



Multisensory Integration and Crossmodal Attention Effects in the Human Brain

John J. McDonald, *et al.*

Science **292**, 1791 (2001);

DOI: 10.1126/science.292.5523.1791a

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of January 20, 2012):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/292/5523/1791.full.html>

This article **cites 23 articles**, 7 of which can be accessed free:

<http://www.sciencemag.org/content/292/5523/1791.full.html#ref-list-1>

This article has been **cited by** 1 article(s) on the ISI Web of Science

This article appears in the following **subject collections**:

Technical Comments

http://www.sciencemag.org/cgi/collection/tech_comment

Multisensory Integration and Crossmodal Attention Effects in the Human Brain

Macaluso *et al.* (1) provided functional magnetic resonance imaging (fMRI) evidence for multisensory processing in the human brain. In their study, a light presented in the right visual field produced a larger neural response in the left fusiform gyrus, a modality-specific area of visual cortex, when a concurrent tactile stimulus was delivered to the right hand (seen next to the right light) than when the light was presented alone or when a concurrent tactile stimulus was delivered to the left hand. These findings dovetail nicely with evidence for spatially specific multisensory effects at the behavioral and neural levels (2, 3) and for bimodal-stimulation effects on activity in auditory cortex (4, 5).

Macaluso *et al.* stated [p. 1206 in (1)] that their finding “provides a neural explanation for crossmodal links in spatial attention.” We view that conclusion as premature, however, because of the stimulus parameters that were chosen. Attributing modulations in behavior or neural activity to a spatial attention mechanism is a nontrivial task in intramodal situations as well as in multimodal ones. A non-predictive visual cue presented at one spatial location might facilitate responses to a subsequent visual target at that location either because it elicited an involuntary shift of attention or simply because the sensory responses to the cue and target were temporally integrated (6, 7). In general, converging evidence must be obtained to rule out the latter, sensory-based explanation for behavioral or neural facilitation.

A minimum condition that is usually imposed to reduce sensory integration is to present the target stimulus some time after the initial cue stimulus has disappeared (7). One might assume that sensory integration would also be reduced by presenting the cue and target in different modalities; however, presenting different-modality stimuli within 100 to 150 ms actually produces a superadditive sensory response from specialized neurons that are capable of responding to both stimuli (3). As a result of this superadditive effect, stimuli that appear at about the same time and place are integrated to form a unified perceptual object, rather than being left as a collection of unrelated sensations. This phenomenon is called multisensory integration, and it provides a neural explanation for several dramatic multisensory perceptual effects such as the ventriloquist’s illusion (8).

Relatively long-lasting visual and tactile stimuli were presented simultaneously in the experiments of Macaluso *et al.* (1); thus, the

neural interactions they observed could have been those involved in multisensory integration of vision and touch. Interestingly, in the Perspectives article that accompanied the Macaluso *et al.* study, de Gelder (9) interpreted their finding in terms of multisensory integration rather than spatial attention and speculated that simultaneous presentation was crucial for the fMRI findings. Clearly, the presentation of visual and tactile stimuli at the same time and place is sufficient to produce multisensory integration at both neural and behavioral levels. However, the questions remain as to whether and how multisensory integration is related to the crossmodal consequences of involuntary spatial attention.

If multisensory integration constituted the neural mechanism that causes involuntary shifts of attention to modulate processing of objects in different modalities, then the conclusions drawn by Macaluso *et al.* about the neurophysiological basis of crossmodal spatial attention effects would be justified. To our knowledge, however, there is no empirical evidence supporting the hypothesis that multisensory integration plays any role in generating crossmodal attention effects. On the contrary, there are convincing lines of evidence that multisensory integration and shifts of spatial attention are independent. (i) Multisensory integration occurs in anaesthetized animals (i.e., without any intent). (ii) The ventriloquism effect, a well-known perceptual consequence of multisensory integration, has been shown to occur preattentively and independently of both voluntary and involuntary spatial attention shifts (10, 11), although it can aid voluntary attentional focusing (12). (iii) Whereas multisensory integration occurs preattentively and without intent, involuntary shifts of attention depend on the attentional goals of the observer (13, 14). (iv) Approximate temporal synchrony is required for multisensory integration but not for involuntary spatial attention effects to occur. Importantly, involuntary spatial attention effects occur even when stimuli are brief (<100 ms) and are separated by 100 to 500 ms (15, 16). Under such conditions, multisensory integration and many of its perceptual consequences (e.g., ventriloquism) are greatly reduced (17, 18).

These findings provide evidence that involuntary shifts of spatial attention arise from stimulus-driven processes that are separate from those involved in multisensory integration. Of course, in some experimental para-

digms, the two effects might co-occur and produce additive facilitation of responses to targets. The fMRI study by Macaluso *et al.* is valuable because it firmly demonstrated an effect of multimodal stimulation on modality-specific cortical processing. Because of the specific experimental procedures used, however, their demonstration did not provide a clear explanation for the neural basis of crossmodal spatial attention effects. Involuntary shifts of spatial attention caused by the appearance of nonvisual stimuli do seem to produce similar enhancements of neural responses to visual stimuli within visual cortex (19, 20), but the details of how this occurs remain to be discovered.

John J. McDonald

Department of Psychology
Simon Fraser University
8888 University Drive
Burnaby, BC V5A 1S6, Canada
E-mail: jmcd@sfu.ca

Wolfgang A. Teder-Sälejärvi

Department of Neurosciences
University of California, San Diego
9500 Gilman Drive
MC 0608
La Jolla, CA 92093-0608, USA
E-mail: wat@sdepl.ucsd.edu

Lawrence M. Ward

Department of Psychology
University of British Columbia
2136 West Mall
Vancouver, BC V6T 1Z4, Canada
E-mail: lward@cortex.psych.ubc.ca

References

1. E. Macaluso, C. D. Frith, J. Driver, *Science* **289**, 1206 (2000).
2. R. B. Welch, D. H. Warren, in *Handbook of Perception and Human Performance*, K. R. Boff, L. Kaufman, J. P. Thomas, Eds. (Wiley, New York, 1986), vol. 1, chap. 25.
3. B. E. Stein, M. A. Meredith, *The Merging of the Senses* (MIT Press, Cambridge, MA, 1993).
4. G. A. Calvert *et al.*, *Science* **276**, 593 (1997).
5. M. Sams *et al.*, *Neurosci. Lett.* **127**, 141 (1991).
6. C. Richard, thesis, University of British Columbia (1999).
7. G. Tassinari, S. Aglioti, L. Chelazzi, A. Peru, G. Berlucchi, *Vision Res.* **34**, 179 (1994).
8. P. Bertelson, in *Cognitive Contributions to the Perception of Spatial and Temporal Events*, G. Aschersleben, T. Bachmann, J. Müssele, Eds. (Elsevier, Amsterdam, 1999), pp. 347-362.
9. B. de Gelder, *Science* **289**, 1148 (2000).
10. P. Bertelson, J. Vroomen, B. de Gelder, J. Driver, *Percept. Psychophys.* **62**, 321 (2000).
11. J. Vroomen, P. Bertelson, B. de Gelder, *Percept. Psychophys.*, in press.
12. J. Driver, *Nature* **381**, 66 (1996).
13. C. L. Folk, R. W. Remington, J. C. Johnston, *J. Exp. Psychol. Hum. Percept. Perform.* **18**, 1030 (1992).
14. L. M. Ward, J. J. McDonald, D. Lin, *Percept. Psychophys.* **62**, 1258 (2000).
15. J. J. McDonald, thesis, University of British Columbia (1999).
16. ———, W. A. Teder-Sälejärvi, S. A. Hillyard, *Nature* **407**, 906 (2000).
17. M. A. Meredith, J. W. Nemitz, B. E. Stein, *J. Neurosci.* **7**, 3125 (1987).

TECHNICAL COMMENTS

18. C. E. Jack, W. R. Thurlow, *Percept. Motor Skills* **37**, 967 (1973).
19. S. Kennett, M. Eimer, C. Spence, J. Driver, *J. Cognit. Neurosci.*, in press.
20. J. J. McDonald, L. M. Ward, *Psychol. Sci.* **11**, 167 (2000).

20 February 2001; accepted 15 May 2001

Response: McDonald *et al.* do not challenge our finding (1) that crossmodal tactile-visual interactions can affect unimodal visual areas of the human brain in a spatially specific manner. Instead, they raise a terminological issue and an empirical issue, both relating mainly to the timing of the stimuli used. The terminological issue is whether the effect we observed should be labeled as reflecting spatial attention or crossmodal integration; we think that both terms may be appropriate. We disagree with the specific suggestion by McDonald *et al.* that crossmodal integration is found only with temporally synchronous stimulation and attentional effects only with asynchronies of 100 ms or more.

In psychology and neuroscience, spatial attention refers to spatially selective internal processing of stimulus information. It is now conventional to distinguish between two forms of spatial attention: endogenous attention, which can be directed voluntarily, and exogenous attention, which is captured automatically by salient stimulus events (2, 3). Our study (1) concerned strictly the latter. A common way to study exogenous spatial attention is to measure how a spatially nonpredictive cue event (such as tactile stimulation on one hand) affects responses to targets (such as a visual event on one side), at the same location or a different one (4–7). The issue in crossmodal cueing studies of exogenous attention is thus how stimulation in one modality can spatially affect responses to another modality. Note that this stimulus-driven issue closely overlaps with that of crossmodal integration, because the latter also concerns how stimulation in one modality can affect responses in another modality (8–10). Although a distinction between crossmodal integration and crossmodal attention may be drawn for the case of endogenous attention (10), this is less straightforward for the stimulus-driven case of exogenous attention.

McDonald *et al.* particularly emphasize one example of crossmodal integration: the interactions found in multimodal neurons when stimulation is presented at a common location in several modalities. Such interactions have been found at the single-cell level in animals for several brain areas, including the superior colliculus (8) plus regions of the cortex (11). Here, neurons respond to stimuli in more than one modality, typically have spatially corresponding receptive fields in these different modalities, and can show overadditive responses to multimodal stimulation at the same location, compared with responses to unimodal stimulation there (8,

11). Neurons of this type could in principle be involved in crossmodal spatial cueing effects of the type found behaviorally in studies of human exogenous spatial attention (4, 5), as previously suggested (12). McDonald *et al.*, however, argue that such effects can reflect only crossmodal integration, never exogenous spatial attention. They base this argument on their claim that these integration effects depend on temporally synchronous multimodal stimulation, whereas behavioral cueing effects can be found even when a spatial cue in one modality precedes a subsequent target by 100 to 500 ms [see (2, 4, 5)].

Yet these cellular integration effects actually can apply for asynchronies extending up to 600 ms (13), well within the range of behavioral cueing effects in humans. Moreover, the temporal window for these interactions depends on the physical characteristics of the stimuli used, plus the particular discharge properties (for example, duration of sustained firing or timing of peak response) for each neuron. Maximal crossmodal interaction within a multimodal neuron is not necessarily observed for synchronous stimulation in the external world, but rather for temporally overlapping discharges in response to the stimuli in the different modalities (13). In the cat superior colliculus, response latencies to visual stimulation can lag behind tactile responses by around 100 ms, in which case maximal crossmodal interaction may be observed with the visual stimulus preceding the tactile, when short bursts of neuronal activity are induced. But note that the time window of interaction will depend on the stimuli used and on the specific multimodal brain area (and the pathways to it from each modality).

We thus cannot agree with the proposal by McDonald *et al.* of a simple timing rule for distinguishing between integration and exogenous attention effects. The 100-ms rule of thumb they propose has arisen within purely visual studies of exogenous attention (14), but only to ensure that intervals between visual cues and visual targets exceed retinal integration time. Stimulation from different modalities (for example, vision and touch) cannot interact on the retina, but only in central multimodal neurons, where interactions arise for substantial asynchronies overlapping with those at which crossmodal cueing effects are found behaviorally. Conversely, attentional effects can arise even for synchronous stimuli [as, for example, in the neurological syndrome of extinction; see (15)].

We suggest that in multimodal brain areas, there may be considerable overlap between the machinery for stimulus-driven crossmodal spatial integration and that for stimulus-driven (exogenous) crossmodal spatial attention. We stress, however, that the main point of our study was that unimodal areas of visual cortex can

also be affected. Our study (1) suggested one possible mechanism for this (feedback from multimodal convergence zones to unimodal areas), and found some evidence that may accord with this from an analysis of coupling between brain areas.

Apart from the terminological issues, the comment of McDonald *et al.* does raise the issue of how the results that we obtained (1) might change if the timing of the stimuli were varied systematically. We doubt that introducing a 100-ms asynchrony between the onset of tactile and visual stimuli would have affected the outcome, given that each stimulus was 300 ms in duration, but this is an empirical issue. By using much briefer stimuli and varying their temporal offset across a considerable range, one could map out the time windows of crossmodal spatial interactions for different areas of the human brain. It would be useful to address this with event-related potentials or magnetoencephalography in addition to the fMRI procedure we used, given the greater temporal resolution of those techniques.

Emiliano Macaluso

*Institute of Cognitive Neuroscience
University College London
London, UK*

E-mail: e.macaluso@fil.ion.ucl.ac.uk

Chris Frith

*Wellcome Department of Cognitive
Neurology
Institute of Neurology
London, UK*

Jon Driver

*Institute of Cognitive Neuroscience
University College London*

References

1. E. Macaluso, C. D. Frith, J. Driver, *Science* **289**, 1206 (2000).
2. M. I. Posner, Q. J. *Exp. Psychol.* **32**, 3 (1980).
3. R. M. Klein, in *Eye Movements and Visual Cognition: Scene Perception and Reading*, K. Rayner, Ed. (Springer-Verlag, New York, 1992), pp. 46–65.
4. C. Spence, J. Driver, *Percept. Psychophys.* **59**, 1 (1997).
5. J. J. McDonald, W. A. Teder-Sälejärvi, S. A. Hillyard, *Nature* **407**, 906 (2000).
6. J. J. McDonald, L. M. Ward, *Psychol. Sci.* **11**, 167 (2000).
7. S. Kennett, M. Eimer, C. Spence, J. Driver, *J. Cognit. Neurosci.*, in press.
8. B. E. Stein, M. A. Meredith, *The Merging of the Senses* (MIT Press, Cambridge, MA, 1993).
9. P. Bertelson, in *Cognitive Contributions to the Perception of Spatial and Temporal Events*, G. Aschersleben, T. Bachmann, J. Musseler, Eds. (Elsevier, Amsterdam, 1999), pp. 347–362.
10. J. Driver, *Nature* **381**, 66 (1996).
11. M. T. Wallace, M. A. Meredith, B. E. Stein, *Exp. Brain Res.* **91**, 484 (1992).
12. J. Driver, C. Spence, *Trends Cognit. Sci.* **2**, 254 (1998).
13. M. A. Meredith, J. W. Nemitz, B. E. Stein, *J. Neurosci.* **7**, 3215 (1987).
14. G. Tassinari, S. Aglioti, L. Chelazzi, A. Peru, G. Berlucchi, *Vision Res.* **34**, 179 (1994).
15. J. Driver, P. Vuilleumier, *Cognition* **79**, 39 (2001).

19 March 2001; accepted 15 May 2001