



Old and new versions of the Molyneux question: A review of experimental answers

Charles Spence^a  (charles.spence@psy.ox.ac.uk)

Nicola Di Stefano^b  (nicola.distefano@istc.cnr.it)

Abstract

The ‘Molyneux problem’ is typically framed in terms of the crossmodal matching of shape information from touch to vision. Indeed, shape along with intensity have commonly been considered amodal stimulus properties/dimensions (at least by developmental researchers). However, it is important to note that what is common, if anything, to the senses differs in the two cases: It is the physical stimulus (and possibly also the associated phenomenology) that is thought to be the same in the case of crossmodal (or intermodal) shape matching between touch and vision, whereas it is the nature of the underlying neural encoding that is said to be similar in the case of crossmodal matching of auditory and visual stimulus intensity. While the first empirical data to have been published on these two forms of putatively amodal crossmodal matching appeared to suggest that they both emerge surprisingly early in the course of human development (i.e., within the first month of life), certain of these seminal findings have proved difficult to replicate. Ultimately, therefore, there is currently little convincing evidence to support the notion that such putatively innate crossmodal matching of amodal stimulus dimensions is actually different in kind from the various other crossmodal correspondences that are seemingly acquired at various points during the course of human development (typically as a result of the internalization of the crossmodal statistics of the environment). As such, there may be nothing particularly special about the type of crossmodal matching thought to underlie the ‘Molyneux problem’, and continued interest in the issue may inadvertently have helped to sustain a misguided account of the differences between different types of crossmodal correspondence.

Keywords

Amodal · Crossmodal correspondences · Crossmodal matching · Development · Innate · Intensity · Intermodal matching · Molyneux problem · Shape

This article is part of a special issue on “Molyneux’s question today”, edited by Gabriele Ferretti and Brian Glenney.

^a University of Oxford.

^b Institute of Cognitive Sciences and Technologies, National Research Council of Italy (CNR), Rome, Italy.

1 Introduction

Based on a letter from Diderot (1777), the ‘Molyneux problem’ (Eilan, 1993; Locke, 1975; Meltzoff, 1993; Morgan, 1977) raises the intriguing question of whether a person born blind would immediately recognize the shape of a sphere/cube by sight were their sight suddenly to be restored (presuming that these shapes had been experienced tactually prior to recovering sight). The question has long remained a lively theoretical debate amongst philosophers (e.g., Campbell, 1996; Evans, 1985; Green, 2021, 2022; Levin, 2008, 2018; Van Cleve, 2007), with early commentators putting forward opposing answers.¹ The philosopher John Locke, as well as Molyneux himself, were both convinced that the question had to be answered negatively (Bruno & Mandelbaum, 2010); in contrast, Leibniz replied in the affirmative, although the reasons behind the latter’s position are not entirely clear. The disagreement has stimulated further discussion amongst a wide range of readers and commentators (Ehli, 2020; Glenney, 2012). One possible explanation here involves considering shape representations as purely formal and modality-independent schema that isomorphically map onto the shapes that are perceived, thus providing the tactile and visual representations with a common underlying structure and the perceiver with a common feature with which to recognize the shapes (Glenney, 2012). Alternatively, however, McRae’s (1976) analysis sets Leibniz’s answer to Molyneux’s question in terms of the possibility of translating between different expressions that target the same perceptual object from different sense modalities (see also Spence & Di Stefano, 2023). Finally, one could put forward the existence of an Aristotelian common sense (Owens, 1982), as Leibniz himself appears to suggest,² thus allowing perceivers to process the information about a limited set of sensory qualities, including shape, through their different senses (see Spence & Di Stefano, 2024).

The most rigorous empirical research to have been published in recent years suggests that shape information is not immediately given by the perceptual array when those individuals who have been blind since birth (namely, congenital cataract patients) have had their cataracts surgically removed (Held et al., 2011). While it is certainly true that matching appears to improve markedly over a matter of days, such observations nevertheless argue against crossmodal matching on the basis of phenomenal quality, at least in the case of shape perception (cf. Campbell, 1996), on differences in the phenomenal quality of the visual and tactile experience of shape). Indeed, such findings can be taken as arguing against the suggestion that shape information is one of the objective properties that is coded amodally (Spence & Di Stefano, 2024). Note here also that the senses of touch and vision only actually provide redundant information about the shape and size of

¹ With some thinkers even changing their mind, such as Condillac, who first gave a positive answer and then opted for a negative one, years later (Anstey, 2023).

² “Ideas which are said to come from more than one sense—such as those of space, figure, motion, rest—come rather from the common sense, that is, from the mind itself” (Leibniz, 1981, p. 135).

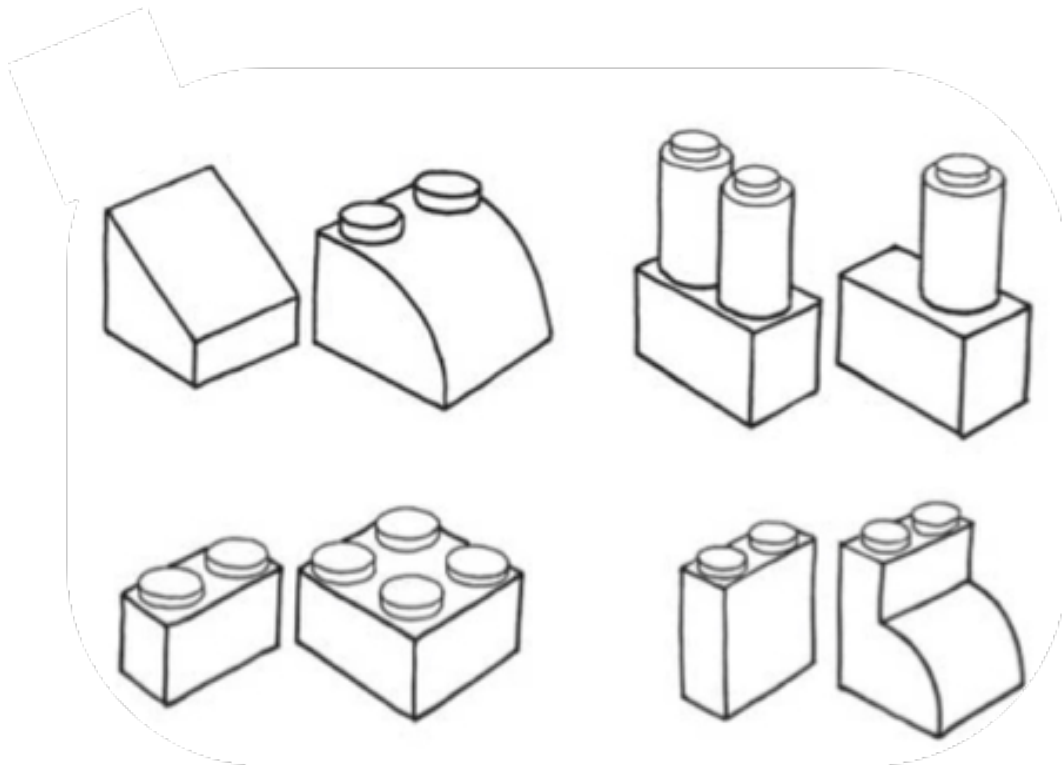


Figure 1: Examples of the 20 pairs of complex shapes presented to the five congenitally blind patients who had just regained their sight in a study by Held et al. (2011).

objects over a very small range of stimuli anyway, and that most of the objects we see (such as planets and skyscrapers etc.) cannot be felt.³

Held et al. (2011) examined cross-modal recognition in a group of five participants who had recently been treated for congenital blindness. The participants first saw or felt a sample 3D Lego-like shape (see Figure 1). Next, two test shapes were presented, with the participants instructed to discriminate which of the test

³ One other question to consider here is whether the congenitally blind are capable of acquiring an understanding of geometry? Marlair et al. (2021) created 3D-printed tactile shapes to investigate the intuitive understanding of basic geometric concepts in blind children and adults. The participants were exposed to four shapes at a time, three instantiating a particular geometric concept (e.g., line vs. curve, parallel vs. secant lines, symmetry) while the remaining one violated it. The participants had to examine all of the shapes in order to detect the one that differed most from the others. The results showed that both the blindfolded sighted and congenitally blind individuals were able to perform this task at above chance levels, thus suggesting that basic geometrical knowledge can develop in the absence of visual experience. However, that said, it should be noted that the blind were significantly worse on this task than the sighted participants performing the same task in the visual modality (i.e., with their eyes open). However, they performed just as well as the sighted performing the task in the tactile modality (i.e., while wearing a blindfold), thus suggesting that the lack of vision somehow impacts on the understanding of geometrical concepts.

shapes matched the sample (thus meaning that chance level performance was 50% correct). Twenty shapes pairs were used in this study. Performance was compared across three conditions: visual–visual, where the sample and test shapes were all presented visually; tactual–tactual, where all of the shapes were presented tactually; and tactual–visual, where the sample shape was presented tactually while the test shapes were presented visually.

The striking results to have emerged from this intriguing study were that the participants (who were tested within 48hrs of surgery on their first eye) recognized the sample shape with a very high degree of accuracy in both the unimodal visual and unimodal tactile/haptic conditions (98% and 92% correct, respectively). However, the five participants performed at close to chance level in the crossmodal tactual–visual condition (58% correct). Held and his colleagues took these results to show that the answer to Molyneux's question is “likely negative” (Held et al., 2011). In a follow-up study, three of the five participants were tested again on the tactual–visual task. Performance improved dramatically (to around 80% accuracy) within five days of initial testing. These researchers summarize the implications of their research findings by suggesting that: “It is interesting to speculate on the possible ecological importance of a learned, rather than innate, mapping between vision and haptics” (Held et al., 2011, p. 552).⁴

Held et al. (2011) used shapes that were quite complex (see Figure 1), combining both angular/round with squared features, instead of simply contrasting a sphere and a cube, as had originally been suggested by Molyneux (Schwenkler, 2012, or a square vs. circle, as suggested by others, to eliminate problems associated with processing depth). It is easy to imagine here how the complexity of the shapes used by Held et al. might have necessitated a longer period of learning (thus preventing immediate crossmodal transfer, Schwenkler, 2012, 2013, 2019; cf. Todd, 2004). Moreover, in a similar task, Fine et al. (2003) found that a congenital cataract patient with vision restored in one eye whom they were studying was able to identify 3D shapes with perfect accuracy when the stimuli included visual cues that simu-

⁴ Held here would appear to equate a yes/no answer to the Molyneux question with the distinction between innate and learned. However, this may be too hasty. Relevant here, Locke points to the fact that people assent to certain propositions upon first hearing them does not prove that those propositions are innately stored in them; it may happen just because those propositions are obvious, or self-evident (*Essay*, I.ii.17-18). Similarly, suppose it were a fact that many newly-sighted individuals were to immediately recognize a circle seen for the first time as the object they previously knew as a circle by touch. That might happen just because there is a manifest similarity between seen circles and felt circles, not because there is an innate connection between the visual and haptic representations of circularity. (See Evans (1985) for the point that similarity and innateness are competing explanations of yes to Molyneux.) At the same time, however, it should be noted that the concept of “manifest similarity” might itself be problematic. One should at least admit that the subject must possess the ability to detect such similarity between seen and felt circles before their initial exposure to these stimuli, as in the case of the Molyneux question. Moreover, while similarity might be straightforwardly applied to stimuli presented within the same sensory domain, it is far less intuitive to talk about similarity across the senses (see Di Stefano & Spence, 2023, for a review).

lated object motion; similarly, the participants in a study by Ostrovsky et al. (2009) were better able to identify photographic images of objects that are frequently seen in motion in everyday life (see also Gregory, 2003; Huber et al., 2015; McKyton et al., 2015; Ostrovsky et al., 2006; Šikl et al., 2013). These findings therefore suggest that sensory information obtained from object motion may be key to the formation of the robust mental, or amodal, representations of shape needed to successfully perform the crossmodal transfer task.

Despite the experimental effort that has been devoted to answering the Molyneux question, the question of the relationship between the visual and tactile representations of shape has not, thus far, been satisfactorily answered empirically (nor will it ever be, at least not according to Jacomuzzi et al., 2003; cf. Degenaar, 1996). Evidence that the surgical restoration of vision provides only a limited capacity for 3D form perception in the period immediately after surgery makes it reasonable to hypothesize that Held et al.'s (2011) participants may have failed to match seen with felt shape because they were unable to form the visual representations of the shapes in the first place.⁵ This hypothesis could be further supported by evidence showing that the newborn's visual system (to which one can compare the newly-sighted one of adults, e.g., Meltzoff, 1993) is characterized both by weak acuity and contrast sensitivity (Allen et al., 1996), as well as by a poor ability to fixate, and uncoordinated saccadic and other ocular movements (Ricci et al., 2008; cf. Ueda & Saiki, 2012). Berkeley (1950) put it a radical way, excluding any possible phenomenological relationship between seen and felt objects: "But if we take a close and accurate view of things, it must be acknowledged that we never see and feel one and the same object. That which is seen is one thing, and that which is felt is another [...] the objects of sight and touch are two distinct things. It may perhaps require some thought to rightly conceive this distinction. And the difficulty seems not a little increased, because the combination of visible ideas hath constantly the same name as the combination of tangible ideas wherewith it is connected [...]" (see also Schwenkler, 2013, 2019).⁶

Given the above, the Molyneux question would seemingly remain open and thus potentially benefit from being considered from a different perspective (cf. Deroy & Auvray, 2013; Piller et al., 2023).⁷ In what follows, we explore the pos-

⁵ That said, an explanation is still needed for how these individuals performed near-perfectly in the intramodal visual matching task.

⁶ The radical import of Berkeley's words should not be forgotten. If the subject is given the question, "which of the things you now see before you is what you previously knew by touch as a sphere," Berkeley would have to say that the correct answer is *neither*. For nothing seen was ever previously touched; visible objects and tangible objects do not even inhabit a common space.

⁷ Here, it is interesting to note that the incredibly robust 'Bouba-Kiki effect', namely the sound symbolic matching of round and angular shapes seems not to occur in the congenitally blind (Sourav et al., 2019), nor immediately in congenital cataract patients once their sight has been restored. Rather, much like the cross modal shape matching reported by Held et al. (2011), appears gradually following the restoration of sight (Piller et al., 2023).

sibility of reframing the Molyneux question in terms of the crossmodal matching of stimulus intensity.

2 Old and new versions of the Molyneux question

What happens if the Molyneux problem were to be considered in terms of the emerging literature on crossmodal correspondences (Spence, 2011)? Crossmodal correspondences have been defined in terms of the often-surprising, yet consensual associations between stimuli, attributes, or dimensions of experience in one sensory modality and another stimulus, attribute, or dimension of experience presented in a different sensory modality (Spence, 2011; Walker-Andrews, 1994). As Guellaï et al. (2019, p. 1) wondered a few years ago: “Are infants able to make cross-sensory correspondences from birth? Do certain correspondences require extensive real-world experience? Some studies have shown that newborns are able to match stimuli perceived in different sense modalities. Yet, the origins and mechanisms underlying these abilities are unclear.”

Importantly, certain researchers have argued that crossmodal matching is possible only following some degree of infant development to establish the relevant neural mechanisms (Ettlinger, 1967). Note here also the much older argument in the developmental literature between those researchers, like Jean Piaget (1937), who assumed that the newborn’s experience of the world begins with separate senses, and Eleanor Gibson (1969) who considered it plausible that the newborn might be capable of unified multisensory perceptual experiences from the get-go. As Meltzoff and Borton (1979, p. 403) put it, one hypothesis “is that the detection of shape invariants across different modalities is a fundamental characteristic of man’s perceptual-cognitive system, available without the need for **learned** correlation.” [emphasis added].

One way to think about the Molyneux question in terms of crossmodal correspondences might be to ask whether an individual born blind, on regaining their sight, would spontaneously match stimulus intensity (i.e., how salient the perceptual experience is) across the senses. This suggestion is based on S. S. Stevens’ (1957, 1971) early intuition that there may be a structural correspondence based on increasing stimulus intensity, possibly represented by increased neural firing in all of the senses. Given the suggestion that such a mapping results from the hard-wired pattern of neural coding (though it should be noted that no empirical evidence has yet been put forward in support of this particular intriguing suggestion), it might be imagined that intensity-based crossmodal correspondences would not have to rely on perceptual experience in order to be established. Studies of crossmodal intensity matching in adults have been reported by a number of psychophysicists (e.g., Root & Ross, 1965; Stevens & Marks, 1965). However, a detailed discussion of whether what is being matched is a relation between two stimuli or an absolute magnitude of a single stimulus falls beyond the scope of the present article (see Mellers & Birnbaum, 1982; Stevens & Marks, 1980).

Relevant here, Lewkowicz and Turkewitz (1980) suggested more than four decades ago that the crossmodal equivalence of auditory loudness and visual brightness is present in babies within the first month of life (see Section [Cross-modal shape matching in human neonates](#) for a full explanation of this study). If true, this would contrast with a number of other crossmodal correspondences that are putatively based on the internalization of the statistics of the environment, and which presumably emerge only after sufficient relevant perceptual experience (Deroy & Auvray, 2013), and/or development of the relevant concepts (see Marks et al., 1987). However, our own intuition is that the matching of stimulus intensity across the senses (say auditory loudness and visual brightness) will likely also only emerge gradually after the restoration of sight (e.g., in congenital cataract patients).⁸

The existence of crossmodal associations (or correspondences) mediated by stimulus intensity has often been considered as proving the existence of some sort of perceptual similarity across the senses. As Marks et al. (1986) noted almost 40 years ago: “The commonality that underlies the perception of intensity seems to emanate from phenomenological, which is perhaps to say from neurophysiological, considerations. It is our view that brightness, loudness, touch, smell, and taste intensity resemble one another because all are mediated by a common mode of neural processing (e.g., greater stimulus intensities act to increase the rate of firing in a given nerve fiber or to recruit greater and greater numbers of nerve fibers). The ultimate source of the cross-modal similarity resides not in the external world but inside the skin.”⁹ However, Marks et al. (1986, p. 517) go on to observe that: “Loudness does not resemble sound pressure any more than pitch resembles sound frequency. But intensity has an interesting feature. Unlike other perceptions of secondary qualities such as pitch, odor quality and color (though like perceptions of primary qualities such as duration and size), intensity is *common* to various modalities. This cross-modal commonality has, it appears, little or nothing to do with common stimulus events. Objects may be large to both sight and touch because the objects themselves *really* are large (primary quality). But brightly illuminated objects need not emit loud sounds.”¹⁰ (Italics in original.) In other words, crossmodal perceptual similarity, if it exists, is suggested to originate from within the perceiver, rather than being a property of the environment (that can be internalized as a result of the relevant associative learning).

⁸ Moving to the chemical senses: One could have sighted participants taste an unfamiliar fruit (under blindfolded conditions) and then have them identify which of two foods they had tasted. The evidence suggests that sighted participants are able to infer at better than chance levels whether a given fruit juice came from a large or small fruit (see Ngo et al., 2013).

⁹ Elsewhere, note, we have taken issue with the idea that there is any kind of perceptual similarity across the senses (Di Stefano & Spence, 2023).

¹⁰ One might choose to take issue with the claim that objects may be ‘large’ to both sight and touch “because the objects themselves *really* are large”. Almost by definition, we are unable to feel large objects, like planets or skyscrapers, only inspect them visually.

The point to note here is that discussion of innate/amodal perceptual qualities/dimensions is somehow intimately tied-up with people's intuitions, assumptions, and/or beliefs about the possibility (or very existence) of the phenomenon of crossmodal perceptual similarity (see Di Stefano & Spence, 2023, for a review). One might wonder whether these questions can be separated, namely, if similarity across (or between) the senses can be explained without the need to assume the existence of some sort of amodal, or intersensory, quality. Philosophers and psychologists have long agreed that similarity is a key organizational principle underlying perceptual experience and grounds the ability of humans to categorize objects into different classes (see Plato's *Parmenides*, as discussed by Allen 1997; and Ryle 1939a, 1939b; see also Gentner and Medina, 1998; Goldstone and Barsalou, 1998; Goldstone and Son, 2012; Goodman, 1972; Quine, 2000; Segundo-Ortin and Hutto, 2021; Tversky, 1977). In turn, this ability is related to the formation of mental categories, or concepts, which group many different things based on some shared feature (Plato, *Republic*, 596a, see Reeve, 2004). If two stimuli presented to different senses cannot be conceived of in terms of mere identity, some degree of similarity needs to be admitted when it comes to accounting for consensual matchings across the senses.¹¹

As Marks (1989, p. 58) has written: "In a cross-modality matching task, for example, virtually all subjects will set higher sound frequencies to match greater visual intensities (Marks, 1974, 1978), thereby revealing a universal appreciation of similarity between the dimension of pitch on the one hand and that of brightness on the other." But why should we admit to the possibility of crossmodal perceptual similarity, given both recent and historic arguments against the very possibility of making such judgments crossmodally (see Di Stefano & Spence, 2023; Helmholtz, 1878). At the same time, one might wonder whether a sensitivity to polar correspondences (e.g., Proctor & Cho, 2006; see also Smith & Sera, 1992) might not provide sufficient grounding for such crossmodal matching without the need to introduce the concept of perceptual similarity.

An additional point to note here concerns the fact that information concerning (increased) neural firing is not directly accessible to the perceiver. Thus, one needs to assume that it is rather the phenomenological, and thus perceivable, effect of the neurophysiological variations that grounds the possibility of establishing similarity across the senses (e.g., the intensity of visual brightness and auditory loudness). However, should this be the case, it would remain unclear what exactly mediates the similarity observed at the phenomenological level, given that sensory informa-

¹¹ An example from Evans (1985, p. 372) in which the manifest similarity of a feature presented in one sensory modality and a feature presented in another sensory modality should make us expect successful matching: 'Continuous' and 'pulsating' (or 'intermittent') are learned by a deaf person through skin stimulations; upon gaining hearing, s/he is then presented with continuous and intermittent tones. Evans' position is that few of us would have a doubt about whether s/he should be able to recognize which is which. (One could reimagine this as involving touch and vision; a broomstick and a row of pebbles of the same diameter as the stick are presented first to touch and then to vision.)

tion regarding intensity across the senses manifests itself in a phenomenologically incommensurable/dissimilar way (e.g., the increase of brightness is not similar to the increase of loudness in any intuitive way), as highlighted by Marks and colleagues in the above quote.

3 Crossmodal matching: Empirical evidence

Having set the scene, we now wish to review the empirical evidence concerning the early (developmental) emergence of the crossmodal matching of shape, intensity, and beyond. While it is becoming increasingly clear that human neonates/infants become sensitive to different crossmodal correspondences at different stages of (early) development (Marks et al., 1987; Meng et al., 2023; Speed et al., 2021; see Spence, 2022, for a review), there would appear to be little convincing evidence for the existence of any innate sensitivity to amodal perceptual qualities/dimensions, contrary to the claims/assumptions of at least some developmental researchers.

3.1 Crossmodal shape matching in human neonates

In a now-classic developmental study, Meltzoff and Borton (1979) had 32 infants aged from 26-33 days old, engage with one of two pacifiers (one smooth and the other nubbled; see Figure 2) for 90 seconds (before removing it) and then presenting the infant with both pacifiers visually. Crucially, the infants demonstrated a significant preference (in terms of their visual fixation) for the pacifier that had been explored orally (attracting 71.8% of fixation time, as compared to a chance level of 50%). This phenomenon was described as indicating some kind of tactual-visual correspondence. This finding was replicated in a second study with the same number of infants (attracting 67.1% of fixation time). However, even were these results to have proved replicable, which unfortunately they have not, one might wonder whether what is actually transferred¹² might not be stimulus hardness (rather than shape, Gibson & Walker, 1984), texture (see Picard, 2007; Sann & Streri, 2007; Streri & Hevia, 2023; cf. Molina & Jouen, 2003), or even numerosity (i.e., of the nubs, Farzin et al., 2009; Izard et al., 2009). Also left unanswered is the question of the orientation-specificity of such crossmodal matching effects (see Huber et al., 2015; Lacey et al., 2009; Ueda & Saiki, 2007), and whether the size of the stimuli matters or merely their shape: Imagine for a moment the infants being shown a pacifier of the same size but different shape versus of different size but the same shape features.¹³ Which would win out in terms of capturing the infant's visual attention?

¹² Note that crossmodal transfer, as opposed to crossmodal integration, has been identified as especially important by a number of researchers (e.g., Guellaï et al., 2019).

¹³ See Cowey and Weiskrantz (1975) for a crossmodal matching study in rhesus monkeys where precisely such a comparison caused problems – in particular, mostly regardless of whether the monkeys had been presented with a small or large edible disc (and a large or small inedible disc) in darkness, they chose a larger disc when presented with them visually.

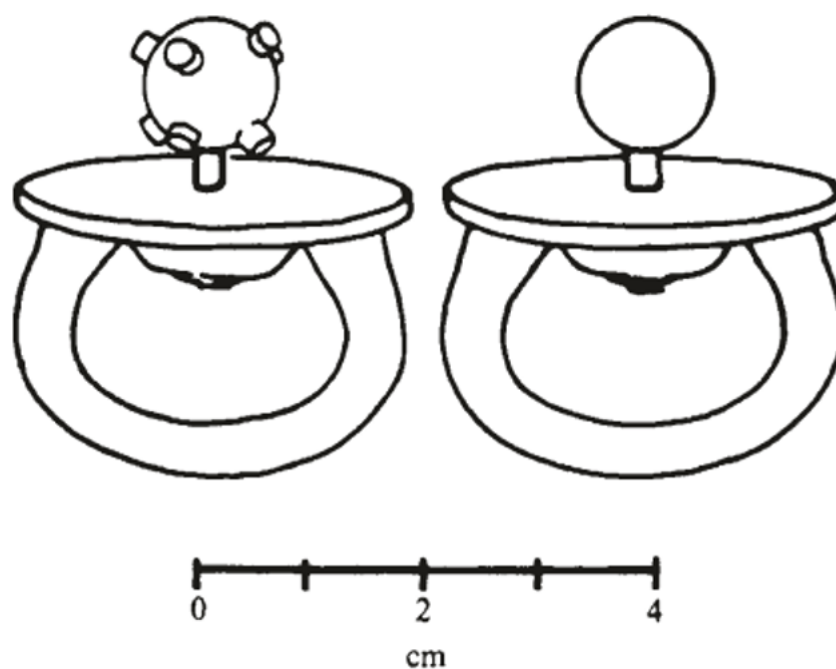


Figure 2: The smooth and nubby pacifiers presented visually to the participants in Meltzoff and Borton's (1979) study of 30-day-old infants.

There is currently no evidence on which to answer this important question (see also Chen et al., 2004, on audio-oral crossmodal matching in neonates).

When thinking about the normally-sighted, it is worth considering the way in which we typically effortlessly recognize shapes through vision and touch, and novel shapes first encountered through one modality can be re-identified through the other (Norman et al., 2004). That said, in the normally sighted, vision plays a more important role than touch in visuo-haptic object recognition (Kassuba et al., 2013; Pietrini et al., 2004). Relevant here, Cao, Kelly, Nyugen, Chow, Leonardo, Sabov, and Ciaramitaro (2024) recently studied the acquisition of audio-haptic matching in sighted and blind children, and pointed to the inefficient haptic exploratory strategies seen in blind children. Simplifying our discussion from the perception of shape to the perception of length, it has also been suggested that there may be differences among varieties of perceived length. Indeed, the non-commensurability of measures of length as presented in different sensory modalities have long discussed (e.g., Jastrow, 1886; Teghtsoonian & Teghtsoonian, 1970; Waterman, 1917).

It is, though, important to stress that Meltzoff and colleagues' (Meltzoff, 1993; Meltzoff & Borton, 1979) infant oral-tactile habituation studies have proved surprisingly difficult to replicate (Maurer et al., 1999; though see also Streri & Gentaz, 2004). For instance, Maurer and her colleagues reported on a series of three studies in which they tested 1-month-olds on their crossmodal matching abilities using exactly the same experimental procedure as described by Meltzoff and Borton. Cru-

cially, however, Maurer et al. also included controls for stimulus preference and side bias (referring to any tendency for neonates to preferentially orient to one side versus the other). In their first experiment (N = 48), infants' looking times to smooth and nubby visual stimuli were not influenced by oral exposure to one of the shapes during the preceding 90 s, except for an effect on the first test trial in one group of participants. However, this result was put down to limited cross-modal transfer, to a Type 1 error, or to side bias, possibly interacting with a small stimulus preference. Furthermore, this effect was not replicated in a new group of 16 infants who exhibited less bias (Experiment 2), thus suggesting that it did not reflect a genuine crossmodal transfer of structural or shape information. Finally, Maurer et al. conducted a third experiment (N = 32), that was an exact replication of Meltzoff and Borton's original study. Once again, though, the results failed to yield any evidence of the crossmodal matching of shape information between oral-touch and vision.

Meanwhile, Streri and Gentaz (2003) conducted a similar study using manual instead of oral exploration. In a haptic familiarization phase, newborns were given an object to explore manually without seeing it; then, in a visual test phase, they were shown the familiar object paired with a novel one for 60 seconds. The participants consisted of 24 newborns (mean age: 62 hours). An experimental group (12 newborns) received the two phases successively (haptic then visual) while a baseline group (12 newborns) received only the visual test phase with the same objects as the experimental group but without the haptic familiarization phase. The results revealed that the newborns in the experimental group looked at the novel object for longer than the familiar one.¹⁴ In contrast, the newborns in the baseline group looked equally at the two objects, thus showing that newborns can transfer shape information from touch to vision before they have had the opportunity to learn the pairing of visual and tactile experiences. Should such results prove replicable, then it might be taken to suggest that manual tactile to visual crossmodal transfer develops prior to oral-visual shape transfer.

A number of other researchers have obtained rather mixed evidence for cross-modal shape transfer in older infants (e.g., Rose et al., 1981; Streri, 1987; Streri & Pecheux, 1986). For example, Rose et al. assessed crossmodal shape transfer between oral or tactile and visual exploration of the same stimuli in 6-month-old infants. However, across a series of three studies (N = 25-30 infants in each study), these researchers obtained no evidence that was deemed consistent with the existence of oral-visual transfer (as had been reported by Meltzoff & Borton, 1979), while tactile/haptic-visual crossmodal transfer was only demonstrated under those conditions where the necessity to retain stimuli in memory was eliminated, i.e., by allowing oral/tactile exposure at the same time as the visual stimuli were presented. However, while this step undoubtedly eliminates the requirement to hold

¹⁴ Note that this is the opposite pattern of behaviour to that reported by Meltzoff and Borton (1979). Inconsistencies might be related to a behavioural paradigm that is based on a single measurement (i.e., looking time), as highlighted by Trehub (2012).

the shape of a stimulus in some kind of memory store, the fact that crossmodal transfer occurred between stimuli that were presented simultaneously from different spatial locations is presumably foregrounded in a way that is simply not the case when the unimodal exposure to the stimuli occurs sequentially. A fourth study ($N = 20$) provided some borderline-significant evidence for intramodal visual transfer in 6-month-olds. In summary, therefore, no evidence of oral-visual transfer in 6-month-olds, only tactile-visual transfer when a suitably long familiarization period (60 seconds long not 30 seconds) and when two unimodal stimuli are presented simultaneously (to eliminate memory). All-in-all, crossmodal shape transfer was less robust than amongst the older infants. Such a pattern of results also question the robustness of Streri and Gentaz's (2003) findings in much younger infants.

Taken together, therefore, the evidence that has been published over the last half century or so would not appear to provide any particularly convincing support for the claim that neonates (more specifically, one-month-olds) can crossmodally match visual and oral-tactile shape properties (see also Brown & Gottfried, 1986; and Pêcheux et al., 1988, for additional failures to show robust crossmodal transfer of shape information in infants). Indeed, one might anyway question just how developed an infant's oral stereognosis abilities and/or their information processing capacities actually are at just one-month of age (Jacobs et al., 1998; see also Bushnell, 1994; Waterman, 1917), not to mention questioning the status of their working memory abilities (cf. Rose et al., 1981), and the extent to which memory systems are shared across the senses at that (st)age (cf. Goodnow, 1971). The latter would presumably be required, given that the pacifier was removed prior to their visual presentation (see Bremner & Spence, 2008; Ernst, 2007).¹⁵ Bear in mind here also the fact that the orally-palpated and seen objects were not presented from the same spatial location either (Filippetti et al., 2015; cf. Lawson, 1980). Indeed, perhaps before the coordination between vision and grasping is well-established, it might be expected that the intermodal transfer between touch and vision would be hard to establish (Streri & Hevia, 2023).

3.2 Crossmodal shape transfer in other species

Shifting now to the crossmodal transfer of shape information in other species (e.g., Cowey & Weiskrantz, 1975; Davenport & Rogers, 1970). Davenport and Rogers studied crossmodal shape matching in a group of chimpanzees and orang-utans ($N = 5$). In particular, these researchers demonstrated that intermodal equivalence (or rather, intermodal transfer, that is, the ability to match one of two felt shapes with the matching visual object) could be acquired in apes (three chimpanzees and

¹⁵ Relevant to this point, newborns can learn arbitrary auditory-visual associations (such as between an oriented coloured line and a syllable), but only when the visual and auditory information are presented synchronously (Slater & Kirby, 1998). At the same time, according to Slater and Kirby, newborns can associate objects and sounds on the basis of temporal synchrony.

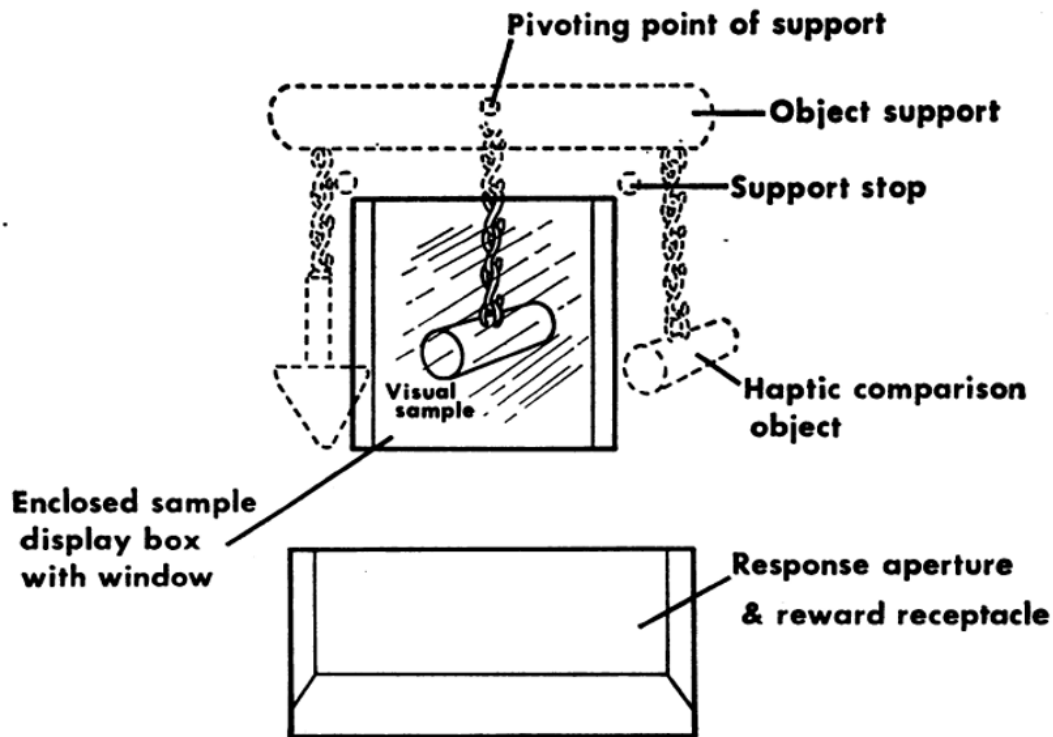


Figure 3: The three seen and felt shapes presented to apes (3 orang-utans and 2 chimpanzees) in a crossmodal matching study reported by Davenport and Rogers (1970).

two orang-utans) (see Figure 3). One does, though, of course, need to be cognizant of the fact that extensive training is needed when working with animal participants to bring them up to the level of even human infants. What is more, the fact that the animals could see the two objects while simultaneously feeling the hidden tactile object once again presumably eliminates the need for the animals to rely on memory to perform the crossmodal matching task. Nevertheless, the results do highlight what appears to be an impressive ability to engage in cross-modal transfer of shape from touch/haptics to vision. That said, the simple fact is that the appropriate answer in this task is to match the tactile/haptic and visual impressions that happen to relate to the same object. One might wonder how performance would change if the apes were tasked with learning an arbitrary mapping between unrelated felt and seen shapes (i.e., mere contingency learning versus picking up on some common external shape property). By using edible and inedible shapes, presented first to rhesus monkeys in darkness, then in the light, Cowey and Weiskrantz (1975) were able to demonstrate cross-modal matching from touch to vision. As such, crossmodal matching should not be considered as unique to apes and humans (see also Zhou & Fuster, 2000).

3.3 Crossmodal intensity matching in human neonates

Lewkowicz and Turkewitz (1980) conducted a seminal study in which they investigated the crossmodal matching of audiovisual stimulus intensity in a group of 3-week-old infants. The infants were repeatedly presented with white-light followed by white-noise stimuli at different intensities. A U-shaped relationship between the magnitude of the infant's cardiac response (CR) and the loudness of the sound was observed. In view of previous findings that without prior visual stimulation a monotonic increase in CR to the same range of auditory stimuli results, this finding of a significant quadratic relationship with loudness was taken to suggest that infants were responding to the auditory stimuli in terms of their similarity¹⁶ to the previously-presented visual stimulus.¹⁷ However, rereading this frequently-cited paper, it is noticeable how the authors initially assume intensity to be an amodal dimension (in that it is a quality that can be perceived in different modalities),¹⁸ but then go on to describe their findings merely as showing that quantitative differences (e.g., in stimulus intensity) may be more salient than qualitative differences (e.g., in modality) in stimulation early in human development, either as a result of a lack of sensory differentiation, or else as a result of selective attention being captured more by intensity changes. Intriguingly, however, the results of a study with 31 adults did not show any systematic relationship between CR and loudness, indicating that unlike infants, adults do not appear to make such crossmodal intensity matches spontaneously.¹⁹

One of the challenges when it comes to evaluating the putatively innate nature of crossmodal intensity matching comes from the confusing definitions of visual stimuli intensity used by at least certain influential developmental researchers. Take, for example, Bahrick et al. (2004, p. 99) who at one point state that: "One type of overlap involves *amodal* information, that is, information that is not specific to a single sense modality, but is completely redundant across more than one sense." So

¹⁶ One might here want to question how 'similarity' is operationally defined in the experimental protocol.

¹⁷ That said, there is obviously much inference required to reach this intriguing conclusion. And, as highlighted elsewhere in the context of claims regarding the putative existence of neonatal synaesthesia, it is difficult to know whether a neonate perceives stimuli presented in different modalities as being similar, versus simply failing to distinguish perceptually between them (see Deroy & Spence, 2013).

¹⁸ They write that: "Amodal features are those that can be used to identify an aspect of an object or an event in more than one modality, whereas modality-specific features can only be used to identify an aspect of a stimulus that is peculiar to a single modality. Thus, intensity, rate, duration, spatial location, spatial extent, rhythm, and shape all represent amodal features of the world that can be specified in more than one modality. They stand in distinction to such modality specific features of stimulation as redness, sweetness, and pitch." (Lewkowicz & Turkewitz, 1980, p. 597).

¹⁹ Intriguingly, there would appear to have been no attempts to replicate this seminal study, despite it having been published only one year after Meltzoff and Borton's (1979) study. The closest may be a conference paper read by Wilson (1969) in which it was apparently reported that animals do not transfer responses to stimulus intensity change from one sensory modality to another, thus seemingly contrasting with Lewkowicz and Turkewitz's (1980) findings.

far, so good. However, it soon becomes unclear what exactly they mean to refer to when talking about visual intensity – certainly not the increased visual brightness, as might legitimately be assumed, and as discussed/studied by Stevens and others (Lewkowicz & Turkewitz, 1980). Just take the following quote from a review article that appeared in *Current Directions in Psychological Science*: “The sights and sounds of a ball bouncing are synchronous, originate in the same location, and share a common rate, rhythm, and intensity pattern. Picking up this redundant, amodal information is fundamental to perceptual development [...] the face and voice of a person speaking share temporal synchrony, rhythm, tempo, and changing intensity” (Bahrick et al., 2004, p. 99). It is entirely unclear (at least to the authors of this review) what visual intensity is meant to refer to in this case.

3.4 Development trends in crossmodal correspondences

Although there has been some disagreement, developmental researchers would now appear to have established when during human development the sensitivity to different crossmodal correspondences typically emerges (e.g., Marks et al., 1987; Spence, 2022; Streri & Hevia, 2023; Walker et al., 2010, 2014).²⁰ Indeed, according to one hypothesis, the early integration hypothesis, cross-sensory integration is already present from birth onwards, while according to another hypothesis, the late integration hypothesis, the role of experience in the development of cross-sensory associations is emphasized (Dionne-Dostie et al., 2015; cf. Schwenkler, 2012).²¹ The stage in human development at which a sensitivity to various different (kinds of) crossmodal correspondence can be demonstrated is currently a particularly active area of empirical research. Looking to the future, it would also be of interest to determine whether different classes of crossmodal correspondence are represented differently in the brain (Guellaì et al., 2019, p. 5). Indeed, it would seem eminently plausible to assume that different classes of correspondence might well emerge at different stages during the course of human development.

4 Conclusions

In conclusion, in the present review, it has been argued that the continued interest in the ‘Molyneux problem’ can be seen as resulting from the intuitive, though we would argue misguided, notion that shape information is coded amodally, at

²⁰ As Lewkowicz and Turkewitz (1980, p. 606) note in their early study: “Our approach further suggests the caveat that care be taken to distinguish between various types of cross-modal equivalence. That is, some equivalences may be based on primitive and undifferentiated functioning, whereas others may represent the highest levels of cross-modal functioning and may involve perceptual and cognitive as well as sensory mechanisms.”

²¹ Guellaì et al. (2019, p. 5) ask whether there: “Are general associative learning mechanisms sufficient to explain how infants come to pair sensory cues across modalities, or do specific learning processes or constraints guide the acquisition of some (or all) cross-sensory correspondences?”

least by the senses of vision and active touch (i.e., haptics). At the same time, however, neither the developmental research, nor the latest findings from congenital cataract patients', whose sight has been restored (Held et al., 2011), would appear to provide any convincing evidence that such crossmodal matching can be achieved in the absence of the relevant sensory experience (i.e., as a result of associative learning/repetitive exposure). Similarly, there is surprisingly little evidence to support the notion that the matching of stimulus intensity across the senses is either innate or amodal either though it may be acquired early in life. However, as we have seen, the evidence on which such a claim is based (Lewkowicz & Turkewitz, 1980), is open to multiple interpretations, and has also yet to be replicated. Nor, we would argue, is there any reason to suspect that audiovisual stimulus intensity necessarily reflects a Gibsonian affordance that needs no integration (Gibson, 1977).²²

Furthermore, considering the distinct nature of sensory impressions and how they are processed by the two sensory systems, it remains unclear how exactly haptic input could be effectively translated into a visual format (or *vice versa*) to generate an amodal representation of shape (or intensity). As such, the notion that shape (and intensity) are somehow special perceptual qualities/dimensions (i.e., amodal and/or innate, see Lewkowicz & Minar, 2014; Maurer et al., 1999) should presumably be abandoned. In its place, it can be argued that it makes more sense simply to accept that different classes of crossmodal correspondence are acquired at different times/rates during the course of human development.²³ We believe that figuring out the relevant constraints on the acquisition of different classes of crossmodal correspondence, which are based on similar neural coding, similar phenomenology, similar hedonic associations, and/or shared lexical descriptors would appear to constitute a more promising approach to furthering our understanding in this area than continuing to pursue an empirical answer to the original Molyneux question which, as has been argued by others, might be best left as a historical thought experiment. To be absolutely clear, the suggestion that is put forward here is that while a prolonged focus on the visual-haptic shape correspondence version of the Molyneux question may have hindered research and theoretical progress in thinking about the origins of crossmodal matching, there are various other versions of the crossmodal matching question that deserve further empirical and theoretical scrutiny.

Crucially, the latter approach is less theoretically laden with regards to the existence of amodal perceptual qualities and assumptions concerning the innate nature of certain perceptual abilities (as promoted, for example, by the work of J. J. Gib-

²² According to Gibson's (1979) ecological theory of perception, amodal information (or information that is not specific to any one modality and that can be conveyed redundantly across several senses), is obtained directly from adaptive interaction between organisms and their environments. It should though be noted that Gibson only proposed that amodal spatial and temporal dimensions were available to all sensory modalities from birth.

²³ It should also be noted that sound symbolism does not appear to be present within the first month of life (Ozturk et al., 2013; Sidhu et al., 2023).

son, see Spence & Di Stefano, 2024). Moreover, it avoids framing the Molyneux's question, as well as many other issues that arise from developmental perception research, from the innatist vs empiricist debate, which would appear to misconceive the nature of biological universals and how they manifest in individuals. Alternatively, we suggest that dichotomies such as innate/learned may be more effectively reconceptualized in a continuous fashion, namely, in terms of the ease and success of the process of acquiring skills rooted in biology as opposed to those that are more imbued by culture. For instance, learning to walk or speak may be more readily, easily, and effectively achievable targets for most people compared to mastering purely cultural skills such as playing the violin or drawing.

Extending the Molyneux problem to the chemical senses could also prove to be a fascinating avenue of investigation. Research in this area could, for example, involve experiments in which the participants are initially exposed to certain spices or ingredients through taste only and then later presented with visual stimuli related to those tastes to assess their ability to make accurate visual identifications (cf. Ngo et al., 2013). Such investigations could presumably contribute to our understanding of crossmodal perception and the integration of sensory information in the human brain.

Acknowledgments

The authors wish to acknowledge the very helpful suggestions made by two anonymous reviewers to earlier versions of this manuscript.

References

- Allen, D., Tyler, C. W., & Norcia, A. M. (1996). Development of grating acuity and contrast sensitivity in the central and peripheral visual field of the human infant. *Vision Research*, 36(13), 1945–1953. [https://doi.org/10.1016/0042-6989\(95\)0257-x](https://doi.org/10.1016/0042-6989(95)0257-x)
- Allen, R. E. (1997). *Plato's parmenides*. Yale University Press.
- Anstey, P. R. (2023). Condillac and the Molyneux problem. In D. Antoine-Mahut & A. Waldow (Eds.), *Condillac and his reception* (pp. 28–43). Routledge.
- Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science*, 13(3), 99–102. <https://doi.org/10.1111/j.0963-7214.2004.00283.x>
- Berkeley, G. (1950). *A new theory of vision*. Dutton.
- Bremner, A. J., & Spence, C. (2008). Unimodal experiences constrain while multisensory experiences enrich cognitive construction. *Behavioural and Brain Sciences*, 31(3), 335–336. <https://doi.org/10.1017/S0140525X0800410X>
- Brown, K. W., & Gottfried, A. W. (1986). Cross-modal transfer of shape in early infancy: Is there reliable evidence? In L. P. Lipsitt & R. Rovée-Collier (Eds.), *Advances in infancy research* (Vol. 4, pp. 163–170). Ablex.
- Bruno, M., & Mandelbaum, E. (2010). Locke's answer to Molyneux's thought experiment. *History of Philosophy Quarterly*, 27(2), 165–180.
- Bushnell, E. W. (1994). A dual-processing approach to cross-modal matching: Implications for development. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 19–38). Erlbaum.
- Campbell, J. (1996). Molyneux's question. *Philosophical Issues*, 7, 301–318. <https://doi.org/10.2307/1522914>
- Cao, S., Kelly, J., Nyugen, C., Chow, H. M., Leonardo, B., Sabov, A., & Ciaramitaro, V. M. (2024). Prior visual experience increases children's use of effective haptic exploration strategies in audio-tactile sound-shape correspondences. *Journal of Experimental Child Psychology*, 241(105856). <https://doi.org/10.1016/j.jecp.2023.105856>
- Chen, X., Striano, T., & Rakoczy, H. (2004). Auditory-oral matching behavior in newborns. *Developmental Science*, 7, 42–47. <https://doi.org/10.1111/j.1467-7687.2004.00321.x>
- Cowey, A., & Weiskrantz, L. (1975). Demonstration of crossmodal matching in rhesus monkeys, macaca mulatta. *Neuropsychologia*, 13(1), 117–120. [https://doi.org/10.1016/0028-3932\(75\)90057-3](https://doi.org/10.1016/0028-3932(75)90057-3)
- Davenport, R. K., & Rogers, C. M. (1970). Intermodal equivalence of stimuli in apes. *Science*, 168(3928), 279–280. <https://doi.org/10.1126/science.168.3928.279>
- Degenaar, M. (1996). *Molyneux's problem* (M. J. Collins, Trans.). Kluwer.

Spence, C., & Di Stefano, N. (2024). Old and new versions of the Molyneux question: A review of experimental answers *Philosophy and the Mind Sciences*, 5.

<https://doi.org/10.33735/phimisci.2024.11337>



- Deroy, O., & Auvray, M. (2013). A new Molyneux's problem: Sounds, shapes and arbitrary crossmodal correspondences. *UNILOG'2013, SHAPES*, 4(2), 61–70.
- Deroy, O., & Spence, C. (2013). Are we all born synaesthetic? Examining the neonatal synaesthesia hypothesis. *Neuroscience & Biobehavioral Reviews*, 37, 1240–1253. <https://doi.org/10.1016/j.neubiorev.2013.04.001>
- Di Stefano, N., & Spence, C. (2023). Perceptual similarity: Insights from the crossmodal correspondences. *Reviews in Philosophy & Psychology*. <https://doi.org/10.1007/s13164-023-00692-y>
- Diderot, D. (1977). Letter on the blind, for the benefit of those who see. In M. J. Morgan (Ed. & Trans.), *Molyneux's questions: Vision, touch and the philosophy of perception* (pp. 25–58). Cambridge University Press.
- Dionne-Dostie, E., Paquette, N., Lassonde, M., & Gallagher, A. (2015). Multisensory integration and child neurodevelopment. *Brain Sciences*, 5, 32–57. <https://doi.org/10.3390/brainsci5010032>
- Ehli, B. (2020). Leibniz and the Molyneux problem. *Journal of Modern Philosophy*, 2(1), 1–18. <https://doi.org/10.32881/jomp.58>
- Eilan, N. (1993). Molyneux's question and the idea of an external world. In N. Eilan, R. McCarthy, & B. Brewer (Eds.), *Spatial representation: Problems in philosophy and psychology* (pp. 236–255). Blackwell.
- Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, 7(5), 1–14. <https://doi.org/10.1167/7.5.7>
- Ettlinger, G. (1967). Analysis of cross-modal effects and their relationship to language. In F. L. Darley & C. H. Millikan (Eds.), *Brain mechanisms underlying speech and language*. Grune & Stratton.
- Evans, G. (1985). Molyneux's question. In A. Phillips (Ed.), *Gareth evans: Collected papers* (pp. 364–399). Clarendon Press.
- Farzin, F., Charles, E. P., & Rivera, S. M. (2009). Development of multimodal processing in infancy. *Infancy*, 14(5), 563–578. <https://doi.org/10.1080/15250000903144207>
- Filippetti, M. L., Orioli, G., Johnson, M. H., & Farroni, T. (2015). Newborn body perception: Sensitivity to spatial congruency. *Infancy*, 20, 455–465. <https://doi.org/10.1111/inf.12083>
- Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., & MacLeod, D. I. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience*, 6(9), 915–916. <https://doi.org/10.1038/nn1102>
- Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition*, 65(2-3), 263–297. [https://doi.org/10.1016/S0010-0277\(98\)00002-X](https://doi.org/10.1016/S0010-0277(98)00002-X)
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts.
- Gibson, E. J., & Walker, A. S. (1984). Development of knowledge of visual-tactual affordances of substance. *Child Development*, 55(2), 453–460.
- Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 67–82). Erlbaum.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.
- Glenney, B. (2012). Leibniz on Molyneux's question. *History of Philosophy Quarterly*, 29(3), 247–264.
- Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition*, 65(2-3), 231–262. [https://doi.org/10.1016/S0010-0277\(97\)00047-4](https://doi.org/10.1016/S0010-0277(97)00047-4)
- Goldstone, R. L., & Son, J. Y. (2012). Similarity. In K. J. Holyoak & R. G. Morrison (Eds.), *The oxford handbook of thinking and reasoning* (pp. 155–176). Oxford University Press.
- Goodman, N. (1972). Seven strictures on similarity. In *Problems and projects* (pp. 437–446). Bobs-Merril.
- Goodnow, J. J. (1971). Eye and hand: Differential memory and its effect on matching. *Neuropsychologia*, 9(1), 89–85. [https://doi.org/10.1016/0028-3932\(71\)90066-2](https://doi.org/10.1016/0028-3932(71)90066-2)
- Green, E. J. (2021). The puzzle of cross-modal shape experience. *Nous*, 56(4), 867–896. <https://doi.org/10.1111/nous.12384>
- Green, E. J. (2022). Representing shape in sight and touch. *Mind & Language*, 37(4), 694–714. <https://doi.org/10.1111/mila.12352>
- Gregory, R. L. (2003). Seeing after blindness. *Nature Neuroscience*, 6(9), 909–910. <https://doi.org/10.1038/nn0903-909>
- Guellaï, B., Callin, A., Bevilacqua, F., Schwarz, D., Pitti, A., Boucenna, S., & Gratier, M. (2019). Sensus communis: Some perspectives on the origins of non-synchronous cross-sensory associations. *Frontiers in Psychology*, 10(523). <https://doi.org/10.3389/fpsyg.2019.00523>
- Held, R., Ostrovsky, Y., Gelder, B., Gandhi, T., Ganesh, S., Mathur, U., & Sinha, P. (2011). The newly sighted fail to match seen with felt. *Nature Neuroscience*, 14(5), 551–553. <https://doi.org/10.1038/nn.2795>
- Helmholtz, H. L. (1878). *Treatise on physiological optics: Vol. II*. Dover Publications.
- Huber, E., Webster, J. M., Brewer, A. A., MacLeod, D. I., Wandell, B. A., Boynton, G. M., & Fine, I. (2015). A lack of experience-dependent plasticity after more than a decade of recovered sight. *Psychological Science*, 26(4), 393–401. <https://doi.org/10.1177/0956797614563957>
- Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences of the u.s.a.*, 106, 10382–10385. <https://doi.org/10.1073/pnas.0812142106>
- Jacobs, R., Serhal, C. B., & Steenberghe, D. (1998). Oral stereognosis: A review of the literature. *Clinical Oral Investigations*, 2(1), 3–10. <https://doi.org/10.1007/s007840050035>
- Jacomuzzi, A. C., Kobau, P., & Bruno, N. (2003). Molyneux's question redux. *Phenomenology and the Cognitive Sciences*, 2(4), 255–280. <https://doi.org/10.1023/B:PHEN.0000007370.68536.2d>
- Jastrow, J. (1886). The perception of space by disparate senses. *Mind*, 11, 539–554.
- Kassuba, T., Klinge, C., Hölig, C., Röder, B., & Siebner, H. R. (2013). Vision holds a greater share in visuo-haptic object recognition than touch. *Neuroimage*, 65, 59–68. <https://doi.org/10.1016/j.neuroimage.2012.09.054>
- Lacey, S., Pappas, M., Kreps, A., Lee, K., & Sathian, K. (2009). Perceptual learning of view-independence in visuo-haptic object representations. *Experimental Brain Research*, 198(2-3), 329–337. <https://doi.org/10.1007/s00221-009-1856-8>
- Lawson, K. R. (1980). Spatial and temporal congruity and auditory-visual integration in infants. *Developmental Psychology*, 16(3), 185–192. <https://doi.org/10.1037/0012-1649.16.3.185>

Spence, C., & Di Stefano, N. (2024). Old and new versions of the Molyneux question: A review of experimental answers *Philosophy and the Mind Sciences*, 5.

<https://doi.org/10.33735/phimisci.2024.11337>



© The author(s). <https://philosophymindscience.org> ISSN: 2699-0369

- Leibniz, G. W. F. (1981). *New essays on human understanding* (P. Remnant & J. Bennett, Eds.). Cambridge University Press.
- Levin, J. (2008). Molyneux's question and the individuation of perceptual concepts. *Philosophical Studies*, 139(1), 1–28.
- Levin, J. (2018). Molyneux's question and the amodality of spatial experience. *Inquiry*, 61(5–6), 590–610. <https://doi.org/10.1080/0020174X.2017.1372306>
- Lewkowicz, D. J., & Minar, N. J. (2014). Infants are not sensitive to synaesthetic cross-modality correspondences: A comment on walker et al. *Psychological Science*, 25(3), 832–834. <https://doi.org/10.1177/0956797613516011>
- Lewkowicz, D. J., & Turkewitz, G. (1980). Cross-modal equivalence in early infancy: Auditory-visual intensity matching. *Developmental Psychology*, 16(6), 597–607. <https://doi.org/10.1037/0012-1649.16.6.597>
- Locke, J. (1975). *An essay concerning human understanding* (P. H. Nidditch, Ed.). Oxford University Press.
- Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *The American Journal of Psychology*, 87(1–2), 173–188.
- Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities*. Academic Press.
- Marks, L. E. (1989). For hedgehogs and foxes: Individual differences in the perception of cross-modal similarity. In G. Ljunggren & S. Dornic (Eds.), *Psychophysics in action* (pp. 55–65). Springer Verlag.
- Marks, L. E., Hammeal, R. J., & Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development*, 52(215), 1–102. <https://doi.org/10.2307/1166084>
- Marks, L. E., Szczesiul, R., & Ohlott, P. (1986). On the cross-modal perception of intensity. *Journal of Experimental Psychology: Human Perception and Performance*, 12(4), 517–534. <https://doi.org/10.1037/0096-1523.12.4.517>
- Marlair, C., Pierret, E., & Crollen, V. (2021). Geometry intuitions without vision? A study in blind children and adults. *Cognition*, 216(104861). <https://doi.org/10.1016/j.cognition.2021.104861>
- Maurer, D., Stager, C. L., & Mondloch, C. J. (1999). Cross-modal transfer of shape is difficult to demonstrate in one-month-olds. *Child Development*, 70(5), 1047–1057. <http://www.jstor.org/stable/1132048>.
- McKyton, A., Ben-Zion, I., Doron, R., & Zohary, E. (2015). The limits of shape recognition following late emergence from blindness. *Current Biology*, 25(18), 2373–2378. <https://doi.org/10.1016/j.cub.2015.06.040>
- McRae, R. (1976). *Leibniz: Perception, apperception, and thought*. University of Toronto Press.
- Mellers, B., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception & Performance*, 8(4), 582–601. <https://doi.org/10.1037/0096-1523.8.4.582>
- Meltzoff, A. N. (1993). Molyneux's babies: Cross-modal perception, imitation and the mind of the preverbal infant. In N. Eilan, R. McCarthy, & B. Brewer (Eds.), *Spatial representation: Problems in philosophy and psychology* (pp. 219–235). Blackwell.
- Meltzoff, A. N., & Borton, R. W. (1979). Intermodal matching by human neonates. *Nature*, 282, 403–404. <https://doi.org/10.1038/282403a0>
- Meng, X., Chen, N., Ishida, J., Watanabe, K., & Murakami, T. (2023). Crossmodal correspondences between visual features and tastes in preschoolers: An exploratory study. *Frontiers in Psychology*, 14(1226661). <https://doi.org/10.3389/fpsyg.2023.1226661>
- Molina, M., & Jouen, F. (2003). Haptic intramodal comparison of texture in human neonates. *Developmental Psychobiology*, 42(4), 378–385. <https://doi.org/10.1002/dev.10111>
- Morgan, M. (1977). *Molyneux's question: Vision, touch and the philosophy of perception*. Cambridge University Press.
- Ngo, M. K., Velasco, C., Salgado, A., Boehm, E., O'Neill, D., & Spence, C. (2013). Assessing crossmodal correspondences in exotic fruit juices: The case of shape and sound symbolism. *Food Quality & Preference*, 28(1), 361–369. <https://doi.org/10.1016/j.foodqual.2012.10.004>
- Norman, J. F., Norman, H. F., Clayton, A. M., Lianekhammy, J., & Zielke, G. (2004). The visual and haptic perception of natural object shape. *Perception & Psychophysics*, 66(2), 342–351. <https://doi.org/10.3758/BF03194883>
- Ostrovsky, Y., Andelman, A., & Sinha, P. (2006). Vision following extended congenital blindness. *Psychological Science*, 17(12), 1009–1014. <https://doi.org/10.1111/j.1467-9280.2006.01827.x>
- Ostrovsky, Y., Meyers, E., Ganesh, S., Mathur, U., & Sinha, P. (2009). Visual parsing after recovery from blindness. *Psychological Science*, 20(12), 1484–1491. <https://doi.org/10.1111/j.1467-9280.2009.02471.x>
- Owens, J. (1982). Aristotle on common sensibles and incidental perception. *Phoenix*, 36(3), 215–236. <https://www.jstor.org/stable/1087890>.
- Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy? Evidence for sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*, 114(2), 173–186. <https://doi.org/10.1016/j.jecp.2012.05.004>
- Pêcheux, M.-G., Lepecq, J.-C., & Salzarulo, P. (1988). Oral activity and exploration in 1- to 2-month-old infants. *British Journal of Developmental Psychology*, 6, 245–256. <https://doi.org/10.1111/j.2044-835X.1988.tb01098.x>
- Piaget, J. (1937). *La naissance de l'intelligence chez l'enfant [the birth of the child's intelligence]*. Delachaux & Niestlé.
- Picard, D. (2007). Tactual, visual, and cross-modal transfer of texture in 5- and 8-year-old children. *Perception*, 36(5), 722–736. <https://doi.org/10.1068/p5575>
- Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W.-H. C., Cohen, L., Guazzelli, M., & Haxby, J. V. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *Proceedings of the National Academy of Sciences*, 101(5), 5658–5663. <https://doi.org/10.1073/pnas.0400707101>
- Piller, S., Senna, I., & O, E. M. (2023). Visual experience shapes the bouba kiki effect and the size weight illusion upon sight restoration from congenital blindness. *Scientific Reports*, 13(11435). <https://doi.org/10.1038/s41598-023-38486-y>
- Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, 132(3), 416–442. <https://doi.org/10.1037/0033-2909.132.3.416>
- Quine, W. V. (2000). Three networks: Similarity, implication, and membership. *The Proceedings of the Twentieth World Congress of Philosophy*, 6, 287–291.

Spence, C., & Di Stefano, N. (2024). Old and new versions of the Molyneux question: A review of experimental answers *Philosophy and the Mind Sciences*, 5.

<https://doi.org/10.33735/phimisci.2024.11337>



© The author(s). <https://philosophymindscience.org> ISSN: 2699-0369

- Reeve, C. D. C. (2004). *Plato. The republic*. Hackett.
- Ricci, D., Cesarini, L., Groppo, M., Carli, A., Gallini, F., Serrao, F., Fumagalli, M., Cowan, F., Ramenghi, L. A., Anker, S., Mercuri, E., & Mosca, F. (2008). Early assessment of visual function in full term newborns. *Early Human Development*, *84*(2), 107–113. <https://doi.org/10.1016/j.earlhumdev.2007.03.010>
- Root, R. T., & Ross, S. (1965). Further validation of subjective scales for loudness and brightness by means of cross-modality matching. *American Journal of Psychology*, *78*, 285–289.
- Rose, S. A., Gottfried, A. W., & Bridger, W. H. (1981). Cross-modal transfer in 6-month-old infants. *Developmental Psychology*, *17*(5), 661–669. <https://doi.org/10.1037/0012-1649.17.5.661>
- Ryle, G. (1939a). Plato's Parmenides (I). *Mind*, *48*(190), 129–151.
- Ryle, G. (1939b). Plato's Parmenides (II). *Mind*, *48*(191), 302–325.
- Sann, C., & Streri, A. (2007). Perception of object shape and texture in human newborns: Evidence from cross-modal transfer tasks. *Developmental Science*, *10*(3), 399–410. <https://doi.org/10.1111/j.1467-7687.2007.00593.x>
- Schwenkler, J. (2012). On the matching of seen and felt shape by newly sighted subjects. *Perception*, *3*(3), 186–188. <https://doi.org/10.1068/i0525ic>
- Schwenkler, J. (2013). Do things look the way they feel? *Analysis*, *73*(1), 86–96. <http://www.jstor.org/stable/23359165>.
- Schwenkler, J. (2019). Molyneux's question within and across the senses. In T. Cheng, O. Deroy, & C. Spence (Eds.), *Spatial senses: Philosophy of perception in an age of science* (pp. 274–283). Routledge.
- Segundo-Ortin, M., & Hutto, D. D. (2021). Similarity-based cognition: Radical enactivism meets cognitive neuroscience. *Synthese*, *198*(5), 5–23. <https://doi.org/10.1007/s11229-019-02505-1>
- Sidhu, D. M., Athanasopoulou, A., Archer, S. L., Czarnecki, N., Curtin, S., & Pexman, P. M. (2023). The maluma/takete effect is late: No longitudinal evidence for shape sound symbolism in the first year. *PLoS ONE*, *18*(11).
- Šikl, R., Šimeček, M., Porubánová-Norquist, M., Bezdíček, O., Kremláček, J., Stodulka, P., Fine, I., & Ostrovsky, Y. (2013). Vision after 53 years of blindness. *Perception*, *4*(8), 498–507. <https://doi.org/10.1068/i0611>
- Slater, A., & Kirby, R. (1998). Innate and learned perceptual abilities in the newborn infant. *Experimental Brain Research*, *123*, 90–94. <https://doi.org/10.1007/s002210050548>
- Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology*, *24*(1), 99–142. [https://doi.org/10.1016/0010-0285\(92\)90004-L](https://doi.org/10.1016/0010-0285(92)90004-L)
- Sourav, S., Kekunnaya, R., Shareef, I., Banerjee, S., Bottari, S., & Röder, B. (2019). A protracted sensitive period regulates the development of cross-modal sound–shape associations in humans. *Psychological Science*, *30*(10), 1473–1482. <https://doi.org/10.1177/0956797619866625>
- Speed, L. J., Croijmans, I., Dolscheid, S., & Majid, A. (2021). Crossmodal associations with olfactory, auditory, and tactile stimuli in children and adults. *I-Perception*, *12*(6), 1–34. <https://doi.org/10.1177/204166952111048513>
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971–995. <https://doi.org/10.3758/s13414-010-0073-7>
- Spence, C. (2022). Exploring group differences in the crossmodal correspondences. *Multisensory Research*, *35*(6), 495–536. <https://doi.org/10.1163/22134808-bja10079>
- Spence, C., & Di Stefano, N. (2023). Sensory translation between audition and vision. *Psychonomic Bulletin & Review*, 1–28. <https://doi.org/10.3758/s13423-023-02343-w>
- Spence, C., & Di Stefano, N. (2024). What, if anything, can be considered an amodal sensory dimension? *Psychonomic Bulletin & Review*. <https://doi.org/10.3758/s13423-023-02447-3>
- Stevens, J. C., & Marks, L. E. (1965). Cross-modality matching of brightness and loudness. *Proceedings of the National Academy of Sciences*, *54*(2), 407–411. <https://doi.org/10.1073/pnas.54.2.407>
- Stevens, J. C., & Marks, L. E. (1980). Cross-modality matching functions generated by magnitude estimation. *Perception & Psychophysics*, *27*, 379–389. <https://doi.org/10.3758/BF03204456>
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, *64*(3), 153–181. <https://doi.org/10.1037/h0046162>
- Stevens, S. S. (1971). Issues in psychophysical measurement. *Psychological Review*, *78*(5), 426–450. <https://doi.org/10.1037/h0031324>
- Streri, A. (1987). Tactile discrimination of shape and inter-modal transfer in 2- to 3-month-old infants. *British Journal of Developmental Psychology*, *5*(3), 213–220. <https://doi.org/10.1111/j.2044-835X.1987.tb01056.x>
- Streri, A., & Gentaz, E. (2003). Cross-modal recognition of shape from hand to eyes in human newborns. *Somatosensory and Motor Research*, *20*(1), 13–18. <https://doi.org/10.1080/0899022031000083799>
- Streri, A., & Gentaz, E. (2004). Cross-modal recognition of shape from hand to eyes and handedness in human newborns. *Neuropsychologia*, *42*(10), 1365–1369. <https://doi.org/10.1016/j.neuropsychologia.2004.02.012>
- Streri, A., & Hevia, M. D. de. (2023). How do human newborns come to understand the multimodal environment? *Psychonomic Bulletin & Review*, 1–16. <https://doi.org/10.3758/s13423-023-02260-y>
- Streri, A., & Pecheux, M.-G. (1986). Vision-to-touch and touch-to-vision transfer of form in 5-month-old infants. *British Journal of Developmental Psychology*, *4*(2), 161–167. <https://doi.org/10.1111/j.2044-835X.1986.tb01007.x>
- Teghtsoonian, R., & Teghtsoonian, M. (1970). Two varieties of perceived length. *Perception & Psychophysics*, *8*(6), 389–392. <https://doi.org/10.3758/BF03207030>
- Todd, J. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences*, *8*(3), 115–121. <https://doi.org/10.1016/j.tics.2004.01.006>
- Trehub, S. E. (2012). Behavioral methods in infancy: Pitfalls of single measures. *Annals of the New York Academy of Sciences*, *1252*(1), 37–42. <https://doi.org/10.1111/j.1749-6632.2012.06448.x>
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*(4), 327–352. <https://doi.org/10.1037/0033-295X.84.4.327>
- Ueda, Y., & Saiki, J. (2007). Viewpoint independence in visual and haptic object recognition. *The Japanese Journal of Psychonomic Science*, *26*, 11–19.

Spence, C., & Di Stefano, N. (2024). Old and new versions of the Molyneux question: A review of experimental answers *Philosophy and the Mind Sciences*, *5*.

<https://doi.org/10.33735/phimisci.2024.11337>



© The author(s). <https://philosophymindscience.org> ISSN: 2699-0369

- Ueda, Y., & Saiki, J. (2012). Characteristics of eye movements in 3-d object learning: Comparison between within-modal and cross-modal object recognition. *Perception*, 41(11), 1289–1298. <https://doi.org/10.1068/p7257>
- Van Cleve, J. (2007). Reid's answer to Molyneux's question. *The Monist*, 90(2), 251–270.
- Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synesthetic cross-modality correspondences. *Psychological Science*, 21(1), 21–25. <https://doi.org/10.1177/0956797609354734>
- Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2014). Preverbal infants are sensitive to cross-sensory correspondences: Much ado about the null results of Lewkowicz and Minar. *Psychological Science*, 25(3), 835–836. <https://doi.org/10.1177/0956797613520170>
- Walker-Andrews, A. (1994). Taxonomy for intermodal relations. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 39–56). Lawrence Erlbaum.
- Waterman, J., C. N. (1917). Hand-tongue space perception. *Journal of Experimental Psychology*, 2(4), 289–294. <https://doi.org/10.1037/h0075127>
- Wilson, W. A. (1969). Cross-modal transfer in animals. In *Paper read at the XIXth international congress of psychology*.
- Zhou, Y. D., & Fuster, J. M. (2000). Visuo-tactile cross-modal associations in cortical somatosensory cells. *Proceedings of the National Academy of Sciences U.S.A.*, 97(17), 9777–9782. <https://doi.org/10.1073/pnas.97.17.9777>

Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

