What you see is what you hear

Traditionally, vision has been considered the dominant modality in our multi-sensory perception of the world. Here we present findings that overturn this established view, by showing that auditory information can change the percept of an unambiguous visual stimulus qualitatively (causing a strong visual illusion). These results indicate that, as with other modalities, our visual percepts are malleable by other modalities.

We have discovered a visual illusion which is induced by sound: when a single visual flash is accompanied by multiple auditory beeps, the single flash is incorrectly perceived as multiple flashes. These results were obtained by flashing a uniform white disk (subtending 2˚ at 5˚ eccentricity) a variable number of times (spaced 50ms apart) on a black background. Flashes were accompanied with a variable number of beeps, each spaced 57ms apart. Observers were asked to judge how many visual flashes were presented on each trial. The trials were randomized and each stimulus combination was run five times on eight naïve observers. Surprisingly, observers consistently and incorrectly reported seeing multiple flashes whenever a single flash was accompanied by more than one beep (Fig. 1a). Control conditions and catch trials (described in figure caption) indicate that the illusory flashing phenomenon is indeed a perceptual illusion and is not due to the difficulty of the task, cognitive bias, etc.. Moreover, as can be seen in Fig. 1b, observers' performance was the same whether a single flash was accompanied with two beeps, or two flashes were accompanied with one or no beeps, suggesting that the illusory double flash is perceptually equivalent to the physical double flash. The post-experiment testimonies of the observers also confirm this statement. The performance of non-naïve subjects (not shown here) indicates that the illusion persisted even when subjects were aware of the fact that the disk was physically flashed only once.

We next investigated the temporal properties of this illusion by varying the relative timing of visual and auditory stimuli. The illusory flashing effect declined from 70ms separation onwards. However, illusory flashing occured so long as beep and flash were within approximately 100ms, consistent with the integration time of polysensory neurons in the brain[2,3].

Our results indicate that the illusory flashing phenomenon is caused by an alteration of the visual perception via auditory stimuli. The modification of the visual percept by sound, however, was not categorical. It was rather selective, as sound did not have a fusing effect when multiple flashes were accompanied with a single beep. We therefore speculate that the direction of cross-modal interactions is partly dependent on the type of stimulus. Consistent with previous observations in other modalities[4], we hypothesize that the percept of a continuous stimulus in one modality is made considerably more malleable by the discontinuous stimulus in another modality, than vice versa.

The influence of auditory cues on visual perception has been demonstrated in other settings in which perceived visual intensity is affected by the presence of an auditory stimulus[5]. This influence, however, is quantitative and does not alter the phenomenological quality of the percept. Others have shown that the perceived direction

of ambiguous visual motion is influenced by auditory stimulation[1]. Our work extends these previous findings by showing that the visual perception can be qualitatively altered by sound even when the visual stimulus is not ambiguous. The conditions under which we have found this alteration to occur—the stimulus configuration and the task—are very simple. The illusion is also surprisingly robust to variations in the many parameters we manipulated (e.g., disk eccentricity and contrast, spatial disparity between sound and flash, shape and texture of the flashing pattern, flash and beep durations). The simplicity and robustness of the illusory flashing phenomenon suggest that it reflects a fundamental and widespread property of polysensory mechanisms in the brain.

Ladan Shams*
Yukiyasu Kamitani*
Shinsuke Shimojo*[†]
* California Institute of Technology, Division of Biology, MC 139-74, Pasadena, CA 91125, USA
[†] NTT Communication Science Laboratories, Human and Information Science Laboratory, Atsugi, Kanagawa, 243-0198 Japan
e-mail: ladan@caltech.edu

References

1. Sekuler, R., Sekuler, A. B. & Lau, R. *Nature* **385**, 308 (1997).
2. Meredith, M. A., Nemitz, J. W. & Stein, B. E. *J. Neurosci.* **10**, 3215-3229 (1987).
3. Wallace, M. T. & Stein, B. E. *J. Neurophysiol.* **71**, 429-432 (1994).
4. Saldaña, H. M. & Rosenblum, L. D. *Percept. Psychophys.* **54**, 406-416 (1993).
5. Stein, B. E., London, N., Wilkinson, L. K. & Price, D. D. *J. Cognit. Neurosci.* **8**, 497-506 (1996).

Figure Captions

Figure 1 Illusory flashing phenomenon. **a,** Perceived number of visual flashes plotted as a function of the number of auditory beeps for eight observers. The number of perceived flashes did not linearly increase with the 3rd and 4th beeps, since they fell outside the optimal window of audiovisual integration, as revealed by our next experiment. **b,** Perceived number of flashes plotted as a function of the actual number of flashes presented, separately for absent sound (broken line), and single beeps corresponding to catch trials (gray line). Observers performed the task very well in the absence of sound (broken line). The results of the catch trials (gray line) confirm that the observers' responses were not determined by their auditory percepts. The curve in (a) is superimposed on (b) for comparison. Further detail on methodology can be obtained from L.S.