

Auditory motion affects visual biological motion processing

A. Brooks^{a,b}, R. van der Zwan^b, A. Billard^c, B. Petreska^c, S. Clarke^d, O. Blanke^{a,*}

^a Laboratory of Cognitive Neuroscience, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

^b Laboratory of Human and Animal Perceptual Processing, School of Psychology, Southern Cross University, Coffs Harbour, NSW 2457, Australia

^c Autonomous Systems Lab 3, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

^d Division Autonome de Neuropsychologie, Centre Hospitalier Universitaire de Lausanne, 1011 Lausanne, Switzerland

Received 26 August 2005; received in revised form 29 November 2005; accepted 18 December 2005

Available online 28 February 2006

Abstract

The processing of biological motion is a critical, everyday task performed with remarkable efficiency by human sensory systems. Interest in this ability has focused to a large extent on biological motion processing in the visual modality (see, for example, Cutting, J. E., Moore, C., & Morrison, R. (1988). Masking the motions of human gait. *Perception and Psychophysics*, 44(4), 339–347). In naturalistic settings, however, it is often the case that biological motion is defined by input to more than one sensory modality. For this reason, here in a series of experiments we investigate behavioural correlates of multisensory, in particular audiovisual, integration in the processing of biological motion cues. More specifically, using a new psychophysical paradigm we investigate the effect of suprathreshold auditory motion on perceptions of visually defined biological motion. Unlike data from previous studies investigating audiovisual integration in linear motion processing [Meyer, G. F. & Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *Neuroreport*, 12(11), 2557–2560; Wuerger, S. M., Hofbauer, M., & Meyer, G. F. (2003). The integration of auditory and motion signals at threshold. *Perception and Psychophysics*, 65(8), 1188–1196; Alais, D. & Burr, D. (2004). No direction-specific bimodal facilitation for audiovisual motion detection. *Cognitive Brain Research*, 19, 185–194], we report the existence of direction-selective effects: relative to control (stationary) auditory conditions, auditory motion in the same direction as the visually defined biological motion target increased its detectability, whereas auditory motion in the opposite direction had the inverse effect. Our data suggest these effects do not arise through general shifts in visuo-spatial attention, but instead are a consequence of motion-sensitive, direction-tuned integration mechanisms that are, if not unique to biological visual motion, at least not common to all types of visual motion. Based on these data and evidence from neurophysiological and neuroimaging studies we discuss the neural mechanisms likely to underlie this effect. © 2006 Elsevier Ltd. All rights reserved.

Keywords: Audiovisual; Crossmodal; Visual sensitivity; Point light walker; Interaction

1. Introduction

The task of detecting and analysing motion on the basis of multisensory and in particular audiovisual input is one that is performed with remarkable success by the human sensory system. Perhaps in order to maximise performance on this ecologically significant task (in modern-day terms it facilitates everyday activities such as catching a ball, crossing the road and navigating a busy footpath), various neural mechanisms for integrating input into the different sensory modalities have evolved. The correlates of such mechanisms are observed at a phenomenological level in illusions such as the one reported by Sekuler, Sekuler,

and Lau (1997), in which perceptions of identical visual motion cues are affected by the introduction of auditory cues. In a bid to explore and identify the mechanisms underlying such effects – a task of increasing focus in the field of multisensory research – behavioural, neurophysiological and other techniques have been employed. The resulting data provide a basis from which to better understand not only the neural mechanisms underlying audiovisual interactions in motion processing in particular, but also multisensory processing within the human sensory system more generally.

The role of subcortical structures in audiovisual integration has been well-documented. Behaviourally it has been demonstrated that, relative to unimodal stimulus presentations, audiovisual presentations can result in reduced saccadic reaction times to a visual target (see Colonius & Arndt, 2001 for a review). These effects, which are observed only when auditory and

* Corresponding author. Tel.: +41 21 6939621; fax: +41 21 6939625.
E-mail address: olaf.blanke@epfl.ch (O. Blanke).

visual stimuli are spatially and temporally coincident, have been attributed to activity in the superior colliculus (see Meredith, Nemitz, & Stein, 1987; Meredith & Stein, 1996). Neurophysiologically it has been demonstrated that some cells within this structure display ‘superadditive’ response characteristics. That is, they exhibit non-linear enhancement or suppression of activity in response to consistent and inconsistent multisensory cues, respectively, (Meredith & Stein, 1996). This, combined with the well-described role of the superior colliculus in orienting visual attention via eye-movement generation, has formed the basis for the proposal that superadditive patterns of cellular activity within the superior colliculus may also play a role in audiovisual integration of *motion* information.

Behavioural tests of this proposal have yielded divergent results (Alais & Burr, 2004; Meyer, Wuerger, Rohrbein, & Zetzsche, 2005; Wuerger, Hofbauer, & Meyer, 2003). In the case of linear motion direction discrimination tasks such as those used by Wuerger et al. and Alais and Burr, the evidence argues against such a model. In both studies, performance in motion discrimination tasks under unisensory conditions was compared to performance under multisensory (audiovisual) conditions. Both sets of authors reported two key findings. The first was that only modest multisensory facilitation was observed. The second was an absence of directional effects: the same moderately facilitative results were observed irrespective of whether auditory and visual motions were in consistent directions. The first of these findings is clearly incompatible with the non-linear response properties of superior colliculus cells, and instead lead Alais and Burr (2004) to argue for a simple probability or general arousal model. The second suggests that audiovisual motion cues, at least for the types of motion tested in these studies, are also not subject to direction-tuned auditory and visual motion processing mechanisms located at higher neural levels (see below).

Multisensory and in particular audiovisual mechanisms have been shown to exist in both cat (Stein, 2002) and monkey (Bremmer et al., 2001; Bruce, Desimone, & Gross, 1981; Perrett, Harries, Benson, Chitty, & Mistlin, 1990) cortex. There is evidence that integration at these loci is not necessarily based on the type of superadditivity observed in multisensory superior colliculus cells, and is in some cases specific to multisensory motion cues (see Bruce et al., 1981). Indeed, Bruce et al. (1981) reported the existence of audiovisual cells in macaque superior temporal polysensory area that show direction-selective motion sensitivity. In order to test specifically for behavioural correlates of multisensory integration in such areas, psychophysical techniques other than the unisensory/multisensory comparisons described above can be used. In the case of a study reported by Meyer and Wuerger (2001), for example, performance on visual detection of linear motion was measured in the presence of suprathreshold auditory motion in consistent and inconsistent directions. In that case, the authors reported an absence of direction-mediated audiovisual motion effects, suggesting their stimulus combinations were *not* driving direction-selective, motion sensitive integration mechanisms located in the cortex. However, consistent with a discussion point raised by Alais and Burr (2004), it remains entirely possible that *other types* of visually defined motion may yield different results.

Neuroimaging evidence in humans suggests visually defined biological motion as one such type. Posterior superior temporal cortex (henceforth pSTS) seems to be of prime importance in the analysis of visually defined biological motion, although several other areas in extrastriate cortex as well as in frontal cortex have been shown to be involved in the analysis of biological motion stimuli (Bonda, Petrides, Ostry, & Evans, 1996; Grossman & Blake, 2002; Howard et al., 1996; Saygin, Wilson, Hagler, Bates, & Sereno, 2004; Servos, Osu, Santi, & Kawato, 2002; Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001). These studies proposed that especially area pSTS is uniquely involved in the processing of visually defined biological motion (to the exclusion of other types of visually defined motion). Additionally and of critical importance in the current context, Howard et al. (1996) reported that pSTS is also involved in the processing of *auditory* cues. Indeed, it has been suggested that pSTS in humans is the human homologue of the macaque superior temporal polysensory area referred to above (Bruce et al., 1981).

On that basis, we hypothesised that perceptions of visually defined *biological motion* may be subject to audiovisual processing interactions. The aim of this study was to test for and investigate such interactions. To that end, we developed a new psychophysical paradigm to test performance detecting visually defined biological motion in the presence of suprathreshold auditory motion. We report evidence of direction-specific (facilitative and inhibitory) audiovisual interactions, and discuss the possible neural mechanisms mediating these effects.

2. Methods

2.1. Subjects

Subjects consisted of six (three males and three females) experienced psychophysical observers with normal or corrected-to-normal vision and normal auditory processing. All subjects were aged between 20 and 45 years, were right-handed, and, in accordance with the 1964 Declaration of Helsinki, gave informed consent to participate in the study. Two of the subjects are authors. The remaining subjects, while aware of the general purposes of the experiment, were naïve to its specific aims. In Experiment 2, data for five of the six original subjects are reported due only to the unavailability of the sixth subject for testing.

2.2. Visual psychophysical paradigm

Visual sensitivity to biological motion in the presence of simultaneously presented, suprathreshold auditory signals was established using a new psychophysical paradigm (see Fig. 1 and van der Zwan, Petreska, Billard, Blanke, & Brooks, in preparation). Within each visual presentation, biological point-light walkers (PLWs; first described by Johansson, 1973) were defined by 11 ‘signal’ dots. These dots moved across space as a group to generate perceptions of coherent biological motion such that the walker appeared to move either left-to-right (LR) or right-to-left (RL). Visible also during each presentation were 300 ‘visual noise’ dots. Each noise dot had its luminance, size, velocity, and net translation matched to one of the signal dots. The starting position of each noise dot was, however, randomly determined so as to produce no coherent biological motion percept. The noise dots were then divided into two groups: one described net translatory movement in a direction *consistent* with the PLW (Noisecon). Another way of thinking of these dots is that their movement is in-phase with the signal dots: a noise dot matched to, say, a signal dot corresponding to one wrist of the walker will move in exactly the same direction as the wrist dot and at exactly the same time. The other group of noise dots had a translatory move-

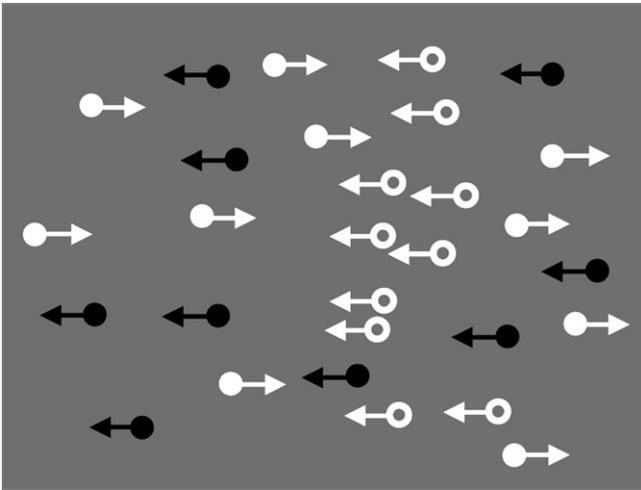


Fig. 1. Diagrammatic representation of a new psychophysical paradigm allowing measurement of sensitivity to visually defined biological motion at consistent stimulus densities. For the purposes of representation only, signal dots are presented in unfilled white, Noisecon dots in solid black, and Noiseopp dots in solid white. Again for the purposes of representation, arrows indicate the direction of net translatory motion of the dot populations across each 2000 ms display.

ment in the opposite direction (Noiseopp). These noise dots can be thought of as being out-of-phase with the signal dots. So, for example, if partnered with a wrist dot that was swinging forward during a step, the Noiseopp dot would move backward at the same moment in time.

The perceptual effect of this manipulation is that difficulty in detecting the direction of motion of the signal dots varies directly with the proportion of Noisecon dots, and inversely with the proportion of Noiseopp dots while maintaining a constant number of individual dots in the stimulus display. That is, difficulty in the task of extracting the PLW is reduced as proportions of Noiseopp are increased (see data below).

On the basis of pilot data, five visual conditions were tested: Noiseopp proportions of 5, 10, 15, 20 and 95% were selected. The concentration of low-percentage values and presence of a single high-percentage value (at which the task is easier) is consistent with signal-to-noise paradigms used in similar vision-based psychophysical paradigms (see, for example, van der Zwan, Badcock, & Parkin, 1999). Total dot density was held constant across frames by re-generating, in random starting positions, noise dots that moved beyond the borders of the viewable field in the previous frame.

2.3. Apparatus and stimuli

Each stimulus presentation was of 2000 ms duration, and consisted of a combined audio/visual presentation in which onset and offset of the two stimuli were simultaneous.

The auditory component of each presentation consisted of binaural white noise bursts (44.1 kHz sampling rate, Butterworth band-pass filter: 400 Hz cut-off and 800 Hz high cut-off) digitised on a P4 3 GHz fitted with an Audigy 2 ZS sound card and Adobe Audition 1.0 software. Stimuli were presented using PHILIPS HP800 headphones at a volume of approximately 80 dB. Auditory stimuli were generated in five different conditions; directed (experimental) conditions during which the auditory stimulus moved in a direction consistent with (DIRcon) or inconsistent with (DIRincon) the visually defined PLW, and stationary (control) conditions during which the perceived spatial position of the auditory stimulus was central (STAcentre), left (STAlleft) or right (STARight) of the midsagittal plane. In directed conditions, the auditory stimuli translated approximately 140° in space (from 70° on one side of the midsagittal plane to 70° on the other). This range exceeded the distance through which the visual target translated (30° of visual angle) but was selected in order to heighten the saliency of auditory motion. In lateralised stationary conditions (STAlleft and STARight), stimuli were located at the most extreme right and left spatial locations described by the auditory motion cues. Interaural time differences (ITDs)

were used to generate the spatial characteristics of the stimuli (see, for example, Ducommun et al., 2004). In addition, 70 ms of fade-in and fade-out was incorporated in order to eliminate audible clicks at stimulus onset and offset.

Visual stimuli were presented on a 19 in. ViewSonic CRT monitor at a viewing distance of 100 cm. The monitor was driven at 100 Hz using a display resolution of 640 pixels × 480 pixels. Visual displays or movies consisted of 50 frames presented for 40 ms each—specifications that, in the absence of visual noise, give rise to strong perceptions of fluid biological motion. Within each display the PLW described an overall translation of approximately 30° of visual angle (from 15° on one side of the midsagittal plane to 15° on the other). The luminance value assigned to each dot population was equivalent at 50 cd/m². Dots were presented on a grey background (luminance value = 43 cd/m²). The low contrast value (Michelson contrast = 0.075) was selected in order to reduce the saliency of the visual signal and thus enhance the possibility of observing audiovisual effects. The diameter of each dot was approximately 18 arcmin of visual angle. No fixation point was used, and subjects were free to make eye-movements during each stimulus presentation. The spatial coordinates for dots were generated using MatLab (Version 7.0.1) software. Purpose-built psychophysical software (RUNSTIM Version 4.1.2) was used to present audiovisual stimulus combinations and for the purposes of data recording.

2.4. Procedure

Subjects were seated in a light- and sound-attenuated room. Following a demonstration of the target (the translating PLW) in the absence of visual noise, subjects completed two practice blocks for familiarisation with the task. Data for each experiment were then collected in a single testing session consisting of five blocks. Each block consisted of a total of 50 trials. The visual stimulus was randomly generated for every trial and the visual target was always present. On half the trials the target was moving in a LR direction and in an RL direction on the remaining trials. Audiovisual combinations were presented in random order and each of the 25 audiovisual conditions (5 visual: 5, 10, 15, 20, 95% Noiseopp × 5 auditory: DIRcon, DIRincon, STAcentre, STAlleft, STARight) was equally probable within a block. The two directions of biological motion (LR and RL) were combined such that each audiovisual condition was tested, for each subject, by 10 trials in a repeated measures design. This design maximised the number of trials that could be run within the testing window available for each subject (approximately 90 min). Inter-stimulus intervals were of 2000 ms duration. Subjects were instructed to indicate by means of a key-press the direction (LR or RL) in which the target stimulus (the PLW) appeared to move as quickly and accurately as possible. On any given trial, there was an equal chance the visual signal would translate in either a LR or RL direction. Both response accuracy and response latency were measured. It was anticipated that if biological motion perception is subject to facilitative and/or inhibitory audiovisual interactions, performance in the experimental (directed) auditory conditions would differ significantly and in predictable ways relative to performance observed in control (stationary) auditory conditions. That is, relative to stationary controls it was expected that auditory motion in a consistent direction with the visual target would improve performance, while auditory motion in the opposite direction would inhibit performance.

2.5. Data analysis

All experiments were run using a repeated-measures, within-subjects design. As the testing window was constrained by subject comfort (limiting each condition to 10 trials per subject for every subject) individual means for each condition were grouped across subjects and the data analysed using repeated-measures ANOVA testing planned, orthogonal contrasts and simple-effects.

3. Results

Here we will focus first on data related to responses latencies (or reaction times), with a discussion of response accuracy to follow.

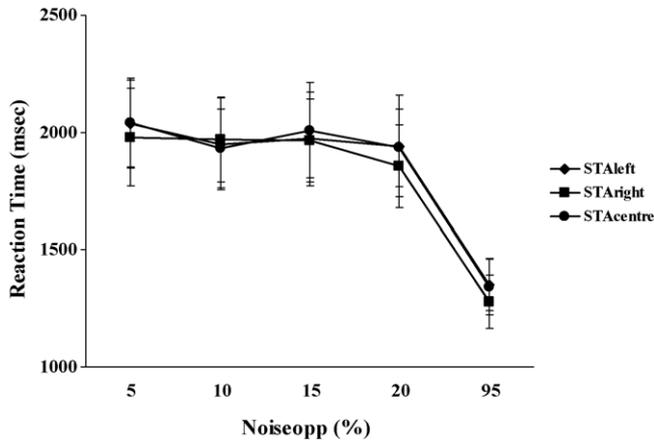


Fig. 2. Mean reaction times for each of the three control conditions: proportion of noise dots moving in the direction opposite to the direction of translation of the PLW (Noiseopp) is shown on the abscissa. Reaction time, in milliseconds, is shown on the ordinate. The only systematic variation arises as a function of the number of Noiseopp dots. Bars show one standard error.

3.1. Experiment 1—the point-light walker

Data for six subjects were collected for both control and experimental conditions across the five “Noiseopp” densities. The mean reaction times for correctly identifying the direction of motion of the PLW, across subjects, for each of the three control conditions (STAleft, STARight and STAcetre) are shown in Fig. 2. Planned contrasts showed that there were no significant differences between STAleft and STARight ($F_{1,14} = 0.22$, $p > 0.05$). Similarly, there was no difference between STAcetre and the other two lateralised control conditions ($F_{1,14} = 0.08$, $p > 0.05$). In other words, there were no systematic variations in reaction times elicited by the different locations of the stationary auditory cues. With this in mind, stationary position was ignored and these scores were treated as if they are from a single condition (see below). There was, however, a significant linear reduction in mean reaction times as Noiseopp increased ($F_{1,14} = 32.45$, $p < 0.01$): that is, response latencies for identifying the direction of translation of the PLW decreased as the number of noise dots moving in the opposite direction to the walker increased. This result suggests subjects became faster at making their discriminations as the PLW became easier to detect.

In comparison to the control data, there were significant differences between the experimental conditions (auditory cue moving in the same direction as the PLW—DIRcon; auditory cue moving in the opposite direction to the PLW—DIRincon). Those data are illustrated here in two ways (Figs. 3 and 4). Fig. 3 shows group means, across noise conditions, for both DIRcon and DIRincon. As in the previous figure, reaction time for correct discriminations of the direction of motion of the PLW is shown on the y-axis, and there is a significant difference between the consistent and inconsistent auditory cue effects ($F_{1,9} = 5.48$, $p < 0.05$). In other words, performance on the visual discrimination task is affected by the direction of motion of an auditory cue: an auditory cue moving in the same direction as the PLW leads to shorter response latencies, at least for some noise con-

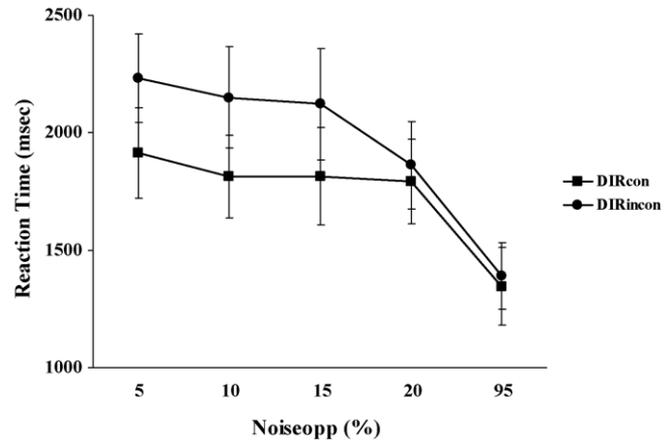


Fig. 3. Mean reaction times for discriminations of the direction of motion of the PLW when an auditory cue is moving in the same (DIRcon) or in the opposite (DIRincon) directions. A consistent auditory cue seems to make discrimination of the PLW easier in high Noiseopp conditions. That advantage disappears as it becomes easier to see the PLW. Bars show one standard error.

ditions, than those observed when an auditory cue is moving in the opposite direction.

To illustrate the nature of the effect of the direction of auditory motion on the discrimination task two sets of difference scores were calculated for each subject. The first set of difference scores compared the reaction times for those conditions in which the auditory and visual cues moved in the same direction (DIRcon) to the corresponding control scores. The other set of difference scores compared those conditions in which the visual and auditory cues moved in opposite directions (DIRincon) to the control scores, and these all are illustrated in Fig. 4. The difference scores are such that positive scores indicate an increase in reaction time in response to a moving auditory cue, while negative difference scores indicate a decrease in reaction time in response to a moving auditory cue (compared to reaction times when the

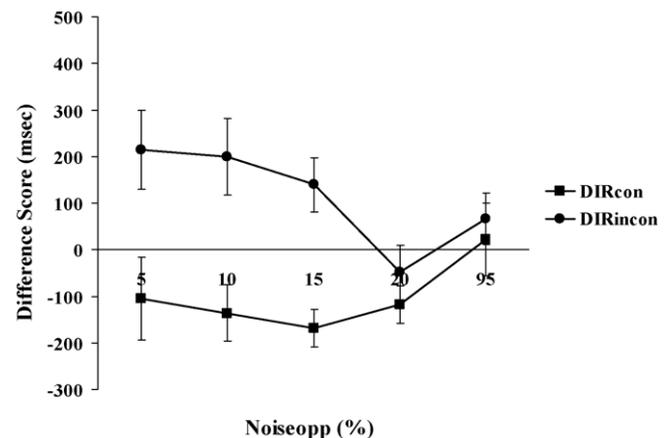


Fig. 4. Mean difference scores. These data were calculated by subtracting from the DIRcon and DIRincon scores corresponding control data (where the auditory cue was stationary). Negative scores indicate facilitation in discrimination, while positive scores indicate inhibition. These data show that having an auditory cue moving in the same direction as the PLW does make detection of the direction of motion of the PLW faster than if the auditory cue is moving in the opposite direction. The facilitatory effect disappears as it becomes easier to see the PLW. Bars show one standard error.

auditory cue is stationary). There was a significant improvement in performance in response to having the auditory cue moving in the same direction as the PLW ($F_{1,9} = 5.84, p < 0.05$). Tests of simple effects showed that there were no significant differences between the consistent and inconsistent scores for the 20 and 95% Noiseopp conditions ($F_{1,9} = 0.12, p > 0.05; F_{1,9} = 0.05, p > 0.05$, respectively).

To check that differences in response latencies were not a consequence of lateralised auditory stimuli simply directing attention to either the side of space on which the visual target first appeared (thereby improving performance as observed in the DIRcon condition) or to the side of space opposite to the one on which the target first appeared (thereby reducing performance as observed in the DIRincon condition) a final comparison between the experimental and control data was conducted. Performances in the moving auditory conditions were compared with control conditions in which the spatial location of the stationary auditory stimulus was either consistent (STAsartcon) or inconsistent (STAsartincon) with the side of space in which the visual target first appeared (a more conservative use of the control data). Again, the comparisons were made as difference scores: if effects observed in the previous analysis were simply the result of directing attention to either the correct or incorrect side of visual space, then the difference scores in this analysis should not be significantly different.

Mean differences are shown in Fig. 5. As for the previous analysis, the difference scores for consistent visual and auditory cues are, together, less than the difference scores for the inconsistent cues ($F_{1,9} = 7.59, p < 0.05$). Tests of simple effects again showed that there were no significant differences between the consistent and inconsistent scores for the 20 and 95% Noiseopp conditions ($F_{1,9} = 0.18, p > 0.05; F_{1,9} = 0.13, p > 0.05$, respectively). These data indicate that the effects of the auditory cues are attributable to their motion, and not to their potential to direct attention towards or away from the side on which the PLW starts.

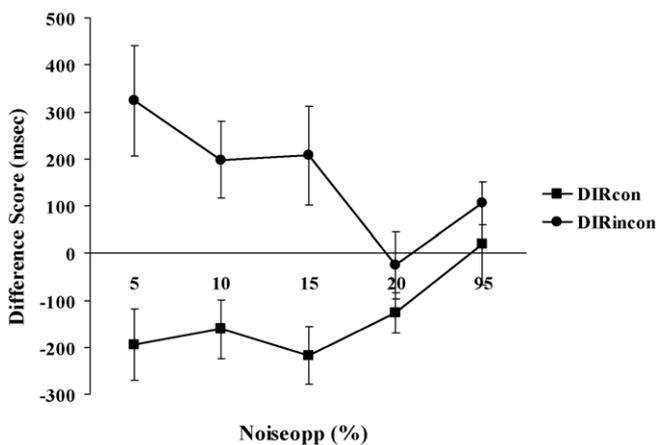


Fig. 5. Mean differences scores. These data were calculated by subtracting from the DIRcon and DIRincon scores corresponding control data where the side of the stationary auditory stimulus was the same side as that on which the PLW started. Again, negative scores indicate facilitation in discrimination, while positive scores indicate inhibition. These data show that it is the direction of motion of the auditory cue that is causing changes in performance. Bars show one standard error.

3.2. Experiment 2—the upside-down point-light walker

In order to confirm that the effects observed in Experiment 1 are attributable to interactions between the motions of the auditory cues and the biological motion of the PLW we conducted a second experiment using exactly the same techniques and conditions as Experiment 1, except that in this case the PLW was inverted (once again, characteristics of the two populations of visual noise dots were matched to randomly assigned signal dots). The use of an upside-down walker was designed to ensure that the effects observed in Experiment 1 were not attributable to changing density cues carried by the walker moving through the noisy array.

Comparisons were made between reaction times only for correct discriminations. As in Experiment 1, control data were collected for stationary auditory cues in three positions (STAlleft, STArigh and STAcenre) and again there were no differences in performance between the STAlleft and the STArigh conditions ($F_{1,14} = 0.35, p > 0.05$). Similarly, there was no difference between those conditions and STAcenre ($F_{1,14} = 0.46, p > 0.05$).

Unlike the data collected in Experiment 1, consistent and inconsistent auditory cues were not different from each other ($F_{1,9} = 0.25, p > 0.05$; see Fig. 6). A trend analysis did show that there was a significant liner reduction in reaction time as a function of noise density ($F_{1,9} = 12.33, p < 0.05$) indicating that the direction of motion of the upside-down walker became more easily discriminated as the visual discrimination task became easier. With those findings in mind, difference scores were again calculated using the control data (see Fig. 7) but unlike Experiment 1, there was no effect on performance of the direction of motion of the auditory cue ($F_{1,9} = 0.30, p > 0.05$).

3.3. Response accuracy

McNemar's test showed that, overall, there was no systematic difference between the error rates across the two experiments ($Z = 0.42, p > 0.05$). Errors across conditions for both Experi-

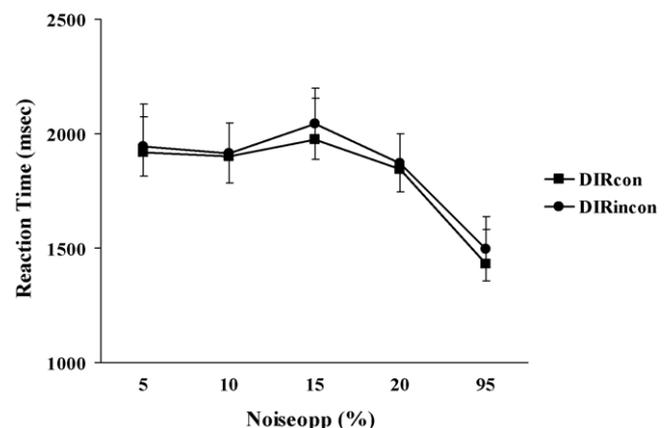


Fig. 6. Mean reaction times for discriminations of the direction of motion of the upside-down PLW when an auditory cue is moving in the same (DIRcon) or in the opposite (DIRincon) directions. Unlike the data obtained for the upright walker (see Fig. 2), there is no difference between the consistent and inconsistent conditions. Bars show one standard error.

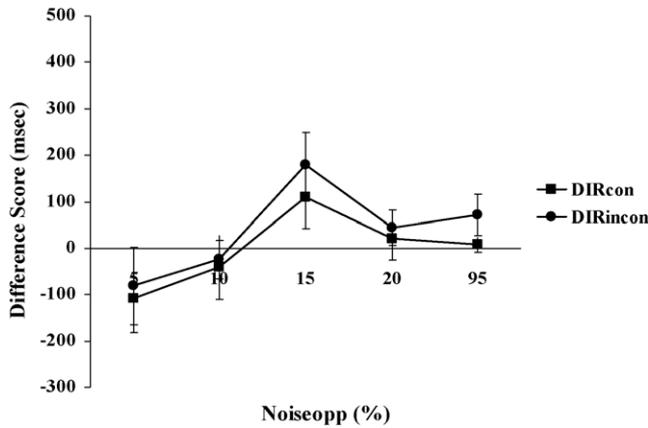


Fig. 7. Mean differences scores for the upside-down walker. These data were in the same as the difference scores in Experiment 1 (see Fig. 4). There was no significant effect of moving auditory cue for these visual stimuli. Bars show one standard error.

ments 1 and 2 are shown in Fig. 8. A repeated-measures analysis of variance showed that there was no effect of the position of a stationary auditory cue (centre, left or right) on errors made in judging the direction of the walker either in Experiment 1 or in Experiment 2 ($F_{2,10} = 0.62, p > 0.05, F_{2,8} = 1.48, p > 0.05$). With that in mind, proportions were collapsed to their mean value for each level of noise and the data for both experiments are presented in Fig. 8a. While there was an apparent decline in error rate as Noiseopp increased, and subjective reports suggested the task of discriminating the direction of the walker got easier, there was no significant effect of the level of noise on error rate for either experiment (Experiment 1: $F_{1,5} = 0.20, p > 0.05$; Experiment 2: $F_{1,4} = 3.58, p > 0.05$).

As Fig. 8b and c shows, error rates when the auditory cues were moving were low and alike for both experiments. Fig. 8b illustrates the error rates (as a mean proportion of the total number of trials) for Experiment 1: repeated-measures analysis of variance showed there was no difference between conditions in which the auditory and visual cues were in consistent and inconsistent directions ($F_{1,5} = 2.36, p > 0.05$). That is, the direction of motion of the auditory cue seems to have had no effect on the errors made. There was a significant change in errors across the noise conditions ($F_{1,5} = 10.31, p < 0.05$), but no interaction between the auditory cue direction of motion and noise ($F_{1,5} = 0.46, p > 0.05$). In other words, fewer errors seem to have been made when Noiseopp increased. Nonetheless, the changes in reaction time observed in Experiment 1 cannot be explained by a speed/accuracy trade-off: the only change in error rate observed was for high Noiseopp conditions (95%) and the reduction in errors probably reflects the ease with which subjects could do the task. Fig. 8c shows a similar pattern of errors for Experiment 2, but without any significant reduction in error rate across different noise conditions: again, there was no difference in error rate between the DIRcon and DIRincon conditions ($F_{1,4} = 1.43, p > 0.05$). There was no effect of noise on error rate ($F_{1,4} = 1.08, p > 0.05$), and no interaction ($F_{1,4} = 1.60, p > 0.05$). As was the case for Experiment 1, these data suggest that the systematic changes in reaction times observed across conditions cannot be explained by systematic changes in speed/accuracy trade-offs by subjects.

4. Discussion

The primary aim of these experiments was to test for and investigate effects of auditory motion on perceptions of visu-

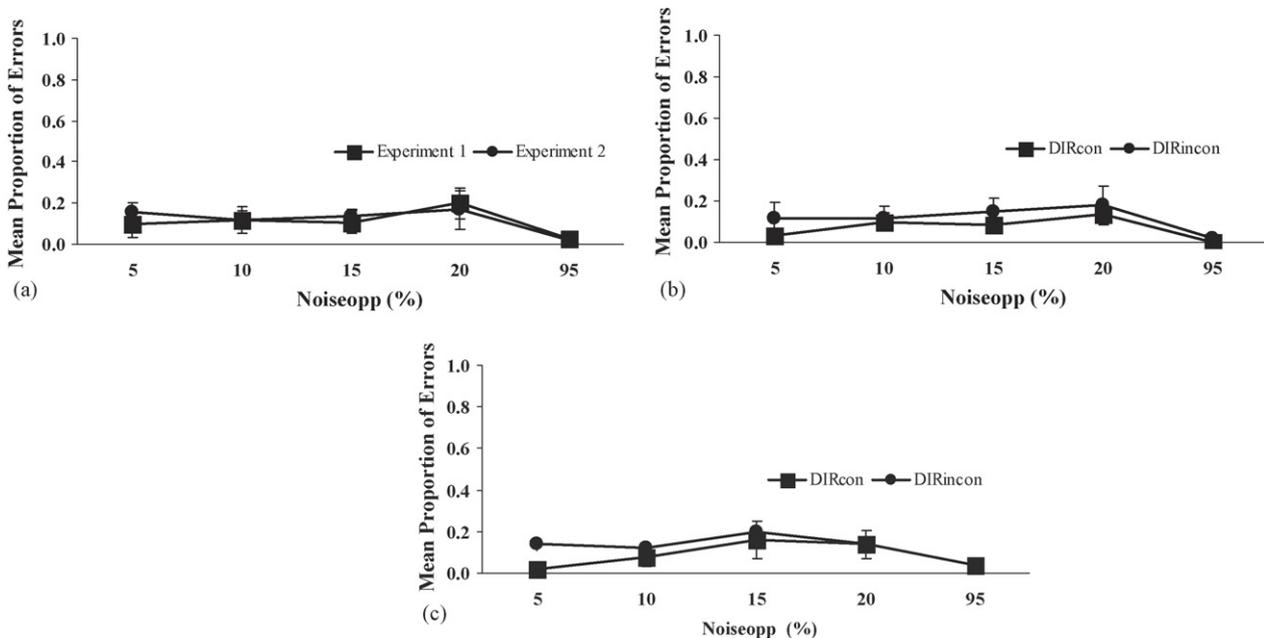


Fig. 8. Group mean error proportions. (a) Mean error rates for each noise condition for Experiments 1 and 2. No significant differences in error rates between the positions of the stationary auditory cues allowed single means to be calculated for each level of noise. (b) The error rates for the conditions in which the auditory cue moved for Experiment 1 and (c) for Experiment 2. Bars show one standard error.

ally defined biological motion. Here we report evidence that such interactions do arise. Moreover, the data indicate that these interactions are mediated by *direction of motion*: data from Experiment 1 constitute evidence that relative to stationary auditory cues, auditory motion in the same direction as the visual target (PLW) resulted in significantly better performance on the visual discrimination task, while auditory motion in the opposite direction gave rise to significantly worse (slower) performance on the task. This was the case for comparisons of performance with stationary cues located in central, left and right auditory space. Additionally, it appears the effects are related to global *biological* visual motion perceptions to the exclusion of at least some other types of visual motion. This is evidenced by analyses of data from Experiment 2 showing that, when applied to PLWs in an upside-down orientation, the same visual discrimination task does not give rise to equivalent facilitative and inhibitory effects—a difference that analyses of error rates across experiments suggests cannot be accounted for on the basis of variations in task difficulty.

Directional effects such as these have not, to the best of our knowledge, been recorded for audiovisual motion processing involving other types of visual motion (see Alais & Burr, 2004; Meyer & Wuerger, 2001; Wuerger et al., 2003). While the paradigm used here is different from previous studies (RT as the dependent variable rather than thresholds), the demonstration of directional effects has an important and novel implication: it suggests that unlike the general, non-direction-selective facilitative effects observed in, for example, visual motion discriminations based on linear trajectories (see again Alais & Burr, 2004), our effects are not compatible with general arousal models. Rather, it seems that in the particular (but not necessarily exclusive) case of biological motion processing, interactions in *direction-tuned* auditory and visual motion processing occur. This represents a novel and important finding in terms of the neural correlates of audiovisual biological motion processing, and the possible neural mechanisms underlying these effects are discussed below.

As noted above, there is evidence that superior colliculus cells are tuned for spatio-temporally coincident audiovisual cues (Meredith & Stein, 1996), and that they mediate multisensory processing via their role in orienting visuo-spatial attention through the generation of eye-movements (see Colonius & Arndt, 2001). Neither characteristic is consistent with the conditions under which audiovisual interactions were observed in the present study. In relation to the first characteristic, our data indicate that the interactions arose in spite of the fact that auditory and visual motions were not spatio-temporally coincident: differences in the objective distance through which each stimulus moved (see Section 2) were such that for almost the entire duration of each presentation, auditory and visual cues were *not* spatially coincident. In relation to the second characteristic, statistical comparisons of experimental and control data from Experiment 1 suggest that the effects were *not* mediated by shifts in visuo-spatial attention: had this been the case, lateralised static auditory cues located on either the consistent or opposite side of space to that on which the visual target first appeared should have yielded the same results as those observed in conditions where auditory motion was in the consistent or inconsistent direction

(relative to the visual target), respectively. While of course superior colliculus mediation of the effects cannot be definitively ruled out on the basis of these behavioural inconsistencies, they do indicate that such a possibility is unlikely.

Based on this evidence and more recent studies on the neural substrates of biological motion processing a more likely scenario is that audiovisual interactions are mediated at a higher, cortical, level of processing. Such a suggestion is consistent with literature suggesting the existence of cortical sites involved in visually defined biological motion perception (monkey: Bruce et al., 1981; Perrett et al., 1990; man: Bonda et al., 1996; Grossman & Blake, 2002; Howard et al., 1996; Saygin et al., 2004; Servos et al., 2002; Vaina et al., 2001). These studies suggest that a network of brain areas – including pSTS, ventral occipital cortex, area KO, area MT+/V5, areas along the intraparietal sulcus (IPS) and prefrontal cortex (Bonda et al., 1996; Grossman & Blake, 2002; Howard et al., 1996; Saygin et al., 2004; Servos et al., 2002; Vaina et al., 2001) – is activated during visually defined biological motion perception. Yet, only two of these ‘biological motion areas’ areas have also been shown to be involved in audiovisual integration: pSTS (monkey: Bruce et al., 1981; Perrett et al., 1990; man: Howard et al., 1996) and IPS (monkey: Andersen, 1997; Schlack, Sterbing-D’Angelo, Hartung, Hoffmann, & Bremmer, 2005; man: Bremmer et al., 2001; Calvert, Campbell, & Brammer, 2000). Given the key role of pSTS in visually defined biological motion perception (Grossman & Blake, 2002; Vaina et al., 2001), and, critically, the demonstration of a convergence of biological motion and auditory *motion* signals in the area (see Bruce et al., 1981; Perrett et al., 1990) as well as a convergence of auditory and visual signals related to hand-held moving tools in pSTS (Beauchamp, Argall, Bodurka, Duyn, & Martin, 2004; Beauchamp, Lee, Argall, & Martin, 2004), we propose it as the most likely candidate for mediating the effects presented here. An additional argument in favour of this cortical, pSTS-based account is that it provides a possible basis upon which to explain the absence of similar audiovisual interactions in the case of more elementary types of visual motion (see Alais & Burr, 2004), the processing of which does not appear to be mediated by the area.

The present study therefore gives rise to a number of important conclusions. Perhaps most critical of these is that we have produced behavioural evidence of facilitatory and inhibitory effects of auditory motion direction on perceptions of visually defined biological motion. Here we argue that these effects might be mediated not by subcortical multisensory processing, but by cortical integration mechanisms that are, if not unique to biological visual motion, at least not common to all types of visual motion. In addition to these findings, we report evidence of the effectiveness of a new psychophysical paradigm in measuring sensitivity to visually defined biological motion across a range of task-related difficulty (for statistical evidence of this efficacy, see analyses of the general downward trend in response latencies across visual conditions). The advantage of this technique is that, unlike earlier paradigms designed to manipulate biological motion detectability (see, for example, Cutting, Moore, & Morrison, 1988), it allows *density* of the visual signal to be held constant across conditions. Using this and other techniques it

will be possible not only to further explore the neural correlates of the audiovisual interactions reported here, but also to explore audiovisual interactions when biological motion is defined in both visual and auditory domains.

References

- Alais, D., & Burr, D. (2004). No direction-specific bimodal facilitation for audiovisual motion detection. *Cognitive Brain Research*, *19*, 185–194.
- Andersen, R. A. (1997). Multimodal integration for the representation of space in the posterior parietal cortex. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences*, *352*, 1421–1428.
- Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., & Martin, A. (2004). Unraveling multisensory integration: Patchy organization within human STS multisensory cortex. *Nature Neuroscience*, *7*(11), 1190–1192.
- Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, *41*(5), 809–823.
- Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *The Journal of Neuroscience*, *16*, 3737–3744.
- Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., & Hoffman, K. (2001). Polymodal motion processing in posterior parietal and premotor cortex: A human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron*, *29*(1), 287–296.
- Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, *46*(2), 369–384.
- Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*, 649–657.
- Colonius, H., & Arndt, P. A. (2001). A two-stage model for visual and auditory interaction in saccadic latencies. *Perception and Psychophysics*, *63*, 126–146.
- Cutting, J. E., Moore, C., & Morrison, R. (1988). Masking the motions of human gait. *Perception and Psychophysics*, *44*(4), 339–347.
- Ducommun, C. Y., Michel, C., Clarke, S., Adriani, M., Seeck, M., Landis, T., et al. (2004). Cortical Motion Deafness. *Neuron*, *43*, 765–777.
- Grossman, E. D., & Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron*, *35*, 1167–1175.
- Howard, R. R., Brammer, M., Wright, I., Woodruff, P. W., Bullmore, E. T., & Zeki, S. (1996). A direct demonstration of functional specialisation within motion-related visual and auditory cortex of the human brain. *Current Biology*, *6*(8), 1015–1019.
- Johansson, G. (1973). Visual motion perception. *Scientific American*, *236*(6), 76–88.
- Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of Neuroscience*, *7*(10), 3215–3229.
- Meredith, M. A., & Stein, B. E. (1996). Spatial determinants of multisensory integrations in cat superior colliculus neurons. *Journal of Neurophysiology*, *75*(5), 1843–1857.
- Meyer, G. F., & Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *Neuroreport*, *12*(11), 2557–2560.
- Meyer, G. F., Wuerger, S. M., Rohrbein, F., & Zetzsche, C. (2005). Low-level integration of auditory and visual motion signals requires spatial co-localisation. *Experimental Brain Research*, *166*, 538–547.
- Perrett, D. I., Harries, M. H., Benson, P. J., Chitty, A. J., & Mistlin, A. J. (1990). Retrieval of structure from rigid and biological motion: An analysis of the visual responses of neurones in the macaque temporal cortex. In A. Blake & T. Troscianko (Eds.), *AI and the eye* (pp. 181–200). John Wiley & Sons Ltd.
- Saygin, A. P., Wilson, S. M., Hagler, D. J., Jr., Bates, E., & Sereno, M. I. (2004). Point light biological motion perception activates human premotor cortex. *The Journal of Neuroscience*, *24*(27), 6181–6188.
- Schlack, A., Sterbing-D'Angelo, S. J., Hartung, K., Hoffmann, K. P., & Bremmer, F. (2005). Multisensory space representations in the macaque ventral intraparietal area. *The Journal of Neuroscience*, *5*, 4616–4625.
- Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*(23), 308.
- Servos, P., Osu, R., Santi, A., & Kawato, M. (2002). The neural substrates of biological motion perception: An fMRI study. *Cerebral Cortex*, *12*, 772–782.
- Stein, B. (2002). Cortex governs multisensory integration in the midbrain. *The Neuroscientist*, *306*–314.
- Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Science, USA*, *98*, 11656–11661.
- van der Zwan, R., Badcock, D. R., & Parkin, B. (1999). Global form perception: Interactions between luminance and texture information. *Australian and New Zealand Journal of Ophthalmology*, *27*, 268–270.
- van der Zwan, R., Reid, R., Petreska, B., Billard, A., Blanke, O., & Brooks, A. (in preparation). A new psychophysical technique for establishing observer sensitivity to biological motion displays. *Journal of Vision*.
- Wuerger, S. M., Hofbauer, M., & Meyer, G. F. (2003). The integration of auditory and motion signals at threshold. *Perception and Psychophysics*, *65*(8), 1188–1196.