Signal reliability modulates auditory–tactile integration for event counting

Jean-Pierre Bresciani and Marc O. Ernst

Max-Planck-Institute for Biological Cybernetics, Tuebingen, Germany

Correspondence to Dr Marc O. Ernst, Max-Planck-Institute for Biological Cybernetics, Spemannstr. 38, 72076 Tuebingen, Germany
Tel: +49 7071 601 566; fax: +49 7071 601 616; e-mail: marc.ernst@tuebingen.mpg.de

Received 12 March 2007; accepted 21 March 2007

Sequences of auditory beeps and tactile taps were simultaneously presented and participants were instructed to focus on one of these modalities and to ignore the other. We tested whether (i) the two sensory channels bias one another and (ii) the interaction depends on the relative reliability of the channels. Audition biased tactile perception and touch biased auditory perception. Lowering the reliability of the auditory channel (i.e. the intensity of the beeps) decreased the effect of audition on touch and increased the effect of touch on audition. These results show that simultaneous auditory and tactile stimuli tend to be automatically integrated in a reliability-dependent manner. *NeuroReport* 18:1157–1161 © 2007 Lippincott Williams & Wilkins.

Keywords: auditory, integration, multisensory, perception, touch

Introduction

Recent experiments showed that the tactile perception of sequences of taps delivered on the index fingertip could be biased if task-irrelevant auditory beeps are simultaneously presented [1,2]. More specifically, the participants perceived more and less taps when more and less beeps were presented [2]. The tendency of the central nervous system is to automatically integrate stimuli provided by different sensory channels when these stimuli are likely to be generated by the same physical event. Such automatic integration has a functional relevance as redundant sensory signals reduce the variance of perceptual estimates [3,4] and enhance stimulus detection [5–7]. These benefits probably result from the existence of multimodal neurons whose firing frequency is more likely to increase when multiple rather than single sensory inputs are available [8]. More specifically, concerning the integration of auditory and tactile afferents, associative cortical areas [9–11] as well as areas traditionally considered as unisensory – like the somatosensory cortices [12,13] or the auditory cortices [12,14–17] – have been found to display a greater neural activation during simultaneous auditory and tactile stimulations than during auditory-alone and tactile-alone stimulations. Assuming reasonably that these neural processes underlie the perception of combined auditory–tactile stimuli, these two sensory channels should influence one another. Specifically, not only should audition bias touch but touch should also bias audition.

Auditory and tactile events in this experiment were simultaneously presented (i.e. maximal temporal overlap between the auditory and tactile sequence) and the participants were instructed to focus on one modality and to ignore the other. The modality that was task relevant varied across sessions, which allowed us to test whether the auditory and tactile modalities biased one another. We also varied the intensity of the auditory events across sessions and measured whether the nature of the auditory–tactile interaction – how much each modality biased the other one – changed accordingly. The aim was to assess whether auditory and tactile signals were integrated in a reliability-dependent manner. The relative reliability of a sensory channel corresponds to the relative uncertainty of the information it conveys, and is inversely proportional to its relative variance [3]. Several experiments suggest that when perceptual estimates are based on multiple redundant modalities, the relative weight of each modality is proportional to its relative variance [3]. We tested here whether the same principle also applies to situations in which one of the channels is explicitly task irrelevant. If this were the case, it would suggest that weighted integration is a ‘generic’ principle by which information uncertainty is accounted for when sensory inputs are processed.

Methods

Twenty-four right-handed participants (aged 18–29, mean=24 years) participated in the experiment. None of these participants had a history of sensorimotor or auditory disorder. All participants gave their informed consent before taking part in the experiment, which was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

The participants were seated, with their right forearm and hand resting palm upwards at belly level on a table (72 cm high). A PHANToM (SensAble Technologies, Woburn, Massachusetts, USA) force-feedback device (see Fig. 1a) fixed to the table was used to generate the tactile stimuli (taps of 1 N indenting the skin of the index fingertip by approximately 2 mm) via a metallic pin of 3 mm diameter. The participants could not see their hands or the force-feedback device. For the whole duration of the experiment, participants wore earphones that continuously emitted a white noise (71 dB), which masked the noise generated by the tactile stimuli. The earphones were also used to present the auditory stimuli (beeps, 790 Hz, 74 dB when loud, 41 dB when quiet). The participants launched the trials and gave their responses using a keypad placed on their laps. The responses were given after each trial. The participants reported on how many events they perceived in the target modality, being free to enter any number as a response.

The experiment was composed of four sessions during which tactile taps and auditory beeps were presented in temporal overlap (see Fig. 1b). For each session, the participants were instructed to count the number of events in one modality (Target) and ignore the other modality (Background). For each trial, a sequence of two to four events was presented in the target modality. The number of events presented in the background modality could be zero (Target alone), one less (#Background=#Target−1), the same number (#Background=#Target) or one more (#Background=#Target + 1). Each session therefore consisted of 12 experimental conditions. Participants performed 10 trials per experimental condition, for a total of 120 trials per session. For each session, all 12 experimental conditions were intermixed and the trials were presented in a random order.

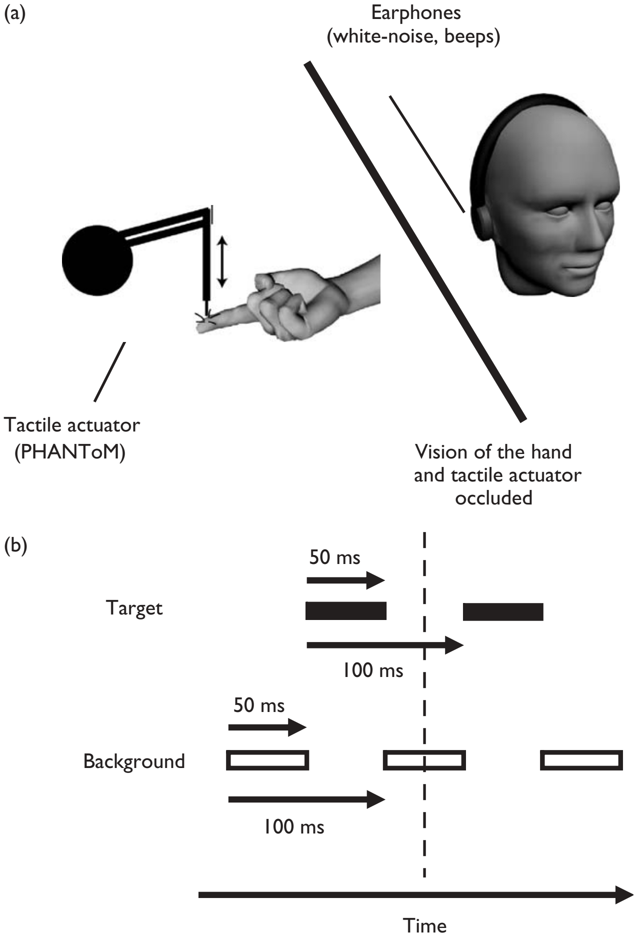

Fig. 1 (a) Experimental setup. (b) Temporal profiles of the stimuli for all sessions. The delay before the onset of the background sequence was systematically adjusted so that the middle of the target and background sequences coincided with respect to time. The example given here corresponds to a trial in which two events were presented in the Target modality and one more event in the background modality.

The target was touch in two sessions (Touch Loud and Touch Quiet) and audition in the other two sessions (Audition Loud and Audition Quiet). The auditory beeps were loud (74 dB, signal-to-noise ratio=3 dB) in two sessions (Touch Loud and Audition Loud) and quiet (41 dB, signal-to-noise ratio=−30 dB) in the other two sessions (Touch Quiet and Audition Quiet). The duration of the taps and beeps was 50 ms, and the delay between the onsets of two successive events in both the tactile and auditory sequences was 100 ms. The delay before the onset of the auditory sequence was systematically adjusted so that the middle portions of the auditory and tactile sequences coincided in time. The auditory and tactile sequences were aligned in time when the two sequences had the same number of events and were, otherwise, temporally interleaved. This adjustment allowed a maximal overlap between the auditory and tactile sequences for trials in which the number of events in the respective sequences differed (i.e. one event less and one event more). All participants participated in the four sessions consecutively, but there was a different order for each participant. The total experiment lasted about 1 h (15 min per session).

For the data analyses, all statistical tests were made using repeated-measures analyses of variance (ANOVAs). Post hoc comparisons using Newman–Keuls tests ($P<0.05$) were performed when necessary.

Results

For each session, we tested whether the background influenced the percept. The individual averages were entered in a 3*4 [number of events in the target (2, 3, 4)*number of events in the background (target alone, one event less, same number of events, one event more)] ANOVA.

The perceived number of events always depended on the actual number of delivered events in the target modality [F (2, 46) = 732.47, 551.78, 903.93 and 520.29 for 'Touch Loud', 'Touch Quiet', 'Audition Loud' and 'Audition Quiet', respectively; P always <0.001]. In each of the four sessions, participants' responses significantly differ between the three Target levels, indicating that the participants always distinguished whether two, three or four events were presented (see Fig. 2). More interestingly though, the perceived number of events also depended on the number of events presented in the background [F (3, 69) = 55.40, 20.92, 8.42 and 51.98 for 'Touch Loud', 'Touch Quiet', 'Audition Loud' and 'Audition Quiet', respectively; P always <0.001]. Participants always perceived significantly fewer events for the 'one event less' than for the 'one event more' condition, except for the 'Audition Loud' session. For the 'Audition Loud' session, the post hoc tests revealed no difference between the four background levels, although, as mentioned above, there was a significant main effect of the background. For the 'Audition Quiet' session, all background levels differed from one another except for 'target alone' and 'one event less'. For the 'Touch Loud' session, all background levels differed from one another except for 'target alone' and 'same amount'. Finally, for the 'Touch

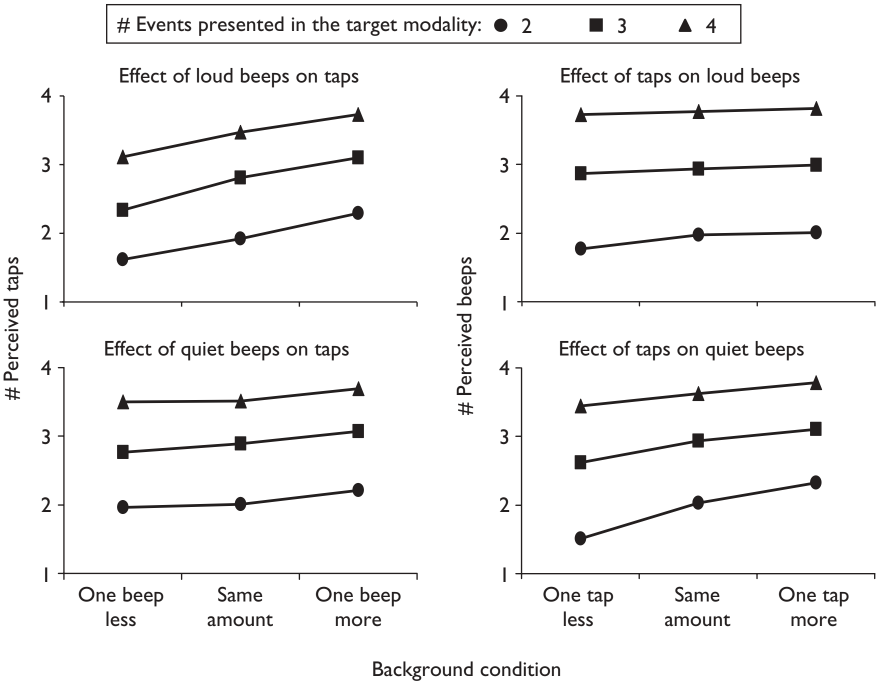

Fig. 2 Number of perceived events in the target modality as a function of both the actual number of delivered events (circle for two, square for three and triangle for four events) and the background condition. The two graphs on the left-hand side correspond to the sessions in which the target modality was touch (i.e. participants had to count taps) and the two graphs on the right-hand side to the sessions in which the target modality was audition (i.e. participants had to count beeps). The two upper graphs correspond to the sessions in which the beeps were loud and the two lower graphs to the sessions in which the beeps were quiet.

Quiet' session, the 'one event more' differed from the other three levels, which did not differ from one another.

For each session, the 'target alone' trials (i.e. no Background was presented) were used as the baseline of participants' perceptions of Target. The variability of these baseline responses was compared between the four different sessions. Individual standard deviations for the target alone trials were entered in a 4*3 [session (Touch Loud, Touch Quiet, Audition Loud, Audition Quiet) * number of events in the target sequence (2, 3, 4)] ANOVA.

A main effect of the session on the variability of the responses [F (3, 69) = 45.63, $P < 0.001$] was seen, and no effect of the number of events on variability of responses could be observed. The post hoc test revealed that the estimates were more variable when the participants had to count quiet beeps than for the other three sessions, which did not significantly differ from one another (see Fig. 3a). This confirmed that lowering sound intensity reduced the reliability of the auditory signals. On the other hand, responses' variability did not depend on the number of events in the target sequence.

Finally, we tested whether changing the intensity of the beeps quantitatively altered touch–audition interaction. To do so, we first computed background-evoked errors for each session and for each participant. The mean responses obtained in the trials in which only the target was presented were subtracted from the mean responses for the trials in which a background was presented (i.e. one event less, same number of events, and one event more background conditions). The individual errors were then averaged across the three 'target' conditions (i.e. two, three and four events presented in the Target modality) and a regression line was fitted to the means. The slope of the regression line represents the influence of the background. A slope of zero would indicate that the background does not influence the percept at all, whereas a slope of one would indicate that the percept is completely determined by the background. The individual slopes were entered in a 4*1 [session (Touch Loud, Touch Quiet, Audition Loud, Audition Quiet)] ANOVA.

As depicted in Fig. 3b, background loud beeps had a stronger influence ($P < 0.001$) on tap perception (average slope of 0.34) than quiet beeps did (average slope of 0.13). Similarly, background taps had a stronger influence ($P < 0.001$) on the perception of quiet beeps (average slope of 0.27) than on the perception of loud beeps (average slope of 0.08).

With loud beeps, the influence of audition on touch was stronger than the influence of touch on audition ($P < 0.001$). With quiet beeps, the pattern of results was reversed, the influence of touch on auditory perception being stronger than the influence of audition on tactile perception ($P < 0.001$).

Discussion

Several studies showed that when stimuli are simultaneously presented in two modalities and the participants instructed to focus on one of these modalities, the perceptual estimates can be biased by the to-be-ignored modality [1,2,18–21]. In line with this, the results of the current experiment confirmed earlier results showing that task-irrelevant auditory beeps can bias the perception of tactile taps [1,2]. Such biases were, however, usually only

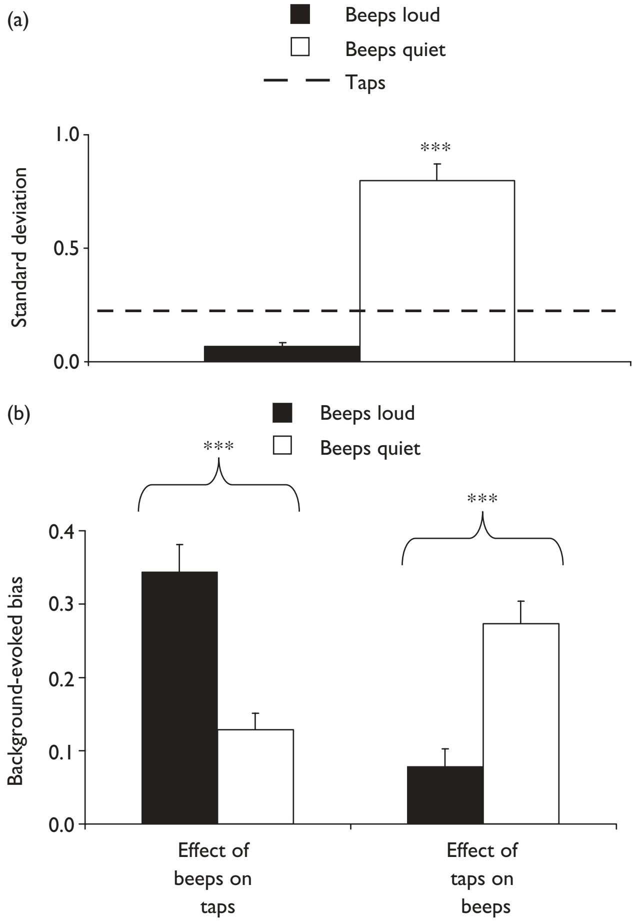

Fig. 3 (a) Average standard deviations of the individual estimates when only the target modality was presented (target alone). The two columns correspond to the sessions in which the target modality was audition (i.e. participants had to count the number of beeps). The filled column corresponds to the session in which the beeps were loud and the empty column to the session in which the beeps were quiet. The dashed line indicates the average standard deviation of the estimates for the two sessions in which the participants had to count the number of taps. The variability was significantly higher when the participants had to count quiet beeps than for any other session. (b) Slopes of the regression lines representing the average effect of the background modality on participants' perception of the target modality for each of the four sessions. For each session, background-evoked errors with respect to the condition where only the focal signal was presented were calculated. Those errors were averaged across the different target conditions (two, three and four events in the sequence) and the regression lines fitted to the means. A slope of zero would indicate that the background modality does not influence the percept at all, whereas a slope of one would indicate that the percept is completely determined by the background modality. The two columns on the left-hand side represent the effect of the background auditory signal on tactile perception of taps and the two columns on the right-hand side represent the effect of the background tactile signal on auditory perception of beeps. Filled columns correspond to the sessions in which the auditory beeps were loud and empty columns to the sessions in which the auditory beeps were quiet.

observed one way, i.e. if modality A biased estimates based on modality B, then modality B failed to bias estimates based on modality A [18–21]. In contrast with these results, we found here a two-way bias between audition and touch, the two modalities biasing one another. The perceived number of tactile taps depended on the number of task-irrelevant auditory beeps and, conversely, the perceived number of beeps depended on the number of task-irrelevant taps. This perceptual two-way bias likely reflects the neural integration of auditory and tactile afferents in brain structures such as the associative cortical areas [9–11], the somatosensory cortices [12,13] and the auditory cortices [12,14–17].

The strength of the relative biases between audition and touch depended on the relative reliability of each modality. With loud beeps, audition was more reliable than touch and the auditory-evoked bias of touch was stronger than the touch-evoked bias of audition. Reducing auditory reliability – as confirmed by the increase in variance of the auditory alone estimates – reversed the pattern of results. In particular, with quiet beeps, the touch-evoked bias of audition was stronger than the auditory-evoked bias of touch. This corresponded to the fact that touch was more reliable than audition with quiet beeps. In other words, which modality dominated in terms of evoked bias depended on the relative reliability of the two modalities. These results are reminiscent of weighted models of multimodal integration, which state that the relative weight allocated to the different channels is inversely proportional to the relative variance of each channel [3,4]. The key idea of these models is that the central nervous system takes into account the relative uncertainty of the information provided by the different sensory channels to come up with a percept that is statistically nearly optimal. This approach has been gaining general acceptance in the past few years both for understanding data collected in behavioral experiments [4,22] and for the development of neural-computational models [23,24]. For example, Ma and colleagues [24] recently proposed that the distribution of neuron populations could 'use' the firing rate variability of individual neurons to code information uncertainty in a statistically optimal way. To date, the behavioral studies that showed that multimodal integration is reliability dependent essentially consisted in providing redundant sensory information about a given stimulus in two different sensory channels and asking the participants to perform an estimate based on both channels [3,24]. In contrast with these studies, in our experiment, only one channel was task relevant. Indeed, our participants were instructed to focus on one modality and to ignore the other one. Our results, nevertheless, show that audition and touch were integrated in a reliability-dependent manner. This is interesting because it suggests that weighted integration is a 'generic' principle by which information uncertainty is accounted for when sensory inputs are processed.

Conclusion

Our results confirmed that task-irrelevant auditory beeps can bias the perception of tactile taps [1,2] and showed that, conversely, task-irrelevant tactile taps can bias auditory perception. With loud beeps, the bias of audition on touch was strong and the bias of touch on audition was small. Reducing the intensity of the beeps reduced the bias of audition on touch and enhanced the bias of touch on audition. Taken together, these results show that (i) auditory and tactile sensory signals tend to be automatically integrated, (ii) the perceptual bias resulting from the integration process is two-way and (iii) the relative reliability of the signals modulates the nature of the integration process.

Acknowledgements

This work was supported by the Max-Planck Society and by the 5th Framework IST Program of the EU (IST-2001-38040, TOUCH-HapSys). The authors thank Christoph Kayser and Roland Fleming for helpful comments on an earlier version of this manuscript.

References

1. Hötting K, Röder B. Hearing cheats touch, but less in congenitally blind than in sighted individuals. *Psychol Sci* 2004; **15**:60–64.
2. Bresciani JP, Ernst MO, Drewing K, Bouyer G, Maury V, Kheddar A. Feeling what you hear: auditory signals can modulate tactile tap perception. *Exp Brain Res* 2005; **162**:172–180.
3. Ernst MO, Bülthoff HH. Merging the senses into a robust percept. *Trends Cogn Sci* 2004; **8**:162–169.
4. Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 2002; **415**:429–433.
5. Hershenson M. Reaction time as a measure of intersensory facilitation. *J Exp Psychol* 1962; **63**:289–293.
6. Bernstein IH, Clark MH, Edelstein BA. Effects of an auditory signal on visual reaction time. *J Exp Psychol* 1969; **80**:567–569.
7. Gielen SC, Schmidt RA, van den Heuvel PJ. On the nature of intersensory facilitation of reaction time. *Percept Psychophys* 1983; **34**:161–168.
8. Stein BE, Wallace MW, Stanford TR, Jiang W. Cortex governs multisensory integration in the midbrain. *Neuroscientist* 2002; **8**:306–314.
9. Hyvarinen J, Poranen A. Function of the parietal associative area 7 as revealed from cellular discharges in alert monkeys. *Brain* 1974; **97**: 673–692.
10. Hikosaka K, Iwai E, Saito H, Tanaka K. Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *J Neurophysiol* 1988; **60**:1615–1637.
11. Leinonen L, Hyvarinen J, Sovijarvi AR. Functional properties of neurons in the temporo-parietal association cortex of awake monkey. *Exp Brain Res* 1980; **39**:203–215.
12. Foxe JJ, Morocz IA, Murray MM, Higgins BA, Javitt DC, Schroeder CE. Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain Res Cogn Brain Res* 2000; **10**:77–83.
13. Lutkenhoner B, Lammertmann C, Simoes C, Hari R. Magnetoencephalographic correlates of audiotactile interaction. *Neuroimage* 2002; **15**:509–522.
14. Schroeder CE, Lindsley RW, Specht C, Marcovici A, Smiley JF, Javitt DC. Somatosensory input to auditory association cortex in the macaque monkey. *J Neurophysiol* 2001; **85**:1322–1327.
15. Fu KM, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, *et al*. Auditory cortical neurons respond to somatosensory stimulation. *J Neurosci* 2003; **23**:7510–7515.
16. Kayser C, Petkov CI, Augath M, Logothetis NK. Integration of touch and sound in auditory cortex. *Neuron* 2005; **48**:373–384.
17. Foxe JJ, Schroeder CE. The case for feedforward multisensory convergence during early cortical processing. *NeuroReport* 2005; **16**:419–423.
18. Shipley T. Auditory flutter-driving of visual flicker. *Science* 1964; **145**:1328–1330.
19. Kitagawa N, Ichihara S. Hearing visual motion in depth. *Nature* 2002; **416**:172–174.
20. Guest S, Spence C. Tactile dominance in speeded discrimination of textures. *Exp Brain Res* 2003; **150**:201–207.
21. Recanzone GH. Auditory influences on visual temporal rate perception. *J Neurophysiol* 2003; **89**:1078–1093.
22. Körding KP, Wolpert DM. Bayesian integration in sensorimotor learning. *Nature* 2004; **427**:244–247.
23. Deneve S, Latham PE, Pouget A. Efficient computation and cue integration with noisy population codes. *Nat Neurosci* 2001; **4**:826–831.
24. Ma WJ, Beck JM, Latham PE, Pouget A. Bayesian inference with probabilistic population codes. *Nat Neurosci* 2006; **9**:1432–1438.