adaptive limitation, because input from discrete events is normally correlated across modalities.

Influences between vision and the chemical senses. Smell intensity and color lightness are correlated, as nonsynesthetic adults associate more intense (concentrated) smells with darker colors and less concentrated smells with lighter colors (Kemp & Gilbert, 1997). Nonsynesthetic adults also match certain smells to specific colors in a nonrandom fashion (Gilbert, Martin, & Kemp, 1996). For example, the smell of bergamot (bee balm) is highly associated with yellow, and caramel is highly associated with brown. Some of these examples can be easily explained by learning. For example, the smell of a lemon is likely to be associated with the color yellow. Not surprisingly, an appropriately colored solution (e.g., light yellow) can facilitate adults' identification of an odor (e.g., lemon; Zellner, Bartoli, & Eckard, 1991). However, the association between more intense smells and darker colors cannot be entirely based upon learning, because ammonia, bleach, and garlic (for example) possess intense smells but have light colors, and eggplant, tree bark, and cola (for example) have mild odors but dark colors. In addition, dark-colored paint or dark-colored animals (for example) are not likely to have more intense smells than light paint or light-colored animals. Nor is there an obvious learning explanation for some of the specific color associations (e.g., grey or black for tarragon oil, brown for patchouli oil). Some of these associations could instead be caused by perseverant cortical connections that link sensory cortical areas; in that case we predict that these associations will be present and stronger in adults with olfactory synesthesia and in young children, in whom they may influence food preferences (e.g., more apprehension about tasting novel brown foods because they are expected to have stronger flavors).

Collectively, these results show that seemingly arbitrary sensory information in one modality can facilitate adults' perception of a stimulus in another modality. Although there is no doubt that experience plays a role, this facilitation cannot be fully explained by learning, because in many of the examples the related stimuli do not typically occur together. Our framework suggests that these effects reflect direct multisensory connections that influence perception throughout development, modified as the child learns individual cross-modal associations (lemons are yellow; Mom has a distinctive face and voice) and the semantics of the particular language.

Methods for Testing the Hypotheses

In the last two sections, we suggested a number of cross-modal and cross-dimensional connections that our framework indicates should be present in early childhood. In each case, there are consistent percepts elicited by specific inducers in adults with synesthesia and/or consistent cross-modal or cross-dimensional effects in typical adults. Although some of the associations may be based on learning to recognize common links in the environment (e.g., larger organisms make lower pitched sounds), others are not readily explained by learning (e.g., lower pitched sounds come from both lighter and darker organisms). To test the predictions—and to evaluate the possible role of learning—it is necessary to test these predictions with infants who have had minimal opportunities to learn common links in the environment and with young children who have not yet learned to read or to understand metaphorical uses of language. (Note that the most common synesthetic percepts—and many of our predictions—involve vision [i.e., color, shape, and/or spatial location] and hence are not likely to have been influenced by experience in utero).

The predictions about the intersensory associations discussed above could be tested with the methodology devised by Bahrick to test infants' understanding of arbitrary cross-modal correspondences (e.g., pitch with shape or color; face with voice): Infants are habituated to two specific cross-modal pairings (the face of Woman A paired with Voice A; the face of Woman B paired with Voice B), and following habituation, are tested for recovery when the pairings are switched (e.g., Face A with Voice B; e.g., Bahrick, 1992, 1994, 2001; Bahrick, Hernandez-Reif, & Flom, 2005). Using this method, Bahrick demonstrated that young infants are sensitive to changes in the amodal property of temporal synchrony between visual and auditory impact at the youngest age tested (4 weeks): After habituating to alternating views of two different object clusters that make distinctive sounds as they hit a surface, they dishabituate when the sound and visual impact are out of sync. However, they do not dishabituate until much later (4–7 months) when arbitrary cross-modal correspondences are switched (e.g., the high-pitched impact of a yellow metal washer versus the lower pitched impact of an orange metal nut), and even then, temporal synchrony facilitates the learning of the correspondence (e.g., Bahrick, 1992, 2001; Bahrick & Pickens, 1994; Bahrick, et al., 2005; see also Flom & Bahrick, 2007). We predict that the associations outlined above will be evident with this method at 4 weeks of age, long before the emergence of sensitivity to learned arbitrary correspondences. This method appears to be more sensitive than preferential looking for revealing cross-modal understanding in young infants: When presented with two visual stimuli, only one of which matches a centrally presented sound, infants often respond randomly, despite showing evidence with the habituation switch method of cross-modal matching (e.g., Bahrick et al., 2005). In addition, some studies have found that babies look at the matching stimulus (e.g., mother's face when hearing mother's voice), whereas other studies have found that babies look at the nonmatching stimulus (e.g., stranger's face when hearing mother's voice; the moving legs from a moment before or of another baby instead of the current movements of the baby's own legs; Bahrick, & Watson, 1985; Rochat & Morgan, 1995; Schmuckler, 1996). Such inconsistent findings from the preferential looking technique make it difficult to draw conclusions about which stimuli the baby links across modalities. The habituation–switch method seems preferable.

A similar habituation–switch design could be used to test hypotheses about natural mappings between properties of objects and of words. When infants are habituated to two word–object pairings and then tested with switched pairings, 14-month-old infants show recovery, indicating that they have learned the arbitrary word–object pairings in the lab (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). When the difference between the words is acoustically minimal ("bih" vs. "dih"), infants do not learn the word–object mappings with this laboratory procedure until 17 months of age (Werker, Fennell, Corcoran, & Stager, 2002). Even with more easily distinguished words, 8- and 12-month-old infants show evidence of remembering the words and remembering the objects, but not of remembering the association between them (Werker et al., 1998). This paradigm could be used to test whether the associations predicted from this framework (e.g., "bouba" as the name of an ameboid shape; "tikey" as the name for a type of fish) are easier for infants to learn, such that they can be learned at a younger age or after shorter training during habituation, or such that they endure for a longer time after training.

It is of theoretical interest to test the predicted cross-modal and cross-dimensional associations in toddlers, rather than infants, when there is no obvious environmental explanation of how the cross-modal association would have developed postnatally. The advantage of testing toddlers over infants is that they understand simple verbal instructions and can be tested with methods that yield more easily interpreted data from more test points. The advantage of testing toddlers over testing older children or adults is that they have been minimally influenced by culturally mediated associations evident in the metaphorical use of language and do not know how to read. On the basis of our experience with toddlers (Maurer et al., 2006; Mondloch & Maurer, 2004; Spector & Maurer, 2008a), we recommend forced choice procedures embedded in a story or game in which the child guesses in which of two containers something is hidden ("I'm looking for my friend the *X*. *X*s look like this. Where do you think my friend is hiding?") or guesses which of two objects a story character is looking for ("One of these toys is called *maluba*; the other is called *takete*. Mr.

Mouse is looking for the *maluba*. Could you find it for him, please?"). In using these methods, it is important to include validity trials that follow the same structure but use real world correspondences the toddler should understand (e.g, "trees are green"); to include the data from experimental trials only if validity trials are passed; to limit the number of test trials so as to collect data only when the toddlers are fresh and attentive (no more than four, in our experience); and to set up the procedure so that the experimenter does not know the location of the expected answer and hence cannot inadvertently cue the child (the Clever Hans phenomenon).

Comparison to Other Approaches

Our framework complements that of Piaget: Like Piaget, we acknowledge a role for learning from cross-modal associations in the environment; but unlike Piaget, we also posit some intrinsic cross-modal links that do not depend on interaction with the environment. Piaget characterized the development of cross-modal perception as the integration of separate sensory systems mediated through the linking of schemas for acting on the world (Piaget, 1952). Piaget emphasized an action-centered development in which the infant learns about the sensory properties of the world by actively exploring it, and intermodal coordination cannot occur until the infant has first learned to act on the world with each sensory modality. Development proceeds by the development of schemas—repetitive patterns of actions on the environment—involving one sensory modality: a looking schema, a grasping schema, a listening schema, etc. Initially, during the first two stages of sensorimotor intelligence, these schemas are linked only by automatic reciprocal assimilation in which the assimilation of an object to one schema (listening to the sound of a rattle) engages a second schema (looking at the rattle) with which it has become linked. According to Piaget, these initial links between senses do not represent any type of cross-modal understanding (e.g, the infant does not understand that the rattle being looked at is the same as the rattle being heard) but rather the automatic linking of motor behaviors. It is only later—during Stage 4 of sensorimotor intelligence, toward the end of the first year—that the flexible linking of schemas develops that reflects intermodal understanding. Consistent with this perspective, improvements in cross-modal recognition have been documented between 6 and 12 months of age. For example, 12-month-olds visually recognized an object after a short period of haptic or oral manipulation, but 6-month-olds failed to show evidence of oral–visual transfer (Gottfried, Rose, & Bridger, 1977; see Rose & Ruff, 1987, for a review). Piaget described a similar developmental trajectory for imitation of complex actions. The ability to imitate involves cross-modal transfer from seeing or hearing someone else's actions (e.g., from vision or audition) to one's own actions (e.g., motor commands; proprioceptive feedback from one's own actions). When Piaget said "Papa" to his 9-month-old infant son, his son responded "Papa" or "baba" (Piaget, 1952). When Piaget said "Papapapapapa," his son responded "papapapapapa." Beyond the association of the sound of a voice to the sight of a mouth movement, Piaget noted that the infant had learned to link those stimuli in a meaningful cross-sensory way to his own schema of phonation. Toward the end of the first year of life, there are similar improvements in the imitation of complex motor actions (e.g., tapping a table; Jones, 2007).

The framework presented here also complements that of Gibson; like Gibson, we posit initial cross-modal associations that do not depend on learning, but unlike Gibson, we posit that those links extend beyond amodal perception. According to the Gibsonian perspective, young infants are sensitive to sensory information that is invariant across modalities, or amodal (i.e., not specific to any one modality; Gibson, 1966). Within this framework, infants are born with the capacity to perceive amodal events, which helps to unite the multisensory aspects of the world through intersensory redundancy (Bahrick & Lickliter, 2004). For example, infants as young as 3 weeks of age appear to understand the temporal synchrony between the sight of an object impacting a surface and the sound of the impact (Bahrick & Lickliter, 2000). The sound of an object's impact is inherent in the sight of the impact and is therefore not specific to the visual or auditory modality, but is instead amodal.

Of course, arbitrary intermodal associations need to be learned. For example, a baby is not born knowing the link between her mother's voice and face but must learn this association through experience. Although Mom's voice is inherently a part of her, the relationship between the sound of her voice and her face is not amodal. Other moms have different voices and faces, and their infants need to learn to link the two in a meaningful way. Similarly, infants are not born knowing the link between certain shapes and certain sounds (e.g., *A* sounds like "ay"), but instead must learn this arbitrary sound–shape correspondence. On the basis of empirical work, Gibsonians argue that infants do not have the cognitive capacity to learn arbitrary associations until 4–7 months of age (Bahrick et al., 2005; Bahrick & Pickens, 1994).

Both the Piagetian and Gibsonian perspectives have inspired studies of cross-modal perception during infancy that have yielded data consistent with their theoretical perspective: the apparent learning of cross-modal connections involving the linking of schemas during the second half of the first year of life, as predicted by Piaget; the early apparent understanding of at least some amodal correspondences and the later acquisition of arbitrary correspondences, as predicted by the Gibsonian view. We do not see our framework as a replacement for either of these theoretical approaches, because they were intended to account for far more than the development of cross-modal perception and there is considerable evidence for the developmental mechanisms they described. Instead, we see our framework as describing an additional developmental mechanism that provides new insights into the development of perception and language. The framework presented here differs from both perspectives in predicting that there are additional, natural correspondences present early in life because of the intrinsic wiring of the nervous system that are not amodal properties of the environment and that are not learned from regularities in the environment. In this article we have illustrated unique predictions that can be derived from a framework rooted in an understanding of synesthesia.

Although our framework generates novel predictions, it obviously does not constitute a complete description of the development of cross-modal perception. Arbitrary cross-modal correspondences present in the environment must be learned as well—but even that learning may be influenced by the presence of cross-sensory connections that facilitate the learning of some arbitrary links: It may be easier for the newborn to learn to link the mother with an unusually large face to her unusually low-pitched voice than to an unusually high-pitched voice. Newborns' understanding of amodal correspondences, as posited by the Gibsonian framework, may be mediated in some cases by the mechanism posited here—inherent links between sensory systems that remain throughout life because they are reinforced by environmental feedback that prevents them from being pruned or inhibited.

Unlike the Gibsonian and Piagetian perspectives, the framework presented here can account for inverted-U-shaped developmental curves in which a particular cross-modal skill is present at birth or early in infancy, then disappears and returns later (for an earlier discussion of this, see Maurer, 1993; Maurer & Maurer, 1988; and Mondloch & Maurer, 2004). A good example is imitation. Newborn babies appear to imitate tongue protrusion: When an adult model sticks out his tongue repeatedly, the baby's tongue is likely to begin to protrude (Meltzoff & Moore, 1977). Yet such imitation declines steadily over the next 6 months, and all evidence of the skill has vanished by 6 months of age (Abravanel & Sigafoos, 1984; Fontaine, 1984; Heimann, Nelson, & Schaller, 1989; Jones, 2007). In fact, toward the end of the first year of life, infants seem to need to learn to imitate tongue protrusion, experimenting with the connection between the sight of someone else's face and the feelings on their own face (Piaget, 1952). Further, there are different time courses for the (re)emergence of different imitative skills. For example, infants imitate the behavior of an adult who is tapping a table or saying "Aah" at 8–10 months, but do not imitate putting a hand on the head or tongue protrusion until 16–18 months (Jones, 2007). If intersensory understanding does not occur until differentiated schemata are linked together in flexible recombinations, as Piaget proposed, then (a) intersensory imitation should not occur at birth; (b) it should not disappear and reappear; and (c) all imitative behaviors involving the same schemas should emerge around the same time. Likewise, according to the Gibsonian perspective, sensory information is (a) amodal and simple enough to be processed at birth; (b) amodal but more complex (e.g., embedded amodal relations such as the sound of many objects versus one object striking a surface), the processing of which emerges postnally; or (c) arbitrary, the processing of which can be learned from about 4 to 7 months of age. This perspective leaves no room for behaviors such as imitation that do not fall easily into either the present-at-birth or learned-later categories. In contrast, the perspective developed here offers a ready explanation of the U-shaped development: The early imitation may be based on direct connections between motor, proprioceptive, and visual sensory areas that are triggered automatically by stimulation of any of the three senses. As the sensory systems differentiate during development and some of those direct connections are pruned, the early imitation disappears, to reemerge later as the baby learns to connect more differentiated sensory systems. As expected, the early imitation is not precise: The model can be a looming pen instead of a tongue (Jacobsen, 1979), and the imitation can occur as lunging fingers rather than a protruding tongue (Gardner & Gardner, 1970). Nevertheless, remnants of the early connections may underlie the influence of the properties of moving objects on the speed and pitch of speech (see above). A similar explanation applies to the documented inverted-U developmental curve for cross-modal transfer between tactile exploration and visual recognition (Streri, 1987; Streri & Pecheux, 1986) and between the sound of a passage being read and the accompanying lip movements (Pickens et al., 1994).

Conclusion

Synesthesia is a phenomenon of intersensory and intrasensory linkage. Its etiology seems to lie in the balance of synaptic pruning, neuronal inhibition, and the endurance of connections between sensory cortical areas characteristic of early childhood. Moreover, synesthesia appears to be a magnification of normative sensory processes. Although nonsynesthetic adults do not experience cross-modally induced sensory percepts, they do associate intersensory information nonrandomly, as synesthetes do. In all cases that have been tested to date, sensory associations are consistent between synesthetic adults, nonsynesthetic adults, and very young children. For example, all three groups associate pitch with lightness (Marks, 1974; Mondloch & Maurer, 2004; Ward et al., 2006) and some letters with colors (Day, 2004; Rich et al., 2005; Simner et al., 2005; Spector & Maurer, 2008a). Such highly consistent sensory associations cannot be fully explained by learning, as the environment does not always follow the same patterns (e.g., lightly colored animals do not make higher pitched sounds, and *O*s are not routinely printed in white ink). Rather, some of these consistencies in sensory associations may reflect perseverant cortical connections, which are present at birth and influence a child's perceptual and language development. Preexistent connections cannot explain all intersensory correspondences, as many associations must be learned from the environment. However, the naturally biased associations that emerge from typical sensory cortical organization may constrain the development and expression of learned sensory associations. Thus, consistent with modern neuroconstructivist theories (e.g., Westermann et al., 2007), the development of cross-modal and intramodal associations is related to a dynamic interplay between learned and naturally biased constraints on the development of neural structures. This framework leads to novel predictions about the cross-modal and cross-dimensional associations likely to be present in infancy and early childhood.

The evidence summarized in this review suggests that there is a systematic way in which sensory information translates across modalities throughout development that influences what we pick up from the environment. Synesthetes may have conscious access to some of this translation process, the origins of which lie in the initial organization of the sensory neural system. Thus, synesthesia is far more than a quirky phenomenon: It is a window into the very nature of sensory processing and development.

References

Abravanel, E., & Sigafoos, A. D. (1984). Exploring the presence of imitation during early infancy. *Child Development, 55,* 381–392.

Aleman, A., Rutten, G. J., Sitskoorn, M. M., Dautzenberg, G., & Ramsey, N. F. (2001). Activation of striate cortex in the absence of visual stimulation: An fMRI study of synesthesia. *Neuroreport, 12,* 2827–2830.

Amedi, A., Merabet, L. B., Bermpohl, F., & Pascual-Leone, A. (2005). The occipital cortex in the blind. *Psychological Science, 14,* 306–311.

Asher, J. E., Aitken, M. R. F., Farooqi, N., Kurmani, S., & Baron-Cohen, S. (2006). Diagnosing and phenotyping visual synesthesia: A preliminary evaluation of the revised test of genuineness (TOG-R). *Cortex, 42,* 137–146.

Bahrick, L. E. (1992). Infants' perceptual differentiation of amodal and modality-specific audio-visual relations. *Journal of Experimental Child Psychology, 53,* 180–199.

Bahrick, L. E. (1994). The development of infants' sensitivity to arbitrary intermodal relations. *Ecological Psychology, 6,* 111–123.

Bahrick, L. E. (2001). Increasing specificity in perceptual development: Infants' detection of nested levels of multimodal stimulation. *Journal of Experimental Child Psychology, 79,* 253–270.

Bahrick, L. E., Hernandez-Reif, M., & Flom, R. (2005). The development of infant learning about specific face–voice relations. *Developmental Psychology, 41,* 541–552.

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36,* 190–201.

Bahrick, L. E., & Lickliter, R. (2004). Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: A developmental test of the intersensory redundancy hypothesis. *Cognitive, Affective and Behavioral Neuroscience, 4,* 137–147.

Bahrick, L. E., & Pickens, J. N. (1994). Amodal relations: The basis for intermodal perception and learning in infancy. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception* (pp. 205–235). Hillsdale, NJ: Erlbaum.

Bahrick, L. E., & Watson, J. S. (1985). Detection of intermodal proprioceptive–visual contingency as a potential basis of self-perception in infancy. *Developmental Psychology, 21,* 963–973.

Baron-Cohen, S., Harrison, J., Goldstein, L. H., & Wyke, M. (1993). Colored speech perception: Is synesthesia what happens when modularity breaks down? *Perception, 22,* 419–426.

Berlin, B. (1994). Evidence for pervasive synesthetic sound symbolism in ethnozoological nomenclature. In L. Hinton, J. Nichols, and J. Ohala (Eds.), *Sound symbolism* (pp. 76–93). New York: Cambridge University Press.

Bulkin, D. A., & Groh, J. M. (2006). Seeing sounds: Visual and auditory interactions in the brain. *Current Opinion in Neurobiology, 16,* 415–419.

Burkhalter, A. (1993). Development of forward and feedback connections between areas V1 and V2 of human visual cortex. *Cerebral Cortex, 3,* 476–487.

Burton, H., McLaren, D. G., & Sinclair, R. J. (2006). Reading embossed capital letters: An fMRI study in blind and sighted individuals. *Human Brain Mapping, 4,* 325–339.

Burton, H., Sinclair, R. J., & McLaren, D. G. (2004). Cortical activity to vibrotactile stimulation: An fMRI study in blind and sighted individuals. *Human Brain Mapping, 23,* 210–228.

Burton, H., Snyder, A.Z., Diamond, J.B. & Raichle, M.E. (2002). Adaptive changes in early and late blind: A fMRI study of verb generation to heard nouns. *Journal of Neurophysiology, 88,* 3359–3371.

Cappe, C., & Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience, 22,* 2886–2902.

Cohen, L. G., Celnik, P., Pascual-Leone, A., Corwell, B., Faiz, L., Dambrosia, J., et al. (1997, September 11). Functional relevance of cross-modal plasticity in blind humans. *Nature, 389,* 180–183.

Cohen, L. G., Weeks, R., Sadato, N., Celnik, P., Ishii, K., & Hallett, M. (1999). Period of susceptibility for cross-modal plasticity in the blind. *Annals of Neurology, 45,* 451–460.

Dale, P. S. (1976). *Language development* (2nd ed.). New York: Holt Rinehart & Winston.

Davis, R. (1961). The fitness of names to drawings: A crosscultural study in Tanganyika. *British Journal of Psychology, 52,* 259–268.

Day, S. (2004). Trends in synesthetically colored graphemes and phonemes—2004 revision. Available from http://home.comcast.net/~sean.day/Trends2004.htm

Day, S. (2007). Types of synesthesia. In *Synesthesia*. Retrieved July 28, 2008, from http://home.comcast.net/~sean.day/html/types.htm

DeCasper, A. J., & Fifer, W. P. (1980, June 6). Of human bonding: Humans prefer their mothers' voices. *Science, 208,* 1174–1176.

Dehay, C., Bullier, J., & Kennedy, H. (1984). Transient projections from the fronto-parietal and temporal cortex to areas 17, 18, and 19 in the kitten. *Experimental Brain Research, 57,* 208–212.

Dehay, C., Kennedy, H., & Bullier, J. (1988). Characterization of transient cortical projections from auditory, somatosensory and motor cortices to visual areas 17, 18, and 19 in the kitten. *Journal of Comparative Neurology, 230,* 576–592.

de Thornley Head, P. (2006). Synesthesia: Pitch-color isomorphism in RGB-space? *Cortex, 42,* 164–174.

Dixon, M. J., Smilek, D., Cudahy, C., & Merikle, P. M. (2000, July 27). Five plus two equals yellow. *Nature, 406,* 365.

Dixon, M. J., Smilek, D., & Merikle, P. (2004). Not all synesthetes are created equal: Projector versus associator synesthetes. *Cognitive, Affective & Behavioral Neuroscience, 4,* 335–343.

Eagleman, D. M., Kagan, A. D., Nelson, S. S., Sagaram, D., & Sarma, A. K. (2007). A standardized test battery for the study of synesthesia. *Journal of Neuroscience Methods, 159,* 130–145.

Edquist, J., Rich, A. N., Brinkman, C., & Mattingley, J. B. (2006). Do synesthetic colors act as unique features in visual search? *Cortex, 42,* 222–231.

Esterman, M., Verstynen, T., Ivry, R. B., & Robertson, L. C. (2006). Coming unbound: Disrupting automatic integration of synesthetic color and graphemes by transcranial magnetic stimulation of the right parietal lobe. *Journal of Cognitive Neuroscience, 18,* 1570–1576.

Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience, 22,* 5749–5759.

Flom, R., & Bahrick, L.E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental Psychology, 43,* 238–252.

Fontaine, R. (1984). Imitative skills between birth and six months. *Infant Behavior and Development, 7,* 323–333.

Gardner, J., & Gardner, H. (1970). A note on selective imitation by a six-week-old infant. *Child Development, 41,* 1209–1213.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gilbert, A. N., Martin, R., & Kemp, S. E. (1996). Cross-modal correspondence between vision and olfaction: The color of smells. *American Journal of Psychology, 109,* 335–351.

Gottfried, A. W., Rose, S. A., & Bridger, W. H. (1977). Cross-modal transfer in human infants. *Child Development, 48,* 118–123.

Gray, J. A., Parslow, D. M., Brammer, M. J., Chopping, S., Vythelingum, G. N., & ffytche, D. H. (2006). Evidence against functionalism from neuroimaging of the alien color effect in synesthesia. *Cortex, 42,* 309–318.

Grossenbacher, P., & Lovelace, G. (2001). Mechanisms of synesthesia: Cognitive and physiological constraints. *Trends in Cognitive Sciences, 5,* 36–41.

Hamilton, R., Keenan, J., Catala, M., & Pascual-Leone, A. (2000). Alexia for braille following bilateral occipital stroke in an early blind woman. *Neuroreport, 11,* 237–240.

Heimann, M., Nelson, K., & Schaller, J. (1989). Neonatal imitation of tongue protrusion and mouth opening: Methodological aspects and evidence of early individual differences. *Scandinavian Journal of Psychology, 30,* 90–101.

Holland, M., & Wertheimer, M. (1964). Some physiognomic aspects of naming, or maluma and takete revisited. *Perceptual and Motor Skills, 19,* 111–117.

Hubbard, E. M., Arman, A. C., Ramachandran, V. S., & Boynton, G. M. (2005). Individual differences among grapheme–color synesthetes: Brain–behavior correlations. *Neuron, 45,* 975–985.

Hubbard, E. M., Manohar, S., & Ramachandran, V. S. (2006). Contrast affects the strength of synesthetic colors. *Cortex, 42,* 184–194.

Huttenlocher, P. R. (1984). Synapse elimination and plasticity in developing human cerebral cortex. *American Journal of Mental Deficiency, 88,* 488–496.

Huttenlocher, P. R. (1994). Synaptogenesis in human cerebral cortex. In G. Dawson & K. Fischer (Eds.), *Human behavior and the developing brain* (pp. 137–152). New York: Guilford.

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neuroscience, 387,* 167–178.

Huttenlocher, P. R., & de Courten, C. (1987). The development of synapses in striate cortex of man. *Human Neurobiology, 6,* 1–9.

Huttenlocher, P. R., de Courten, C., Garey, L. J., & Van der Loos, H. (1982). Synaptogenesis in human visual cortex: Evidence for synapse elimination during normal development. *Neuroscience Letters, 13,* 247–252.

Jacobsen, S. W. (1979). Matching behavior in the young infant. *Child Development, 50,* 425–430.

Johnson, J., & Clydesdale, F. M. (1982). Perceived sweetness and redness in colored sucrose solutions. *Journal of Food Science, 47,* 87–103.

Jones, S. S. (2007). Imitation in infancy. *Psychological Science, 18,* 593–599.

Kemp, S. E., & Gilbert, A. N. (1997). Odor intensity and lightness are correlated sensory dimensions. *American Journal of Psychology, 110,* 35–46.

Kennedy, H., Batardiere, A., Dehay, C., & Barone, P. (1997). Synesthesia: Implications for developmental neurobiology. In S. Baron-Cohen & J. Harrison (Eds.), *Synaesthesia: Classic and contemporary readings* (pp. 243–256). Oxford, England: Wiley-Blackwell.

Kim, C.-Y., Blake, R., & Palmieri, T. J. (2006). Perceptual interaction between real and synesthetic colors. *Cortex, 42,* 195–203.

Kohler, W. (1929). *Gestalt psychology.* Liveright, New York.

Koriat, A., & Levy, I. (1979). Figural symbolism in Chinese ideographs. *Journal of Psycholinguistic Research, 8,* 353–365.

Laeng, B., Svartdal, F., & Oelmann, H. (2004). Does color synesthesia pose a paradox for early-selection theories of attention? *Psychological Science, 15,* 277–281.

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

Lindauer, M. (1990). The effects of the physiognomic stimuli *takete* and *maluma* on the meanings of neutral stimuli. *Bulletin of the Psychonomic Society, 28,* 151–154.

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *American Journal of Psychology, 87,* 173–188.

Marks, L. E. (1975). On colored–hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin, 82,* 303–331.

Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 177–193.

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.

Marks, L. E. (1996). On perceptual metaphors. *Metaphor and Symbolic Activity, 11,* 39–66.

Mattingley, J. B., Payne, J. M., & Rich, A. N. (2006). Attentional load attenuates synesthetic priming effects in grapheme–color synesthesia. *Cortex, 42,* 213–221.

Mattingley, J. B., Rich, A. N., Yelland, G., & Bradshaw, J. L. (2001, March 29). Unconscious priming eliminates automatic binding of color and alphanumeric form in synesthesia. *Nature, 410,* 580–582.

Maurer, D. (1993). Neonatal synaesthesia: Implications for the processing of speech and faces. In B. Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, and J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 109–124). Dordrecht, The Netherlands: Kluwer.

Maurer, D., Lewis, T. L., & Mondloch, C. J. (2005). Missing sights: Consequences for visual cognitive development. *Trends in Cognitive Sciences, 9,* 144–151.

Maurer, D., & Maurer, C. (1988). *The world of the newborn.* New York: Basic Books.

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

Meltzoff, A. N., & Moore, M. K. (1977, October 7). Imitation of facial and manual gestures by human neonates. *Science, 198,* 75–78.

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,* 1633–1641.

Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience, 27,* 4120–4131.

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

Muggleton, N., Tsakanikos, E., Walsh, V., & Ward, J. (2007). Disruption of synesthesia following TMS of the right posterior parietal cortex. *Neuropsychologia, 45,* 1582–1585.

Mulvenna, C. M., & Walsh, V. (2006). Synesthesia: Supernormal integration? *Trends in Cognitive Sciences, 10,* 350–352.

Neville, H. (1995). Developmental specificity in neurocognitive development in humans. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 219–231). Cambridge, MA: Bradford.

Nikolié, D., Lichti, P., & Singer, W. (2007). Color opponency in synesthetic experiences. *Psychological Science, 18,* 481–486.

Nuckolls, J. (1999). The case for sound symbolism. *Annual Reviews of Anthropology, 28,* 225–252.

Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., et al. (2002). Functional magnetic resonance imaging of synesthesia: Activation of V4/V8 by spoken words. *Nature Neuroscience, 5,* 371–375.

Ockleford, E. M., Vince, M. A., Layton, C., & Reader, M. R. (1988). Responses of neonates to parents' and others' voices. *Early Human Development, 18,* 27–36.

Odgaard, E. C., Arieh, Y., & Marks, L. E. (2004). Brighter noise: Sensory enhancement of perceived loudness by concurrent visual stimulation. *Cognitive, Affective, and Behavioral Neuroscience, 4,* 127–132.

Palmieri, T. J., Blake, R., Marois, R., Flanery, M. A., & Whetsell, W. (2002). The perceptual reality of synesthetic colors. *Proceedings of the National Academy of Sciences, USA, 99,* 4127–4131.

Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. *Progress in Brain Research, 134,* 427–445.

Paulesu, E., Harrison, J., Baron-Cohen, S., Watson, J. D. G., Goldstein, L., Heather, J., et al. (1995). The physiology of colored-hearing: A PET activation study of color-word synesthesia. *Brain, 118,* 661–676.

Paulsen, H. G., & Laeng, B. (2006). Pupillometry of grapheme-color synesthesia. *Cortex, 42,* 290–294.

Piaget, J. (1952). *Origins of intelligence in children*. New York: International Universities Press.

Pickens, J., Field, T., Nawrocki, T., Martinez, A., Soutullo, D., & Gonzalez, J. (1994). Full-term and pre-term infants' perception of face–voice synchrony. *Infant Behavior and Development, 17,* 447–455.

Ramachandran, V. S., & Azoulai, S. (2006). Synesthetically induced colors evoke apparent motion perception. *Perception, 35,* 1557–1560.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synethesia: A window into perception, thought, and language. *Journal of Consciousness Studies, 12,* 3–34.

Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale systematic study of synesthesia: Implications for the role of early experience in lexical-color associations. *Cognition, 98,* 53–84.

Rich, A., & Mattingley, J. (2003). The effects of stimulus competition and voluntary attention on color–graphemic synesthesia. *NeuroReport, 14,* 1793–1798.

Rich, A. N., Williams, M. A., Puce, A., Syngeniotis, A., Howard, M. A., McGlone, F., & Mattingley, J. B. (2006). Neural correlates of imagined and synesthetic colors. *Neuropsychologia, 44,* 2918–2925.

Rochat, P., & Morgan, R. (1995). Spatial determinants in the perception of self-produced leg movements by 3- to 5-month-old infants. *Developmental Psychology, 31,* 626–636.

Rockland, K. S., & Ojima, H. (2003). Multisensory interactions in calcarine visual areas in macaque monkey. *Journal of Psychophysiology, 50,* 19–26.

Röder, B., Rösler, F., Hennighausen, E., & Nacker, F. (1996). Event-related potentials during auditory and somatosensory discrimination. *Cognitive Brain Research, 4,* 77–93.

Röder, B., Stock, O., Bien, S., Neville, H., & Rösler, F. (2002). Speech processing activates visual cortex in congenitally blind humans. *European Journal of Neuroscience, 16,* 930–936.

Rose, S. A., & Ruff, H. A. (1987). Cross-modal abilities in human infants. In J. Osofky (Ed.), *Handbook of infant development* (2nd ed.; pp. 316–362). New York: Wiley.

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme–color synesthesia. *Nature Neuroscience, 10,* 792–797.

Sadato, N., Okado, T., Honda, M., & Yonekura, Y. (2002). Critical period for cross-modal plasticity in blind humans: A functional MRI study. *NeuroImage, 16,* 389–400.

Sadato, N., Pascual-Leone, A., Grafman, J., Deiber, M.-P., Ibañez, V., & Hallett, M. (1998). Neural networks for Braille reading by the blind. *Brain, 121,* 1213–1229.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M. P., Dold, G., & Hallet, M. (1996, April 11). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature, 380,* 526–528.

Sagiv, N., Heer, J., & Robertson, L. (2006). Does binding of synesthetic color to the evoking grapheme require attention? *Cortex, 42,* 232–242.

Sagiv, N. & Ward, J. (2006). Crossmodal interactions: Lessons from synesthesia. *Progressive Brain Research, 155,* 259–271.

Schmuckler, M. A. (1996). Visual–proprioceptive intermodal perception in infancy. *Infant Behavior and Development, 19,* 221–232.

Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived note duration. *Perception, 36,* 888–897.

Shams, L., Kamitani, Y., & Shimojo, S. (2000, December 14). Illusions: What you see is what you hear. *Nature, 408,* 788.

Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language, 55,* 167–177.

Simner, J., Sagiv, N., Mulvenna, C., Tsakanikos, E., Witherby, S., Fraser, C., et al. (2006). Synesthesia: The prevalence of atypical cross-modal experiences. *Perception, 35,* 1024–1033.

Simner, J., Ward, J., Lanz, M., Jansari, A., Noonan, K., Glover, L., & Oakley, D. (2005). Non-random associations of graphemes to colors in the synaesthetic and non-synaesthetic populations. *Cognitive Neuropsychology, 22,* 1–17.

Smilek, D., Callejas, A., Dixon, M. J., & Merikle, P. M. (2007). Ovals of time: Time–space associations in synesthesia. *Consciousness and Cognition, 26,* 507–519.

Smilek, D., Dixon, M. J., & Merikle, P. M. (2003). Synesthetic photisms guide attention. *Brain and Cognition, 53,* 364–367.

Smilek, D., Dixon, M. J., Cudahy, C., & Merikle, P. M. (2001). Synesthetic photisms influence visual perception. *Journal of Cognitive Neuroscience, 13,* 930–936.

Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology, 24,* 99–142.

Spector, F., & Maurer, D. (2007a). [Color associations to the letters *B* and *Y* in literate children and adults]. Unpublished raw data.

Spector, F., & Maurer, D. (2007b, January). *The color of O's: Naturally-biased shape–color associations.* Paper presented at the meeting of the American Synesthesia Association, St. Petersburg, FL.

Spector, F., & Maurer, D. (2007c, April). *The color of O's: Naturally-biased shape–color associations.* Poster presentation at the meeting of the Society for Research in Child Development, Boston, MA.

Spector, F. & Maurer, D (2007d). [Influences on color mapping to the letters *Z* and *I* in toddlers]. Unpublished raw data.

Spector, F., & Maurer, D. (2008a). The color of *O*s: Naturally biased associations between shape and color. *Perception, 37,* 841–847.

Spector, F., & Maurer, D. (2008b). [Influence of consonant sound on shape mapping in toddlers]. Unpublished raw data.

Sperling, J. M., Prvulovic, D., Linden, D. E. J., Singer, W., & Stirn, A. (2006). Neuronal correlates of color–grapheme synesthesia: A fMRI study. *Cortex, 42,* 295–303.

Stein, B. E., London, N., Wilkinson, L. K., & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *Journal of Cognitive Neuroscience, 8,* 497–506.

Stevens, M. S., Hansen, P. C., & Blakemore, C. (2006). Activation of color-selective areas of the visual cortex in a blind synesthete. *Cortex, 42,* 304–308.

Streri, A. (1987). Tactile discrimination of shape and intermodal transfer in 2- to 3-month-old infants. *British Journal of Developmental Psychology, 5,* 213–220.

Streri, A., & Pecheux, M.-G. (1986). Tactual habituation and discrimination of form in infancy: A comparison with vision. *Child Development, 57,* 100–104.

Tanz, C. (1971). Sound symbolism in words relating to proximity and distance. *Language and Speech, 14,* 266–276.

Tzourio-Mazoyer, N., de Schonen, S., Crivello, F., Reutter, B., Aujard, Y., & Mazoyer, B. (2002). Neural correlates of woman face processing by 2-month-old infants. *NeuroImage, 15,* 454–461.

Violentyev, A., Shimojo, S., & Shams, L. (2005). Touch-induced visual illusion. *NeuroReport, 16,* 1107–1110.

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound–color synesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex, 42,* 264–280.

Ward, J., & Sagiv, N. (2007). Finger counting and dice patterns: A case of higher synesthesia? *Neurocase, 13,* 86–93.

Ward, J., & Simner, J. (2003). Lexical–gustatory synesthesia: Linguistic and conceptual factors. *Cognition, 89,* 237–261.

Ward, J., & Simner, J. (2005). Is synesthesia an X-linked dominant trait with lethality in males? *Perception, 34,* 611–623.

Weiss, P. H., Zilles, K., & Fink, G. R. (2005). When visual perception causes feeling: Enhanced cross-modal processing in grapheme–color synesthesia. *NeuroImage, 28,* 859–868.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology, 34,* 1289–1309.

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy, 3,* 1–30.

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language, 93,* 10–19.

Westermann, G., Mareschal, D., Johnson, M. H., Sirois, S., Spratling,

W. W., & Thomas, M. S. C. (2007). Neuroconstructivism. *Developmental Science, 10,* 75–83.

Winawer, J., & Witthoft, N. (2004, April). *Anticolors: Early visual mechanisms in color–grapheme synesthesia.* Poster presented at the meeting of the Cognitive Neuroscience Society, San Francisco.

Witthoft, N., & Winawer, J. (2006). Synesthetic colors determined by having colored refrigerator magnets in childhood. *Cortex, 24,* 175–183.

Wolff, P., Matsumiya, Y., Abrohms, I. F., van Velzer, C., & Lombroso, C. T. (1974). The effect of white noise on the somatosensory evoked responses in sleeping newborn infants. *Electroencephalography and Clinical Neurophysiology, 37,* 269–274.

Yaka, R., Yinon, U., & Wollberg, Z. (1999). Auditory activation of cortical visual areas in cats after early visual deprivation. *European Journal of Neuroscience, 11,* 1301–1312.

Zellner, D. A., Bartoli, A. M., & Eckard, R. (1991). Influence of color on odor identification and liking ratings. *American Journal of Psychology, 104,* 547–561.

Zellner, D. A., & Kautz, M. A. (1990). Color affects perceived odor intensity. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 391–397.

Received October 15, 2007
Revision received July 30, 2008
Accepted August 20, 2008 ■