Research Report

Vision Merges With Touch in a Purely Tactile Discrimination

Ehsan Arabzadeh, Colin W.G. Clifford, and Justin A. Harris
University of Sydney

ABSTRACT—To construct a coherent percept of the world, the brain continuously combines information across multiple sensory modalities. Simple stimuli from different modalities are usually assumed to be processed in distinct brain areas. However, there is growing evidence that simultaneous stimulation of multiple modalities can influence the activity in unimodal sensory areas and improve or impair performance in unimodal tasks. Do these effects reflect a genuine cross-modal integration of sensory signals, or are they due to changes in the perceiver’s ability to locate the stimulus in time and space? We used a behavioral measure to differentiate between these explanations. Our results demonstrate that, under certain circumstances, a noninformative flash of light can have facilitative or detrimental effects on a simple tactile discrimination. These findings reveal that sensory signals from different modalities can be integrated, even for perceptual judgments within a single modality.

The development of multiple sensory-specific brain areas increases an organism’s chances of survival, but confronts researchers with the question of how different senses are combined to produce a coherent, unified percept of the world. Recent evidence shows that different modalities can interact even in early sensory cortices that are conventionally considered unimodal areas (Fu et al., 2003; Giard & Peronnet, 1999; Macaluso & Driver, 2005; Taylor-Clarke, Kennett, & Haggard, 2002). In addition, behavioral experiments show that subjects’ performance in unimodal tasks can be affected by the simultaneous presentation of a stimulus in another modality (Harris, Arabzadeh, Moore, & Clifford, 2007; Kennett, Taylor-Clarke, & Haggard, 2001; McDonald, Teder-Salejarvi, & Hillyard, 2000). These performance changes may reflect an intersensory interaction whereby a stimulus in one modality modifies how a stimulus in another modality is perceived, as in the McGurk effect (seeing the movement of lips changes the perception of a speech sound; McGurk & MacDonald, 1976) and the Shams illusion (a single visual flash is perceived as a double flash if accompanied by two auditory beeps; Shams, Kamitani, & Shimojo, 2000). However, the extent to which these effects are due to changes in sensory processing or to subjective decisions about the output of that processing is not clear. In the experiments reported here, we adopted a design that eliminated the role of the decision criterion in the perceptual measure, and thus tested for a genuine effect on sensory processing.

Early processing of unimodal sensory stimuli can be characterized by a sigmoidal transfer function, which describes the conversion of the intensity of physical stimuli into neural responses (Laughlin, 1989; see Fig. 1a). This function exhibits an accelerating nonlinearity at low stimulus values (near detection threshold) and a compressive nonlinearity at higher stimulus values. This nonlinearity is apparent when measuring the smallest stimulus difference detectable by the subject (the just-noticeable difference, or JND) across a range of stimulus intensities. Specifically, a stimulus with a fixed base amplitude (pedestal) can have a decreased (for small pedestal values) or increased (for large pedestal values) JND relative to the absolute detection threshold (the JND obtained when the pedestal has zero amplitude). As a result, the function relating a subject’s sensitivity (JND) to stimulus intensity (pedestal magnitude) shows a characteristic biphasic profile (and hence is referred to as the “dipper” function; Graham, 1989; Nachmias & Kocher, 1970; see Fig. 1b); thus, low pedestals improve stimulus discriminability relative to the absolute detection threshold (the JND obtained when the pedestal has zero amplitude). As a result, the function relating a subject’s sensitivity (JND) to stimulus intensity (pedestal magnitude) shows a characteristic biphasic profile (and hence is referred to as the “dipper” function; Graham, 1989; Nachmias & Kocher, 1970; see Fig. 1b); thus, low pedestals increase stimulus discriminability relative to the absolute threshold, and high pedestals impair stimulus discriminability, a relationship known as the pedestal effect. If the introduction of a stimulus in a second modality influences performance in a unimodal task, the nature of the change in the dipper function indicates the underlying mechanism (see Fig. 1c). A lateral shift is consistent with merging of the signals from the two modalities, whereas a vertical shift is consistent with a change in the uncertainty regarding the spatiotemporal location of the stimulus. The study we report here...
demonstrates that a visual stimulus causes a lateral shift in the dipper function for a tactile discrimination, a result indicative of merging between visual and tactile signals.

**METHOD**

In four experiments, subjects sat in front of the test apparatus (Fig. 2a) with the pad of their index finger touching a steel rod (5 mm in diameter). The rod was driven by a vibrotactile device (Type 4810 MiniShaker; Brüel & Kjær, Nærum, Denmark). The timing and waveform of the vibrations were controlled using MATLAB (MathWorks, Natick, MA) and a National Instruments (Austin, TX) interface board. In all four experiments, the subjects fixated a dot perceived at the spatial location of the tactile stimulation. Each trial contained two intervals (see Fig. 2b), the beginnings of which were marked by a sound. One interval (when the pedestal was set to zero) or both intervals (when the pedestal was greater than zero; shown here) contained a tactile vibration (tactile stimulus, or TS); in some conditions, a visual stimulus (VS) appeared as well. At the end of the second interval, subjects made a forced-choice judgment to indicate the interval containing the stronger tactile stimulus (when only one interval contained a vibration). The visual stimulus, if present, was the same in both intervals, and therefore could not inform the subject how to respond on the tactile task.

![Fig. 1.](image1.png)  
**Fig. 1.** The sigmoidal transfer function and psychophysical dipper function. In the presence of a constant noise level, a minimum perceptual difference (threshold, labeled “Th”) is required for a perceiver to discriminate reliably between two stimulus intensities. The graph in (a) shows the typical sigmoidal shape of the perceptual transfer function (i.e., perceptual intensity as a function of stimulus intensity). Because of this sigmoidal shape, the just-noticeable difference (JND) decreases with increasing stimulus intensity at low stimulus intensities (solid green lines) and increases with increasing stimulus intensity at high stimulus intensities (dotted green lines). The latter compressive nonlinearity is equivalent to Weber’s law. Thus, JND is a biphasic function of stimulus intensity and has a characteristic “dipper” shape (b). In the experiments reported here, we examined how the dipper function for a purely tactile stimulus changed when a visual stimulus was presented as well (c). The nature of the change was expected to reveal the form of interaction between the tactile and visual modalities. A vertical shift in the function (blue) would indicate that the visual stimulus reduced uncertainty regarding the spatiotemporal location of the tactile stimulus, whereas a horizontal shift (red) would indicate genuine multimodal integration.

![Fig. 2.](image2.png)  
**Fig. 2.** The experimental setup and design. Subjects sat in front of a computer monitor that was positioned face down above a mirror and held their hand palm up so that their index finger touched a vibrotactile device (a). The mirror was placed halfway between the monitor and the vibrotactile device, so that the image reflected from the monitor appeared to be coming from the same spatial location as the index finger. Subjects could not see their hand, but a fixation point always appeared on the reflected black background at the spatial location of tactile stimulation. Each trial contained two intervals (b), the beginnings of which were marked by a sound. One interval (when the pedestal was set to zero) or both intervals (when the pedestal was greater than zero; shown here) contained a tactile vibration (tactile stimulus, or TS); in some conditions, a visual stimulus (VS) appeared as well. At the end of the second interval, subjects made a forced-choice judgment to indicate the interval containing the stronger tactile stimulus (when only one interval contained a vibration). The visual stimulus, if present, was the same in both intervals, and therefore could not inform the subject how to respond on the tactile task.
RESULTS

First, in Experiment 1, we investigated the pedestal effect within the tactile modality, measuring each subject’s JNDS for vibration pedestals of different amplitudes. Despite the high variability in absolute detection threshold across subjects (range = 2.5–10 μm), all showed the biphasic profile characteristic of the dipper function (Fig. 3a), with the lowest JNDS obtained when the base pedestal was at an amplitude close to the detection threshold (Fig. 3b). The biphasic shape of the function was confirmed by a significant quadratic trend across all subjects, $F(1, 4) = 128.7, p < .001$.

If vision and touch are integrated at early levels of sensory processing, could a visual stimulus induce a pedestal effect in a tactile task? In Experiment 2, we measured tactile detection thresholds in three conditions: In the no-flash condition, subjects received no visual stimulus except the fixation point present in both intervals and therefore was noninformative for the tactile detection task. Figure 4 shows the average detection thresholds of two different visual stimuli: (a) the bright visual stimulus (129 cd/m²) at the same spatiotemporal location as the tactile stimulus. The tactile stimulus was present in only one interval, whereas the visual stimulus was always present in both intervals and therefore was noninformative for the tactile detection task. Figure 4 shows the average detection thresholds (pedestal = 0) in each of the three conditions. Across all subjects, paired $t$ tests showed that detection thresholds differed significantly between the no-flash and dim-flash conditions ($p < .001$) and between the dim-flash and bright-flash conditions ($p < .01$). However, the no-flash and bright-flash conditions did not differ ($p = .49$).

Did the dim visual stimulus improve detection of near-threshold tactile stimuli via uncertainty reduction due to temporal cuing (cf. Pelli, 1985; Perez, Cohn, Medina, & Donoso, 2007)? Two lines of evidence indicate that this was not the case. First, the bright flash provided the same temporal information as the dim flash, yet the bright flash did not improve detection. Second, in Experiment 3, we tested the potential advantage provided by temporal cuing from a dim visual stimulus in 6 new subjects. As in Experiment 2, the tactile stimulus was present in only one interval, whereas the visual stimulus was always present in both intervals. We examined the effect on tactile detection thresholds of two different visual stimuli: (a) the appearance of a dim spot of light at the onset of the tactile

...
stimulus (as in Experiment 2) and (b) the removal of the same dim spot of light at the onset of the tactile stimulus (see Fig. 5a).

The tactile detection thresholds in these two visual conditions were compared with thresholds in a control condition in which no visual stimulus was present.

As Figure 5b shows, the appearance of the dim spot reduced detection thresholds, but the removal of that spot had no effect. Across all subjects, paired t tests showed that detection thresholds in the timed-appearance condition differed significantly from detection thresholds in both the timed-removal condition ($p < .01$) and the no-flash control condition ($p = .026$), and that detection thresholds in these latter two conditions did not differ ($p = .377$).

The difference between results for the timed-appearance condition and the timed-removal condition is important because the visual transients in these two conditions provided equal information about the timing of the tactile stimulus. Thus, the improvement caused by the presence of the dim flash in Experiment 2 was not due to a reduction in temporal uncertainty.

In Experiment 4, to distinguish further between the alternative multimodal interactions outlined in Figure 1c, we tested 4 subjects extensively in conditions combining a dim flash with a range of tactile pedestals, as well as in conditions with the same pedestals and no flash. In all subjects, the dim flash induced a leftward shift in the biphasic tactile dipper function (see Fig. 6). Paired t tests compared JNDs, across subjects, in the with-flash and without-flash conditions at each tactile pedestal. These tests confirmed that for pedestal values below the detection threshold, JNDs were consistently smaller with the flash than without it (all $p s < .05$), and, conversely, that for pedestal values equal to or
above the detection threshold, JNDs were consistently larger with the flash than without it (all $p < .05$). This double dissociation demonstrates a genuine visuo-tactile integration in which the dim visual stimulus acted as an additional pedestal for the tactile discrimination. The biphasic dipper function is a powerful psychophysical tool for testing sensory integration because fusion of sensory inputs will produce opposite effects on discrimination performance depending on whether stimuli are of low or high amplitude. In contrast, other types of sensory interactions (e.g., masking, attentional cuing, and distraction) will produce the same effect (impairment or improvement) on discrimination performance at any position along the stimulus dimension.

**DISCUSSION**

Previous multisensory research has shown that when two modalities both carry information, people are capable of combining
that information optimally, taking into account the reliability of each source (Calvert, Spence, & Stein, 2004; Ernst & Banks, 2002). Even when one modality is noninformative in the perceptual task, it can still affect performance by modulating sensory gain in the task-relevant modality (Harris et al., 2007; Kennett et al., 2001; McDonald et al., 2000). Intersensory integration has previously been implied by multimodal illusions (McGurk & MacDonald, 1976; Shams et al., 2000). However, according to a signal detection theory analysis, the perceptual change that defines these illusions reflects a two-stage process: First, the signal (and accompanying noise) undergoes sensory processing, and then the output is submitted to a decision process (that is subject to bias). Therefore, these illusions could reflect either a change in sensory processing within the task-relevant modality or a change in subjective bias. Our experimental design eliminated the opportunity for a decision bias to affect perceptual measurement, and thus our findings constitute unambiguous evidence that sensory input to one modality can change sensitivity in another modality. The cross-modal pedestal effect provides a powerful means of quantifying the merging of one sense with another.

Hillis, Ernst, Banks, and Landy (2002) reported that whereas cues from within a single modality are combined in a mandatory fashion, resulting in the loss of single-cue information, combination of cues from separate modalities (vision and touch) is not mandatory. Here, we have shown that inputs from different modalities can undergo fusion even when fusion is detrimental to performance, but only under specific conditions (i.e., this effect was observed when a tactile stimulus was combined with a dim visual stimulus, but not when it was combined with a bright visual stimulus). The absence of any change in detection threshold when the flash was bright could indicate either that the weak tactile and strong visual stimuli were not fused or that the bright flash was a pedestal sufficiently large to shift performance beyond the dip. Future research could systematically manipulate the properties of the visual and tactile stimuli (e.g., the position or brightness of the flash) to provide a detailed understanding of the conditions necessary for sensory fusion.

The shift in the dipper function could reflect the population response of a group of multimodal neurons located in subcortical (Stein & Meredith, 1993) or cortical (Graziano & Gross, 1995) areas or changes in response of unimodal neurons receiving feedback from multimodal areas (Schroeder et al., 2001). For example, electrophysiological recordings in the posterior parietal cortex of monkeys have identified populations of neurons that respond to both visual and tactile stimulation and that have a visual receptive field that can track the spatial location of the hand (Avillac, Deneve, Olivier, Pouget, & Duhamel, 2005; Duhamel, Colby, & Goldberg, 1998; Iriki, Tanaka, & Iwamura, 1996; Ohayashi, Tanaka, & Iriki, 2000). These bimodal neurons may contribute directly to tactile perception or modulate gain within somatosensory regions of anterior parietal cortex (Fiorio & Haggard, 2005; Harris et al., 2007; Ro, Wallace, Hagedorn, Farne, & Pienkos, 2004; Taylor-Clarke et al., 2002). Alternatively, direct connections between primary sensory cortices (Falchier, Clavagnier, Barone, & Kennedy, 2002) might be involved in multimodal interactions. The evidence that primary sensory cortices can be functionally rewired (Von Melchner, Pallas, & Sur, 2000) points to a commonality in the way they process sensory inputs that could permit the merging of modalities in human perception that we have demonstrated here.

Acknowledgments—This work was supported by grants from the Australian Research Council to J.A.H. and C.W.G.C. and by a long-term fellowship from the Human Frontiers Science Program to E.A.

REFERENCES


