20 Influences of Top-Down Attention on Multisensory Processing

DURK TALSMA, DANIEL SENKOWSKI, SALVADOR SOTO-FARACO, AND MARTY G. WOLDORFF

Traditionally, multisensory integration has been considered to be a preattentive, automatic process (e.g., de Gelder & Bertelson, 2003). Moreover, multisensory processing was also considered to adhere closely to well-established rules, with the temporal and spatial proximity of sensory inputs from different modalities being major determinants for integration to occur (e.g., Stein & Meredith, 1993; Welch & Warren, 1986). In addition, the law of inverse effectiveness seemed to straightforwardly predict the strength of integration, positing that when the responsiveness to individual sensory stimuli decreases, multisensory integration is stronger (Holmes, 2009). Further, in part because of the putatively general adherence of multisensory integration to such rules, it has been argued that such integration is essentially immune to cognitive modulation and independent of the availability of processing resources.

Despite this original interpretation of multisensory integration as being largely preattentive, there are several lines of reasoning to expect that top-down attention could indeed affect such integration processes, along with a number of recent studies providing direct evidence for such influences. First, multisensory integration and attention serve similar purposes, in particular to enhance perceptual clarity and reduce stimulus ambiguity. Based on these similarities in function, we and others have previously hypothesized that multisensory integration and attention would be likely to interact (e.g., Alsius, Navarra, Campbell, & Soto-Faraco, 2005; Senkowski, Talsma, Herrmann, & Woldorff, 2005; Soto-Faraco & Alsius, 2009; Talsma & Woldorff, 2005). Second, the existence of cross-modal spatial attention effects (e.g., directing visual attention to a location in space improves the processing of stimuli in other modalities at that location, even when these stimuli are completely task-irrelevant) demonstrates that attention can be oriented in parallel across multiple modalities in a supramodal fashion (Eimer & van Velzen, 2002; Weissman, Warner, & Woldorff, 2004). Third, it has been argued that attention is facilitatory, if not crucial, for top-down feature-binding processes, either within a modality (Treisman & Gelade, 1980) or across modalities (Cinel, Poli, & Humphreys, 2002; Treisman & Davies, 1973). And, finally, as described below, a number of recent studies have provided strong evidence for the robust influence of attention on multisensory integration processes.

Following a brief introduction to the most relevant attentional mechanisms, this chapter provides an overview of some recent studies indicating that top-down voluntary attention can affect multisensory processing. Findings from these studies will also be contrasted with other work that has indicated that multisensory integration can occur preattentatively.

ATTENTIONAL MECHANISMS

Attention is an essential neurobiological function that allows us to continually and dynamically select the most important and/or salient stimuli or events in our environment, both external and internal, so that greater neural resources can be devoted to their processing. Attentional selection mechanisms are often broadly divided into two main categories. First, top-down control refers to a voluntary mode of orienting attention toward behaviorally relevant stimuli, objects, events, or locations on the basis of internal states such as goals, motivation, or expectations (Serences & Boynton, 2007; Theeuwes, Atchley, & Kramer, 2000; Yantis & Jonides, 1984). Top-down attention is thought to be goal-directed when attentional priority is given to events and objects that are aligned with the observer's goals (Theeuwes et al., 2000). Moreover, top-down attention can be directly allocated to specific locations in the external world; in this case attention is said to be oriented in a spatially selective fashion (e.g., Corbetta & Shulman, 2002; Hillyard, Woldorff, Mangun, & Hansen, 1987; Luck, Woodman, & Vogel, 2000), thereby giving priority to the processing of stimuli presented at certain (attended) locations over stimuli occurring at other locations. Second, bottom-up, stimulus-driven attentional control

Q

refers to a much more automatic mechanism of attention in which salient stimuli and events in the environment capture processing resources based on stimulus properties such as saliency, even when they are irrelevant to, or even counter to, current goals or expectations (Theeuwes, 1991; Yantis & Jonides, 1984). Here, we argue that top-down attention can impart major influence on multisensory processing, and we review several important mechanisms by which this can occur.

A major topic in attention research, especially from a multisensory perspective, concerns the question of how attention is allocated and coordinated across sensory modalities. Considerable overlap can be found in frontal and parietal brain areas responsible for the top-down orienting of visual and auditory attention (cf. Woldorff et al., 2004; Wu, Weissman, Roberts, & Woldorff, 2007 in closely parallel paradigms). In addition, attending to a specific location in one modality also enhances the processing of stimuli from another sensory modality presented at that same location, and, conversely, voluntarily directing attention to different locations in each modality tends to lead to costs in task performance (Eimer & Driver, 2001; McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2003; Spence & Driver, 1994, 1996; Talsma & Kok, 2002). Overall, these cross-modal attentional findings make it unlikely that there are completely independent attentional control mechanisms for the different modalities. Rather, they suggest that the sensory modalities operate synergistically in orienting to new locations in space. The tendency for such a synergy supports the view that attention, particularly spatial attention, can operate in a supramodal fashion.

It has been proposed that top-down spatial selective attention operates by increasing the sensitivity of neurons responsive to the attended stimulus feature or location (Hillyard, Vogel, & Luck, 1998; Khayat, Spekreijse, & Roelfsema, 2004). Presumably during supramodal attention, the biasing signals that are sent from the frontoparietal attentional-control network are projected to several of the unisensory cortices in parallel. Such a mechanism would make it possible for attention to be oriented in parallel across modalities. In contrast, it has been proposed that attentional resolving—that is, the processing of relevant information within each modality—may be carried out in a more independent manner within each modality (Klemen, Büchel, & Rose, 2009; Talsma, Doty, Strowd, & Woldorff, 2006).

INFLUENCES OF TOP-DOWN ATTENTION ON MULTISENSORY INTEGRATION

The number of studies showing that top-down attention can affect multisensory integration has been steadily increasing. In particular, evidence is now available to show that attention can affect multisensory processes at various stages, from relatively early perceptual stages to later, more complex, perception-related stages involving the analysis of audiovisual congruency or related to audiovisual illusions.

Spatial and Nonspatial Selective Attention

A relatively straightforward demonstration of attentional influence on multisensory integration processes is provided by an EEG experiment in which participants covertly attended to one of two possible spatial locations (Talsma & Woldorff, 2005). The responses evoked by a multisensory stimulus were compared with the sum of the responses evoked by its unisensory components separately (and correcting analytically for any possible differential overlap in such a contrast). By contrasting the evoked responses at attended and unattended locations Talsma and Woldorff were able to investigate both the effects of multisensory integration on stimulus processing and how such effects are influenced by spatial attention. The results of this manipulation showed that the neural responses associated with multisensory integration, including ones occurring at relatively early latencies, were larger at the attended location than at the unattended one (see figure 20.1 [plate 26]). Although these results show that multisensory integration can be affected by spatial selective attention, the onset of the earliest multisensory integration effects occurred at somewhat longer latencies (>90 msec) than some previously reported under conditions in which stimuli were presented centrally (Giard & Peronnet, 1999; Molholm et al., 2002). To investigate the possibility that the peripheral stimulation in the Talsma and Woldorff (2005) study resulted in a relatively low sensitivity of ERPs for detecting the early multisensory integration effects, a follow-up study was devised in which all the stimuli were presented centrally (Talsma, Doty, & Woldorff, 2007). Attention was manipulated by requiring participants to perform a secondary digit detection task. In this study, an early-latency frontocentral deflection in the ERP (i.e., on the P50 component) was observed, apparently reflecting early multisensory interaction processes. Interestingly, although this effect was observed in response to multisensory stimuli, it was significant only when attention was directed to both the visual and auditory parts of the multisensory stimulus. This latter result is consistent with recent findings of Wu et al. (2009), who reported that multisensory integration effects occurred relatively late when participants were attending to the visual modality only.

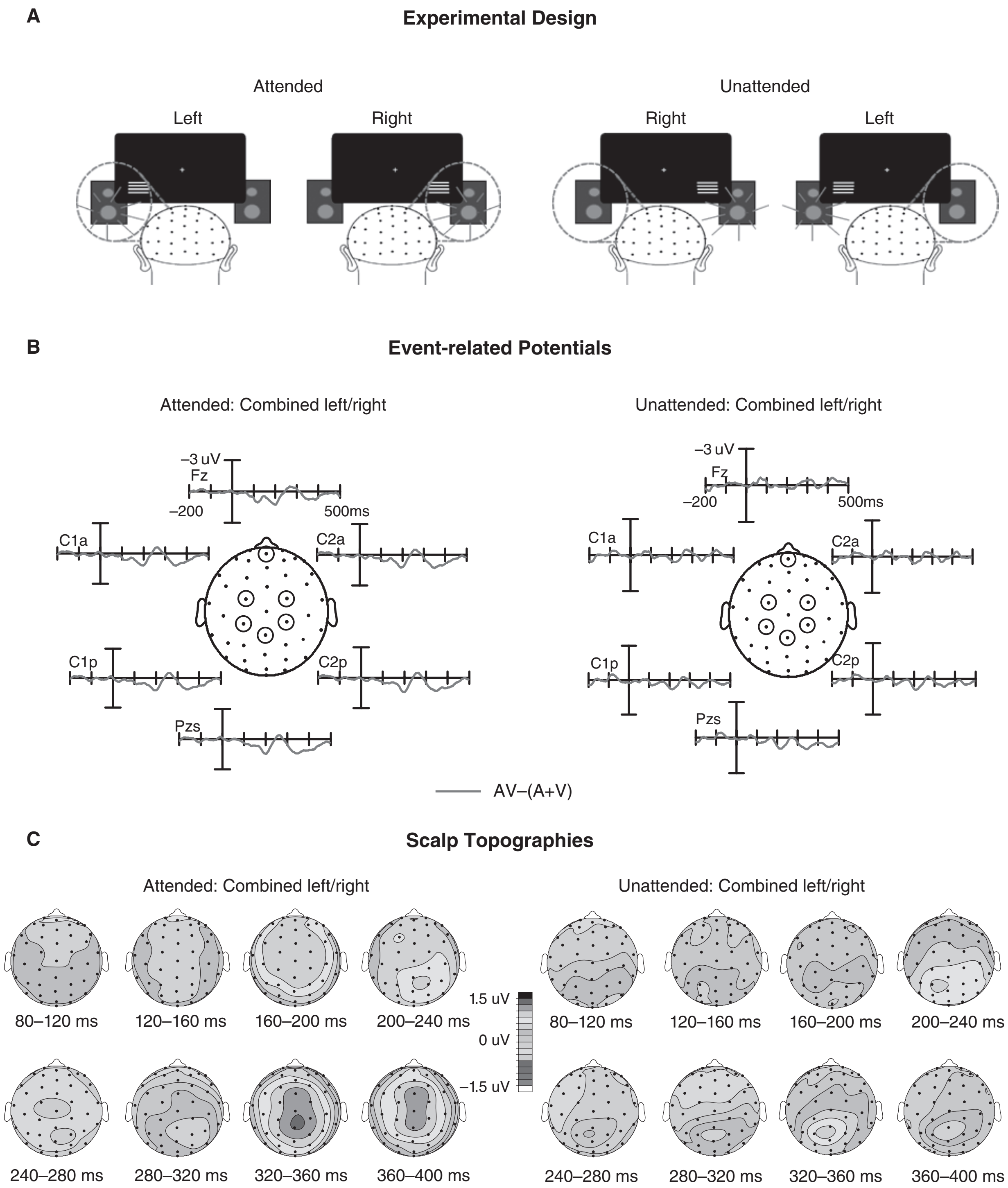

FIGURE 20.1 (PLATE 26) Effects of spatial selective attention on multisensory integration. (A) Experimental design. During a block of trials, participants were selectively attending to stimuli presented in one hemifield while ignoring stimuli presented in the opposite hemifield. Stimuli consisted of visual (square-wave gratings), auditory (tone-pips), and audiovisual stimuli, presented in rapid and random succession. Audiovisual stimuli consisted of the simultaneous presentation of the visual and auditory stimuli, and data were subsequently collapsed across the left and right hemifields. (B) Multisensory integration effects as reflected in an ERP difference wave that was obtained by subtracting the summated ERP response to the auditory (A) and visual (V) stimuli from the ERP response to the audiovisual (AV) stimuli. Note that the effects represented in the resulting [AV – (A + V)] difference wave are more pronounced when the stimuli are attended (left), as compared to when they are unattended (right). (C) Scalp topographies of the effects shown in B. (Redrawn from Talsma & Woldorff, 2005.)

Q

Multisensory Congruency

Another example of the top-down influence of attention on multisensory processing is provided by a functional magnetic resonance imaging (fMRI) study by Fairhall and Macaluso (2009). In this study a speech fragment was played acoustically from a central location, together with two side by side video-clips of visual speech (lip-moments). One of the visual streams was congruent with the auditory speech signals while the other stream was incongruent. Spatial attention was covertly oriented to either the left or the right visual stream while fixation was kept in the center. As a main result, the authors found increased activity in several brain areas, including in the superior temporal sulcus, the striate and extrastriate retinotopic visual cortices, and the superior colliculus, when visual spatial attention was directed to lip movements that were congruent with the auditory speech signals compared to when attention was directed to the simultaneously presented incongruent lip movements. These results can be taken to indicate that attention can not only boost the processing of stimuli presented at specific locations but may also aid in resolving conflicts in audiovisual stimulus congruency.

A similar conclusion concerning congruency was reached by Senkowski et al. (2008), who used EEG data to show that allocating visual attention toward irrelevant lip movements in the visual scene interferes with the recognition of audiovisual speech signals from an attended speaker (figure 20.2). In this study participants were presented with a continuous stream of audiovisual speech stimuli in which either a center speaker alone produced a syllable (i.e., no-interference trials) or in which three horizontally aligned speakers (including the center speaker) simultaneously produced syllables (i.e., interference trials). The participant's task was to detect a target syllable by the center speaker, while ignoring syllables from two flanking speakers. The faces of the center speaker and the flanking speakers (figure 20.2) were amplitude modulated at different visual stimulation frequencies. This enabled the steady-state visual evoked potentials (SSVEPs) in response to the center speaker and the flanking speakers to be separately extracted (figure 20.2C), providing a real-time index of visuospatial attention toward the three speakers. A main finding of the study was that highly distracted participants (reflected by longer reaction times in interference than in no-interference trials) showed larger SSVEPs during omitted trial periods in response to flanking speakers in comparison to participants displaying little distraction. Similarly, the amplitude of the SSVEPs in response to the flanking speakers was positively correlated with the participant's distractibility during speech processing (figure 20.2D). This suggests that the allocation of spatial attention to task-irrelevant sensory input interferes with the recognition of task-relevant multisensory speech signals under noisy environmental conditions.

In another study using a binocular rivalry paradigm consisting of a visual stimulus with looming motion to one eye and radial motion to the other, van Ee and colleagues (2009) demonstrated that participants were able to hold on to one of the two percepts longer by means of attention. Moreover, this attentional gain for one of the percepts was prolonged when the attended visual stimulus was accompanied by a sound that matched the temporal characteristics of the attended visual stimulus. This pattern of results thus also suggests an intimate interactional relationship between attention and multisensory integration. Although the exact neural mechanisms involved in this process are not yet clear, the findings suggest that attention may boost the neural response to one of the competing visual signals and that this boost, in turn, facilitates integration with the matching auditory signal. Interestingly, van Ee et al. (2009) also demonstrated that the mere presence of such a matched sound was insufficient. Additional attention to the auditory modality was needed to facilitate the effect that the congruent sound could have on the attentional facilitation of one of the two visual percepts.

Spreading of Attention

A third type of cross-modal process involving attention and multisensory integration has been described as a "spread of attention" between modalities. As reported by Busse et al. (2005), it was shown that a task-irrelevant auditory stimulus elicited a different electrophysiological brain response when it was paired in time with an attended versus an unattended visual stimulus, even when the visual stimulus arose from a completely different spatial location. This difference appeared as a prolonged negative-polarity deflection over frontocentral scalp areas, beginning at ~200 msec poststimulus and lasting for hundreds of milliseconds, as well as being reflected by a corresponding enhancement of fMRI activity in auditory cortex. The prolonged negative-polarity electrophysiological effect, variants of which have also been observed in several other studies (Donohue, Roberts, Grent-'t-Jong, & Woldorff, 2011; Fiebelkorn, Foxe, & Molholm, 2010; Talsma et al., 2007), resembles an activation known as the *late processing negativity* observed during unisensory auditory selective attention (Näätänen, 1982). This resemblance is particularly intriguing in that it is induced by visuospa-

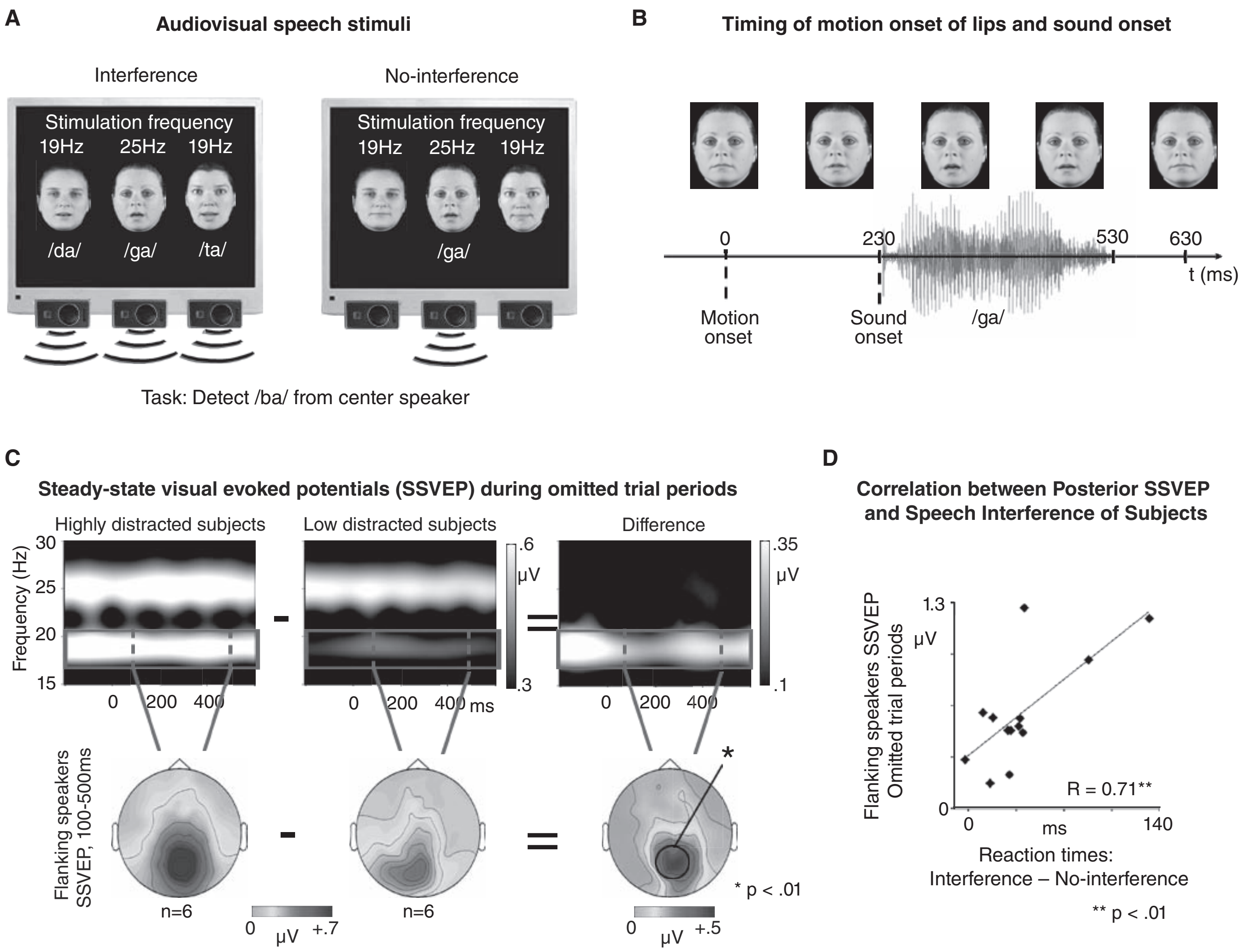

FIGURE 20.2 Visual processing of flanking speakers interferes with the recognition of multisensory speech signals under noisy environmental conditions. (A) Participants were presented with a continuous stream of audiovisual speech stimuli in which either syllables were produced by a single center speaker (no interference trials; left panel) or three syllables were simultaneously produced by the three speakers (interference trials; right panel). The participant's task was to detect the target syllable /ba/ by the center speaker, while ignoring syllables from two flanking speakers. Faces of the center speaker and the flanking speakers were presented at different visual stimulation frequencies, i.e., 25 Hz and 19 Hz, respectively, which enabled the separate analysis of steady-state evoked potentials (SSVEP) in response to the center and the flanking speakers. (B) The motion onset of the lips preceded the sound onset of the syllables on average by 230 msec. (C) The upper panel shows time-frequency representations of a channel located over central posterior scalp during omitted trial periods (periods without speech stimulation but with the presentation of flickering faces) for participants who were highly distracted in speech interference trials, i.e., has longer reaction times (RTs) compared to participants who were less distracted in these trials. Highly distracted participants showed larger SSVEP in response to flanking speakers over occipital scalp regions in comparison to less-distracted participants (lower panel). (D) The SSVEP in response to flanking speakers during omitted trial periods correlated positively with the distractibility of participants during speech processing (i.e., RTs of interference minus RTs of no-interference trials). (Redrawn from Senkowski et al., 2008.)

tial selective attention. Moreover, this cross-modal attentional spread process can be interpreted as a systems-level cascade of interactions between top-down and bottom-up influences. Specifically, it appears to begin with selectively focused top-down visuospatial attention that determines which stimulus in the visual modality is preferentially processed. Then, presumably by means of an automatic (bottom-up) binding mechanism derived from the temporal coincidence of the multisensory stimulus components, attention spreads across modalities to encompass the auditory component despite its being completely task-irrelevant and not even in the same spatial location. In other words, it appears that choosing to attend to a specific aspect of the visual world can determine saliency in the auditory world.

Q

Another set of recent studies used both EEG (Zimmer, Itthipanyanan, Grent-'t-Jong, & Woldorff, 2010) and fMRI (Zimmer, Roberts, Harshburger, & Woldorff, 2010) with audiovisual letter combinations to study how the spreading of attention across modalities and space (as in Busse et al., 2005) would vary as a function of the congruency of the unisensory stimulus components. These studies found enhanced visual-to-auditory spreading of attention activity associated with the auditory cortex when the task-irrelevant auditory stimulus was incongruent relative to when it was congruent, consistent with a greater distractibility by the auditory stimuli when they conflicted with the task-relevant visual stimulus at a representational level. The incongruent stimulus trials were also associated with enhanced fMRI activity in the anterior cingulate cortex (ACC) and in the visual cortices, consistent with a detection of the incongruency by the ACC followed by an attempt to focus more attention on the visual stimulus in the face of the presence of conflicting auditory input. Similar to the findings on audiovisual incongruency, as discussed above, these results underscore the close interactive relationship between attention and multisensory integration processes.

Audiovisual Illusions

Further evidence for the impact of top-down attention on multisensory processing derives from studies using audiovisual illusions (Alsius et al., 2005; Alsius, Navarra, & Soto-Faraco, 2007; Mishra, Martinez, & Hillyard, 2010). For instance, during the well-known McGurk effect (McGurk & MacDonald, 1976), an auditory phoneme dubbed onto incongruent visual lip movements can lead to an illusory auditory percept. However, endogenously directing attention away from the stimulus (e.g., to a demanding concurrent task) reduces susceptibility to the illusion (Alsius et al., 2005; 2007). Interestingly, Alsius et al. also reported that this reduced susceptibility to the McGurk effect (i.e., reduced visual influence on phoneme perception) occurred above and beyond any unisensory attention effects, and regardless of the sensory modality (visual, auditory, or tactile) in which the concurrent task was performed (Alsius et al., 2005; Alsius et al., 2007) (see figure 20.3). This suggests that the reduction of the illusory effect was being accomplished by the pulling away of resources from the multisensory integration processing mechanisms.

Another example of the interaction of top-down attention and audiovisual illusions concerns the sound-induced double-flash illusion, in which two sounds in quick succession can induce the perception of a double flicker of a concurrent continuous visual stimulus (Shams, Kamitani, & Shimojo, 2000). More specifically, the perception of this effect has been found to be larger when spatial attention was directed to, versus away from, the stimuli eliciting the illusion (Mishra et al., 2010). Thus, these various findings provide further evidence that directed top-down attention can influence, and/or interact with, a broad range of multisensory perceptual processes.

IS ATTENTION NECESSARY FOR MULTISENSORY INTEGRATION?

Although the studies reviewed above have shown that attention can robustly influence multisensory integration processes, it should be noted that this does not imply that attention is necessary for multisensory integration. For instance, van der Burg et al. (2008) used a difficult-to-detect visual target-detection task (target stimuli were presented among an array of similar distractors) and found that search times increased with an increasing number of display items, which suggested that participants were using an item-by-item serial search strategy (Wolfe, 2003) in this visual task. However, when a transient sound was presented together with an irrelevant color change of the target stimulus, search times became much shorter, which occurred relatively independent of the number of distractor items, providing evidence that the concurrent auditory event helped make the visual target pop out from the background of distractors (i.e., target selection was no longer based on a serial attentional process). Furthermore, it was found that visual distractor stimuli could also capture attention when they were synchronized with an auditory or tactile accessory stimulus, as reflected by a decreased ability to detect competing visual targets (van der Burg et al., 2008; van der Burg, Olivers, Bronkhorst, & Theeuwes, 2009). These results suggest that multisensory integration can take the form of a highly stimulus-driven, relatively automated process that can increase the saliency of visual stimuli sufficiently to make them quickly detectable (Noesselt, Bergmann, Hake, Heinze, & Fendrich, 2008). In addition, this also indicates that the resulting saliency increase can then capture or otherwise modulate attentional orienting processes. It should be noted, however, that some studies (e.g., Fujisaki, Koene, Arnold, Johnston, & Nishida, 2006; Alsius & Soto-Faraco, 2011) using similar search paradigms did not find that a synchronously presented sound was sufficient to make a visual stimulus pop out from the background. Presumably, this is because these studies used auditory stimuli that were not particularly

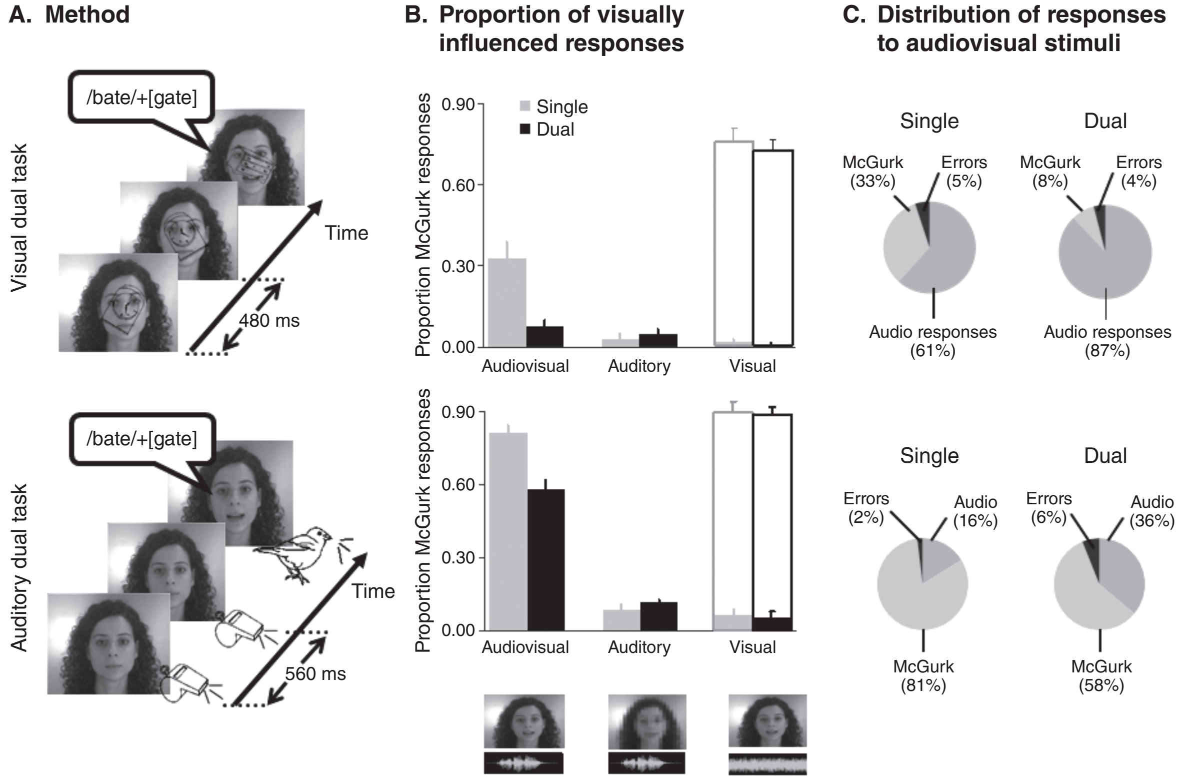

FIGURE 20.3 Diverting attention from audiovisual speech reduces the prevalence of the McGurk illusion. (A) The dual-task paradigms used in the visual (top) and auditory (bottom) attention tasks. In both cases, concurrent stimuli (drawings or sounds of everyday-life objects, respectively) were superimposed on the original video clips. Participants were asked to simply report the words occasionally uttered by the speaker (single task) or, in addition, to detect stimulus repetition in the stream of objects (dual task). (B) Proportion of visually influenced responses under single (pale gray) and dual (darker gray) task when participants were presented with the audiovisual stimuli, with the sound track alone, or with the silent video clips. The empty bars in the visual-alone condition correspond to scores based on the words' phonemic equivalence classes (Mattys, Bernstein, & Auer, 2002) so that visual performance could gauged while avoiding floor effects. (C) Distribution of visually influenced (McGurk), auditory, or other types (errors) of responses in the audiovisual conditions under each task. Note that the reduction in McGurk responses in dual-task conditions is concomitant with an increase in auditory responses, not in erroneous responses. This was true even when the concurrent dual task was being made in the auditory modality. (Redrawn from Alsius et al., 2005.)

salient, which might have rendered them incapable of triggering a strongly automatic multisensory integration process.

To summarize, the literature reviewed above shows that top-down attention can influence multisensory integration processes in a variety of ways and at different levels in the processing hierarchy. Even though there is ample evidence that multisensory integration can also occur in a preattentive fashion, recent literature indicates that attention can robustly influence multisensory integration in ways that can sometimes be facilitatory and sometimes inhibitory. Attentional facilitation of multisensory integration appears to occur specifically when either stimulus delivery rates are high (Talsma et al., 2007; Talsma & Woldorff, 2005) or there is conflict between congruent and incongruent streams of stimuli across modalities (Talsma, Senkowski, Soto-Faraco, & Woldorff , 2010). Along similar lines, when attention is directed away from salient stimuli, such as with audiovisual speech stimuli, it can reduce effective integration of these stimuli (Alsius et al., 2005; Alsius et al., 2007; Alsius & Soto-Faraco, 2011).

CONTROVERSIAL ISSUES

Mixed Empirical Findings

The disparity between the more recent results described above and those of earlier studies that failed to find a top-down effect of attention (Bertelson, Vroomen, de Gelder, & Driver, 2000; Colin et al. 2002) is still a highly debated issue. This discrepancy between studies that do not find an influence of attention on multisensory integration and those that do seems likely to be due to the former studies having used a relatively low perceptual load, which may have left enough processing capacity to still effectively integrate all or most of the multisensory stimuli (see Navarra, Alsius, Soto-Faraco, & Spence, 2010, for a discussion). By contrast, under high perceptual demands, including those involving a high degree of conflict between relevant and irrelevant stimuli, the perceptual integration system would tend to become overloaded unless the sensory inputs with higher priority are allocated sufficient attentional capacity for accomplishing effective and useful integration (Alsius et al., 2007; Talsma et al., 2007). To counter this claim, it has been argued that audiovisual cues remain effective in capturing spatial attention in a bottom-up, stimulus-driven way, even under high-load conditions (Santangelo & Spence, 2007). In this latter study, however, the cues by themselves were still relatively rare and therefore perhaps salient enough to induce robust multisensory interactions preattentively, thereby leading to capture of attention in a bottom-up fashion, even under high-load conditions. Taken together, there is currently substantial evidence for the notion that attention can influence multisensory integration processing, particularly when perceptual demands are high. Conversely, however, sufficiently salient bottom-up influences can nevertheless still compete and interact with these top-down factors.

Unisensory or Multisensory Effects?

One important point of debate when assessing the influence of attention on multisensory integration is whether the attentional modulation is actually affecting the multisensory integration process per se or whether the effect is at a single-modality level and then carries over to subsequent, multisensory, processing stages. Not many studies have attempted to differentiate between these two accounts despite the important consequences in terms of interpretation. Alsius and colleagues (2005, 2007) included a measurement of single-modality processing, thereby enabling measurement of the costs of diverting attention in multisensory processing, over and above any costs in unisensory processing. More specifically, the costs of diverting attention were measured in terms of prevalence of illusory (McGurk effect) percepts as well as in terms of visual-alone and auditory-alone recognition. These analyses indicated that the effect of attention on unisensory perception (for each modality separately) was weaker than the effect of attention on multisensory integration. This was true even with compensation for the naturally reduced performance in visual-alone word recognition (see figure 20.3 for more detail).

FUTURE DIRECTIONS

Delineating the exact conditions that determine how attention and integration affect one another will likely constitute one of the major future challenges in this field. Recent studies have provided evidence for at least two distinct stages in multisensory integration (Magnée, de Gelder, van Engeland, & Kemner, 2008a, 2008b; Talsma et al., 2007; Talsma, Senkowski, & Woldorff, 2009). For instance, we recently reported a distinction between an early (semi)automatic multisensory analysis process affecting the early-latency ERPs such as the visual P1 component and later stages of higher-level processing, as reflected in a longer-latency occipital selection negativity in the ERP waveform (Talsma et al., 2009). A similar distinction of levels of multisensory processing was identified by Magnée et al. (2008a, 2008b), who demonstrated that in a group of individuals diagnosed with pervasive developmental disorders, the early stages of multisensory integration were still intact, whereas the later stages of multisensory integration were impaired. Because disrupted multisensory processing is considered to be a symptom of various clinical populations, including autism spectrum disorder (Iarocci & McDonald, 2006) and schizophrenia (Ross et al., 2007), a closer understanding of the relation between attention and multisensory integration can be a viable tool to study the interactions between cognitive function and multisensory integration processing in these patient groups.

Additionally, a limited number of studies are now reporting that older adults are benefiting more from combining perceptual cues across sensory modalities than younger adults do (Peiffer, Mozolic, Hugenschmidt, & Laurienti, 2007). This observation is interesting because other research has suggested that older adults typically have weaker top-down attentional control (Wecker, Kramer, Wisniewski, Delis, & Kaplan, 2000; West, 1996; but see Madden et al., 2007) and are more easily distracted (Machado, Devine, & Wyatt, 2009), including by stimuli in different modalities (Campbell, Al-Aidroos, Fatt, Pratt, & Hasher, 2010; Talsma, Kok, & Ridderinkhof, 2006). This last observation suggests that older adults may receive greater benefits from congruent auditory and visual information but are more susceptible to interference from incongruent audiovisual information. This hypothesis leads to the question as to what degree the aforementioned interactions between attention and multisensory integration change across the life span. Addressing these issues of how the interactions between attention and multisensory integration differ in psychiatric populations as well as how they change in normal development and aging seems likely to be an important direction to be explored in the next decade.

ACKNOWLEDGMENTS

The effort for this work was supported by funding from the Institute for Behavioral Research at the University of Twente to D.T., from the German Research Foundation (SE 1859/1–2) and the EU (ERC-2010-StG_20091209) to D.S., by grants SEJ2007–64103/PSIC—CDS2007–00012 from MICINN (Spanish government) and ERC-StG 2010-263145 to S.S.-F., and U.S. NIH grant NS-R01–051048 to M.G.W.

REFERENCES

Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, *15*, 839–843.

Alsius, A., Navarra, J., & Soto-Faraco, S. (2007). Attention to touch weakens audiovisual speech integration. *Experimental Brain Research*, *183*, 399–404.

Alsius, A., & Soto-Faraco, S. (2011). Searching for audiovisual correspondence in multiple speaker scenarios. *Experimental Brain Research*. doi:10.1007/s00221-011-2624-0.

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, *62*, 321–332.

Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 18751–18756.

Campbell, K. L., Al-Aidroos, N., Fatt, R., Pratt, J., & Hasher, L. (2010). The effects of multisensory targets on saccadic trajectory deviations: eliminating age differences. *Experimental Brain Research*, *201*, 385–392.

Cinel, C., Poli, R., & Humphreys, G. W. (2002). Cross-modal illusory conjunctions between vision and touch. *Journal of Experimental Psychology. Human Perception and Performance*, *28*, 1243–1266.

Colin, C., Radeau, M., Soquet, A., Demolin, D., Colin, F., & Deltenre, P. (2002). Mismatch negativity evoked by the McGurk-MacDonald effect: a phonetic representation within short-term memory. *Clinical Neurophysiology*, *113*, 495–506.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*, 201–215.

de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, *7*, 460–467.

Donohue, S. E., Roberts, K. C., Grent-'t-Jong, T., & Woldorff, M. G. (2011). The cross-modal spread of attention reveals different constraints for the temporal and spatial linking of visual and auditory stimulus events. *Journal of Neuroscience*, *31*, 7982–7990.

Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: Evidence from event-related brain potential studies. *Neuroscience and Biobehavioral Reviews*, *25*, 497–511.

Eimer, M., & van Velzen, J. (2002). Crossmodal links in spatial attention are mediated by supramodal control processes: Evidence from event-related potentials. *Psychophysiology*, *39*, 437–449.

Fairhall, S. L., & Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. *European Journal of Neuroscience*, *29*, 1247–1257.

Fiebelkorn, I. C., Foxe, J. J., & Molholm, S. (2010). Dual mechanisms for the cross-sensory spread of attention: how much do learned associations matter? *Cerebral Cortex*, *20*, 109–120.

Fujisaki, W., Koene, A., Arnold, D., Johnston, A., & Nishida, S. (2006). Visual search for a target changing in synchrony with an auditory signal. *Proceedings. Biological Sciences*, *273*, 865–874.

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.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society B. Biological Sciences*, *353*, 1257–1270.

Hillyard, S. A., Woldorff, M., Mangun, G. R., & Hansen, J. C. (1987). Mechanisms of early selective attention in auditory and visual modalities. *Electroencephalography and Clinical Neurophysiology. Supplement*, *39*, 317–324.

Holmes, N. P. (2009). The principle of inverse effectiveness in multisensory integration: some statistical considerations. *Brain Topography*, *21*, 168–176.

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

Khayat, P. S., Spekreijse, H., & Roelfsema, P. R. (2004). Visual information transfer across eye movements in the monkey. *Vision Research*, *44*, 2901–2917.

Klemen, J., Büchel, C., & Rose, M. (2009). Perceptual load interacts with stimulus processing across sensory modalities. *European Journal of Neuroscience*, *29*, 2426–2434.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, *4*, 432–440.

Machado, L., Devine, A., & Wyatt, N. (2009). Distractibility with advancing age and Parkinson's disease. *Neuropsychologia*, *47*, 1756–1764.

Madden, D. J., Spaniol, J., Whiting, W. L., Bucur, B., Provenzale, J. M., Cabeza, R., et al. (2007). Adult age differences in the functional neuroanatomy of visual attention: a combined fMRI and DTI study. *Neurobiology of Aging*, *28*, 459–476.

Magnée, M. J. C. M., de Gelder, B., van Engeland, H., & Kemner, C. (2008a). Atypical processing of fearful face-voice pairs in pervasive developmental disorder: an ERP study. *Clinical Neurophysiology*, *119*, 2004–2010.

Magnée, M. J. C. M., de Gelder, B., van Engeland, H., & Kemner, C. (2008b). Audiovisual speech integration in pervasive developmental disorder: Evidence from event-related potentials. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *49*, 995–1000.

Q

Mattys, S., Bernstein, L. E., & Auer, E. T., Jr. (2002). Stimulus-based lexical distinctiveness as a general word recognition mechanism. *Perception & Psychophysics, 64*, 667–679.
McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience, 15*, 10–19.
McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746–748.
Mishra, J., Martinez, A., & Hillyard, S. A. (2010). Effect of attention on early cortical processes associated with the sound-induced extra flash illusion. *Journal of Cognitive Neuroscience, 22*, 1714–1729.
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.
Näätänen, R. (1982). Processing negativity: an evoked-potential reflection. *Psychological Bulletin, 92*, 605–640.
Navarra, J., Alsius, A., Soto-Faraco, S., & Spence, C. (2010). Assessing the role of attention in the audiovisual integration of speech. *Information Fusion, 11*, 4–11.
Noesselt, T., Bergmann, D., Hake, M., Heinze, H. J., & Fendrich, R. (2008). Sound increases the saliency of visual events. *Brain Research, 1220*, 157–163.
Peiffer, A. M., Mozolic, J. L., Hugenschmidt, C. E., & Laurienti, P. J. (2007). Age-related multisensory enhancement in a simple audiovisual detection task. *Neuroreport, 18*, 1077–1081.
Ross, L. A., Saint-Amour, D., Leavitt, V. M., Molholm, S., Javitt, D. C., & Foxe, J. J. (2007). Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophrenia Research, 97*, 173–183.
Santangelo, V., & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology. Human Perception and Performance, 33*, 1311–1321.
Senkowski, D., Saint-Amour, D., Gruber, T., & Foxe, J. J. (2008). Look who's talking: the deployment of visuo-spatial attention during multisensory speech processing under noisy environmental conditions. *NeuroImage, 43*, 379–387.
Senkowski, D., Talsma, D., Herrmann, C. S., & Woldorff, M. G. (2005). Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. *Experimental Brain Research, 166*, 411–426.
Serences, J. T., & Boynton, G. M. (2007). Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron, 55*, 301–312.
Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature, 408*, 788.
Soto-Faraco, S., & Alsius, A. (2009). Deconstructing the McGurk-MacDonald illusion. *Journal of Experimental Psychology. Human Perception and Performance, 35*, 580–587.
Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: exogenous and endogenous mechanisms. *Journal of Experimental Psychology. Human Perception and Performance, 20*, 555–574.
Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology. Human Perception and Performance, 22*, 1005–1030.
Stein, B., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.
Talsma, D., Doty, T. J., Strowd, R., & Woldorff, M. G. (2006). Attentional capacity for processing concurrent stimuli is larger across sensory modalities than within a modality. *Psychophysiology, 43*, 541–549.
Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cerebral Cortex, 17*, 679–690.
Talsma, D., & Kok, A. (2002). Intermodal spatial attention differs between vision and audition: an event-related potential analysis. *Psychophysiology, 39*, 689–706.
Talsma, D., Kok, A., & Ridderinkhof, K. R. (2006). Selective attention to spatial and non-spatial visual stimuli is affected differentially by age: effects on event-related brain potentials and performance data. *International Journal of Psychophysiology, 62*, 249–261.
Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences, 14*, 400–410.
Talsma, D., Senkowski, D., & Woldorff, M. G. (2009). Intermodal attention affects the processing of the temporal alignment of audiovisual stimuli. *Experimental Brain Research, 198*, 313–328.
Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience, 17*, 1098–1114.
Theeuwes, J. (1991). Exogenous and endogenous control of attention—the effect of visual onsets and offsets. *Perception & Psychophysics, 49*, 83–90.
Theeuwes, J., Atchley, P., & Kramer, A. (2000). On the time course of top-down and bottom-up control of visual attention. In S. Monsel & J. Driver (Eds.), *Attention and performance 18* (pp. 105–124). Cambridge, MA: MIT Press.
Treisman, A. M., & Davies, A. (1973). Divided attention to ear and eye. In S. Kornblum (Ed.), *Attention & performance IV* (pp. 101–117). New York: Academic Press.
Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.
van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology. Human Perception and Performance, 34*, 1053–1065.
van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2009). Poke and pop: tactile-visual synchrony increases visual saliency. *Neuroscience Letters, 450*, 60–64.
van Ee, R., van Boxtel, J. J. A., Parker, A. L., & Alais, D. (2009). Multisensory congruency as a mechanism for attentional control over perceptual selection. *Journal of Neuroscience, 29*, 11641–11649.
Wecker, N. S., Kramer, J. H., Wisniewski, A., Delis, D. C., & Kaplan, E. (2000). Age effects on executive ability. *Neuropsychology, 14*, 409–414.
Weissman, D. H., Warner, L. M., & Woldorff, M. G. (2004). The neural mechanisms for minimizing cross-modal distraction. *Journal of Neuroscience, 24*, 10941–10949.
Welch, R. B., & Warren, D. H. (1986). Intersensory interactions. In K. R. Kauffman & J. P. Thomas (Eds.), *Handbook*

of perception and human performance, volume 1: sensory processes and perception (pp. 1–36). New York: John Wiley & Sons.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120*, 272–292.

Woldorff, M. G., Hazlett, C. J., Fichtenholtz, H. M., Weissman, D. H., Dale, A. M., & Song, A. W. (2004). Functional parcellation of attentional control regions of the brain. *Journal of Cognitive Neuroscience, 16*, 149–165.

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences, 7*, 70–76.

Wu, C. T., Weissman, D. H., Roberts, K. C., & Woldorff, M. G. (2007). The neural circuitry underlying the executive control of auditory spatial attention. *Brain Research, 1134*, 187–198.

Wu, J., Li, Q., Bai, O., & Touge, T. (2009). Multisensory interactions elicited by audiovisual stimuli presented peripherally in a visual attention task: a behavioral and event-related potential study in humans. *Journal of Clinical Neurophysiology, 26*, 407–413.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention—evidence from visual-search. *Journal of Experimental Psychology. Human Perception and Performance, 10*, 601–621.

Zimmer, U., Itthipanyanan, S., Grent-'t-Jong, T., & Woldorff, M. G. (2010). The electrophysiological time course and interaction of stimulus conflict and the multisensory spread of attention. *European Journal of Neuroscience, 31*, 1744–1754.

Zimmer, U., Roberts, K. C., Harshburger, T. B., & Woldorff, M. G. (2010). Multisensory conflict modulates the spread of visual attention across a multisensory object. *NeuroImage, 52*, 606–616.

Q