



Published in final edited form as:

Optom Vis Sci. 2009 June ; 86(6): 595–598. doi:10.1097/OPX.0b013e3181a72999.

Visual-Haptic Mapping and the Origin of Crossmodal Identity

Richard Held, PhD

Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts

Abstract

We found that the congenitally blind person who gains sight initially fails to identify seen objects with their felt versions: a negative answer to the Molyneux question. However, s(he) succeeds in doing so after a few days of sight. We argue that this rapid learning resembles that of adaptation to rearrangement in which the experimentally produced separations of seen and felt perceptions of objects are rapidly reunited by the process called capture. Moreover, the original ability to identify objects across modalities by the neonate may be assured by the same process.

Keywords

crossmodal; sight; haptics; prism adaptation; infant performance

The Question

In 1688 William Molyneux sent a message¹ to the philosopher John Locke asking: “Suppose a man born blind, and now adult, and taught by his touch to distinguish between a cube and a sphere of the same metal ... Suppose then the cube and sphere placed on a table, and the blind man be made to see: query, whether by his sight, before he touched them he could now distinguish and tell which is the globe, which the cube ...?”

Taken literally in contemporary terms, will the newly sighted, formerly blind patient identify by sight objects s(he) has experienced only by touch and feel (haptics) ? Although the history of cases of recovery from blindness goes back to antiquity² and continues in recent times^{3,4}, Molyneux's question has not yet been convincingly answered. The reason for this lack is failure to meet the stringent requirements for definitively answering the question. First of all, the blindness must have been verifiably congenital and continuous. Otherwise there may have been an opportunity for acquisition of the crossmodal transfer through visual experience whose exclusion is the purpose of testing the previously blind. When still blind the patient should demonstrate light sense, in order to assure function of the retina, optic nerve and beyond, but resolution no better than light and dark. After successful surgery the patient must exhibit acuity sufficient to discriminate visually among the objects used for testing. Selected cases of blindness from occlusion of both eyes, such as may be caused by occlusive cataracts or corneal scarring, best satisfy these conditions. Post-op testing should begin as soon after surgery as possible - ideally when bandages are first removed. Crossmodal recognition has been reported in some cases long after surgery. But if it has not been tested immediately we cannot exclude the possibility of acquisition by experience and, as we have found (read below) acquisition seems to occur within days after return of normal conditions of seeing.

Recently the Prakash group, defined in reference ⁵, reported the outcome of tests performed in India and specifically designed to answer the classic Molyneux question.⁶ They managed to conduct critical tests on three successfully operated patients. Not only did they satisfy requirements specified above, but also the structure of the tests further assured compliance with the requirements.

Three two-alternative forced choice tests were administered. Initially one of two target stimuli, either a small object to be palpated but not seen or the object seen but not touched, was presented within arm's length for a short period of familiarization. Following immediately upon removal of the haptic target, either a pair of haptic or of visible objects, one of which was the target, the other novel, was presented and the patient instructed to choose which one of the two s(he) had previously experienced. Following removal of the visible target a pair of such objects, one of which was the target were presented for choice. The crossmodal choice tested crossmodal transfer. The two intramodal choices tested the discriminability of the stimuli in each modality as well as the patient's understanding of the task. As reported at VSS 2008 ⁶ all three patients performed at chance on the crossmodal choice but were clearly capable of the intramodal discriminations. Hence, the answer to our version of the Molyneux question is "no" within the confidence limits of the procedures. However, a few days after the patients were released the tests were repeated yielding the result of crossmodal transfer better than chance. We are left with the surprising finding of acquisition of the ability to perform the cross-modal task within a few days after surgery.

The rapidity of acquisition of crossmodal transfer implies that we must attribute the result to some process, which acts quickly so as to allow the newly sighted to perceive the identity between seen and felt impressions of objects. We conjecture that this same process could operate in the normally sighted infant from birth so as to enable the crossmodal identification reported in an extensive review to occur as early as two months of age ⁷ (but see below). In the following we try to establish the case that the sought-for process is the same as that which accounts for adaptation to perturbations of visual-haptic spatial mapping.

Visual-Haptic Mapping

What is involved in performing the crossmodal transfer? Haptic and visual sensorineural channels are independent at early processing levels but at later levels coupling between their signals must be established to account for transfer when it occurs. It is the origin of this coupling (innate vs empirical) that has always underlain recurrent raising of the Molyneux question. Since, apart from time, spatial loci are the simplest perceptual properties common to both vision and haptics, it would seem that the coupling must be based on congruency of spatiotemporal signals. Congruency could in turn be determined from a map of haptic on visual spatiotemporal signals. The following argues for a process by which this mapping may be accomplished.

The performer of the crossmodal task first palpates the invisible object to gain its haptic representation (neuronal) in terms of a sequence of discrete movements with their tactual and proprioceptive accompaniments. The haptic representation must in turn elicit a visual representation for comparison with a visual target. We infer that in memory haptic representations are mapped against congruent visual representations that once activated may be compared with the visible samples to establish either identity or difference. The same must happen in opposite sequence when viewing an object without touching it and then recognizing it by touch. The seen and felt representations mapped together are then the presumed basis of crossmodal identification. According to our findings the congenitally blind lack such mappings. But they must rapidly develop after vision is achieved. Since it is the hand and its movements that create the haptic representation, one might as well restrict attention to the hand itself and consider the palpated representations of objects to be derived from a concatenation

of the hand's positions and movements in space. Consideration is then reduced to that of how the hand itself achieves and retains the map of its seen and felt representations. This map should also account for the control of positioning and repositioning of the hand in body space as in reaching, for example, specification of the goal of a reach would be the mapped pair of representations of the hand at some remove - the goal of the reach -- from that of its initial position. One way of studying the mapping process is to perturb the habitual map and to observe the conditions under which recovery is achieved. These conditions are then candidates for accounting for the original mapping as it may occur in cases of recovered vision and perhaps in infancy.

Perturbing the Map

Normally the habitual seen-felt mapping is such that a visual stimulus can elicit an accurate reach (or other response) to the locus of the source of the stimulus. But it is easy to alter this orderly relation by any number of means: optical, mechanical, and electronic. Accordingly, for more than a century the use of procedures that do so has constituted a small industry of experiments known variously as sensory recombinations, rearrangements, perceptual compensation, adaptation of space perception etc, (see Welch⁸ for an excellent but dated review). Rearrangement experiments have been performed in vision, hearing, and haptic sense modalities. The classic rearrangements are optical. Over a hundred years ago George Stratton studied the consequences of optically rotating his field of vision by 180 deg.^{9,10} He also wore a mirror device designed to have him view his own body from above¹¹. In time, at least some of the disturbing consequences, such as mislocalizations, were reduced if not eliminated and body parts were perceived as outside his own body. Stratton's findings were recently replicated using video systems to displace the subject's view of his own body^{12,13}. Even earlier in time,¹⁴ von Helmholtz discussed the wedge prism type of rearrangement that produces a lateral displacement of the image of the hand with erroneous localization followed in time by correction. This type of rearrangement was made popular beginning in the 1960s following which innumerable variations were performed⁸ and continue to be performed. Mechanical analogues of the Helmholtz prism experiment have proliferated¹⁵⁻¹⁸, as have electronic analogues^{12,13,19,20,21}. The essential feature of these types of perturbation is the production of a visible stimulus spatially displaced but closely coupled in time with haptic stimulation from a body part, often a limb.

Consider the classic Helmholtz experiment of seeing through refracting wedge prisms, which displace images on the retinae. It is as if the images of objects were displaced around the head without moving the objects. The observer, who may be initially unaware of the change, uses his habitual mapping of seen and felt versions of objects with the result that reaching for a visible target *without sight* of the hand (open loop), will be in error by approximately the prism-induced optical displacement. This displacement error has been widely termed a *discrepancy* in the literature⁸ and regarded as the cause of consequent adaptation by error correction without much further explanation. But a further step in this experiment is revelatory of how compensation really occurs. The moment the observer sees his hand through the prism the spatial separation of seen and felt loci is either not sensed at all or is at least reduced from what might be predicted on the basis of physical separation²². Despite the prism-induced geometric separation of haptic and visual signals, the immediate effect of viewing the hand through a wedge prism tends to eliminate the discrepancy that might be predicted on the basis of physical separation. Even very brief exposures to near-simultaneous seen and felt signals from the hand/arm remap their combination so as to restore correct localization. Moreover, a similar case can be made for a perturbation of vision that alters the shape of objects as in the experiments of Rock et al,²³. The prism-produced displacement supplies a parameter value, which remaps the entire range of representation of the felt hand in body-centered space²⁴. Harris and others²⁵ have attributed these changes to an altered position sense of the limb leaving open the question

of whether it is a cause or an effect. Since the haptic localization tends to approximate the visual, it is referred to as “captured” after Tastevin¹⁸. When complete capture is achieved, haptic and visual representations tend to be congruent and reaching for a visible target or other mode of visual-haptic localization becomes accurate. In summary, this observation implies that spatial *discrepancy*, instead of being just an error signal, actually supplies the parameter value for remapping the relation between visual and haptic space thereby restoring spatial congruence between the perturbed sensory systems. The term near-simultaneous can be quantified from the effects of time delay on adaptation^{19,26}.

Partial Capture

In most rearrangement experiments the capture-displacement reported is partial, a consequence, we believe, of the residual influence of the previous habitual mapping between vision and haptics. There appears to be an averaging process between the full capture remapping and the previous habitual mapping. This account is borne out by several observations. The aftereffect, following removal of the prism and return to habitual conditions of seeing, can be taken as evidence for persistence of the prism-induced remapping however short-lived it may be. Just as the habitual mapping persists during the remapping consequence of prism exposure, so the remapping now persists after prism removal and continues to be averaged with the habitual mapping. With continued exposure to this condition, the new mapping becomes stronger, the effective adaptation increases, and the aftereffects will more and more clearly outlast the generating condition on removal of the perturbation. When experimenters (and subjects) are sufficiently persistent, full and exact adaptation can be produced as in some of the experiments of Ivo Kohler, Hein, Held & Bossom, Mikaelian & Held²⁷⁻³⁰ and others. In a few cases dual localizations - one showing complete adaptation, the other zero -- have been reported^{28,31}. Finally, the adaptive shift never significantly exceeds the initial differential between visual and haptic loci. The discrepancy between visual and haptic representations determines the extent of the remapping, which may be reduced by the influence of prior mapping, but never significantly increased beyond full correction of the induced error.

Adaptation

Capture is the prototype of adaptation to rearrangement. Analysis of many of the various types of rearrangement reveals that after the initial sight of the rearranged scene, continuing exposure constitutes iteration of the capture conditions. Many prism experiments have repeated short-lived conjunctions of visual and haptic stimulations^{16,32} which are, in effect, capture opportunities. Others entail continuous movement of arm and hand. Since the hand movements, as observed by the performer, obligatorily entail a conjunction of near-simultaneous changes in vision and proprioception they also constitute a series of capture opportunities. Iteration of movements with their simultaneous paired signals constitutes synchronous signals. Active movement ensures synchronicity among the sensory accompaniments of movement: visual, tactual (haptic), and auditory. Desynchronization by passively imposed movement³³, time delay²¹, mechanical interference³⁴ or other source of noise¹³ reduces or eliminates adaptation. It has also been pointed out that active movement entails a more selective proprioceptive response than does passive movement with a probable decrease in noise.³⁵

The Answer

Now it may be obvious how the process that adapts the perturbation-induced discrepancy may account for the rapid acquisition of crossmodal mapping. For the blind there are no visual representations. On acquiring vision and returning to normal activities -- viewing the hand for example -- sources of visual stimulation should rapidly establish neural representations in the normal course of post-surgical activity. Here we must recognize that such representations may

very well not be as well calibrated as the habitually available ones involved in prism adaptation. Nonetheless we press on. Initially there should be no map of combined visual-haptic space. Just as in prism perturbation, the spatial representations are likely to be discrepant and crossmodal identification absent. But the crossmodal interaction via the process we have explored above, especially capture, should quickly map common loci for objects detected by both senses. In the case of the former patient there can be no previous mapping. Hence the new mapping should be accurate and serve as a basis for the crossmodal matching capability.

The case for acquisition by the neonate is similar. However, recently Streri and Gentaz³⁶ in ingenious studies find that newborns show crossmodal recognition of shape from hand to eyes and, consequently, claim that they have demonstrated amodal perception of shape and have answered “yes” to the Molyneux question. However, the authors mention a caveat to their conclusion whose interpretation they characterize as “very difficult”: that is, the crossmodal transfer is demonstrable only with the right hand although both hands were tested. Perhaps another caveat is in order. What the authors call a “newborn” is in fact 54 hours old on average. Consider the propensity of infants to assume the ATNR (asymmetrical tonic neck reflex) posture placing a hand in front of their eyes. Given the speed of capture and its hand/arm selectivity, the conditions are ripe for an interpretation of the infants' crossmodal transfer using the right hand in terms of the mapping process discussed above. Moreover, this interpretation would relieve the embarrassment of having to claim that an amodal perception is specific to one hand. But irrespective of the newborn's native capacity, the adaptation process should at least provide an assurance of correct performance by both hands.

ACKNOWLEDGMENTS

Funding was received for this work from the the National Institutes of Health; NIH R21EU015521

REFERENCES

1. Morgan, MJ. Molyneux's Question: Vision, Touch, and the Philosophy of Perception. Cambridge University Press; Cambridge: 1977.
2. von Senden, M. Space and Sight: The Perception of Space and Shape in Congenitally Blind Before and After Operation. Methuen; London: 1960.
3. Fine I, Smallman HS, Doyle P, MacLeod DI. Visual function before and after the removal of bilateral congenital cataracts in adulthood. *Vision Res* 2002;42:191–210. [PubMed: 11809473]
4. Ostrovsky Y, Andalman A, Sinha P. Vision following extended congenital blindness. *Psychol Sci* 2006;17:1009–14. [PubMed: 17201779]
5. Mandavilli A. Visual neuroscience: look and learn. *Nature* 2006;441:271–2. [PubMed: 16710386]
6. Held R, Ostrovsky Y, deGelder B, Sinha P. Revisiting the molyneux question. *J Vis* 2008;8:523a. Available at: <http://www.journalofvision.org/8/6/523/>. Accessed March 11, 2009
7. Kellman, PJ.; Arterberry, ME. *The Cradle of Knowledge: Development of Perception in Infancy*. MIT Press; Cambridge, MA: 1998.
8. Welch, RB. Adaptation to space perception. In: Boff, KR.; Kaufman, L.; Thomas, JP., editors. *Handbook of Perception and Performance, vol 1: Sensory Processes and Perception*. Wiley-Interscience; New York: 1986. p. 24.1-24.45.
9. Stratton GM. Upright vision and the retinal image. *Psychol Rev* 1897;4:182–7.
10. Stratton GM. Some preliminary experiments on vision without inversion of the retinal image. *Psychol Rev* 1896;3:611–7.
11. Stratton GM. The spatial harmony of touch and sight. *Mind* 1899;8:492–505.
12. Ehrsson HH. The experimental induction of out-of-body experiences. *Science* 2007;317:1048. [PubMed: 17717177]
13. Lenggenhager B, Tadi T, Metzinger T, Blanke O. Video ergo sum: manipulating bodily self-consciousness. *Science* 2007;317:1096–9. [PubMed: 17717189]

14. von Helmholtz, H. Helmholtz's Treatise on Physiological Optics. Vol. 3rd. German ed.. Southall, JPC., editor. Dover Publications; New York: 1962.
15. Botvinick M, Cohen J. Rubber hands 'feel' touch that eyes see. *Nature* 1998;391:756. [PubMed: 9486643]
16. Lackner JR. Adaptation to displaced vision: role of proprioception. *Percept Mot Skills* 1974;38:1251–6. [PubMed: 4424453]
17. Nielsen TL. Volition: a new experimental approach. *Scand J Psychol* 1963;4:225–30.
18. Tastevin J. En partant de l'expérience d'aristote. *L'Encephale* 1937;1:57–84. 140–58.
19. Held R, Efstathiou A, Greene M. Adaptation to displaced and delayed visual feedback from the hand. *J Exper Psychol* 1966;72:887–91.
20. Smith, KU.; Smith, WK. Perception and Motion: An Analysis Of Space-Structured Behavior. Saunders; Philadelphia: 1962.
21. Held, R.; Durlach, N. Telepresence, time delay and adaptation. In: Ellis, SR., editor. Pictorial Communication in Virtual and Real Environments. Vol. 2nd ed.. Taylor & Francis; New York: 1993. p. 232-46.
22. Hay JC, Pick HL Jr, Ikeda K. Visual capture produced by prism spectacles. *Psychonom Sci* 1965;2:215–6.
23. Rock I, Mack A, Adams L, Hill AL. Adaptation to contradictory information from vision and touch. *Psychonom Sci* 1965;3:435–6.
24. Bedford F. Constraints on learning new mapping between perceptual dimensions. *J Exper Psychol Hum Percep Perform* 1989;15:517–30.
25. Harris, CS. Insight or out of sight? Two examples of perceptual plasticity in the human adult. In: Harris, CS., editor. Visual Coding and Adaptability. L. Erlbaum Associates; Hillsdale, NJ: 1980. p. 95-149.
26. Held, R.; Durlach, N. Telepresence, time delay and adaptation. In: Ellis, SR., editor. Pictorial Communication in Virtual and Real Environments. Vol. 2nd ed.. Taylor & Francis; New York: 1993. p. 232-46.
27. Kohler, I. The Formation and Transformation of the Perceptual World. International Universities Press; New York: 1964.
28. Held R, Bossom J. Neonatal deprivation and adult rearrangement: complementary techniques for analyzing plastic sensory-motor coordinations. *J Comp Physiol Psychol* 1961;54:33–7. [PubMed: 13713069]
29. Mikaelian H, Held R. Two types of adaptation to an optically-rotated visual field. *Am J Psychol* 1964;77:257–63. [PubMed: 14141479]
30. Hein, A. Acquiring components of visually guided behavior. In: Pick, A., editor. Minnesota Symposium on Child Development. Vol. vol. 6. University of Minnesota Press; Minneapolis, MN: 1972. p. 53-68.
31. Held R. Shifts in binaural localization after prolonged exposure to atypical combinations of stimuli. *Am J Psychol* 1955;68:526–48. [PubMed: 13275600]
32. Bedford FL. Perceptual and cognitive spatial learning. *J Exp Psychol Hum Percept Perform* 1993;19:517–30. [PubMed: 8331313]
33. Held R, Hein A. Adaption to disarranged hand-eye coordination contingent upon reafferent stimulation. *Percept Motor Skills* 1958;8:87–90.
34. Abplanalp, P.; Held, R. Effects of de-correlated visual feedback on adaptation to wedge prisms. Eastern Psychological Association meeting; Atlantic City, NJ. April 1965;
35. Lackner JR. Adaptation to visual and proprioceptive rearrangement: origin of the differential effectiveness of active and passive movements. *Percept Psychophys* 1977;21:55–9.
36. Streri A, Gentaz E. Cross-modal recognition of shape from hand to eyes and handedness in human newborns. *Neuropsychologia* 2004;42:1365–9. [PubMed: 15193944]