

EDITORIAL

Making sense of multisensory integration

Sophie Molholm and John J. Foxe

The Cognitive Neurophysiology Laboratory, Departments of Pediatrics and Neuroscience, Children's Evaluation and Rehabilitation Center (CERC), Albert Einstein College of Medicine, Van Etten Building – 1C, 1300 Morris Park Avenue, Bronx, NY 10461, USA

A pseudo-aquatic parasite, voiceless as a fish, yet constructing within itself an instrument of voice against the time when it will talk. Organs of skin, ear, eye, nose, tongue, superfluous all of them in the watery dark where formed, yet each unhaltingly preparing to enter a daylight, airy, object-full manifold world which they will be wanted to report on. A great excrescence at one end of a nerve-tube, an outrageously outsized brain, of no avail at the moment but where the learning of a world which is to be experienced will go forward.

– Sherrington (1940)

Apart from the obvious appeal of his prose, which must surely give rise to an element of regret at the rather stolid constraints of modern science writing, one hears in this quotation of Sherrington's his profound fascination for the very nature of sensory development in the human fetus. Elsewhere in his wonderful book 'Man on his Nature' he turns specifically to the issue of multisensory integration: "The naïve would have expected evolution in its course to have supplied us with more various sense organs for ampler perception of the world...The policy has rather been to bring by the nervous system the so called 'five' into closer touch with one another...A central clearing house of sense has grown up...Not new senses, but better liaison between old senses is what the developing nervous system has in this respect stood for".

Clearly scientists have pondered the seamlessness of liaison between the senses since the very early days of modern neuroscience. And it is surely one of the great remaining scientific puzzles just how it is that signals from such completely separated and wholly dissimilar sensory epithelia as the haircells of the cochlea, the photoreceptors of the retina and the corpuscles of the skin can be integrated centrally to form such a seamless unitary perceptual world. Our own predecessors here at The Albert Einstein College of Medicine were already studying the development of intersensory abilities in pediatric populations in the 1960s (e.g. Birch & Lefford, 1963, 1967). Yet despite these early research efforts, it is also fair to say that multisensory research was not at the forefront of the scientific effort across the early decades of the 20th century. Given the complexities of working with isolated sensory systems, many surely turned away from these issues because of the seeming intractability of trying to control more than just one system. But multisensory research has evolved considerably over the past 25 years. For this we owe a large debt to a few stalwarts of the field who forged ahead with this work when there was little to go on and when there was as yet limited interest in elucidating multisensory influences on information processing (e.g. Stein & Arigbede, 1972; Seltzer & Pandya, 1980; King & Palmer, 1985; see Stein & Meredith, 1993). From their extensive work on the multisensory properties of cells in the superior colliculus, Stein and Meredith defined a set of governing principles of multisensory integration at the cellular level.

By providing a set of hypotheses to test in a domain for which we are still considerably lacking in solid models, these principles have given important fodder to those examining multisensory integration under different stimulus and task configurations. Their substantial role in generating research questions is clearly signified not only by the extraordinary number of citations to their work but also by the many studies aimed at challenging these principles (see Foxe, 2008). Even before there was hard evidence of anatomical plausibility, we and others made the case that multisensory integration was not deferred to higher-order processing areas, but also occurred at very early stages of cortical information processing in a feedforward manner and in regions considered to be unisensory (Molholm *et al.*, 2002; Schroeder & Foxe, 2002). This commenced a marked change in how multisensory influences on neural processing are conceptualized. Falchier *et al.* (2002) documented monosynaptic connectivity between hierarchically early unisensory cortical areas, supplying anatomical plausibility for this thesis (see also Rockland & Ojima, 2003), and a significant volume of psychophysical work, much from the laboratories of John Driver and of Charles Spence, but many others as well, has made absolutely clear the ubiquity of multisensory influences on perception (e.g. Driver & Spence, 2000; Spence & Squire, 2003).

In this special issue we see that the question of when and where cortical multisensory integration first occurs remains a topic of considerable interest. Using magnetoencephalographic recordings, Raj *et al.* (2010) examined the temporal dynamics of cross-sensory activation and auditory–visual multisensory integration in human primary auditory and visual cortices. With this approach they find activation of visual cortex in response to auditory stimuli and activation of auditory cortex in response to visual stimuli. They show that these 'cross-sensory' activations are delayed by 30–40 ms relative to the onset of responses to the native inputs of these same sensory regions (i.e. auditory cortical responses to auditory input). Further, they show that multisensory integration occurs in sensory cortical areas only a few milliseconds following the onset of cross-sensory activity. Raj *et al.* (2010) make the case that this cross-sensory activity reflects transmission between the sensory cortical regions, as opposed to inputs from subcortical regions or higher-order association areas. However, this argument from timing, while compelling, cannot definitively resolve this issue, and the transmission pathways leading to early cross-sensory activation of sensory cortex remain to be firmly established. This may be addressed with the use of additional analytic techniques than have been applied thus far, or more probably require the use of methodologies that can reveal causal links combined with the timecourse of involvement, such as transcranial magnetic stimulation (TMS). In this issue we find two papers in which direct manipulation of neural processing in a cortical region is deployed to understand the functional architecture of

a particular process. With the use of anodal transcranial direct current stimulation, Bolognini *et al.* (2010) show that inducing an increase in the excitability of neurons in right posterior parietal cortex facilitates spatial orienting toward both auditory and visual stimuli presented to the contralateral field. This is consistent with a key role for posterior parietal cortex in the deployment of so-called supramodal spatial attention. Using single-pulse TMS to produce quite the opposite effect, Pasalar *et al.* (2010) show that the posterior parietal cortex near the junction of the anterior intraparietal sulcus and the postcentral sulcus is involved in the integration of visual and tactile information. Disruption of neuronal activity in this region leads to a decrease in multisensory facilitation whilst leaving unisensory performance intact.

In addition to neurophysiological and neuroanatomical studies spearheading changes in how we think about the multisensory sciences and the mediating neural mechanisms, behavioral studies have been key to gaining insight into the consequences of the multisensory brain. Welch & Warren (1980, 1986) proposed the modality-appropriateness hypothesis to explain why information conveyed in certain sensory modalities seems to exert a disproportionate influence on certain types of judgments. They argued that particular sensory systems best conveyed specific types of information and accordingly dominated related judgments under multisensory conditions, and even exerted an influence when ancillary to the task. For example, auditory stimuli are reliably mislocalized toward a concurrently presented but task-irrelevant visual stimulus, reflecting the dominance of visual stimulation in spatial judgment (see e.g. Warren *et al.*, 1981). More recent work has generated an appealing explanation of sensory dominance that differs from this original thesis only in that the dominant sensory modality can be altered by manipulating the integrity, or quality, of the respective multisensory signals (Ernst & Banks, 2002). Here it is argued that it is the reliability of the estimates that the signal provides that affect its relative contribution to performance. Of course this fits well with the notion that there is sensory-specific dominance for some types of judgments as, for example, it is seldom that an auditory signal will provide better spatial information than a visual signal. This can be well-accounted for using Bayesian modeling or variants thereof. In this issue Fetsch, DeAngelis and Angelaki describe how Bayesian decision theory accounts for the contributions of visual and vestibular information to estimations of one's translation in space (Fetsch *et al.*, 2010). Perception of heading is intrinsically multisensory insofar as presenting both visual and vestibular inputs tend to result in superior performance compared to presenting just one. In their review, Fetsch and colleagues describe corresponding cellular activity that acts in accordance with behavioral measures of visual-vestibular cue integration. Along a similar vein, in this issue van Wanrooij *et al.* (2010) demonstrate that implicit expectations impact the degree to which ancillary multisensory information is used to facilitate performance. In their study subjects made heading movements to the perceived location of a visual stimulus along the vertical plane in meridian space. In blocks of trials in which ancillary auditory stimuli were always spatially aligned with the visual stimuli, reaction times were speeded compared to visual-alone presentations. In contrast, in blocks of trials in which the auditory stimuli were misaligned on 50% or 90% of trials, reaction time facilitation was not seen; follow-up analysis of sequential effects showed that reaction times for spatially congruent trials were faster when preceded by a spatially congruent trial compared to when preceded by a spatially incongruent trial, and that there was no effect of the preceding trial on reaction time for the spatially incongruent trials. These data show a dynamic system that is continually updated as a function of the statistical probabilities of the

local environment. This signifies an adaptive perceptual-cognitive system that continues to be highly flexible even into adulthood.

Such a flexible system precludes the possibility that multisensory processes are fully automatic, nor would one expect this to uniformly be the case. The question remains, however, as to just how malleable multisensory processes are, and which processes are more or less susceptible to higher-order influences such as attention and learning. Most of the foundational physiological work on multisensory integration has been done in anesthetized animals and it is clear that significant multisensory integration proceeds in the absence of attention (Stein & Meredith, 1993). Yet data from human scalp electrical recordings (Talsma *et al.*, 2007) suggest that even the earliest evidence of cortical audiovisual effects, which one might expect to proceed in a fully bottom-up manner, can be affected by attention. Attention effects have been demonstrated for a number of multisensory situations, including integration of seen and heard speech (e.g., Soto-Faraco *et al.*, 2004; Alsius *et al.*, 2005). In this issue, Heron *et al.* (2010) show that attention to the temporal relationship between lagged auditory and visual inputs affects the degree of temporal adaptation, and thus appear to resolve some discrepancies in the literature on multisensory temporal adaptation effects (Fujisaki *et al.*, 2004; Vroomen *et al.*, 2004; Keetels & Vroomen, 2007; Hanson *et al.*, 2008a; Harrar & Harris, 2008). Subjects were exposed to audiovisual stimulus pairs with onsets that were consistently misaligned by 120 ms during an adaptation phase. When subjects engaged in a task in which they were to detect infrequent targets in which the temporal order of the stimuli differed, there was a substantial increase in the amount of temporal adaptation observed during test as compared to when subjects engaged in tasks in which they detected either infrequent changes in the luminance of a fixation cross or infrequent changes in the visual stimulus. Thus inconsistencies in the literature with regard to the magnitude of such recalibration effects may well be related to the task that is employed and whether or not it is related to the adapting feature.

In the 'unisensory' visual attention literature there is strong evidence for an inherent bias to process all of the features of an object, even when attention is explicitly focused on just a single feature (e.g., Stroop, 1935; Egly *et al.*, 1994; Martinez *et al.*, 2006). Until recently this had only been demonstrated for visual features, and an obvious next question was whether this principle would also apply to the multisensory features of an object. A series of recent studies has indeed shown that, when attention is directed toward just a single feature of a multisensory object, there is a strong bias to process the unattended features of that object in other sensory modalities, even when those features are completely irrelevant to the task at hand (Busse *et al.*, 2005; Molholm *et al.*, 2007; Fiebelkorn *et al.*, 2010a). In this issue, work from our own laboratory using event-related potentials (ERPs) suggests that this automatic spread of attention across sensory systems occurs only after the spatial boundaries of the visual object have first been resolved [Fiebelkorn *et al.*, 2010a, 2010b; this volume]. It remains to be seen whether this will turn out to be a fundamental principle of object recognition under differing multisensory conditions.

Kennett *et al.* (2001) made the observation that participants demonstrated consistently greater tactile acuity when they could view the stimulated body site, even though there was no explicit visual information available at that site. That is, it was the simple act of bringing the 'felt' receptive fields into the center of gaze that caused an improvement in tactile abilities. With the use of ERPs, they went on to show a corresponding and dramatic amplification of the somatosensory evoked potential (Taylor-Clarke *et al.*, 2002) and

suggested that this reflected feedback from tactile–visual multisensory neurons in posterior parietal cortex to primary somatosensory cortex. The implications of course are great, indicating that the ability to recruit visuotactile neurons actually influences how we experience touch. In this special issue we see a bevy of related papers, emphasizing that this is an area of keen interest for multisensory scientists. Work from the laboratory of Olaf Blanke in Lausanne (Aspell *et al.*, 2010; this volume), for example, adds a particularly interesting twist to this visualization story. They asked their participants to judge the elevation of four vibratory stimulators that were attached to their backs while they played unrelated auditory stimuli from speakers placed both in front of and behind the observers. It is often the case that automatic multisensory integrations will cause some spatial shifting under such circumstances, a form of audiotactile ventriloquism if you will. However, the auditory stimuli had no effect on the tactile elevation judgments under these circumstances. Nevertheless, when the participants were provided with a real-time but completely uninformative view of their own back using a head-mounted display fed by a camera placed 2 m behind them, incongruent auditory spatial information began to interfere with the tactile elevation judgments: a trisensory effect. And so, the interaction of the senses is clearly profound but also complex. This area will not lend itself to very simple models or principles and there is much fertile ground to be tilled before we are likely to arrive at any true understanding.

In closing, we would like to thank the many multisensory scientists who submitted their work to this special issue. As in past years, the competition for limited places was fierce and the quality ran deep. With the production of this issue, an annual tradition since the founding of the International Multisensory Research Forum (IMRF) conference, we hand the reins over to the very capable hands of Dr Georg Meyer who will chair the organizing committee of the 11th annual meeting. He will host this year's meeting from June 16th to June 19th 2010 at The University of Liverpool (<http://imrf.mcmaster.ca/IMRF/ocs2/index.php/imrf/2010/>). Finally, we wish to extend our heartfelt appreciation once again to the local organizing committee of IMRF 2009 in New York, who made the meeting such a tremendous success. We look forward to what is sure to be a most engaging and stimulating IMRF 2010, and to meeting in Liverpool with colleagues and colleagues-to-be who share our enthusiasm for the multisensory perspective.

Acknowledgements

The work of S.M. and J.J.F. on multisensory integration is supported by a grant from the US National Institute of Mental Health (MH85322). The authors would like to thank Professor Herb Cohen of the CERC at Einstein for helping us rediscover the important work of Birch & Lefford (1963, 1967).

References

- Alsius, A., Navarra, J., Campbell, R. & Soto-Faraco, S. (2005) Audiovisual integration of speech falters under high attention demands. *Curr. Biol.*, **15**, 839–843.
- Aspell, J., Lavanchy, T., Lenggenhager, B. & Blanke, O. (2010) Seeing the body modulates audiotactile integration. *Eur. J. Neurosci.*, **31**, 1868–1873.
- Birch, H.G. & Lefford, A. (1963) Intersensory development in children. *Monogr. Soc. Res. Child Dev.*, **28**, 1–47.
- Birch, H.G. & Lefford, A. (1967) Visual differentiation, intersensory integration, and voluntary motor control. *Monogr. Soc. Res. Child Dev.*, **32**, 1–82.
- Bolognini, N., Olgiati, E., Rossetti, A. & Maravita, A. (2010) Enhancing multisensory spatial orienting by brain polarization of the parietal cortex. *Eur. J. Neurosci.*, **31**, 1800–1806.
- Busse, L., Roberts, K.C., Crist, R.E., Weissman, D.H. & Woldorff, M.G. (2005) The spread of attention across modalities and space in a multisensory object. *Proc. Natl Acad. Sci. USA*, **102**, 18751–18756.
- Driver, J. & Spence, C. (2000) Multisensory perception: beyond modularity and convergence. *Curr. Biol.*, **10**, R731–R735.
- Egely, R., Driver, J. & Rafal, R.D. (1994) Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J. Exp. Psychol. Gen.*, **123**, 161–177.
- Ernst, M.O. & Banks, M.S. (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, **415**, 429–433.
- Falchier, A., Clavagnier, S., Barone, P. & Kennedy, H. (2002) Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.*, **22**, 5749–5759.
- Fetsch, C.R., DeAngelis, G.C. & Angelaki, D.E. (2010) Visual-vestibular cue integration for heading perception: applications of optimal cue integration theory. *Eur. J. Neurosci.*, **31**, 1721–1729.
- Fiebelkorn, I.C., Foxe, J.J. & Molholm, S. (2010a) Dual mechanisms for the cross-sensory spread of attention: how much do learned associations matter? *Cereb. Cortex*, **20**, 109–120.
- Fiebelkorn, I.C., Foxe, J.J., Schwartz, T.H. & Molholm, S. (2010b) Staying within the Lines: the formation of Visuospatial Boundaries influences multisensory feature integration. *Eur. J. Neurosci.*, **31**, 1737–1743.
- Foxe, J.J. (2008) Toward the end of a “principled” era in multisensory science. *Brain Res.*, **1242**, 1–3.
- Fujisaki, W., Shimojo, S., Kashino, M. & Nishida, S. (2004) Recalibration of audiovisual simultaneity. *Nat. Neurosci.*, **7**, 773–778.
- Hanson, J.V.M., Heron, J. & Whitaker, D. (2008a) Recalibration of perceived time across sensory modalities. *Exp. Brain Res.*, **185**, 347–352.
- Harrar, V. & Harris, L.R. (2008) The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Exp. Brain Res.*, **186**, 517–524.
- Heron, J., Roach, N., Whitaker, D.J. & Hanson, J. (2010) Attention regulates the plasticity of multisensory timing. *Eur. J. Neurosci.*, **31**, 1755–1762.
- Keetels, M. & Vroomen, J. (2007) No effect of auditory-visual spatial disparity on temporal recalibration. *Exp. Brain Res.*, **182**, 559–565.
- Kennett, S., Taylor-Clarke, M. & Haggard, P. (2001) Noninformative vision improves the spatial resolution of touch in humans. *Curr. Biol.*, **11**, 1188–1191.
- King, A.J. & Palmer, A.R. (1985) Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp. Brain Res.*, **60**, 492–500.
- Martinez, A., Teder-Salejari, W., Vazquez, M., Molholm, S., Foxe, J.J., Javitt, D.C., Di Russo, F., Worden, M.S. & Hillyard, S.A. (2006) Objects are highlighted by spatial attention. *J. Cogn. Neurosci.*, **18**, 298–310.
- Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E. & Foxe, J.J. (2002) Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res. Cogn. Brain Res.*, **14**, 115–128.
- Molholm, S., Martinez, A., Shpaner, M. & Foxe, J.J. (2007) Object-based attention is multisensory: co-activation of an object's representations in ignored sensory modalities. *Eur. J. Neurosci.*, **26**, 499–509.
- Pasalar, S., Ro, T. & Beauchamp, M. (2010) TMS of Posterior parietal cortex disrupts visual tactile multisensory integration. *Eur. J. Neurosci.*, **31**, 1783–1790.
- Raij, T., Ahveninen, J., Lin, F.-H., Witzel, T., Jääskeläinen, I.P., Letham, B., Israeli, E., Sahyoun, C., Vasios, C., Stufflebeam, S., Hämäläinen, M. & Belliveau, J.W. (2010) Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *Eur. J. Neurosci.*, **31**, 1772–1782.
- Rockland, K.S. & Ojima, H. (2003) Multisensory convergence in calcarine visual areas in macaque monkey. *Int. J. Psychophysiol.*, **50**, 19–26.
- Schroeder, C.E. & Foxe, J.J. (2002) The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res. Cogn. Brain Res.*, **14**, 187–198.
- Seltzer, B. & Pandya, D.N. (1980) Converging visual and somatic sensory cortical input to the intraparietal sulcus of the rhesus monkey. *Brain Res.*, **192**, 339–351.
- Sherrington, C.S. 1940. *Man on His Nature*. Cambridge University Press, Cambridge, UK.
- Soto-Faraco, S., Navarra, J. & Alsius, A. (2004) Assessing automaticity in audiovisual speech integration: evidence from the speeded classification task. *Cognition*, **92**, B13–B23.
- Spence, C. & Squire, S. (2003) Multisensory integration: maintaining the perception of synchrony. *Curr. Biol.*, **13**, R519–R521.
- Stein, B.E. & Arigbede, M.O. (1972) Unimodal and multimodal response properties of neurons in the cat's superior colliculus. *Exp. Neurol.*, **36**, 179–196.

- Stein, B.E. & Meredith, M.A. (1993) *The Merging of the Senses*. MIT Press, Cambridge, MA.
- Stroop, J.R. (1935) Studies of interference in serial verbal reactions. *J. Exp. Psychol.*, **18**, 643–662.
- Talsma, D., Doty, T.J. & Woldorff, M.G. (2007) Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cereb. Cortex*, **17**, 679–690.
- Taylor-Clarke, M., Kennett, S. & Haggard, P. (2002) Vision modulates somatosensory cortical processing. *Curr. Biol.*, **12**, 233–236.
- Vroomen, J., Keetels, M., de Gelder, B. & Bertelson, P. (2004) Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cogn. Brain. Res.*, **22**, 32–35.
- van Wanrooij, M., Bremen, P. & Van Opstal, J. (2010) Acquired prior knowledge modulates audiovisual integration. *Eur. J. Neurosci.*, **31**, 1763–1771.
- Warren, D.H., Welch, R.B. & McCarthy, T.J. (1981) The role of visual-auditory “compellingness” in the ventriloquism effect: implications for transitivity among the spatial senses. *Percept. Psychophys.*, **30**, 557–564.
- Welch, R.B. & Warren, D.H. (1980) Immediate perceptual response to intersensory discrepancy. *Psychol. Bull.*, **88**, 638–667.
- Welch, R.B. & Warren, D.H. (1986) Intersensory interactions. In Boff, K.R., Kaufman, L. & Thomas, J.P. (Eds), *Handbook of Perception and Human Performance*, chap. 25. Wiley, New York, pp. 1–25.