

Top-down control and early multisensory processes: chicken vs. egg

Rosanna De Meo¹, Micah M. Murray^{1,2}, Stephanie Clarke¹ and Pawel J. Matusz^{3,4,5*}

¹ Neuropsychology and Neurorehabilitation Service, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland

² Electroencephalography Brain Mapping Core, Center for Biomedical Imaging (CIBM), Lausanne and Geneva, Switzerland

³ The Laboratory for Investigative Neurophysiology, Neuropsychology and Neurorehabilitation Service and Department of Radiology, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland

⁴ Faculty in Wroclaw, University of Social Sciences and Humanities, Wroclaw, Poland

⁵ Attention, Brain and Cognitive Development Group, Department of Experimental Psychology, University of Oxford, Oxford, UK

*Correspondence: pawel.matusz@chuv.ch

Edited by:

Salvador Soto-Faraco, Universitat Pompeu Fabra, Spain

Reviewed by:

Mark T. Wallace, Vanderbilt University, USA

Durk Talsma, UGent, Belgium

Keywords: attention, control processes, top-down control, bottom-up, multisensory, EEG/ERP, crossmodal

Traditional views contend that behaviorally-relevant multisensory interactions occur relatively late during stimulus processing and subsequently to influences of (top-down) attentional control. In contrast, work from the last 15 years shows that information from different senses is integrated in the brain also during the initial 100 ms after stimulus onset and within low-level cortices. Critically, many of these early-latency multisensory interactions (hereafter *eMSI*) directly impact behavior. The prevalence of *eMSI* substantially advances our understanding of how unified perception and goal-related behavior emerge. However, it also raises important questions about the dependency of the *eMSI* on top-down, goal-based attentional control mechanisms that bias information processing toward task-relevant objects (hereafter *top-down control*). To date, this dependency remains controversial, because *eMSI* can occur independently of top-down control, making it plausible for (some) multisensory processes to directly shape perception and behavior. In other words, the former is not necessary for these early effects to occur and to link them with perception (see **Figure 1A**). This issue epitomizes the fundamental question regarding direct links between sensation, perception, and behavior (*direct perception*), and also extends it in a crucial way to incorporate the multisensory nature of everyday experience. At

the same time, the emerging framework must strive to also incorporate the variety of higher-order control mechanisms that likely influence multisensory stimulus responses but which are not based on task-relevance. This article presents a critical perspective about the importance of top-down control for *eMSI*: In other words, who is controlling whom?

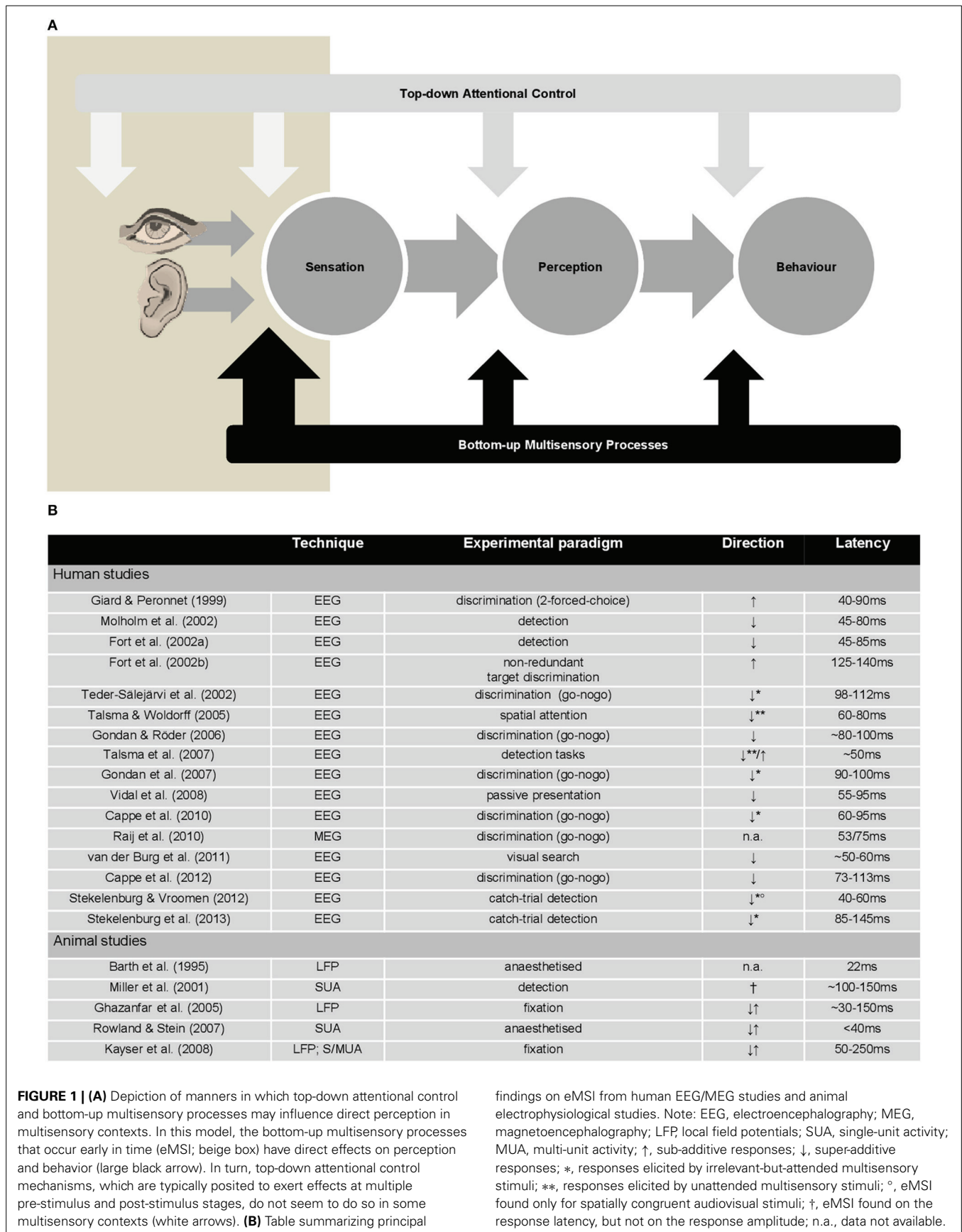
THE UBIQUITY OF *eMSI*

For the purposes of this article we focus exclusively on auditory-visual interactions and define *eMSI* as those multisensory processes that occur within the first 100 ms post-stimulus onset (but see (Giard and Peronnet, 1999); Giard and Peronnet, who qualified effects <200 ms as early-latency). This definition is in keeping with influential models of visual perception and attentional selection, positing that top-down and recursive inputs manifest after the initial 100 ms of stimulus-driven brain activity, which is believed to be sensory-perceptual and bottom-up in nature (e.g., Luck et al., 1997; Lamme and Roelfsema, 2000). It is likewise important to distinguish between *integration* effects, which are responses elicited by a combination of inputs to different senses, and *cross-modal* effects, which refer to influences of inputs to one sense on activity associated with another sense (e.g., Stein et al., 2010).

The typical perceptual outcome of multisensory integration is that stimulus processing is facilitated (as shown by

faster and/or more accurate responses) in contexts where inputs to different senses are carrying similar (*redundant*) information and are presented close in time. This behavioral facilitation is typically accompanied by brain responses to multisensory stimuli that diverge from the summed brain responses to the constituent unisensory signals (*nonlinear responses*; **Figure 1B**). Given the growing evidence for links between the brain and the behavioral responses (reviewed in Murray et al., 2012), one mechanism may be that the temporal co-occurrence of multisensory information lowers the threshold for neural activity that in turn drives perception and action (e.g., Rowland and Stein, 2007).

Based on the extant literature, we argue that these particular multisensory processes, which are reflected by *eMSI*, are stimulus-driven, *bottom-up* in nature and affect perception and behavior in a direct manner and largely independently of top-down control (**Figure 1A**). The idea of a variety, or even a range of multisensory processes, where some are “automatic” while others dependent on one’s current behavioral goals, has until now been systematically investigated mainly in the context of attentional selection of objects in space, rather than their perception *per se* (e.g., Matusz and Eimer, 2011, 2013; Matusz et al., 2015; Talsma and Woldorff, 2005; but see Murray et al., 2004; Soto-Faraco et al., 2004; Tiippana et al., 2004;



Alsus et al., 2005, 2007; Thelen et al., 2012, 2014; Matusz et al., in press). However, control processes are likely to be important for both cognitive functions (e.g., Günseli et al., 2014); this should hold for both unisensory and multisensory processes, and bottom-up and top-down processes alike (i.e., multisensory processes are not mechanistically “special”; van Atteveldt et al., 2014).

It is difficult to argue with the idea that early responses are a hallmark of bottom-up multisensory processes in the service of perception, if one considers how ubiquitous and context-independent they are in both humans and in the animal models (see **Figure 1B**; reviewed in Murray et al., 2012; Kajikawa et al., 2012). The eMSI in local field potentials as well as spiking activity have been measured in the primary and secondary auditory fields of fixating monkeys (Ghazanfar et al., 2005; Kayser et al., 2008; see also Lakatos et al., 2008; Wang et al., 2008 for cross-modal effects). Importantly, these eMSI occurred for both ethological objects (conspecific communication signals) and simple audiovisual stimuli, though modulated according to bottom-up stimulus salience and neural efficacy. Moreover, non-linear interactions mirroring the behavioral gains in stimulus detection have been recorded in single neurons in the area 4 of the monkey motor cortex within 100–150 ms post-stimulus (Miller et al., 2001).

Electroencephalography and magnetoencephalography (EEG/MEG) studies in humans have likewise demonstrated eMSI across a variety of tasks, ranging from simple detection (Fort et al., 2002a; Molholm et al., 2002) and discrimination (Giard and Peronnet, 1999; Fort et al., 2002b; Teder-Sälejärvi et al., 2002; Gondan and Röder, 2006; Gondan et al., 2007; Raij et al., 2010; Cappe et al., 2010, 2012; Stekelenburg and Vroomen, 2012; Stekelenburg et al., 2013) tasks to multi-stimulus/ multi-stream paradigms necessitating selection (Talsma and Woldorff, 2005; Talsma et al., 2007; van der Burg et al., 2008, 2011). Importantly, eMSI were observed irrespective of whether the multisensory stimuli were targets (e.g., Giard and Peronnet, 1999; Pérez-Bellido et al., 2013), attended but task-irrelevant stimuli (e.g., Cappe et al., 2010) or were presented passively

(Vidal et al., 2008). As will be detailed below, data from brain stimulation studies allow causal inference regarding behavioral consequences of eMSI (see below).

The interpretability of the eMSI in terms of bottom-up vs. top-down mechanisms critically depends on their localization. Despite the ubiquity of the eMSI in extant EEG/MEG studies, only few have applied the requisite signal analysis and source reconstruction methods. Localization results support the predominant role of low-level cortices in the eMSI (Cappe et al., 2010; Raij et al., 2010). While the localization of the eMSI to low-level cortices could be taken as evidence for their strictly bottom-up nature, their latency at ~50–100 ms is sufficiently “late” to provide ample opportunity for recursive processing (Musacchia and Schroeder, 2009; also Moran and Reilly, 2006 for modeling results). This may involve top-down modulation or the extraction and disambiguation of stimulus features (Lamme and Roelfsema, 2000). Thus, care is warranted in regarding all eMSI as indicative of bottom-up multisensory integration. For example, *the pip and pop* effect (van der Burg et al., 2008) triggers eMSI-like responses, but only in the case of targets, not distractors (van der Burg et al., 2011). Thus, dependency of the eMSI on top-down control can be assessed only by analyzing studies where the latter is directly manipulated¹.

THE CHICKEN: TOP-DOWN CONTROL AND ITS LIMITED ROLE IN eMSI

The strongest evidence for the dependence of eMSI on top-down control comes from studies where attended and unattended multisensory stimuli were directly compared (e.g., Alsus et al., 2005, 2007; Talsma and Woldorff, 2005; Talsma et al., 2007). However, the literature seems prone to misconstruing the full breadth of the results. In one study participants detected infrequent targets in one of two central streams of rapidly presented alphanumeric symbols or combinations of beeps and flashes (Talsma et al., 2007). When attended, audiovisual stimuli triggered early enhanced (*super-additive*) nonlinear responses.

¹To our knowledge semantic congruence does not modulate eMSI (Fort et al., 2002b; Molholm et al., 2004; Yuval-Greenberg and Deouell, 2007).

But, when the competing stream was attended, these nonlinear interactions changed polarity, becoming suppressed (*sub-additive*). One interpretation of these results is that top-down control regulates multisensory integration, from its magnitude and quality to its very presence (Koelewijn et al., 2010). We believe this viewpoint should perhaps be more nuanced. The top-down control manipulations modulated the eMSI, but did not eliminate them. Additionally, the eMSI were observed despite the paradigm manipulating in fact *multiple* top-down mechanisms (inter-modal, but also spatial, feature-based, and object-based). While further research is required to fully characterize the mechanistic underpinnings of super- vs. sub-additive interactions, the results of this study are in line with the importance of top-down control processes revealed by unisensory studies, wherein responses to stimuli are enhanced according to the task-relevance of their location, features or identity (reviewed in Nobre and Kastner, 2013). Talsma et al. (2007) was the first to demonstrate the pivotal role of the task-relevance of multisensory pairings for the *quality* of the eMSI they trigger. However, the *presence* of the eMSI in this study was independent of task-relevance, though some evidence would suggest that the eMSI are preferentially observed in *unattended* contexts (Table 2 in Talsma and Woldorff, 2005). This latter evidence is in line with the eMSI being a hallmark of stimulus-driven processing.

It is difficult to ignore that in these few studies, where top-down control mechanisms were directly manipulated, the eMSI were sub-additive in nature. What is striking is that this is precisely the direction of effects reported in the literature irrespective of whether responses to targets, non-targets or passively presented stimuli are considered (**Figure 1B**). Historically, sub-additive effects were dismissed as confounds related to common activity across both unisensory and multisensory conditions. More recently, they have been increasingly recognized as a canonical mechanism that can convey information particularly efficiently (Kayser et al., 2009; Altieri et al., in press; reviewed in Stevenson et al., 2014). The issue of the quantification of the eMSI is further complicated by the fact that the overwhelming

majority of the human EEG studies have used relative, reference-dependent measures of amplitude (cf., Murray et al., 2008).

THE EGG: eMSI AS A BOTTOM-UP PHENOMENON

Several independent lines of research across various species provide converging evidence for the bottom-up nature of the eMSI. On the one hand, there are reports of eMSI in anesthetized animals (e.g., rats, Barth et al., 1995; cats, Rowland and Stein, 2007; see also reviews in Sarko et al., 2012; Rowland and Stein, 2014), where top-down modulations are blocked². On the other hand, sounds have been shown to enhance the excitability of low-level visual cortices, as measured via phosphene perception. Several aspects of this effect demonstrated by TMS studies in humans support the bottom-up nature of the eMSI and the causal links between eMSI and behavior (Romei et al., 2007, 2009, 2013; Spierer et al., 2013).

First, it is modulated by low-level sound features, with greater excitability increases observed for narrowband and higher pitch sounds. Visual cortex excitability is furthermore enhanced selectively by structured approaching (looming) sounds versus stationary or receding sounds as well as non-structured white-noise versions of these sounds. Second, the effect is delimited in time, occurring when sounds precede the TMS by 30–150 ms, in correspondence with the eMSI identified using EEG/MEG. Third, the sound-induced enhancements of visual cortex excitability transpire before subjects can explicitly differentiate between the sounds, i.e., at pre-perceptual processing stages. Relatedly, increases in the occipital excitability occur with sounds that themselves fail to elicit startle responses, arguing against an alerting explanation. Fourth, evidence against a top-down account of these effects comes from studies

demonstrating that individuals' attentional preference (as independently measured in an auditory-visual divided attention task) affect late, but not early, stages of the excitability changes.

Finally, the TMS-driven visual cortex activity is behaviorally relevant. Occipital TMS delivered 60–90 ms post-stimulus has opposing effects of roughly equal magnitude (~15 ms) on reaction times to unisensory auditory and visual stimuli (speeding and slowing, respectively) and has no measurable effect on reaction times to simultaneous auditory-visual multisensory stimuli. Critically, the response speed facilitation obtained from the combination of occipital TMS and an external auditory stimulus was as great as and correlated with that obtained from presenting participants with genuine multisensory stimuli. The TMS-induced cross-modal effects seem to emulate those observed with multisensory stimuli.

CONCLUSIONS AND FUTURE DIRECTIONS

We demonstrated that the eMSI are robust phenomena, observable across species, experimental paradigms and measures of neural activity (Figure 1B). To refer more explicitly to the Research Topic of this issue, we subscribe to a view of multiple multisensory processes: The eMSI are a hallmark of bottom-up multisensory processes that facilitate perception and behavior directly, independently of top-down control (Figure 1A).

We focused here exclusively on stimulus-locked brain activity. Thus, temporal dynamics complement the understanding of the interplay between bottom-up and top-down mental processes as hitherto provided from the vantage-point of brain oscillations, which assay both intra-population excitability as well as inter-population communication (Thut et al., 2012; van Atteveldt et al., 2014).

A critical next step will be the detailed mechanistic characterization of the eMSI. The sub-additive archetype of the eMSI goes together with the evidence from unisensory research linking reduced responses with more efficient and information-rich processing akin to the repetition suppression phenomena and the predictive coding accounts (e.g.,

Grill-Spector et al., 2006; Summerfield and Egner, 2009). When and why do top-down control processes flip the sub-additive eMSI to become super-additive? If top-down control affects the nature, rather than the presence, of multisensory processes, then what are the consequences for our understanding of perception? Paradoxically, while the eMSI are on the one hand upturning somewhat dogmatic views on the brain functional organization, they simultaneously are entrenching a classic model of perceptual processing positing direct links between sensation, perception, and behavior. An accurate picture of the nature of perceptual processes is thus provided by studying them in naturalistic, multisensory contexts and where the task demands dynamically vary.

AUTHOR CONTRIBUTIONS

All authors have contributed to all aspects of this work. All authors have approved the final version of the manuscript and agreed to be held accountable for all aspects of the work.

ACKNOWLEDGMENTS

Financial support was provided by the Swiss National Science Foundation (grant 320030-149982 to MMM, grant 320030B-141177 to SC), National Centre of Competence in Research project "SYNAPSY, The Synaptic Bases of Mental Disease" [project 51AU40-125759] and the Swiss Brain League (2014 Research Prize to MMM).

REFERENCES

- Alsius, A., Navarra, J., Campbell, R., and Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Curr. Biol.* 15, 839–843. doi: 10.1016/j.cub.2005.03.046
- Alsius, A., Navarra, J., and Soto-Faraco, S. (2007). Attention to touch weakens audiovisual speech integration. *Exp. Brain Res.* 183, 399–404. doi: 10.1007/s00221-007-1110-1
- Altieri, N., Stevenson, R. A., Wallace, M. T., and Wenger, M. J. (in press). Learning to associate auditory and visual stimuli: behavioral and neural mechanisms. *Brain Topogr.* 28. doi: 10.1007/s10548-013-0333-7
- Barth, D. S., Goldberg, N., Brett, B., and Di, S. (1995). The spatiotemporal organization of auditory, visual, and auditory-visual evoked potentials in rat cortex. *Brain Res.* 678, 177–190. doi: 10.1016/0006-8993(95)00182-P
- Cappe, C., Thelen, A., Romei, V., Thut, G., and Murray, M. M. (2012). Looming signals reveal synergistic principles of multisensory integration. *J. Neurosci.* 32, 1171–1182. doi: 10.1523/JNEUROSCI.5517-11.2012

²We would hasten to remind the reader that convergent anatomical input is necessary but in and of itself insufficient for eMSI as defined in this opinion piece. It is true that the anatomic pathways/connectivities as well as their shaping by experiences are prerequisites for multisensory processes. However, the activation of these physical substrates in relation to the cascade of sensory-evoked responses must be sufficiently early so as to influence perception and behavior directly and thus be qualified as eMSI.

- Cappe, C., Thut, G., Romei, V., and Murray, M. M. (2010). Auditory–visual multisensory interactions in humans: timing, topography, directionality, and sources. *J. Neurosci.* 30, 12572–12580. doi: 10.1523/JNEUROSCI.1099-10.2010
- Fort, A., Delpuech, C., Pernier, J., and Giard, M. H. (2002a). Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. *Cereb. Cortex* 12, 1031–1039. doi: 10.1093/cercor/12.10.1031
- Fort, A., Delpuech, C., Pernier, J., and Giard, M. H. (2002b). Early auditory–visual interactions in human cortex during nonredundant target identification. *Cogn. Brain Res.* 14, 20–30. doi: 10.1016/S0926-6410(02)00058-7
- Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., and Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J. Neurosci.* 25, 5004–5012. doi: 10.1523/JNEUROSCI.0799-05.2005
- Giard, M. H., and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* 11, 473–490. doi: 10.1162/089892999563544
- Gondan, M., and Röder, B. (2006). A new method for detecting interactions between the senses in event-related potentials. *Brain Res.* 1073, 389–397. doi: 10.1016/j.brainres.2005.12.050
- Gondan, M., Vorberg, D., and Greenlee, M. W. (2007). Modality shift effects mimic multisensory interactions: an event-related potential study. *Exp. Brain Res.* 182, 199–214. doi: 10.1007/s00221-007-0982-4
- Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23. doi: 10.1016/j.tics.2005.11.006
- Gunseli, E., Meeter, M., and Olivers, C. N. (2014). Is a search template an ordinary working memory? Comparing electrophysiological markers of working memory maintenance for visual search and recognition. *Neuropsychologia* 60, 29–38. doi: 10.1016/j.neuropsychologia.2014.05.012
- Kajikawa, Y., Falchier, A., Musacchia, G., Lakatos, P., and Schroeder, C. E. (2012). “Audiovisual integration in nonhuman primates: a window into the anatomy and physiology of cognition,” in *The Neural Bases of Multisensory Processes*, ed M. M. Murray and M. T. Wallace (Boca Raton, FL: CRC Press), 65–98.
- Kayser, C., Montemurro, M. A., Logothetis, N. K., and Panzeri, S. (2009). Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* 61, 597–608. doi: 10.1016/j.neuron.2009.01.008
- Kayser, C., Petkov, C. I., and Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cereb. Cortex* 18, 1560–1574. doi: 10.1093/cercor/bhm187
- Koelewijn, T., Bronkhorst, A., and Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: a review of audiovisual studies. *Acta Psychol.* 134, 372–384. doi: 10.1016/j.actpsy.2010.03.010
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113. doi: 10.1126/science.1154735
- Lamme, V. A., and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579. doi: 10.1016/S0166-2236(00)01657-X
- Luck, S. J., Chelazzi, L., Hillyard, S. A., and Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.* 77, 24–42.
- Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, R., and Scerif, G. (2015). Multimodal distraction: insights from children’s limited attention. *Cognition* 136, 156–165. doi: 10.1016/j.cognition.2014.11.031
- Matusz, P. J., and Eimer, M. (2011). Multisensory enhancement of attentional capture in visual search. *Psychon. B. Rev.* 18, 904–909. doi: 10.3758/s13423-011-0131-8
- Matusz, P. J., and Eimer, M. (2013). Top–down control of audiovisual search by bimodal search templates. *Psychophysiology* 50, 996–1009. doi: 10.1111/psyp.12086
- Matusz, P. J., Thelen, A., Geiser, E., Anken, J., and Murray, M. M. (in press). The role of auditory cortices in the retrieval of single-trial auditory-visual memories. *Eur. J. Neurosci.*
- Miller, J., Ulrich, R., and Lamarre, Y. (2001). Locus of the redundant-signals effect in bimodal divided attention: a neurophysiological analysis. *Percept. Psychophys.* 63, 555–562. doi: 10.3758/BF03194420
- Molholm, S., Ritter, W., Javitt, D. C., and Foxe, J. J. (2004). Multisensory visual–auditory object recognition in humans: a high-density electrical mapping study. *Cereb. Cortex* 14, 452–465. doi: 10.1093/cercor/bhh007
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., and Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cogn. Brain Res.* 14, 115–128. doi: 10.1016/S0926-6410(02)00066-6
- Moran, R. J., and Reilly, R. B. (2006). “Neural mass model of human multisensory integration,” in *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (New York, NY: IEEE), 5559–5562.
- Murray, M. M., Brunet, D., and Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr.* 20, 249–264. doi: 10.1007/s10548-008-0054-5
- Murray, M. M., Cappe, C., Romei, V., Martuzzi, R., and Thut, G. (2012). “Auditory–visual multisensory interactions in humans: a synthesis of findings from behavior, ERPs, fMRI, and TMS,” in *The New Handbook of Multisensory Processes*, ed B. E. Stein (Cambridge, MA: MIT Press), 223–238.
- Murray, M. M., Michel, C. M., de Peralta, R. G., Ortigue, S., Brunet, D., Andino, S. G., et al. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage* 21, 125–135.
- Musacchia, G., and Schroeder, C. E. (2009). Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. *Hear. Res.* 258, 72–79. doi: 10.1016/j.heares.2009.06.018
- Nobre, K., and Kastner, S. (eds.). (2013). *The Oxford Handbook of Attention*. Oxford: Oxford University Press.
- Pérez-Bellido, A., Soto-Faraco, S., and López-Moliner, J. (2013). Sound-driven enhancement of vision: disentangling detection-level from decision-level contributions. *J. Neurophys.* 109, 1065–1077. doi: 10.1152/jn.00226.2012
- Raij, T., Ahveninen, J., Lin, F. H., Witzel, T., Jääskeläinen, I. P., Letham, B., et al. (2010). Onset timing of cross–sensory activations and multisensory interactions in auditory and visual sensory cortices. *Eur. J. Neurosci.* 31, 1772–1782. doi: 10.1111/j.1460-9568.2010.07213.x
- Romei, V., Murray, M. M., Cappe, C., and Thut, G. (2009). Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Curr. Biol.* 19, 1799–1805. doi: 10.1016/j.cub.2009.09.027
- Romei, V., Murray, M. M., Cappe, C., and Thut, G. (2013). The contributions of sensory dominance and attentional bias to cross-modal enhancement of visual cortex excitability. *J. Cogn. Neurosci.* 25, 1122–1135. doi: 10.1162/jocn_a_00367
- Romei, V., Murray, M. M., Merabet, L. B., and Thut, G. (2007). Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: implications for multisensory interactions. *J. Neurosci.* 27, 11465–11472. doi: 10.1523/JNEUROSCI.2827-07.2007
- Rowland, B. A., and Stein, B. E. (2007). Multisensory integration produces an initial response enhancement. *Front. Integr. Neurosci.* 1:4. doi: 10.3389/neuro.07.004.2007
- Rowland, B. A., and Stein, B. E. (2014). A model of the temporal dynamics of multisensory enhancement. *Neurosci. Biobeh. Rev.* 41, 78–84. doi: 10.1016/j.neubiorev.2013.12.003
- Sarko, D. K., Nidiffer, A. R., Powers, I. I. I., A. R., Ghose, D. and Wallace, M. T. (2012). “Spatial and temporal features of multisensory processes,” in *The Neural Basis of Multisensory Processes*, ed M. M. Murray and M. T. Wallace (Boca Raton, FL: CRC Press), 191–215.
- Soto-Faraco, S., Navarra, J., and Alsius, A. (2004). Assessing automaticity in audiovisual speech integration: evidence from the speeded classification task. *Cognition* 92, B13–B23. doi: 10.1016/j.cognition.2003.10.005
- Spierer, L., Manuel, A. L., Bueti, D., and Murray, M. M. (2013). Contributions of pitch and bandwidth to sound-induced enhancement of visual cortex excitability in humans. *Cortex* 49, 2728–2734. doi: 10.1016/j.cortex.2013.01.001
- Stein, B. E., Burr, D., Constantinidis, C., Laurienti, P. J., Alex Meredith, M., Perrault, T. J., et al. (2010). Semantic confusion regarding the development of multisensory integration: a practical solution. *Eur. J. Neurosci.* 31, 1713–1720. doi: 10.1111/j.1460-9568.2010.07206.x
- Stekelenburg, J. J., Maes, J. P., Van Gool, A. R., Sitskoorn, M., and Vroomen, J. (2013). Deficient multisensory integration in schizophrenia: an event-related potential study. *Schizophr. Res.* 147, 253–261. doi: 10.1016/j.schres.2013.04.038
- Stekelenburg, J. J., and Vroomen, J. (2012). Electrophysiological correlates of predictive coding of auditory location in the perception of

- natural audiovisual events. *Front. Integr. Neurosci.* 6:26. doi: 10.3389/fnint.2012.00026
- Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., et al. (2014). Identifying and quantifying multisensory integration: a tutorial review. *Brain Topogr.* 27, 707–730. doi: 10.1007/s10548-014-0365-7
- Summerfield, C., and Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends Cogn. Sci.* 13, 403–409. doi: 10.1016/j.tics.2009.06.003
- Talsma, D., Doty, T. J., and Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cereb. Cortex* 17, 679–690. doi: 10.1093/cercor/bhk016
- Talsma, D., and Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *J. Cogn. Neurosci.* 17, 1098–1114. doi: 10.1162/0898929054475172
- Teder-Sälejärvi, W. A., McDonald, J. J., Di Russo, F., and Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Cogn. Brain Res.* 14, 106–114. doi: 10.1016/S0926-6410(02)00065-4
- Thelen, A., Cappe, C., and Murray, M. M. (2012). Electrical neuroimaging of memory discrimination based on single-trial multisensory learning. *NeuroImage* 62, 1478–1488. doi: 10.1016/j.neuroimage.2012.05.027
- Thelen, A., Matusz, P. J., and Murray, M. M. (2014). Multisensory context portends object memory. *Curr. Biol.* 24, R734–R735. doi: 10.1016/j.cub.2014.06.040
- Thut, G., Miniussi, C., and Gross, J. (2012). The functional importance of rhythmic activity in the brain. *Curr. Biol.* 22, R658–R663. doi: 10.1016/j.cub.2012.06.061
- Tiippana, K., Andersen, T. S., and Sams, M. (2004). Visual attention modulates audiovisual speech perception. *Eur. J. Cogn. Psychol.* 16, 457–472. doi: 10.1080/09541440340000268
- van Atteveldt, N., Murray, M. M., Thut, G., and Schroeder, C. E. (2014). Multisensory integration: flexible use of general operations. *Neuron* 81, 1240–1253. doi: 10.1016/j.neuron.2014.02.044
- van der Burg, E., Olivers, C. N., Bronkhorst, A. W., and Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1053–1065. doi: 10.1037/0096-1523.34.5.1053
- van der Burg, E., Talsma, D., Olivers, C. N., Hickey, C., and Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *Neuroimage* 55, 1208–1218. doi: 10.1016/j.neuroimage.2010.12.068
- Vidal, J., Giard, M. H., Roux, S., Barthelemy, C., and Bruneau, N. (2008). Cross-modal processing of auditory-visual stimuli in a no-task paradigm: a topographic event-related potential study. *Clin. Neurophysiol.* 119, 763–771. doi: 10.1016/j.clinph.2007.11.178
- Wang, Y., Celebrini, S., Trotter, Y., and Barone, P. (2008). Visuo-auditory interactions in the primary visual cortex of the behaving monkey: electrophysiological evidence. *BMC Neurosci.* 9:79. doi: 10.1186/1471-2202-9-79
- Yuval-Greenberg, S., and Deouell, L. Y. (2007). What you see is not (always) what you hear: Induced gamma band responses reflect cross-modal interactions in familiar object recognition. *J. Neurosci.* 27, 1090–1096. doi: 10.1523/JNEUROSCI.4828-06.2007

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 October 2014; accepted: 13 February 2015; published online: 03 March 2015.

Citation: De Meo R, Murray MM, Clarke S and Matusz PJ (2015) Top-down control and early multisensory processes: chicken vs. egg. *Front. Integr. Neurosci.* 9:17. doi: 10.3389/fnint.2015.00017

This article was submitted to the journal *Frontiers in Integrative Neuroscience*.

Copyright © 2015 De Meo, Murray, Clarke and Matusz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.