

NEUROSCIENCE:

More to Seeing Than Meets the Eye

Beatrice de Gelder*

The outer layer of the cerebral cortex is divided into different areas specialized for detecting and processing sensory signals from the eyes and ears and from receptors for touch, taste, and smell. Differences between these sensory areas may reflect variations in the rate of evolution of the five senses and the special information processing requirements for each type of sensory signal. Everyday experience illustrates that, despite their differences, the sensory regions of the cortex must be cooperating with each other by integrating the sensory stimuli they receive from the outside world. Now, on page [1206](#) of this issue ([1](#)), Macaluso *et al.* report an elegant example of this cooperation and provide empirical justification for the aphorism that there is more to seeing than meets the eye. They show that the administration of a tactile (touch) stimulus and a visual stimulus to human volunteers at the same time and on the same side of the body enhanced neural activity in the lingual gyrus of the visual cortex, above that achieved with the visual stimulus alone. The authors propose that neurons in the somatosensory (touch) area of the cortex project back to the visual cortex, thus keeping the visual cortex informed about touch stimuli that are received simultaneously with visual stimuli.

How widespread is the interaction of one sensory area of the cerebral cortex with another (cross-modal impact), and how general is the underlying neural mechanism? Cross-modal information exchange between the auditory and visual cortex has been found in speech perception and in a few other cases. In ventriloquism ([2](#)), the apparent direction of sound is attracted toward the displaced location of a simultaneous visual stimulus--the sight of the speaker's lip movements influences the hearing of speech ([3](#)). Similarly, a facial expression, even if not consciously perceived, modifies the perception of emotion in the voice of the speaker ([4](#)).

Our experience tells us that in nature, simultaneous signals from different sensory organs are the rule rather than the exception. But, in fact, most

connections between different sensory signals are irrelevant, such as hearing the call of a seagull as we watch the waves crashing against the rocks. So, how does the brain discern what sounds go with what sights? The cross-modal interactions that produce the unified objects and events that we perceive around us require a very high degree of selection. Too many interactions in the brain would create an internal booming, buzzing confusion to match the one surrounding us. But it is only biologically important combinations of sensory stimuli that are likely to be endowed with hard-wired neural pathways in the brain. When it comes to packaging individual sensory stimuli together into a single event (see the figure), the brain, like a good playwright, is likely to ask "when" (time), "where" (space), "what" (identity), and "why" (why does the stimulus matter to the organism).

Integration of different but related sensory stimuli does not require the glue of attention or awareness (5, 6). Recently, multisensory neurons that receive more than one type of sensory signal have been found in different areas of human and monkey brains, for example, in the parietal areas (vision, hearing, and touch) and in the superior colliculus (vision and hearing) (7). This has led to the popular notion that cross-modal connections simply reflect the existence of multisensory neurons. Macaluso *et al.* go beyond this explanation, proposing that cross-modal effects are the result of signals--carried by multisensory neurons projecting from the parietal areas of the somatosensory cortex back to the primary visual cortex--that modulate the activity of visual neurons. Their proposal is similar to that put forward to explain the modulation of auditory cortical activity by visual signals from moving lips (8) or from facial expressions (4) during speech perception. It is unlikely that multisensory neurons by themselves could account for all cross-modal effects without some feedback from the visual (or in some cases the auditory) cortex. In the sensory cortical architecture proposed by Macaluso *et al.*, multisensory neurons and their back-projections each have their own distinct functions. Multisensory neurons--or structures that establish connections between different sensory signals, such as the amygdala (9)--alert the organism to possible coincidences among sensory stimuli (by detecting similarities among the when, where, what, and why for each

stimulus) and so behave as possible event detectors (see the figure). Presumably, simultaneous administration of visual and tactile stimuli to human volunteers by Macaluso *et al.* was crucial to their finding that the lingual gyrus was activated by the integration of both sensory signals. If the two stimuli had been administered at slightly different times, it is possible that activation of the lingual gyrus would not have been observed. Also, the human volunteers were only shown very simple objects. It would be interesting to know whether more complex visual stimuli would have resulted in lingual gyrus activation. Synchrony between visual and auditory stimuli ([4](#), [8](#), [10](#)) as well as object identity (moving lips, facial expressions) is crucial for the cross-modal integration of different sensory signals.

Feeling is seeing. Two independent sensory stimuli, light and touch, are processed in the visual cortex and somatosensory cortex, respectively. Each sensory signal carries the information of where, when, what, and why to the brain. An event-detection system in the brain alerts the organism to the co-occurrence of the two stimuli and to the fact that they may be connected. Confirmation that the signals are indeed connected is provided by the event-detection system when it receives two simultaneous sensory signals. In this case, the event-detection system is the bundle of neurons that projects from the parietal areas of the somatosensory cortex back to the visual cortex and provides the cross-modal effect.

At last, we are beginning to understand how the brain detects only the biologically important combinations of sensory stimuli that emanate from our complex world.

References

1. E. Macaluso, C. Frith, J. Driver, *Science* **289**, [1206](#) (2000).
2. P. Bertelson and G. Aschersleben, *Psychol. Bull. Rev.* **5**, 482 (1998).
3. H. McGurk and J. McDonald, *Nature* **264**, 746 (1976) [[Medline](#)].
4. B. de Gelder *et al.*, *Neurosci. Lett.* **260**, 133 (1999) [[Medline](#)].
5. P. Bertelson *et al.*, *Percept. Psychophys.* **29**, 57 (2000).
6. R. Kawashima *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5969 (1995) [[Medline](#)].
7. J. Driver and C. Spence, *Trends Cognit. Sci.* **2**, 254 (1998).
8. G. Calvert *et al.*, *Science* **276**, [593](#) (1998).
9. F. Nahm *et al.*, *Neuropsychologia* **31**, 727 (1993) [[Medline](#)].
10. J. Vroomen and B. de Gelder, *JEP/HPP* **26**, 2030 (2000).