

COGNITIVE NEUROSCIENCE

Dissociation of perception and action in audiovisual multisensory integration

Lynnette M. Leone and Mark E. McCourt

Center for Visual and Cognitive Neuroscience, Department of Psychology, North Dakota State University, Fargo, ND 58108, USA

Keywords: audiovisual, multisensory integration, race model, reaction time, simultaneity judgment, temporal order judgment

Edited by Sophie Molholm

Received 15 April 2015, revised 16 September 2015, accepted 23 September 2015

Abstract

The 'temporal rule' of multisensory integration (MI) proposes that unisensory stimuli, and the neuronal responses they evoke, must fall within a window of integration. Ecological validity demands that MI should occur only for physically simultaneous events (which may give rise to non-simultaneous neural activations), and spurious neural response simultaneities unrelated to environmental multisensory occurrences must somehow be rejected. Two experiments investigated the requirements of simultaneity for facilitative MI. Experiment 1 employed a reaction time (RT)/race model paradigm to measure audiovisual (AV) MI as a function of AV stimulus-onset asynchrony (SOA) under fully dark adapted conditions for visual stimuli that were either rod- or cone-isolating. Auditory stimulus intensity was constant. Despite a 155-ms delay in mean RT to the scotopic vs. photopic stimulus, facilitative AV MI in both conditions occurred exclusively at an AV SOA of 0 ms. Thus, facilitative MI demands both physical and physiological simultaneity. Experiment 2 investigated the accuracy of simultaneity and temporal order judgements under the same stimulus conditions. Judgements of AV stimulus simultaneity or temporal order were significantly influenced by stimulus intensity, indicating different simultaneity requirements for these tasks. The possibility was considered that there are mechanisms by which the nervous system may take account of variations in response latency arising from changes in stimulus intensity in order to selectively integrate only those physiological simultaneities that arise from physical simultaneities. It was proposed that separate subsystems for AV MI exist that pertain to action and perception.

Introduction

One measure of facilitative multisensory integration (MI) is decreased reaction time (RT) to the presentation of multisensory stimuli vs. their unisensory counterparts (Leone & McCourt, 2013). The integration of unisensory signals depends on their relative timing (Hershenson, 1962; Raab, 1962; Miller, 1986; Diederich & Colonius, 2004) such that afferent signals arising from the stimulation of separate sensory modalities must converge at a neural locus at nearly the same time in order to be integrated (Stein & Meredith, 1993). However, systematic variations in the latency of stimulus-evoked neural activity can occur for a variety of reasons, such as differences in receptor transduction time (due either to intrinsic factors or to differential stimulus intensity), and/or to differences in neuronal transmission time due to unequal axonal length, conduction velocity and/or number of relay synapses (Piéron, 1952; Roufs, 1963; Celesia & Puletti, 1971; Prestrude, 1971; Jeffreys & Axford, 1972; Barlow *et al.*, 1978; Clark & Hillyard, 1996; Maunsell *et al.*, 1999; Foxe & Simpson, 2002; Foxe *et al.*, 2008).

Several reports have suggested that the simultaneous arrival of afferent signals at some neural locus is critical for facilitative MI because the greatest decrease in RT occurred when the unisensory stimulus-onset asynchrony (SOA) corresponded to the difference in RT to the unisensory stimuli (Hershenson, 1962; Raab, 1962; Miller, 1986; Diederich & Colonius, 2004). In neurons of the superior colliculi of domestic cats, Stein & Meredith (1993) reported that audiovisual (AV) MI was maximal when the post-stimulus time histograms (PSTHs) of responses to the unisensory auditory (A) and visual (V) stimuli overlapped. However, the overlap in the PSTHs did not necessarily correspond to simultaneous stimulus presentation. Thus, according to this logic, facilitative MI should require unisensory SOAs, which roughly correspond to the difference in simple RT to the two unisensory stimuli, where this difference is presumed to reflect differences in afferent neural processing time.

Facilitative MI can be indexed by an increase in the speed of response to multisensory vs. unisensory stimuli, and the decreased RT to multisensory stimuli is the 'redundant signals effect', or RSE (Miller, 1982; see Appendix B for a detailed explanation of Miller's inequality).

Miller (1982) contrasted two different models to explain the RSE. The first is called the 'race model'. It assumes separate activation and independent processing of each unisensory (redundant) signal.

Correspondence: Dr M.E. McCourt, as above.
E-mail: mark.mccourt@ndsu.edu

Thus, on each trial, the unisensory signal arriving soonest at some central site ‘wins’ the race and initiates the response. Redundancy gains according to the race model accrue from statistical probability summation, and do not require (but do not rule out) an interaction between the unisensory signals. The second model assumes an interaction between the unisensory channels, which begins at stimulus onset and continues until a response criterion is met. This can occur sooner than the same criterion can be reached via separate activation. Miller (1982) developed an inequality to describe the race model, which identifies an upper limit for the probability of obtaining RTs to redundant stimuli (see Appendix A). Probability summation cannot explain violations of Miller’s inequality, and such violations are interpreted to imply neural coactivation.

The present study

While one potential cause for physiological simultaneity is true physical stimulus co-occurrence, there are adventitious causes as well. Most environmental stimuli relevant to survival arise from within praxis space (Previc, 1998) where, for example, A and V unisensory stimuli will reach their respective sensory receptors nearly simultaneously despite the large difference in speed of propagation. It was suggested, therefore, that preferentially tuning the MI apparatus to physically simultaneous multisensory occurrences will produce an ecological advantage with regard to survival.

In a previous study, A and V unisensory stimulus intensity as well as SOA_{AV} were manipulated in order to measure the effect of varying physiological simultaneity on facilitative MI (Leone & McCourt, 2013). Using violations of the race model (Miller’s inequality) as an indicator of facilitative MI, it was found that despite significant differences in mean RT to the unisensory stimuli that resulted from variations in stimulus intensity, in all cases facilitative MI occurred exclusively over a very narrow range of SOA_{AV} , with the largest effect occurring at physical simultaneity (0 ms). Those results were interpreted to support the idea that both physiological and physical simultaneity were required for optimal facilitative MI. The present experiments examine the question of whether these stimulus simultaneity requirements would also be necessary for optimal multisensory facilitation in the case of more extreme differences in physiological response latencies of rod vs. cone photoreceptors (see Appendix B for a detailed description of these differences).

Rod- or cone-isolating visual stimuli were paired with an auditory stimulus at a variety of AV SOAs in order to determine the temporal offset giving rise to optimal facilitative MI with respect to RT (Experiment 1), and also with respect to perceptual judgements of temporal order and simultaneity (Experiment 2).

Experiment 1: RT measures

Experiment 1 used a RT/race model paradigm where RSE was measured under fully dark adapted conditions for A, V and AV stimuli, where the V stimulus wavelengths were selected to stimulate either rod or cone photoreceptors. It was sought to ascertain whether the very different temporal response properties of the rod and cone systems would cause facilitative MI to occur at different values of SOA_{AV} .

Condition 1: RT and photopic visual stimuli

The A stimulus and a photopic (cone-isolating) V stimulus were paired across a range of SOA_{AV} in fully dark adapted participants where mean RT to the unisensory A stimulus was 26 ms faster than to unisensory V stimulus.

Method

Participants. Participants ($N = 5$, 3 female; 22–42 years old) possessed normal or corrected-to-normal spatial vision, normal colour vision, and normal hearing. The experiment was undertaken with the understanding and written consent of all participants, subsequent to review and approval by the North Dakota State University Institutional Review Board, consistent with the World Medical Association Declaration of Helsinki.

Stimuli and apparatus. Figure 1 illustrates the stimulus arrangement used for both Condition 1 (photopic visual stimulus; a) and Condition 2 (scotopic visual stimulus; b). In both conditions the fixation light (F) was a steadily illuminated 630-nm LED (0.25° diameter) positioned 6° to the left of the test stimulus (V). The test stimuli, the photopic and scotopic visual stimulus conditions, were 630-nm or 525-nm LEDs (0.25° diameter). Test stimuli were presented on each trial at a fixed suprathreshold intensity for 100 ms. Viewing distance was 114 cm. The A stimulus was a 1000-Hz pure tone, 100 ms duration, 33 dB_A SPL as measured at the ears, delivered via a speaker approximately co-localized with the V stimulus.

Procedure. Participants adapted in complete darkness for 40 min, after which the experiment began. A single-interval yes–no signal detection paradigm was employed, in which participants were instructed to continuously gaze at the fixation light (F) and to make a speeded manual response (by hand-held button press with finger) to any stimulus (VA, V or AV).

The general trial time-line is illustrated in Fig. 2. On unisensory trials a stimulus (A or V) occurred 1000–2000 ms after trial onset (flat probability distribution). On multisensory trials another stimulus (A or V) occurred at a variable SOA_{AV} . Trials ended on participant response or timed out after 1500 ms. Once a block of trials had begun, participants were not given indication of individual trial onset or inter-trial interval, but rather continuously responded to stimuli as they were presented until the block ended (indicated verbally by the experimenter).

Stimuli were presented in blocks containing 10 unisensory V trials, 10 unisensory A trials, 25 catch (no signal) trials and 10 trials each of multisensory (AV) combinations, numbering 135 trials in

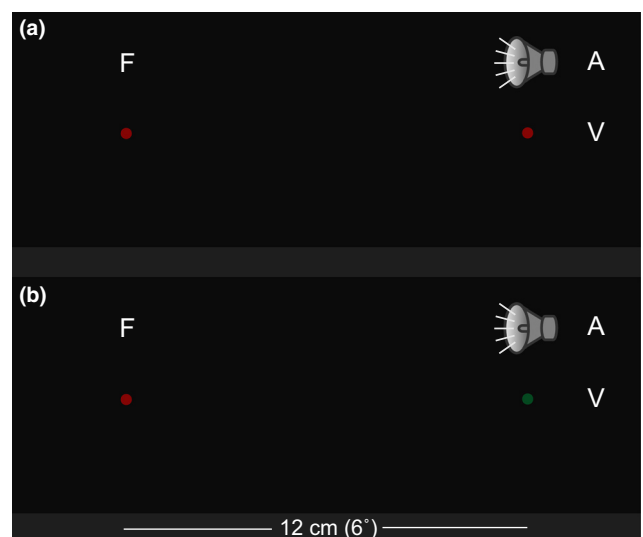


FIG. 1. Stimulus arrangement for Condition 1 (photopic visual stimuli) and Condition 2 (scotopic visual stimuli).

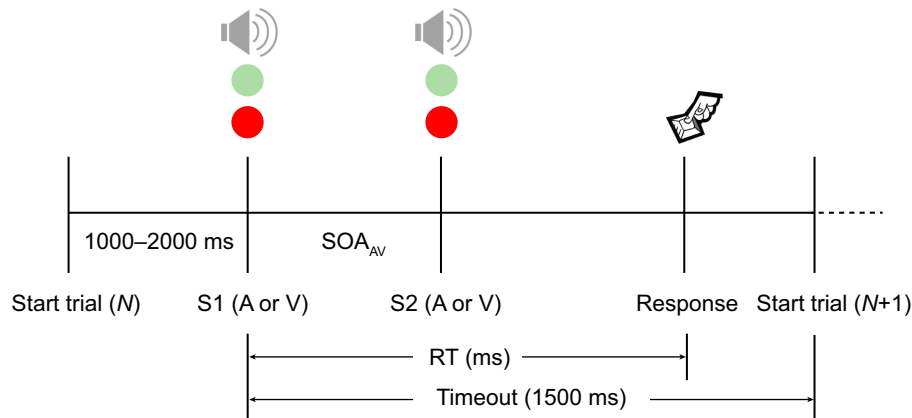


FIG. 2. Trial time-line.

total. SOAs on multisensory (AV) trials ranged between -200 and $+200$ ms at 50-ms intervals (negative SOA_{AV} means that the A stimulus preceded the V stimulus). In a single session each participant finished a total of 10 blocks of trials. Participants were given breaks between blocks.

Analysis. RT was recorded to the nearest millisecond, and outliers ($100 \text{ ms} > \text{RT} > 1000 \text{ ms}$) were trimmed. Independent samples *t*-tests compared mean RT to the fastest unisensory stimulus with mean RT to each multisensory stimulus. Cumulative RT density functions in each condition were bootstrapped using 1000 iterations (Foster & Bischof, 1991) to obtain 95% confidence intervals. When Miller's inequality is positive, it indicates that the probability of obtaining an RT faster than (*t*) ms in response to a multisensory stimulus is 'greater' than the sum of the probabilities of obtaining RTs faster than (*t*) ms to the unisensory components. Values of Miller's inequality were significantly greater than zero index facilitation via neural coactivation.

Results

Response enhancement. Figure 3 displays the mean RT as a function of SOA_{AV} in the photopic V stimulus conditions. Bars coloured green indicate conditions where significant multisensory response enhancement occurred. The mean RT in the unisensory auditory and visual conditions was 330 ms and 356 ms, respectively. The mean RT in multisensory conditions was significantly faster than the fastest unisensory stimulus (A) at three SOA_{AV} : -100 ms ($t_{993} = 3.78$, $P < 0.001$); -50 ms ($t_{994} = 3.89$, $P < 0.001$); and 0 ms ($t_{995} = 5.65$, $P < 0.001$).

MI. The left panel of Fig. 4 shows the mean bootstrapped Miller's inequality (spectrum-coded) as a function of RT and SOA_{AV} . The rectangular region outlined in black in the left panel, which contained significant positive values of Miller's inequality, is shown in magnified view in the right panel. The bootstrapped data (analysed at both individual and group levels) revealed that significant violations of Miller's inequality occurred exclusively at an SOA_{AV} of 0 ms (i.e. physical simultaneity).

Condition 2: RT and scotopic visual stimuli

Condition 1 disclosed that violations of the race model occurred for a photopic (cone-isolating) visual stimulus only when it was simultaneously paired with an A stimulus. In Condition 2, similarly the range of SOA_{AV} was measured where violations of the race model

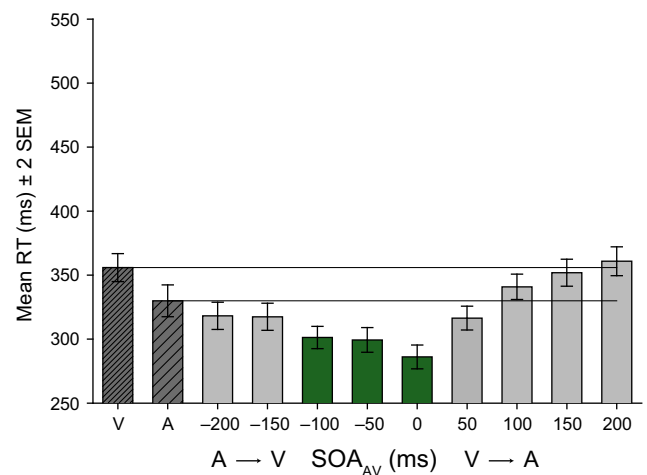


FIG. 3. Mean A, V and AV RTs for Condition 1 as a function of SOA_{AV} (photopic visual stimuli). Mean RT SOA_{AV} values of -100 ms, -50 ms and 0 ms are significantly faster than the fastest unisensory RT (A).

occurred when using scotopic (rod-isolating) V stimuli, to which RT is significantly slower than to photopic stimuli.

Method

Participants. As in Experiment 1, Condition 1.

Stimuli. See Fig. 1b.

Pretest procedure. The method of adjustment was used to obtain the absolute and colour thresholds of each participant to the 525-nm LED V test stimulus. In the experiment itself the intensity of this stimulus was midway between the absolute and colour thresholds for each participant.

Procedure and analysis. As in Experiment 1, Condition 1.

Results

Response enhancement. Figure 5 displays the mean RT as a function of SOA_{AV} in the scotopic V stimulus conditions. Bars coloured green indicate conditions where significant multisensory response enhancement occurred. The mean RT in the unisensory auditory and

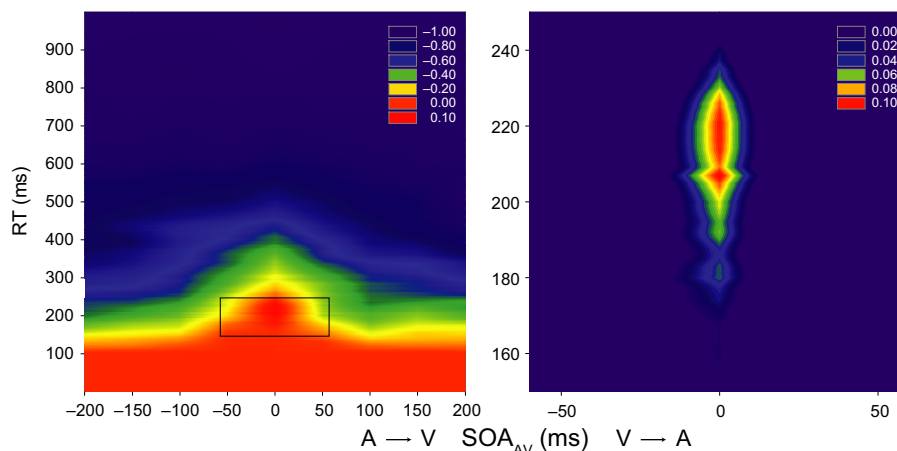


FIG. 4. Miller's inequality (spectrum coded) as a function of RT and SOA_{AV} for Condition 1 (photopic visual stimuli). The right-hand panel shows exploded view of area defined by the box in the left-hand panel. Violations of the race model that imply neural coactivation occur exclusively at an SOA_{AV} of 0 ms.

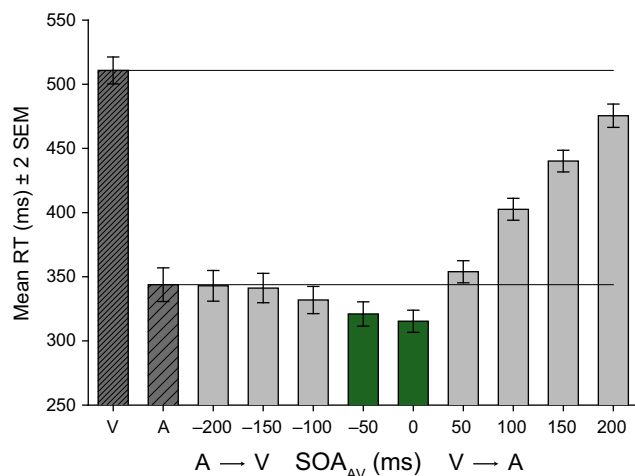


FIG. 5. Mean A, V and AV RTs for Condition 2 as a function of SOA_{AV} (scotopic visual stimuli). Mean RT SOA_{AV} values of -50 ms and 0 ms are significantly faster than the fastest unisensory RT (A).

visual conditions was 344 ms and 511 ms, respectively. The mean RT in multisensory conditions was significantly faster than the fastest unisensory stimulus (A) at two SOA_{AV} : -50 ms ($t_{1003} = 2.84$, $P = 0.005$); and 0 ms ($t_{1000} = 3.64$, $P < 0.001$).

MI. The left panel of Fig. 6 shows the mean bootstrapped Miller's inequality (spectrum-coded) as a function of RT and SOA_{AV} . The rectangular region outlined in black in the left panel, which contained significant positive values of Miller's inequality, is shown in magnified view in the right panel. The bootstrapped data (analysed at both individual and group levels) revealed that significant violations of Miller's inequality occurred exclusively at an SOA_{AV} of 0 ms.

In both conditions of Experiment 1, it was found that violations of the race model occurred exclusively at an SOA_{AV} of 0 ms. While this result seems at odds with previous research indicating that optimal AV integration occurs when the V stimulus precedes the A stimulus by 50–100 ms (Lewald & Guski, 2003), it was noted that the current results also show that simple facilitation of RT does occur in the subjects at values of SOA_{AV} of -50 or -100 ms (Figs 3 and 5).

Experiment 2: temporal order and simultaneity judgements

Experiment 2 addressed whether the stimulus simultaneity requirements for facilitative MI with regard to RT extend to judgements of temporal order and/or simultaneity?

Task 1: temporal order judgements

Method

Participants. Participants ($n = 6$, three female; 22–42 years old) possessed normal or corrected-to-normal spatial vision, normal colour vision, and normal hearing. Four of the participants had previously participated in Experiment 1, which was conducted between 1 and 3 months prior to Experiment 2. The long interval separating the two experiments greatly diminishes the likelihood of significant order effects. The experiment was undertaken with the understanding and written consent of all participants, subsequent to review and approval by the North Dakota State University Institutional Review Board, consistent with the World Medical Association Declaration of Helsinki.

Stimuli. As in Experiment 1.

Procedure. Participants were given two response buttons (one held in each hand). When making temporal order judgements, participants were instructed to respond via button press with the thumb or forefinger of their left hand if the auditory stimulus 'started first', or with their right hand if the visual stimulus 'started first'.

Stimuli were presented in blocks consisting of 10 quasi-randomly interleaved trials per level of SOA for nine values of SOA_{AV} , ranging from -200 to $+200$ ms at 50-ms intervals; negative SOA_{AV} values denote that the A stimulus preceded the V stimulus. Each participant completed all 900 trials (10 blocks of 90 trials each) in a single session with rest between blocks.

Analysis. For each participant, the proportion of 'visual first' responses was plotted as a function of SOA_{AV} . Individual psychometric functions were fit to cumulative normal distributions by the method of least-squares, and PSS was taken as the SOA_{AV} yielding a 50% 'visual first' response rate. Paired samples t -tests were conducted to determine whether PSS differed across V stimulus

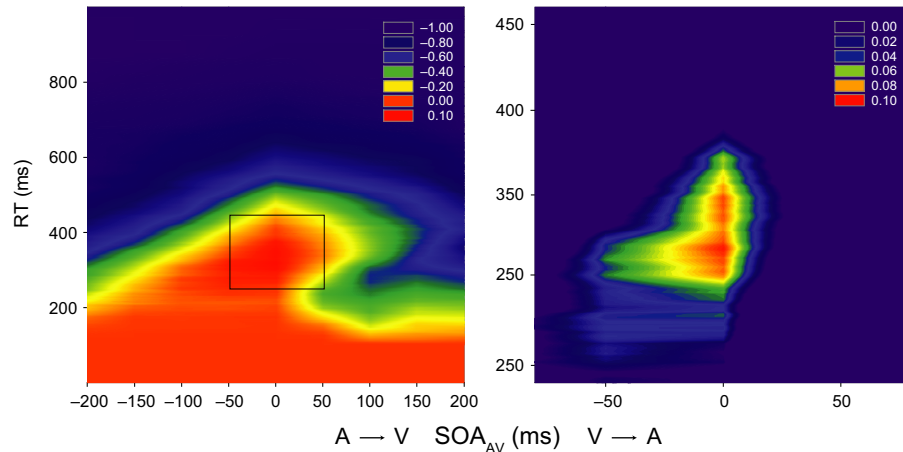


FIG. 6. Miller's inequality (spectrum coded) as a function of RT and SOA_{AV} for Condition 2 (scotopic visual stimuli). The right-hand panel shows exploded view of area defined by the box in the left-hand panel. Violations of the race model that imply neural coactivation occur exclusively at an SOA_{AV} of 0 ms.

conditions. Figure 7a illustrates curves fitted for responses to the scotopic (green) and photopic (red) stimuli. The bar graph in Fig. 7b shows group mean PSS (± 1 SEM) in each stimulus condition.

Results and discussion

The group mean PSS (where positive values indicate that the V stimulus preceded the A stimulus) was 70.7 ms in the scotopic condition and 30.6 ms in the photopic condition. A paired-samples t -test showed this difference to be significant ($t_5 = 5.07$, $P = 0.004$).

Task 2: simultaneity judgements

Method

Participants. As in Experiment 2, Condition 1.

Stimuli. As in Experiment 1.

Procedure. Participants were given two response buttons (one held in each hand). When making simultaneity judgements, participants

responded with the thumb or forefinger of their right hand if the stimuli 'started at the same time', and with their left hand if 'one stimulus started before the other'. Stimuli were presented as in Condition 1.

Analysis. For each participant, the proportion of 'simultaneous' responses was plotted as a function of SOA_{AV} . Individual data were fit to a three-parameter Gaussian function using the method of least-squares, and PSS was taken as the SOA_{AV} corresponding to the peak of the fitted function. Figure 8a shows an example of how curves were fitted for the scotopic (green) and photopic (red) stimuli. The bar graph in Fig. 8b shows group mean PSS (± 1 SEM) in each stimulus condition.

Results and discussion

Similar to temporal order judgements, the group mean PSS (where positive values indicate that the V stimulus preceded the A stimulus) was 113.9 ms in the scotopic condition and 39.1 ms in the photopic condition. A paired-samples t -test showed this difference to be highly significant ($t_5 = 4.56$, $P = 0.006$).

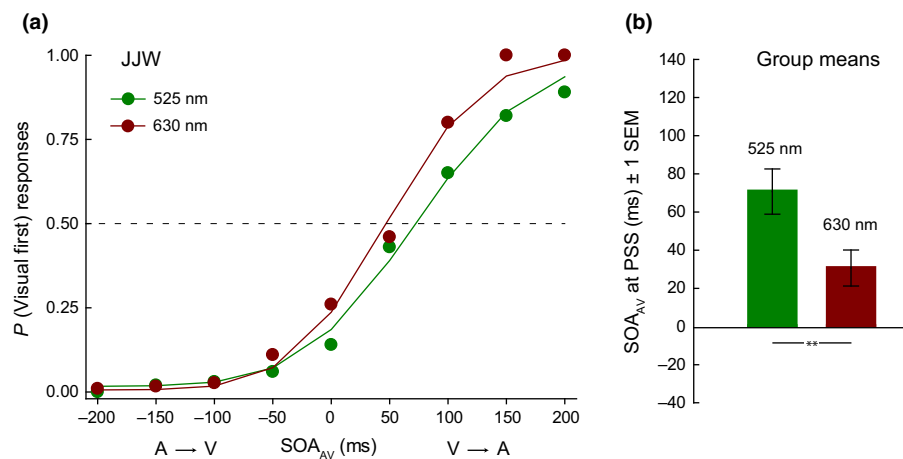


FIG. 7. (a) Representative example (participant JJW) of data and fitted psychometric functions for scotopic (green) and photopic (red) stimuli in the temporal order judgement task. (b) Bar chart showing significant difference in group mean PSS (± 2 SEM) for scotopic (green) and photopic (red) stimuli in the temporal order judgement task.

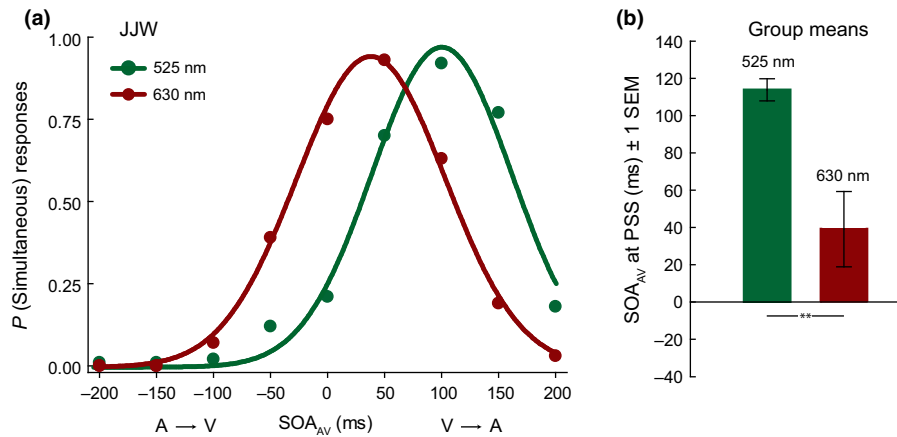


FIG. 8. (a) Representative example (participant JJW) of data and fitted psychometric functions for scotopic (green) and photopic (red) stimuli in the simultaneity judgement task. (b) Bar chart showing significant difference between group mean PSS (± 2 SEM) for scotopic (green) and photopic (red) stimuli in the simultaneity judgement task.

Discussion

Physical and physiological simultaneity

In Experiment 1 (Condition 1), where the V stimulus was photopic and the mean visual RT was relatively fast (356 ms), physical simultaneity of the A and V stimuli was required for facilitative MI. Although the average RT to multisensory AV combinations was significantly faster than to the fastest unisensory (A) stimulus at simultaneity, as well as for several non-zero values of SOA_{AV}, redundancy gains exceeded those accounted for by probability summation only at an SOA_{AV} of 0 ms. In Experiment 1 (Condition 2), where the V stimulus was scotopic and the mean visual RT (511 ms) was 155 ms slower than in the photopic condition, redundancy gains likewise exceeded those predicted by probability summation only at an SOA_{AV} of 0 ms. These results are consistent with those reported by Leone & McCourt (2013).

While it is generally accepted that the afferent signals from two unisensory events must simultaneously converge on 'coincidence detectors' for facilitative MI to occur (Stein & Meredith, 1993), the current results show that facilitation occurs only when the physical stimuli are simultaneously presented, notwithstanding other factors that can significantly influence convergence latency such as stimulus intensity or receptor response. The ecological validity of this result is self-evident. Of paramount interest to organisms is the detection and/or response to singular physical events in their close environment. The mechanism by which stimulus intensity can be taken into account, however, is unclear.

Let us stipulate that facilitative MI occurs only at coincidence detectors when the postsynaptic potentials generated by two or more unisensory stimuli temporally overlap. Varying the relative stimulus intensity (i.e. the speed of transduction) of these simultaneous stimuli will, however, cause the resulting afferent signals to converge asynchronously, and yet the current experiments show that physically simultaneous stimuli integrate despite differences in relative intensity. Conversely, the asynchronous occurrence of unisensory stimuli can, depending on their relative intensity, result in spuriously coincident physiological convergence that nevertheless fails to result in facilitation. It therefore appears that multisensory systems 'take into account' differences in convergence latency that are not caused by physical non-simultaneity (i.e. differences caused by unequal pathway length and/or changes in relative stimulus intensity) in order to reject false correspon-

dences and integrate only multisensory events that are truly simultaneous.

Leone & McCourt (2013) noted that unisensory afferent signals must possess information concerning stimulus intensity (or else intensity information would be lost) and that the preponderance of veridically simultaneous AV multisensory events occurring within praxis space (Previc, 1998) are likely to be self-generated (e.g. tapping a surface). These self-generated multisensory events originate from locations sufficiently close to sensory receptors to make the large difference between the speed of propagation of light and sound energy trivial.

Leone & McCourt (2013) argued that organisms have routine access to the neural responses evoked by self-generated physically simultaneous multisensory events in their environment. Over their development organisms establish Bayesian priors with respect to multisensory convergence based on the class of multisensory stimulation for which exact latency/intensity information is known, viz., self-generated events. Such events might include tapping an object with the hand (or a tool), or throwing a projectile whose impact with a nearby surface produces both visual and auditory consequences, and which produces visual, auditory and/or haptic/kinesthetic signals of known common origin. Because such events will vary in intensity, the latencies at which these signals converge at multisensory coincidence detectors can, over repeated stimulation, result in the accumulation of intensity-adjusted probability distributions upon which the integration of novel multisensory events can be based.

Hence, through early exposure to instances of multisensory ground truth organisms may learn the trading relations between intensity and latency, and thereby establish Bayesian priors that are then applied to AV stimuli more generally. The potential role of early exposure and learning is highlighted by results that AV MI in neurons of the cat superior colliculus is severely disrupted when cats are reared under environmental conditions where access to multisensory ground truth is perturbed, such as by dark-rearing (Yu *et al.*, 2013), or through continuous exposure to auditory noise (Xu *et al.*, 2014).

MI for perception vs. action

In Experiment 2, perceived simultaneity (PSS) in both temporal order and simultaneity judgement tasks never corresponded to physical simultaneity, but always required the V stimulus to precede the

A stimulus. These results replicate previous findings for judgements of temporal order and simultaneity tasks (Lewkowicz, 1996; Slutsky & Recanzone, 2001; Lewald & Guski, 2004; Fugisaki & Nishida, 2009), where 'visual-lead' stimuli most often induced the perception of multisensory simultaneity (Vroomen & Keetels, 2010). Further, both types of judgements were affected by alterations in visual processing speed caused by manipulations of signal strength and photoreceptor type: the mean difference in PSS in the photopic vs. scotopic stimulus conditions was 40.08 ms and 74.79 ms for temporal order and simultaneity judgements, respectively.

The current results thus disclose a significant discrepancy between the temporal requirements for MI with respect to action vs. perception, and are reminiscent of the dissociation between visual subsystems specialized for perception vs. action introduced by Goodale and Milner (Goodale & Milner, 1992; Goodale & Westwood, 2004). Evidence for the existence of two subsystems has come from neuropsychological studies of patients with optic ataxia (errors of misreaching that usually accompany damage to dorsal stream structures) and visual form agnosia (the inability to recognize/identify objects and/or faces that are associated with damage to the ventral stream) who show opposite patterns of deficit and preservation of function (for reviews, see Milner & Goodale, 2008; Goodale, 2011; Kravitz *et al.*, 2011). Examples of such dissociations include that hand grip aperture is relatively unaffected by a variety of visual size illusions, such as the Ponzo, Ebbinghaus and Müller-Lyer (Goodale, 2011).

RT tasks are visuomotor, whereas judgements of temporal order and simultaneity are perceptual in nature. If facilitative MI affords a survival advantage, then ecological validity demands that visuomotor action should be strongly coupled to the physical environment. MI realized via the action subsystem therefore requires a mechanism to compensate for adventitious changes in the latency of component unisensory afferent signals at multisensory coincidence detectors caused by differences in stimulus intensity. There appear to be few such demands on the perceptual subsystem, and examples of how vision for perception can be influenced by changes in stimulus intensity abound, as in the Pulfrich and Hess effects (Pulfrich, 1922; Williams & Lit, 1983). The Pulfrich effect occurs when a dark filter over one eye lowers stimulus intensity and delays transduction, causing zero-disparity binocular stimuli to appear shifted in depth. The Hess effect refers to a Vernier-like shift in the position of two aligned moving lines when they possess dissimilar luminances, where the darker line lags behind the brighter one. Both effects illustrate that intensity-dependent differences in transduction latency can and do alias as differences in spatial position and are not compensated for by the perceptual subsystem.

Finally, judgements of temporal order and simultaneity are similarly vulnerable to influences such as stimulus distance (Stone *et al.*, 2001; Lewald & Guski, 2004; Arnold *et al.*, 2005), spatial disparity (Zampini *et al.*, 2003), stimulus intensity and duration (Boenke *et al.*, 2009), and adaptation to stimulus asynchrony (Fugisaki *et al.*, 2004; Vroomen *et al.*, 2004; Vatakis *et al.*, 2007; Harrar *et al.*, 2008; Harrar & Harris, 2008; Di Luca *et al.*, 2009).

In conclusion, the current experiments disclose that the temporal requirements for MI differ for tasks that are primarily visuomotor (RT) vs. perceptual. Visuomotor tasks demand physical stimulus simultaneity regardless of stimulus intensity, whereas MI in perceptual tasks depends on both stimulus intensity and physical SOA.

Acknowledgements

The research described in this paper was supported by NIH P20 GM103505. The National Institute of General Medical Sciences (NIGMS) is a component

of the National Institutes of Health (NIH). The contents of this report are solely the responsibility of the authors, and do not necessarily reflect the official views of the NIH or NIGMS. The authors thank Enrique Alvarez and Huanzhong (Dan) Gu for expert technical assistance. The authors declare that there is no conflict of interest.

Abbreviations

AV, audiovisual; MI, multisensory integration; PSTH, post-stimulus time histogram; RSE, redundant signals effect; RT, reaction time; SOA, stimulus-onset asynchrony.

References

- Arnold, D.H., Johnston, A. & Nishida, S. (2005) Timing sight and sound. *Vision Res.*, **45**, 1275–1284.
- Barbur, J.L. (1982) Reaction-time determination of the latency between visual signals generated by rods and cones. *Ophthalm. Physiol. Opt.*, **2**, 179–185.
- Barlow, R.B., Snodderly, D.M. & Swadlow, H.A. (1978) Intensity coding in primate visual system. *Exp. Brain Res.*, **31**, 163–177.
- Boenke, L.T., Deliano, M. & Ohl, F.W. (2009) Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. *Exp. Brain Res.*, **198**, 233–244.
- Celesia, G.G. & Puletti, F. (1971) Auditory input to the human cortex during states of drowsiness and surgical anesthesia. *Electroen. Clin. Neuro.*, **31**, 603–609.
- Clark, V.P. & Hillyard, S.A. (1996) Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *J. Cognitive. Neurosci.*, **8**, 387–402.
- Di Luca, M., Machulla, T.-K. & Ernst, M.O. (2009) Recalibration of multisensory simultaneity: cross-modal transfer coincides with a change in perceptual latency. *J. Vision*, **9**, 1–16.
- Diederich, A. & Colonius, H. (2004) Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Percept. Psychophys.*, **66**, 1388–1404.
- Foster, D.H. & Bischof, W.F. (1991) Bootstrap estimates of the statistical accuracy of thresholds obtained from psychometric functions. *Spatial Vision*, **11**, 135–139.
- Foxe, J.J. & Simpson, G.V. (2002) Flow of activation from V1 to frontal cortex in humans: a framework for defining "early" visual processing. *Exp. Brain Res.*, **142**, 139–150.
- Foxe, J.J., Strugstad, E.C., Sehatpour, P., Molholm, S., Pasielka, W., Schroeder, C.E. & McCourt, M.E. (2008) Parvocellular and magnocellular contributions to the initial generators of the visual evoked potential: high-density electrical mapping of the "C1" component. *Brain Topogr.*, **21**, 11–21.
- Freiberg, C., Thomas, M.M. & Lamb, T.D. (2001) Time course of the flash response of dark- and light-adapted human rod photoreceptors derived from the electroretinogram. *Jpn. J. Physiol.*, **53A**, 217–242.
- Fugisaki, W. & Nishida, S. (2009) Audio-tactile superiority over visuo-tactile and audio-visual combinations in the temporal resolution of synchrony perception. *Exp. Brain Res.*, **198**, 245–259.
- Fugisaki, W., Shimojo, S., Kashino, M. & Nishida, S. (2004) Recalibration of audiovisual simultaneity. *Nat. Neurosci.*, **7**, 773–778.
- Goodale, M.A. (2011) Transforming vision into action. *Vision. Res.*, **51**, 1567–1587.
- Goodale, M.A. & Milner, A.D. (1992) Separate visual pathways for perception and action. *Trends Neurosci.*, **15**, 20–25.
- Goodale, M.A. & Westwood, D.A. (2004) An evolving view of duplex vision: separate but interacting cortical pathways for perception and action. *Curr. Opin. Neurobiol.*, **14**, 203–211.
- Hanson, J.V.M., Heron, J. & Whitaker, D. (2008) Recalibration of perceived time across sensory modalities. *Exp. Brain Res.*, **185**, 347–352.
- Harrar, V. & Harris, L.R. (2008) The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Exp. Brain Res.*, **186**, 517–524.
- Hershenson, M. (1962) Reaction time as a measure of intersensory facilitation. *J. Exp. Psychol.*, **63**, 289–293.
- Jeffreys, D.A. & Axford, J.G. (1972) Source locations of pattern-specific components of human visual evoked potentials. II. Component of extrastriate cortical origin. *Exp. Brain Res.*, **16**, 22–40.
- Kravitz, D.J., Saleem, K.S., Baker, C.I. & Mishkin, M. (2011) A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.*, **12**, 217–230.

- Leone, L. & McCourt, M.E. (2013) The roles of physical and physiological simultaneity in audiovisual multisensory facilitation. *i-Perception*, **4**, 213–228.
- Lewald, J. & Guski, R. (2003) Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Brain Res. Cogn. Brain Res.*, **16**, 468–478.
- Lewald, J. & Guski, R. (2004) Auditory-visual temporal integration as a function of distance: no compensation for sound-transmission time in human perception. *Neurosci. Lett.*, **357**, 119–122.
- Lewkowicz, D.J. (1996) Perception of auditory-visual temporal synchrony in human infants. *J. Exp. Psychol. Hum. Percept. Perform.*, **22**, 1094–1106.
- Maunsell, J.H.R., Ghose, G.M., Assad, J.A., McAdams, C.J., Boudreau, C.E. & Noerager, B.D. (1999) Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Vis. Neurosci.*, **16**, 1–14.
- Miller, J. (1982) Divided attention: evidence for coactivation with redundant signals. *Cogn. Psychol.*, **14**, 247–279.
- Miller, J. (1986) Timecourse of coactivation in bimodal divided attention. *Percept. Psychophys.*, **40**, 331–343.
- Milner, A.D. & Goodale, M.A. (2008) Two visual systems re-reviewed. *Neuropsychologia*, **46**, 774–785.
- Piéron, H. (1952) *The Sensations: Their Functions, Processes and Mechanisms*. Frederick Muller Ltd, London.
- Prestrude, A.M. (1971) Visual latencies at photopic levels of retinal illumination. *Vision. Res.*, **11**, 351–361.
- Previc, F.H. (1998) The neuropsychology of 3-D space. *Psychol. Bull.*, **124**, 123–164.
- Pulfrich, C. (1922) Die stereoskopie im Dienste der isochromen und heterochromen photometrie. *Die Naturwissenschaften*, **10**, 553–64.
- Raab, D.H. (1962) Statistical facilitation of simple reaction times. *T. New York Acad. Sci.*, **24**, 574–590.
- Roufs, J.A.J. (1963) Perception lag as a function of stimulus luminance. *Vision. Res.*, **3**, 81–91.
- Rutschmann, R. (1966) Perception of temporal order and relative visual latency. *Science*, **152**, 1099–1101.
- Slutsky, D.A. & Recanzone, G.H. (2001) Temporal and spatial dependency of the ventriloquism effect. *NeuroReport*, **12**, 7–10.
- Stein, B.E. & Meredith, M.A. (1993) *The Merging of the Senses*. MIT Press, Cambridge.
- Stone, J.V., Hunkin, N.M., Porril, J., Wood, R., Keeler, V., Beanland, M., Port, M. & Porter, N.R. (2001) When is now? Perception of simultaneity. *P. R. Soc. London, B*, **268**, 31–38.
- Vatakis, A., Navarra, J., Soto-Faraco, S. & Spence, C. (2007) Temporal recalibration during asynchronous audiovisual speech perception. *Exp. Brain Res.*, **181**, 173–181.
- Verdon, W.A., Schneck, M.E. & Haegerstrom-Portnoy, G. (2003) A comparison of three techniques to estimate the human dark-adapted cone electroretinogram. *Vision. Res.*, **43**, 2089–2099.
- Vroomen, J. & Keetels, M. (2010) Perception of intersensory synchrony: a tutorial review. *Atten. Percept. Psycho.*, **72**, 871–884.
- Vroomen, J., Keetels, M., de Gelder, B. & Bertelson, P. (2004) Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Res.*, **22**, 32–35.
- Williams, J.M. & Lit, A. (1983) Luminance-dependent visual latency for the Hess effect, the Pulfrich effect, and simple reaction time. *Vision. Res.*, **23**, 171–179.
- Xu, J., Yu, L., Rowland, B.A., Stanford, T.R. & Stein, B.E. (2014) Noise-rearing disrupts the maturation of multisensory integration. *Eur. J. Neurosci.*, **39**, 602–613.
- Yu, L., Xu, J., Rowland, B.A. & Stein, B.E. (2013) Development of cortical influences on superior colliculus multisensory neurons: effects of dark-rearing. *Eur. J. Neurosci.*, **37**, 1594–1601.
- Zampini, M., Shore, D.I. & Spence, C. (2003) Audiovisual temporal order judgments. *Exp. Brain Res.*, **152**, 198–210.

Appendix A

The RSE

Miller's inequality asserts that the cumulative probability (CP) of obtaining the fastest responses to redundant signals must be less than or equal to the CP of obtaining the fastest responses to individual stimuli. Thus, for a pair of A and V stimuli, at some AV SOA (SOA_{AV}), and at an observed response latency (t), the inequality states that:

$$CP(RT < t | AV) \leq CP(RT < t - SOA_{AV} | A) + CP(RT < t | V)$$

where the A stimulus is delayed relative to the V stimulus, or

$$CP(RT < t | AV) \leq CP(RT < t | A) + CP(RT < t - SOA_{AV} | V)$$

where the V stimulus is delayed relative to the A stimulus.

The expression $CP(RT < t | AV)$ is the CP of obtaining an RT faster than response latency (t) in response to the presentation of the A 'and' V stimuli. This CP must be less than or equal to the sum of the CPs of obtaining RTs faster than time in response to the individual unisensory stimuli:

$$CP(RT < t - SOA_{AV} | A) + CP(RT < t | V) \text{ or } CP(RT < t | A) + CP(RT < t - SOA_{AV} | V).$$

Appendix B

The duplex retina

The human retina contains two types of photoreceptors: rods and cones. The rods operate under fully dark adapted conditions and are disabled under fully light adapted conditions due to response saturation. The scotopic (rod-based) luminous efficiency function, $V'(\lambda)$, peaks at a wavelength of 507 nm, and the photopic (cone-based) luminous efficiency function, $V(\lambda)$, peaks at 555 nm. Long-wavelength sensitive (red) cones can be stimulated by long wavelength (630 nm) light to which rods are insensitive. Thus, both the rod and cone systems can be stimulated under fully dark adapted conditions by carefully selecting stimulus wavelengths.

Rods and cones have very different temporal response characteristics. Rod responses to moderately intense stimuli peak at about 120 ms (Freiberg *et al.*, 2001), whereas the responses of cones peak at 10–30 ms (Verdon *et al.*, 2003), and manual RT to photopic stimuli is about 80 ms faster than to scotopic stimuli (Rutschmann, 1966; Barbur, 1982).