Selective integration of auditory-visual looming cues by humans

Céline Cappe [a,*], Gregor Thut [b], Vincenzo Romei [b], Micah M. Murray [a,c,**]

[a] The Functional Electrical Neuroimaging Laboratory, Neuropsychology and Neurorehabilitation Service and Radiology Service,
Centre Hospitalier Universitaire Vaudois and University of Lausanne, rue du Bugnon 46, 1011 Lausanne, Switzerland
[b] Centre for Cognitive Neuroimaging, Department of Psychology, University of Glasgow, Glasgow G12 8QB, United Kingdom
[c] Electroencephalography Brain Mapping Core, Center for Biomedical Imaging of Lausanne and Geneva, rue du Bugnon 46, 1011 Lausanne, Switzerland

ARTICLE INFO

ABSTRACT

An object's motion relative to an observer can confer ethologically meaningful information. Approaching or looming stimuli can signal threats/collisions to be avoided or prey to be confronted, whereas receding stimuli can signal successful escape or failed pursuit. Using movement detection and subjective ratings, we investigated the multisensory integration of looming and receding auditory and visual information by humans. While prior research has demonstrated a perceptual bias for unisensory and more recently multisensory looming stimuli, none has investigated whether there is integration of looming signals between modalities. Our findings reveal selective integration of multisensory looming stimuli. Performance was significantly enhanced for looming stimuli over all other multisensory conditions. Contrasts with static multisensory conditions indicate that only multisensory looming stimuli resulted in facilitation beyond that induced by the sheer presence of auditory-visual stimuli. Controlling for variation in physical energy replicated the advantage for multisensory looming stimuli. Finally, only looming stimuli exhibited a negative linear relationship between enhancement indices for detection speed and for subjective ratings. Maximal detection speed was attained when motion perception was already robust under unisensory conditions. The preferential integration of multisensory looming stimuli highlights that complex ethologically salient stimuli likely require synergistic cooperation between existing principles of multisensory integration. A new conceptualization of the neurophysiologic mechanisms mediating real-world multisensory perceptions and action is therefore supported.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

An organism's evolutionary success partially depends on both the ability to reliably detect and discriminate between predators and prey in the environment and also to appropriately respond to them. When encountering an approaching or looming object, one must determine whether to avoid it (a defensive action) or confront it (an aggressive action). Similarly, when encountering a distancing or receding object, one can on the one hand be more assured of one's own safety or can alternatively use this information to determine whether or not pursuit would be worthwhile. In these (and other) ways, simple spatial cues can confer ethologically meaningful information. Given the potentially mortal cost of missing or misinterpreting looming signals, it is unsurprising that ethologists and neuroscientists consider preferential responsiveness to looming signals to be an evolved capacity (Ghazanfar, Neuhoff, & Logothetis, 2002; Graziano & Cooke, 2006; Maier, Chandrasekaran, & Ghazanfar, 2008; Maier, Neuhoff, Logothetis, & Ghazanfar, 2004; Neuhoff, 1998, 2001; Schiff, 1965; Schiff, Caviness, & Gibson, 1962; Seifritz et al., 2002). Moreover, these situations, like many perceptual events, can likely be facilitated by the integration of multisensory cues to enhance perception and render behavior quicker and/or more accurate (Stein & Meredith, 1993; Welch & Warren, 1980).

Multisensory interactions are a fundamental feature of brain organization (Calvert, Spence, & Stein, 2004; Ghazanfar & Schroeder, 2006; Stein & Meredith, 1993; Stein & Stanford, 2008). Studies are increasingly revealing how the brain achieves such multisensory integration. Anatomic evidence now exists for direct projections between unisensory, even primary, cortices (Cappe & Barone, 2005; Falchier, Clavagnier, Barone, & Kennedy, 2002; Rockland & Ojima, 2003). At a functional level, auditory-visual multisensory interactions occur early in time post-stimulus onset and also within areas typically considered unisensory, again including even primary cortices (e.g. Giard & Peronnet, 1999; Martuzzi

et al., 2007; Molholm et al., 2002; Romei, Murray, Merabet, & Thut, 2007). From such findings, new models of brain organization are being developed that incorporate the occurrence of multisensory interactions and integration both at low and high levels of processes and also at early and late time periods following stimulus presentation (Driver & Noesselt, 2008; Ghazanfar & Schroeder, 2006; Stein & Stanford, 2008; Wallace, Ramachandran, & Stein, 2004).

Given this shift in our conceptualization of brain organization, it is increasingly important to understand the functional significance of multisensory interactions as well as the circumstances governing their occurrence. The seminal works of Stein and Meredith (1993) offer several 'rules' of multisensory processing based on receptive field properties of single neurons. More recent data nuance these rules by showing that patterns of interactive effects can be impacted developmentally or through experience (Wallace, Carriere, Perrault, Vaughan, & Stein, 2006; Wallace & Stein, 2007) or even by the spatial heterogeneity within single neurons' receptive fields (Carriere, Royal, & Wallace, 2008). To date, the overwhelming majority of studies have investigated the influences of spatial information on multisensory processing using variation in azimuth or elevation (i.e. 2-dimensional variation in location with respect to the observer). There is comparatively sparse evidence regarding the integration of signals across spatial positions towards versus away from an observer.

Notable exceptions have demonstrated that rhesus monkeys preferentially looked at a looming visual stimulus when presented with a looming, but not receding, sound (Maier et al., 2004). Similarly, 5-month-old infants preferentially looked at matching visual stimuli when presented either with a looming or receding sound (Walker-Andrews & Lennon, 1985). Even though effects were selective for structured sounds instead of noises, the results were only qualitatively suggestive of integrative processes and they did not reveal whether neural response interactions need forcibly be evoked. Likewise, the measurement of looking time cannot differentiate effects occurring at a perceptual level from those driven by biases in attention. Studies of multisensory distance perception by adult humans have predominantly focused on the estimation of time to arrival and remain controversial as whether (and how) auditory and visual distance cues interact and whether or not there is a benefit from multisensory stimulation (Gordon & Rosenblum, 2005; Lewald & Guski, 2004; Sugita & Suzuki, 2003). Moreover, the interpretation of such studies in terms of a neurophysiologic mechanism of either temporal or spatial perception is made complicated by the consistent finding that listeners overestimate the loudness and underestimate the distance of looming sounds (Neuhoff, 1998; Seifritz et al., 2002).

As such, it remains unknown whether multisensory looming/receding signals are integrated to facilitate behavior. Our study addressed this question in humans using a go/no-go motion detection paradigm with unisensory (visual or auditory) and multisensory (simultaneous auditory-visual) stimuli. The perception of visual motion in depth was induced with a central disc that contracted, expanded, or remained constant (i.e. static). The perception of auditory motion in depth was induced with a complex tone that fell or rose in intensity or remained constant (Fig. 1). To ensure that observers used dynamic information in the stimuli, all conditions were initially of the same size/intensity. We assessed multisensory integration of motion perception as measured by reaction times for motion detection (irrespective of its direction or congruence between the senses) and subjective ratings of movement intensity (using a 5-point Likert scale). Performance on multisensory conditions was then compared with that from the constituent unisensory conditions to determine if performance was significantly facilitated to a degree consistent with integrative processes. Finally, the comparison of performance across different multisensory conditions

allowed us to determine whether there is selective facilitation for processing multisensory looming signals by humans.

2. Methods

Sixteen healthy individuals (aged 18–32 years: mean = 25 years; 7 women and 9 men) with normal hearing and normal or corrected-to-normal vision participated. All participants provided written informed consent to the procedures that were approved by the Ethics Committee of the Faculty of Biology and Medicine of the University of Lausanne. The main experiment involved the go/no-go detection of moving versus static stimuli that could be auditory, visual, or multisensory auditory-visual (A, V, and AV, respectively). To induce the perception of movement, visual stimuli changed in size and auditory stimuli changed in volume so as to give the impression of either looming or receding (denoted by L and R, respectively). Static stimuli were of constant size/volume (hereafter denoted by S). We chose these stimulus features as they have been previously shown to be the dominant cue for motion-in-depth. The stimulus conditions are schematized in Fig. 1. Specific multisensory conditions were generated using the full range of combinations of movement type (L, R, and S) and congruence between the senses. For convenience we use shorthand to describe experimental conditions such that, for example, ALVL refers to the multisensory combination of auditory looming and visual looming and ARVL refers to the multisensory combination of auditory receding and visual looming. There were fifteen configurations of stimuli in total (6 unisensory and 9 multisensory). Go trials (i.e. those on which either or both sensory modalities contained moving stimuli) occurred on 80% of the time. Each of the 15 conditions was repeated 252 times across 18 blocks of randomly intermixed trials.

The visual stimulus consisted of a centrally presented disc (either black on a white background or white on a black background, counterbalanced across blocks of trials) that symmetrically expanded (from 7° to 13° diameter with the radius increasing linearly at a constant rate over the 500 ms duration of the stimulus) in the case of looming or contracted (from 7° to 1° diameter) in the case of receding. Auditory stimuli were 1000 Hz tones composed of a triangular waveform and generated with Adobe Audition software (Adobe Systems Inc., San Jose, CA, USA). Auditory stimuli were presented over insert earphones (Etymotic model ER4S). The tones were 500 ms in duration and either rose (looming stimulus) or fell (receding stimulus) 10 dB in intensity approximately linearly over this duration (from 77 dB to 87 dB SPL for the looming sound and from 77 dB to 67 dB SPL for the receding sound). They were sampled at 44.1 kHz, had 10 ms onset and offset ramps (to avoid clicks). Prior research has shown that tonal stimuli produce more reliable perceptions of looming and receding (Neuhoff, 1998) and may also be preferentially involved in multisensory integration (Maier et al., 2004). The particular sounds we used were selected after a pilot study of six participants who used a 5-point Likert scale to rate the strength of movement perceived for each of 12 different pairs of looming and receding sounds that differed in their spectral composition (400 Hz, 1000 Hz), waveform type (triangular, square), and manner of intensity modulation (linear, exponential, and variants of the two). We selected the pair of sounds with the strongest ratings.

The rating experiment involved 14 of the original 16 participants (aged 18–32 years; mean = 25 years; 6 women and 8 men). Their task was to indicate the perceived strength of movement, using a 5-point Likert scale, of each of the 15 conditions from the main experiment. Each condition was presented 48 times and was randomly intermixed within a block of trials.

The follow-up experiment included five naïve participants (aged 26–35 years; mean = 29 years; all women). As in the main experiment, their task was to indicate as fast and as accurately as possible whether or not they perceived movement. There were auditory, visual, and multisensory conditions. Movement types were looming, receding, or static. There were two types of static conditions – low and high – so as to physically match the initial states of the looming and receding conditions, respectively (see Fig. 6A for a schematic including details of stimulus sizes and amplitudes). In contrast to the main experiment, multisensory conditions here always involved the same type of movement (or lack thereof) such that there was never any incongruent multisensory combination. Each of the 12 experimental conditions was repeated 64 times across 4 blocks of randomly intermixed trials.

For all experiments stimulus delivery and response collection were controlled by E-prime software (Psychology Software Tools, Pittsburgh, USA). The stimuli were 500 ms in duration as mentioned above and were presented randomly with an interstimuli interval between stimulus presentations varying from 800 ms to 1400 ms.

3. Results

3.1. Multisensory integration of perceived motion in depth

In a first set of analyses, we evaluated if there was evidence for multisensory integration of looming and receding auditory-visual stimulus pairs and if such was affected by the congruence in the direction of perceived motion between the senses. This was done by testing for a redundant signals effect (RSE) (Giard & Peronnet, 1999; Martuzzi et al., 2007; Miller, 1982; Molholm et al., 2002; Raab,
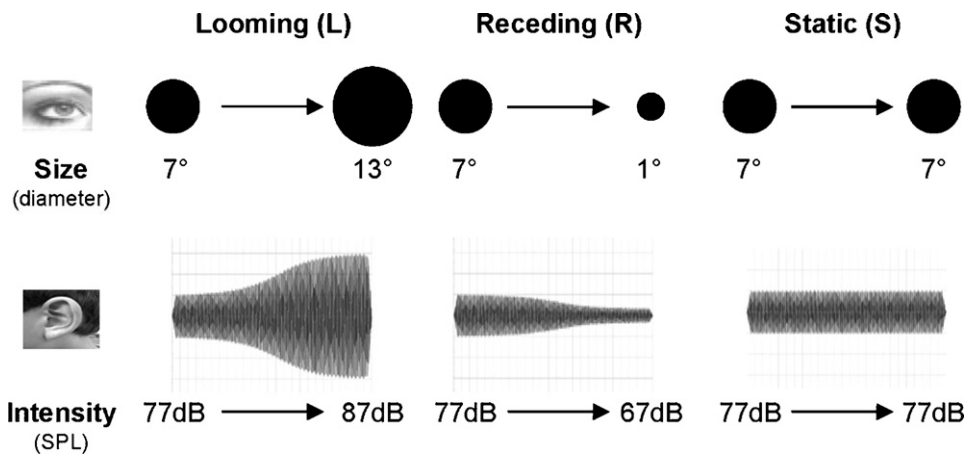
Fig. 1. Experimental paradigm. Participants performed a go/no-go movement detection task with looming, receding, or constant stimuli (500 ms duration) that were either visual, auditory, or multisensory (auditory-visual). All stimuli were initially of the same size/intensity. In the visual modality the perception of movement was induced by changing the size of the centrally displayed disk. In the auditory modality the perception of movement was induced by changing the intensity of the complex tone that was binaurally presented over headphones.

1962; Romei et al., 2007; Schröger & Widmann, 1998) on reaction times (RTs) and movement ratings, which would manifest as performance improvements for multisensory versus either constituent unisensory condition. As participants were able to reliably detect when stimuli changed size/amplitude (i.e. they reliably perceived motion in depth), with a mean percentage of correct responses across conditions of $88 \pm 1.1\%$, we do not discuss accuracy rates further.[1]

RTs and movement ratings for congruent multisensory conditions were compared with those from the constituent unisensory conditions first with a 2×3 repeated measures analysis of variance (ANOVA) using perceived motion direction (looming, receding) and stimulus type (multisensory, auditory, visual) as the within subjects factors and second (in the event of a significant main effect of perceived motion direction) with separate repeated measures ANOVAs with stimulus type. For RTs there were significant main effects of perceived motion direction ($F_{(1,15)} = 47.09$; $p < 0.001$) and stimulus type ($F_{(2,14)} = 144.99$; $p < 0.001$), as well as a significant interaction between these factors ($F_{(2,14)} = 12.62$; $p = 0.001$). Additional follow-up ANOVAs confirmed that RTs were significantly facilitated for multisensory looming ($F_{(2,14)} = 115.51$; $p < 0.001$) as well as for multisensory receding stimuli ($F_{(2,14)} = 152.01$; $p < 0.001$) versus their constituent unisensory stimuli (Fig. 2A, left). This is indicative of a RSE for both looming and receding auditory-visual stimulus pairs when the direction of perceived motion was congruent. Likewise, for movement ratings there were significant main effects of perceived motion direction ($F_{(1,13)} = 31.94$; $p < 0.001$) and stimulus type ($F_{(2,12)} = 24.91$; $p < 0.001$), as well as a significant interaction between these factors ($F_{(2,12)} = 20.48$; $p < 0.001$). Follow-up ANOVAs confirmed that movement was rated significantly higher for multisensory looming ($F_{(2,12)} = 23.04$; $p < 0.001$) as well as for multisensory receding stimuli ($F_{(2,12)} = 37.74$; $p < 0.001$) versus their constituent unisensory stimuli (Fig. 3A, left). As above, this is indicative of a RSE for both looming and receding auditory-visual stimulus pairs when the direction of perceived motion was congruent.

RTs and movement ratings for incongruent multisensory conditions were compared with those from the constituent unisensory conditions first with a 2×3 repeated measures ANOVA, as above. For this analysis, the data were coded in the ANOVA such that the level 'looming' included the auditory looming, visual receding and their multisensory combination (denoted AL, VR, and ALVR, respectively) and the level 'receding' included the auditory receding, visual looming, and their multisensory combination (denoted AR, VL, and ARVL, respectively). For RTs, there were significant main effects of perceived motion direction ($F_{(1,15)} = 34.29$; $p < 0.001$) and stimulus type ($F_{(2,14)} = 132.61$; $p < 0.001$), as well as a signif-
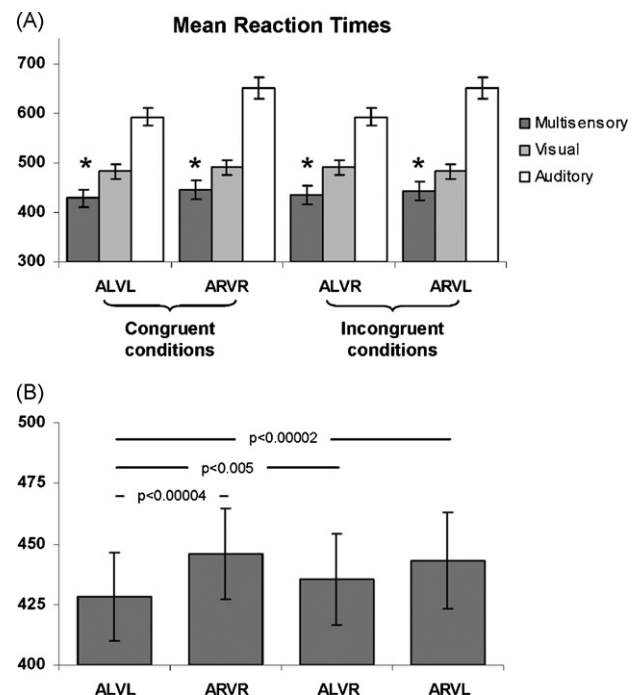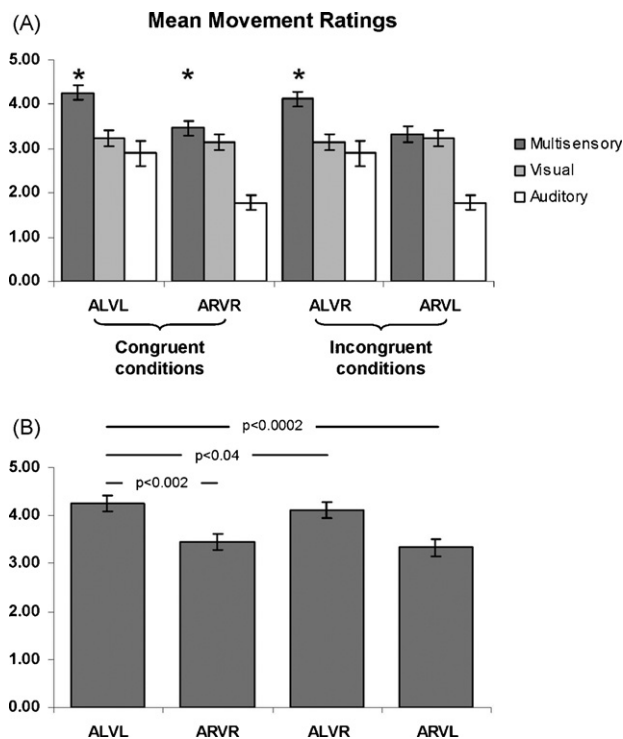

Fig. 2. Multisensory facilitation of reaction times. Group-averaged ($N = 16$; S.E.M. indicated) reaction times (RTs) are plotted for each experimental condition. (A) In all multisensory conditions reaction times were significantly faster than in either of the constituent unisensory conditions (asterisks). This was the case both when the movement direction was congruent as well as when it was incongruent between the senses. (B) Direct comparison of RTs to multisensory conditions revealed that performance with multisensory looming stimuli (ALVL) was selectively facilitated beyond that for other multisensory conditions (p-values indicated).

---

[1] We also performed a series of paired t-tests to assess whether accuracy in movement detection differed between looming and receding conditions. In no case was there a significant difference in accuracy rates between looming and receding conditions, except in the case of unisensory auditory stimulation where performance was more accurate with looming stimuli ($p < 0.000002$). Despite this difference, we would hasten to note that such cannot account for the differences in RTs observed across the different multisensory conditions in this study, which did not significantly differ in their accuracy rates of perceived motion detection.
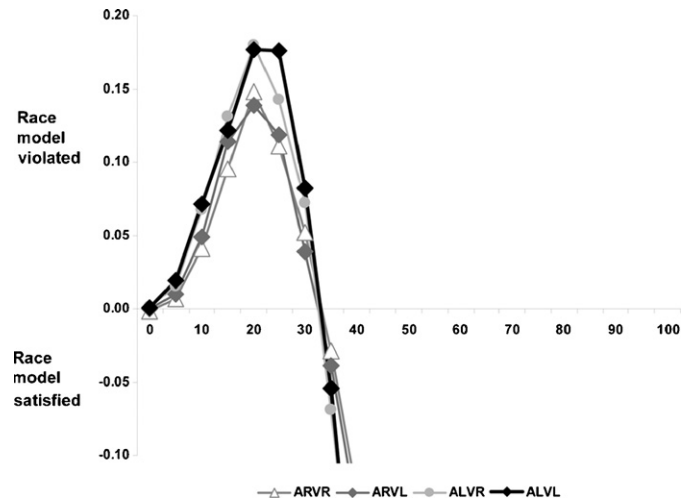
icant interaction between these factors ($F_{(2,14)} = 49.35$; $p < 0.001$). Follow-up ANOVAs confirmed that RTs were significantly facilitated for multisensory conditions (vs. their constituent unisensory conditions) even when the direction of perceived motion was incongruent between the senses (ALVR: $F_{(2,14)} = 97.79$; $p < 0.001$ and ARVL: $F_{(2,14)} = 123.54$; $p < 0.001$; Fig. 2A, right). Likewise, for movement ratings there were significant main effects of perceived motion direction ($F_{(1,13)} = 35.18$; $p < 0.001$) and stimulus type ($F_{(2,12)} = 24.60$; $p < 0.001$), as well as a significant interaction between these factors ($F_{(2,12)} = 15.43$; $p < 0.001$). Follow-up ANOVAs confirmed that movement ratings were significantly facilitated for the ALVR multisensory condition (vs. its constituent unisensory conditions; $F_{(2,12)} = 23.34$; $p < 0.001$) even though the direction of perceived motion was incongruent between the senses. By contrast, although there was a significant main effect of stimulus condition for the ALVR multisensory pair ($F_{(2,12)} = 37.59$; $p < 0.001$), ratings did not reliably differ between the multisensory and constituent visual condition (Fig. 3A, right).

We would already mention here that the observation of an RSE for multisensory conditions wherein the perceived direction of motion was incongruent is not altogether surprising. Prior studies examining the impact of spatial position in azimuth of auditory-visual stimuli likewise found RT facilitation even when the auditory and visual stimuli were presented to different hemifields (Gondan, Niederhaus, Röder, & Rösler, 2005; Teder-Sälejärvi, Di Russo, McDonald, & Hillyard, 2005). We return to this point below when contrasting multisensory dynamic stimuli with multisensory conditions wherein either the visual or auditory component remained static.

Finally, we also assessed whether the observed RSEs with RTs could be fully explained by probability summation or instead are

Fig. 4. Test of integrative interactions. Integrative effects revealed by Miller's race model inequality are plotted as a function of the percentile of the RT distribution for each condition. Positive values indicate violation of the model and negative values its satisfaction. Note that some instantiations of this model assume statistical independence between the senses and therefore also subtract the joint probability (i.e. the product) from the constituent unisensory conditions (c.f. Colonius and Diederich, 2006 for discussion). No such assumption was applied here and if anything would have constituted a more conservative test.

consistent with integrative processes. To do this, we applied Miller's race model inequality (Miller, 1982), which compares the cumulative probability distribution from the multisensory condition (actual distribution) with the sum of those from the constituent unisensory conditions (modeled distribution). Results from this analysis are depicted in Fig. 4 and show that all conditions exhibited facilitation in excess of probability summation over the fastest 25% of the RT distribution (i.e. significant differences between the actual and modeled probability distributions using a 2-tailed t-test for each 5%-wide bin). These results provide an indication that integrative processes contribute to the RSEs.

3.2. Selective facilitation of multisensory looming

Given the above evidence for multisensory integration of perceived motion in depth, a third set of analyses was conducted to determine if the mean RTs were faster and movement ratings were higher for congruent looming AV stimulus pairs relative to other multisensory movement conditions. For this ANOVA, the within-participant factor was multisensory condition (ALVL, ARVR, ALVR, and ARVL). For RTs, there was a significant main effect of condition ($F_{(3,13)} = 20.874$; $p < 0.001$) that was due to faster responses to looming AV stimulus pairs relative to all other multisensory conditions (Fig. 2B). This was the case for 12 of the 16 participants. Likewise, for movement ratings there was a significant main effect of condition ($F_{(3,11)} = 13.597$; $p < 0.001$) that was due to higher ratings for looming AV stimulus pairs relative to all other multisensory conditions (Fig. 3B). This was the case for 9 of the 14 participants. These results indicate that there is selective facilitation of multisensory looming both in terms of the speed with which movement is detected and in terms of the perceived strength of movement. That is, the specific combination of looming stimuli in both modalities led to a further enhancement of behavior beyond that observed for other multisensory movement combinations. Likewise, there was no evidence to indicate that multisensory receding stimuli were reliably slower than the 'incongruent' multisensory conditions, which would be expected had the physical properties of the stimuli been directly driving behavioral responses. Consequently, it is difficult to readily explain the results as following from either a straightforward

Fig. 3. Multisensory facilitation of subjective ratings. Group-averaged ($N = 14$; S.E.M. indicated) subjective movement ratings based on a 5-point Likert scale are plotted for each experimental condition. (A) As was the case for reaction times (Fig. 2), movement ratings were significantly higher for multisensory than for either of the constituent unisensory conditions, with the exception of the ARVL condition (asterisks). (B) Direct comparison of movement ratings revealed that subjective ratings for multisensory looming stimuli (ALVL) was selectively greater than that for other multisensory conditions.

difference in the physical attributes of the unisensory stimuli or a general effect of attention/arousal to a particular sensory modality of perceived motion direction.

Further evidence favoring selective enhancement of multisensory looming stimuli is provided by comparing performance on the multisensory conditions with the conditions wherein one sensory modality provided movement information while the other was static (denoted by S). RTs and movement ratings were submitted to separate 2×3 repeated measures analysis of variance (ANOVA) using perceived motion direction (looming, receding) and modality of movement (multisensory, auditory, visual) as the within subjects factors (Fig. 5). For RTs, there was a significant main effect of modality of movement ($F_{(2,14)} = 31.18$; $p < 0.0001$) as well as an interaction between perceived motion direction and modality of movement ($F_{(2,14)} = 11.92$; $p < 0.001$). Subsequent ANOVAs were conducted for looming and receding stimuli, separately, and in both cases revealed a significant main effect of modality of movement ($F_{(2,14)} = 36.65$; $p < 0.0001$ and $F_{(2,14)} = 28.57$; $p < 0.0001$, respectively). In the case of looming stimuli this was because RTs to the ALVL condition were reliably faster than to both the ALVS and ASVL conditions ($p < 0.0002$ for all pair-wise contrasts). Note that all 16 participants had equal ($N = 2$) or faster ($N = 14$) RTs for the ALVL than either ALVS or ASVL condition. By contrast, in the case of receding stimuli this was because RTs to the ARVS condition were significantly slower than to the ARVR and ASVR conditions; the latter of which did not reliably differ ($p > 0.24$). For movement ratings, there were significant main effects of direction of perceived movement ($F_{(1,13)} = 24.42$; $p < 0.0001$) and of modality of movement ($F_{(2,12)} = 32.72$; $p < 0.0001$) as well as an interaction between perceived motion direction and modality of movement ($F_{(2,14)} = 16.83$; $p < 0.0001$). Subsequent ANOVAs were conducted for looming and receding stimuli, separately, and in both cases revealed a significant main effect of modality of movement ($F_{(2,12)} = 25.91$; $p < 0.0001$ and $F_{(2,12)} = 84.47$; $p < 0.0001$, respectively). In the case

of looming stimuli this was because ratings to the ALVL condition were reliably higher than to both the ALVS and ASVL conditions ($p < 0.0002$ for all pair-wise contrasts). Note that 13 of the 14 participants had higher ratings for the ALVL than either ALVS or ASVL condition. By contrast, in the case of receding stimuli this was because ratings to the ARVS condition were significantly lower than to the ARVR and ASVR conditions; the latter of which did not reliably differ ($p > 0.77$). More generally, this pattern of results would suggest that in the case of multisensory receding stimuli (ARVR), the facilitation of RTs and ratings is not reliably different from that occurring when a static auditory stimulus is simultaneously presented with a receding visual stimulus. In other words, performance benefits for the ARVR condition are not particular to there being multisensory movement. This is consistent with the abovementioned findings of Teder-Sälejärvi et al. (2005) and Gondan et al. (2005) who obtained RSEs for spatially incongruent (in terms of azimuthal position) auditory-visual stimuli. By contrast, we would emphasize that in the case of multisensory looming stimuli (ALVL), performance is enhanced by there being multisensory movement beyond that due to there simply being multisensory stimulation.

To more directly address whether the above advantage for multisensory looming stimuli followed from physical differences across conditions, we performed a follow-up experiment with five naïve participants who performed the same go/no-go motion detection task. Here, the looming and receding stimuli were physically identical except in their dynamics (i.e. looming was the reverse in time of receding; Fig. 6). As above, there were auditory, visual, and multisensory AV conditions as well as catch trials when the stimuli remained constant in their size/amplitude (see Section 2). RTs were submitted to a 2 perceived motion direction (looming vs. receding) $\times$ 3 stimulus type (A, V, AV) ANOVA. We obtained significant main effects of perceived motion direction ($F_{(1,4)} = 14.39$; $p = 0.019$) and stimulus type ($F_{(2,3)} = 32.22$; $p = 0.009$), as well as a significant interaction between these factors ($F_{(2,3)} = 165.26$; $p < 0.001$). Follow-up contrasts (ANOVAs) showed an RSE for both looming ($p < 0.01$) and receding stimuli ($p < 0.05$), and additional t-tests showed that this was because the AV condition was faster than either A or V condition (all p-values less than 0.025). Most germane is that multisensory looming stimuli resulted in faster RTs than multisensory receding stimuli (507 ms vs. 523 ms; $t_{(4)} = 2.83$; $p < 0.05$). This was the case for each of the 5 participants. These results rule out an account of the selective facilitation of looming signals based on differences in physical energy, because the stimuli were identical except in their dynamics. Moreover, the initial state for the receding conditions had more physical energy than in the looming conditions, which if anything would bias effects in favor of receding being faster than looming.

4. Discussion

Our demonstration of selective integration of multisensory looming signals that affects both reaction times and subjective experience has direct implications for how longstanding principles of multisensory integration, established through parametric variation of position, timing and effectiveness, are to be considered alongside ethologically salient stimuli such as looming signals. The 'spatial rule' (Stein & Meredith, 1993) is based on relative superposition of a neuron's excitatory zones, such that facilitative integration is limited to situations where stimuli of both (all) sensory modalities fall within these zones. This principle has been extensively tested, and to the best of our knowledge no evidence contradicting this rule exists (though data that nuance the varieties of interactions one can record are increasingly numerous (Allman, Keniston, & Meredith, 2008; Wallace et al., 2006)). We therefore turn to
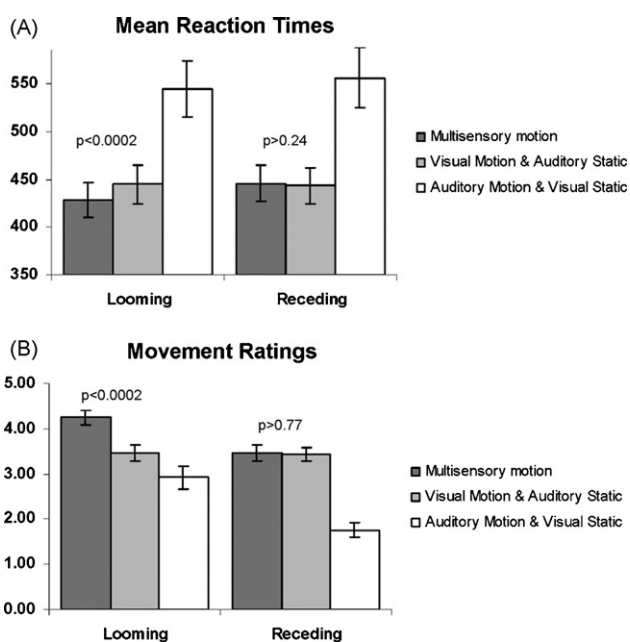
Fig. 5. Selective facilitation of multisensory looming. (A) Group-averaged ($N = 16$; S.E.M. indicated) reaction times (RTs) are shown as a function of experimental condition and movement direction. (B) Group-averaged ($N = 14$; S.E.M. indicated) movement ratings are shown as in (A). Multisensory motion that is looming significantly facilitated RTs and ratings over other multisensory conditions with movement in only one sensory modality. This was not the case for multisensory motion that was receding. p-Values reflect the result of the contrast (t-test) between multisensory motion and visual motion.
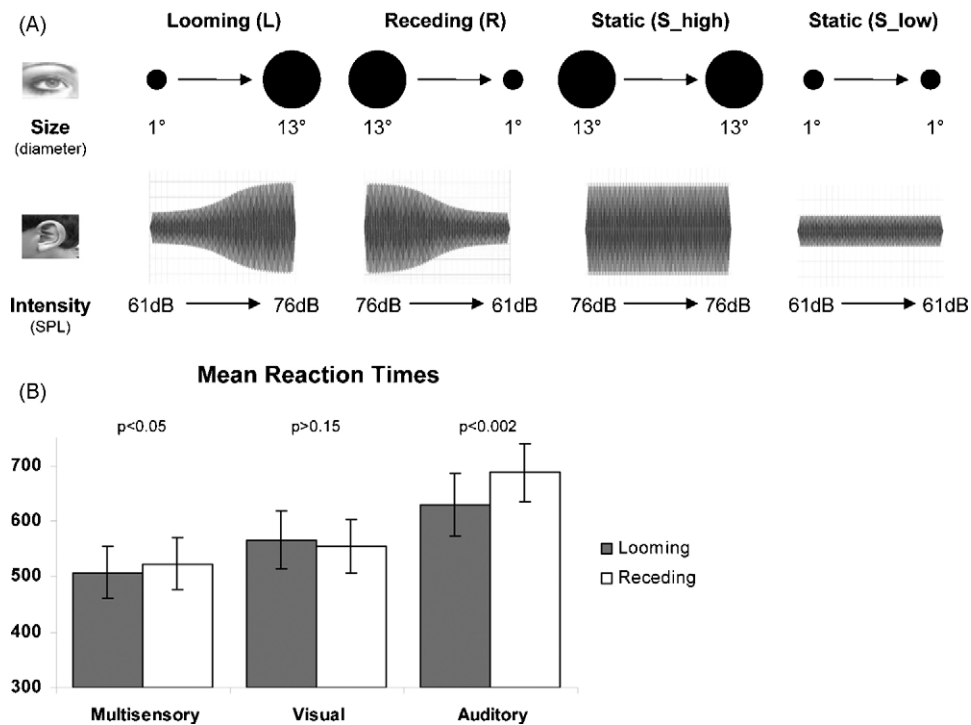
Fig. 6. Schematic and results of the follow-up experiment. (A) The follow-up experiment followed a highly similar paradigm to that of the main experiment, except that looming and receding conditions were physically identical and inversed in their dynamics. (B) Mean RTs (S.E.M. indicated) are plotted as a function of stimulus condition and movement direction. Both looming and receding stimuli resulted in an RSE. Most importantly, RTs to multisensory looming were faster than those to multisensory receding (*p*-values indicated), replicating our findings in the main experiment and excluding a role for differences in total physical energy or initial size/volume.

the issue of incorporating the present results, which are based on dynamic variation of stimulus features, alongside this principle as well as that concerning inverse effectiveness, wherein maximal enhancement would likely be induced by the most minimally effective stimuli (i.e. those producing the weakest neural responses). By controlling for total variation in size/amplitude and by contrasting different multisensory conditions we can rule out a major role of physical features in directly driving the present effects. As such, it is unlikely that the selective facilitation for looming signals constitutes a simple instantiation of the principle of inverse effectiveness. In terms of spatial representations, the stimuli used here were always presented from the same location in azimuth (i.e. directly in front of the observer), and our contrasts controlled for total variation in the visual size of the stimuli. Thus, our results once again cannot be readily explained by a simple instantiation of the spatial rule—at least when considered only in terms of signaling azimuthal positions. Given such and given the overwhelming support for the abovementioned rules, our data would thus suggest there to be an accounting for the dynamic features of stimuli (here, in terms of their perceived spatial position in depth) in the neurophysiologic mechanisms of multisensory interactions and integration. That is, our results are not readily described by any one of the abovementioned rules in isolation, but rather likely reflect a combination of their instantiation (Carriere et al., 2008) that in turn engenders sensitivity to ethologically salient stimuli such as when signaled by looming cues. Still, we would be remiss not to highlight that the abovementioned principles have been established based on myriad single-unit recordings in animal models, whereas our conjectures here are based on psychophysical findings in humans. Comparable studies providing single-unit response profiles to the kinds of stimuli presented here would be critical for a fuller understanding of the applicability of the spatial rule and inverse effectiveness principle to interactions between approaching/looming multisensory stimuli. More generally, our results call for targeted neurophysiologic

research in this domain and for the likely expansion of the existing rules governing multisensory processing to include synergistic, second-order effects.

4.1. Correlation between RT and rating enhancement indices

In a final, exploratory analysis, we determined the percentage of performance enhancement for both RT and movement ratings, using an index of multisensory enhancement typical of microelectrode studies in animals (Stein & Meredith, 1993) and of some functional imaging studies in humans (Beauchamp, 2005). This index equals the ratio of the difference between the multisensory and best constituent unisensory condition relative to the best unisensory condition. We then related these indices to determine whether there was a systematic relationship between the facilitation in the speed of movement detection and the enhancement in the perceived strength of movement following multisensory stimulation (Fig. 7). It is important to recall that RT and movement ratings were provided by participants during separate experimental sessions, thereby excluding the possibility that correlations stem from systematic strategic differences or cognitive biases across participants. In the case of multisensory looming stimuli, there was a significant negative correlation between these enhancement indices ($r_{(12)} = -0.590$; $p = 0.026$). No other condition exhibited a reliable positive or negative correlation (all *p*-values >0.20). This negative correlation is suggestive of a tradeoff between multisensory effects on the strength movement perception and the speed of detection of movement. That is, an observer who receives a large multisensory benefit on RTs exhibits a small benefit on perceived movement strength and conversely an observer who receives a large benefit on movement strength exhibits a small RT facilitation. Taking this a step further, one can speculate that an observer need already have a strong perception of movement under unisensory conditions (and therefore be less prone to a gain in such following
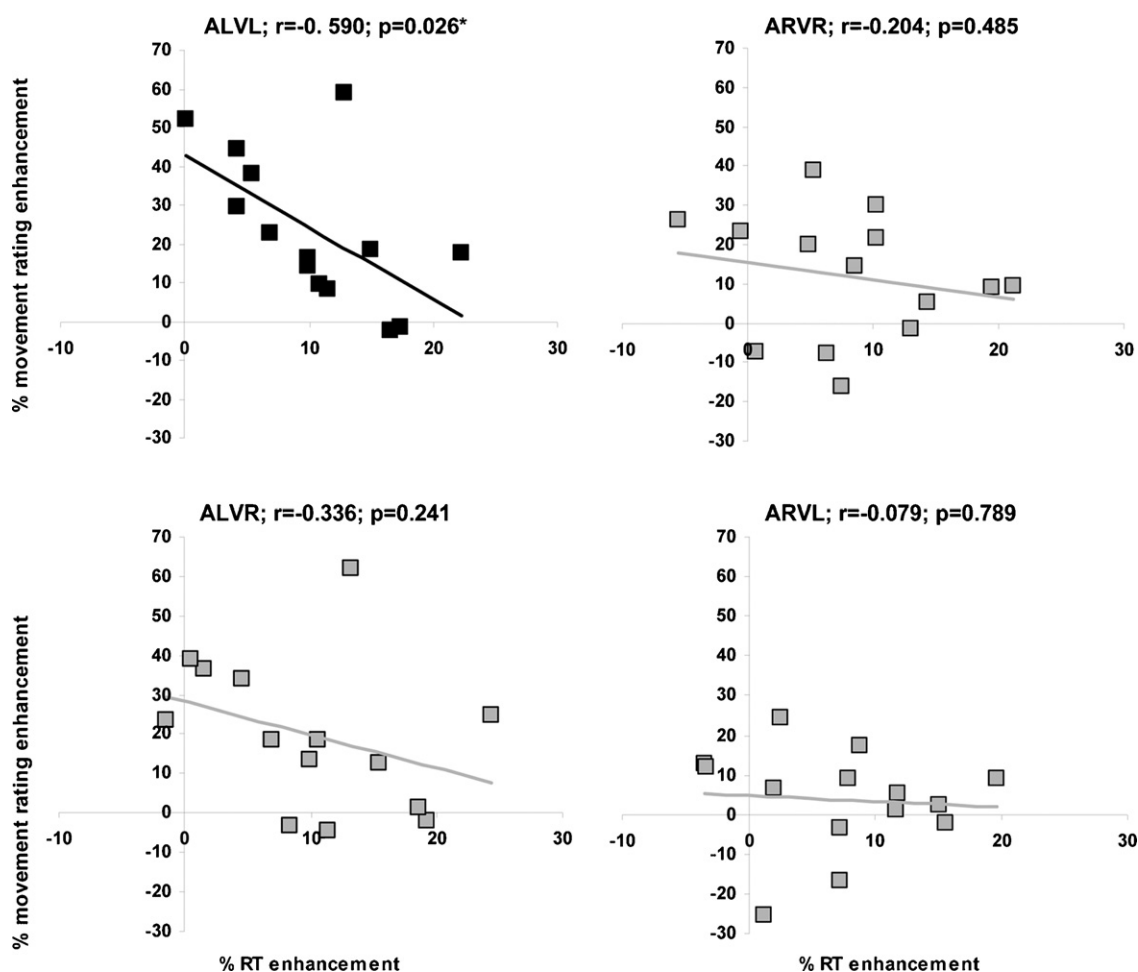
Fig. 7. Relation between multisensory enhancement indices. These scatter plots relate the percentage of RT enhancement to the percentage of movement rating enhancement (*x*-axis and *y*-axis, respectively) for each of the multisensory conditions. Percentage of enhancement was calculated as the difference between the multisensory and best unisensory condition divided by the best unisensory condition for each participant. Only in the case of multisensory looming (ALVL) was there a reliable linear correlation.

multisensory stimulation) in order to express the maximal multisensory facilitation of RTs. As this was an exploratory analysis, we are reticent to draw overly strong conclusions. Plus, additional studies that entail detection and ratings on each trial (as opposed to the separate sessions used in this study) will be required to determine if this observation persists at the level of single participants.

In summary, we provide the first evidence that multisensory auditory-visual looming signals are selectively integrated by humans to facilitate behavior and enhance the impression of movement. Prior studies (Gordon & Rosenblum, 2005; Maier et al., 2004; Walker-Andrews & Lennon, 1985) have generally emphasized either the attentional or perceptual benefits of multisensory looming stimuli without demonstrating there to be integration. While there is ample evidence highlighting the role of unisensory looming stimuli in eliciting avoidance or defensive responses consistent with the interpretation of looming as a signal of imminent danger (Ghazanfar et al., 2002; Graziano & Cooke, 2006; Maier et al., 2004, 2008; Neuhoff, 1998, 2001; Schiff, 1965; Schiff et al., 1962; Seifritz et al., 2002), more recent data shows that such might be supported by enhanced neural responsiveness within a distributed network of low-level cortices (e.g. non-primary auditory cortex (Maier & Ghazanfar, 2007; Seifritz et al., 2002)), limbic structures (e.g. the amygdala (Bach et al., 2008)), as well as higher-level association areas, including the superior temporal sulcus, intraparietal cortex, and premotor cortex (Bach et al., 2008; Maier & Ghazanfar, 2007; Mathiak et al., 2003; Seifritz et al., 2002).

Whether, when, and how such structures are involved in the multisensory integration of looming stimuli is the topic of our ongoing research, with very recent findings in monkeys suggesting there to be a role of coherent gamma band activity (Maier et al., 2008). Of parallel importance will be to use looming/receding stimuli as a means of tracking the developmental time course of multisensory integration (Neil, Chee-Ruiter, Scheier, Lewkowicz, & Shimojo, 2006) and its potential disruption in clinical populations exhibiting deficits in communication and social interactions (Iarocci & McDonald, 2006; Kern et al., 2007). However, the present results do already expand how the well-known 'rules' governing multisensory integration are to be conceived when complex stimuli are encountered that vary in their location, dynamics, and effectiveness. For one, our results highlight the importance of incorporating the encoding of (dynamic) distance information, thereby extending the spatial rule from azimuth to depth. More generally, our results, like other recent findings (Carriere et al., 2008) indicate that second-order and synergistic descriptions of the operation of these rules need be established that account for both behavioral and neurophysiologic responses to complex ethological stimuli.

Acknowledgement

References

Allman, B. L., Keniston, L. P., & Meredith, M. A. (2008). Subthreshold auditory inputs to extrastriate visual neurons are responsive to parametric changes in stimulus quality: Sensory-specific versus non-specific coding. *Brain Research*, doi:10.1016/j.brainres.2008.03.086. Published online 11 April

Bach, D. R., Schächinger, H., Neuhoff, J. G., Esposito, F., Di Salle, F., Lehmann, C., et al. (2008). Rising intensity: an intrinsic warning cue activating the amygdala. *Cerebral Cortex*, *18*, 145–150.

Beauchamp, M. S. (2005). Statistical criteria in fMRI studies of multisensory integration. *Neuroinformatics*, *3*, 93–113.

Calvert, G., Spence, C., & Stein, B. E. (Eds.). (2004). *The handbook of multisensory processes*. Cambridge, MA: MIT Press.

Cappe, C., & Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience*, *22*, 2886–2902.

Carriere, B. N., Royal, D. W., & Wallace, M. T. (2008). Spatial heterogeneity of cortical receptive fields and its impact on multisensory interactions. *Journal of Neurophysiology*, *99*, 2357–2368.

Colonius, H., & Diederich, A. (2006). The race model inequality: interpreting a geometric measure of the amount of violation. *Psychological Review*, *113*(1), 148–154.

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron*, *57*, 11–23.

Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience*, *22*, 5749–5759.

Ghazanfar, A. A., Neuhoff, J. G., & Logothetis, N. K. (2002). Auditory looming perception in rhesus monkeys. *Proceedings of the National Academy of Sciences USA*, *99*, 15755–15757.

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*, 278–285.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, *11*, 473–490.

Gondan, M., Niederhaus, B., Röder, B., & Rösler, F. (2005). Multisensory processing in the redundant target effect: A behavioral and ERP study. *Perception & Psychophysics*, *67*, 713–726.

Gordon, M. S., & Rosenblum, L. D. (2005). Effects of intrastimulus modality change on audiovisual time-to-arrival judgments. *Perception & Psychophysics*, *67*, 580–594.

Graziano, M. S., & Cooke, D. F. (2006). Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia*, *44*, 845–859.

Iarocci, G., & McDonald, J. (2006). Sensory integration and the perceptual experience of persons with autism. *Journal of Autism and Developmental Disorders*, *36*, 77–90.

Kern, J. K., Trivedi, M. H., Grannemann, B. D., Garver, C. R., Johnson, D. G., Andrews, A. A., et al. (2007). Sensory correlations in autism. *Autism*, *11*, 123–134.

Lewald, J., & Guski, R. (2004). Auditory-visual temporal integration as a function of distance: no compensation for sound-transmission time in human perception. *Neuroscience Letters*, *357*, 119–122.

Maier, J. X., Chandrasekaran, C., & Ghazanfar, A. A. (2008). Integration of bimodal looming signals through neuronal coherence in the temporal lobe. *Current Biology*, *18*, 963–968.

Maier, J. X., & Ghazanfar, A. A. (2007). Looming biases in monkey auditory cortex. *Journal of Neuroscience*, *27*, 4093–4100.

Maier, J. X., Neuhoff, J. G., Logothetis, N. K., & Ghazanfar, A. A. (2004). Multisensory integration of looming signals by rhesus monkeys. *Neuron*, *43*, 177–181.

Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J. P., Maeder, P. P., Clarke, S., et al. (2007). Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cerebral Cortex*, *17*, 1672–1679.

Mathiak, K., Hertrich, I., Kincses, W. E., Riecker, A., Lutzenberger, W., & Ackermann, H. (2003). The right supratemporal plane hears the distance of objects: neuromagnetic correlates of virtual reality. *Neuroreport*, *14*, 307–311.

Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247–279.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Research Cognitive Brain Research*, *14*, 115–128.

Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science*, *9*, 454–464.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*, 123–124.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology*, *13*, 87–110.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*, 574–590.

Rockland, K. S., & Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology*, *50*, 19–26.

Romei, V., Murray, M. M., Merabet, L. B., & Thut, G. (2007). Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: implications for multisensory interactions. *Journal of Neuroscience*, *27*, 11465–11472.

Schiff, W. (1965). Perception of impending collision: A study of visually directed avoidant behavior. *Psychological Monographs: General and Applied*, *79*, 1–26.

Schiff, W., Caviness, J. A., & Gibson, J. J. (1962). Persistent fear responses in rhesus monkeys to the optical stimulus of "looming". *Science*, *136*, 982–983.

Schröger, E., & Widmann, A. (1998). Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology*, *35*, 755–759.

Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schachinger, H., et al. (2002). Neural processing of auditory looming in the human brain. *Current Biology*, *12*, 2147–2151.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses.* Cambridge, MA: MIT Press.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*, 255–266.

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*, *421*, 911.

Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., & Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *Journal of Cognitive Neuroscience*, *17*, 1396–1409.

Walker-Andrews, A. S., & Lennon, E. M. (1985). Auditory-visual perception of changing distance by human infants. *Child Development*, *56*, 544–548.

Wallace, M. T., Carriere, B. N., Perrault, T. J., Jr., Vaughan, J. W., & Stein, B. E. (2006). The development of cortical multisensory integration. *Journal of Neuroscience*, *26*, 11844–11849.

Wallace, M. T., Ramachandran, R., & Stein, B. E. (2004). A revised view of sensory cortical parcellation. *Proceedings of the National Academy of Sciences USA*, *101*, 2167–2172.

Wallace, M. T., & Stein, B. E. (2007). Early experience determines how the senses will interact. *Journal of Neurophysiology*, *97*, 921–926.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*, 638–667.