Acta Psychologica 162 (2015) 20–28

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement

N. Van der Stoep [a,*], C. Spence [b], T.C.W. Nijboer [a,c], S. Van der Stigchel [a]

[a] Utrecht University, Department of Experimental Psychology, Helmholtz Institute, Utrecht, The Netherlands
[b] Oxford University, Department of Experimental Psychology, Oxford, United Kingdom
[c] Brain Center Rudolf Magnus, and Center of Excellence for Rehabilitation Medicine, University Medical Center Utrecht and De Hoogstraat Rehabilitation, The Netherlands

ARTICLE INFO

Article history:
Received 13 May 2015
Received in revised form 22 September 2015
Accepted 24 September 2015
Available online 2 October 2015

Keywords:
Multisensory response enhancement
Multisensory integration
Exogenous spatial attention
Race model
Cueing
Crossmodal

ABSTRACT

Two processes that can give rise to multisensory response enhancement (MRE) are multisensory integration (MSI) and crossmodal exogenous spatial attention. It is, however, currently unclear what the relative contribution of each of these is to MRE. We investigated this issue using two tasks that are generally assumed to measure MSI (a redundant target effect task) and crossmodal exogenous spatial attention (a spatial cueing task). One block of trials consisted of unimodal auditory and visual targets designed to provide a unimodal baseline. In two other blocks of trials, the participants were presented with spatially and temporally aligned and misaligned audiovisual (AV) targets (0, 50, 100, and 200 ms SOA). In the integration block, the participants were instructed to respond to the onset of the first target stimulus that they detected (A or V). The instruction for the cueing block was to respond only to the onset of the visual targets. The targets could appear at one of three locations: left, center, and right. The participants were instructed to respond only to lateral targets. The results indicated that MRE was caused by MSI at 0 ms SOA. At 50 ms SOA, both crossmodal exogenous spatial attention and MSI contributed to the observed MRE, whereas the MRE observed at the 100 and 200 ms SOAs was attributable to crossmodal exogenous spatial attention, alerting, and temporal preparation. These results therefore suggest that there may be a temporal window in which both MSI and exogenous crossmodal spatial attention can contribute to multisensory response enhancement.

1. Introduction

It is now commonly acknowledged that our senses do not operate independently and that what is perceived via one sense will often (for better or for worse) influence what is perceived via another. For example, when a sound attracts attention to the perceived location of its source, it can facilitate the processing of any visual information that happens to be presented from that location as compared to other locations (i.e., crossmodal exogenous spatial attention; e.g., Driver & Spence, 1998; McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2003, 2005; McDonald & Ward, 2000; Spence & Driver, 2004). It is often suggested that what we hear can also be integrated with what we see (i.e., multisensory integration (MSI); e.g., Meredith & Stein, 1986; Molholm et al., 2006; Stein & Meredith, 1993; Stein & Stanford, 2008), often resulting in improved sensory information processing (e.g., Laurienti, Burdette, Maldjian, & Wallace, 2006; Leone & McCourt, 2013; Miller, 1982; Stevenson, Fister, Barnett, Nidiffer, & Wallace, 2012; Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008). Thus, both crossmodal exogenous spatial attention and multisensory integration can facilitate sensory information processing. It is, however, currently unclear what the relative contributions of crossmodal shifts of exogenous spatial attention and multisensory integration are to multisensory response enhancement (MRE; i.e., shorter RTs to multisensory stimuli as compared to the shortest RT to either of the unimodal component stimuli).

What both the crossmodal exogenous spatial attention and the multisensory integration accounts have in common is the suggestion that the benefits of multisensory stimulation are most pronounced when the unimodal components of a multisensory stimulus are spatially aligned (i.e., presented from the same spatial location, the spatial rule; McDonald, Teder-Sälejärvi, & Hillyard, 2000; Spence & Driver, 2004; Spence & McDonald, 2004; Störmer, McDonald, & Hillyard, 2009; though see Spence, 2013) as compared to when they are spatially misaligned (that is, presented from different spatial positions). The principle of spatial alignment seems to hold true for both the horizontal and depth dimension in the case of both crossmodal exogenous spatial attention (e.g., Ngo & Spence, 2010; Van der Stoep, Nijboer, & Van der Stigchel, 2014) and multisensory integration (e.g., Canzoneri, Magosso, & Serino, 2012; Sambo & Forster, 2009; for a review, see Van der Stoep, Nijboer, Van der Stigchel, & Spence, 2015).

It is not surprising to find that there is a debate, here, about whether these processes are essentially the same or not (see McDonald, Teder-

* Corresponding author at: Experimental Psychology, Helmholtz Institute, Heidelberglaan 1, 3584 CS, Utrecht, The Netherlands.
E-mail address: N.VanderStoep@uu.nl (N. Van der Stoep).

http://dx.doi.org/10.1016/j.actpsy.2015.09.010
0001-6918/

Sälejärvi, & Ward, 2001; and Spence, 2010, pp. 183–184), given the similarities between crossmodal exogenous spatial attention and multisensory integration. Interestingly, several ways of differentiating between them have already been proposed in a technical commentary by McDonald et al. (2001). Despite this commentary, studies directly investigating the difference between crossmodal spatial attention and MSI are currently still lacking. McDonald et al. (2001) argued that one of the ways in which to differentiate between them is by looking at the time-course of their behavioral effects. In terms of the temporal alignment/misalignment of sound and light, crossmodal exogenous spatial attention and multisensory integration show very different temporal profiles behaviorally. The beneficial effects of crossmodal exogenous shifts of attention are often most pronounced when there is an interval between the presentation of the auditory and the visual stimulus (i.e., at stimulus onset asynchronies (SOA) of between ~50 and ~300 ms; e.g., Berger, Henik, & Rafal, 2005; McDonald & Ward, 1999, 2000; Spence & Driver, 1997; Spence & McDonald, 2004). In contrast, the behavioral benefits of multisensory integration are often most pronounced when the auditory and visual stimuli are presented in close temporal alignment (SOAs between 0 and ±50 ms; e.g., Leone & McCourt, 2013; Stevenson et al., 2012; though see, for example, King & Palmer, 1985) with the behavioral benefits decreasing more or less symmetrically as the SOA increases (e.g., Leone & McCourt, 2013; though see Vroomen & Keetels, 2010). Thus, time is needed for crossmodal exogenous spatial attention to shift to the location of the cue in order to facilitate the processing of the target, whereas there is a more specific (narrow) time window within which auditory and visual stimuli need to be presented for multisensory integration to occur.[1] The differing temporal profiles of MSI and crossmodal exogenous spatial attention provide support for the notion that multisensory integration and crossmodal exogenous spatial attention are fundamentally different processes.

Further support for this distinction comes from those studies that have indicated that multisensory integration can occur pre-attentively: as multisensory integration can occur before attention has had its effect, this indicates that they are indeed two separate processes (e.g., Soto-Faraco, Navarra, & Alsius, 2004; Spence & Driver, 2000; Vroomen, Bertelson, & De Gelder, 2001a; see McDonald et al., 2001). Furthermore, it has also been shown recently that exogenous crossmodal spatial attention modulates multisensory integration (Van der Stoep, Van der Stigchel, & Nijboer, 2015). When an exogenous spatial auditory cue was presented some time before (SOA: 200–250 ms) and at the same location as a multisensory target, multisensory integration was reduced as compared to when the cue was presented from a different location. This result indicates that exogenous spatial attention can act independently of multisensory integration when there is enough time for exogenous spatial attention to shift to the location of the cue (cf. Vroomen, Bertelson, & de Gelder, 2001b). Lastly, integrated auditory and visual cues can attract spatial attention to their location even under conditions of high perceptual load, whereas unimodal exogenous cues do not (see Spence, 2010, and Spence & Santangelo, 2009, for reviews). Taken together, behavioral effects of multisensory integration and crossmodal exogenous spatial attention not only have different temporal profiles, but also can act independently of and modulate each other.

Generally, two different types of tasks are used to measure the effect of multisensory integration and crossmodal exogenous spatial attention: the redundant target effect (RTE) task and crossmodal spatial cueing tasks (e.g., the orthogonal spatial cueing task, Driver & Spence, 1998; or the implicit spatial discrimination task, McDonald & Ward, 1999; Ward, McDonald, & Lin, 2000). These two paradigms are sometimes referred to as the crossmodal signals paradigm (RTE task) and the focused attention paradigm (crossmodal spatial cueing task; e.g., Colonius & Diederich, 2012). In previous studies, the effects of multisensory stimulation on (saccadic) response times (RTs) for different SOAs in the spatial cueing and RTE paradigms have been modeled within the Time Window of Integration (TWIN) framework (Colonius & Diederich, 2004, 2012; Diederich & Colonius, 2008, 2011). The TWIN model predicts the pattern of multisensory response enhancement (MRE) for a broad range of different SOAs for both paradigms.

Although insights from the TWIN model are specifically helpful in thinking about the optimal time window of multisensory integration under different conditions, it does not provide information about the relative contributions of various crossmodal processes that might contribute to MRE (i.e., crossmodal exogenous spatial attention and multisensory integration). The aim of the present study was therefore to systematically investigate the relative contribution of multisensory integration and crossmodal exogenous spatial attention to MRE at different temporal intervals between an auditory and a visual stimulus (at SOAs of 0, 50, 100, and 200 ms, auditory lead). To do this, two tasks were used that are generally considered to measure the effects of either crossmodal exogenous spatial attention (the implicit spatial discrimination task, e.g., McDonald & Ward, 1999) or multisensory integration (an RTE task; e.g., Laurienti et al., 2006; Miller, 1986; Stevenson et al., 2012; Van der Stoep, Van der Stigchel, & Nijboer, 2015) using exactly the same auditory and visual stimuli. By comparing the results from the two tasks, it was possible to explore the stimulus intervals at which MRE was caused by multisensory integration, exogenous crossmodal spatial attention, or both processes. Based on the above-mentioned literature, it was hypothesized that MRE is caused by multisensory integration at the shortest SOAs (0 ms), by crossmodal exogenous spatial attention and multisensory integration at intermediate SOAs (50 ms), and by crossmodal exogenous spatial attention at longer SOAs (100–200 ms).

2. Materials and methods

2.1. Participants

Twenty-four participants were tested in this experiment (mean age = 26.6 years, *SD* = 3.3, 10 male, 14 female). All of the participants reported a normal sense of hearing and normal or corrected-to-normal visual acuity. They all signed an informed consent form prior to taking part in the study and were rewarded with £10 sterling for their participation. All of the participants took part in the current study and another study of multisensory interactions in one session that lasted for approximately 1.5 h. The order in which the experiments were conducted was counterbalanced across participants. The study was reviewed and approved by the Central University Research Ethics Committee of the University of Oxford.

2.2. Apparatus and stimuli

A custom-built audiovisual stimulus generator (see also Van der Stoep, Van der Stigchel, & Nijboer, 2015) connected to a PC running MATLAB was used to present the auditory stimuli through different loudspeakers (e-audio black 4″ Full Range Mini Box Speaker, dimensions: 120 × 120 × 132 mm, frequency response: 80–20,000 Hz) and visual stimuli through different Light Emitting Diodes (LEDs, Forge Europa, bulb size: 5 mm, viewing angle: 65°, tri-colored LED color: red, green, and blue). The loudspeaker array consisted of three loudspeakers placed at eye-level. One loudspeaker was positioned directly in front of the participant at eye-level at a distance of 64 cm, and two loudspeakers were positioned 26.1° to the left and right of the central loudspeaker. The auditory targets consisted of a 100 ms white noise burst [15 ms rise and fall time, ~65 dB(A)]. The use of tri-colored LEDs

[1] For multisensory integration to occur, it seems especially important that the responses to auditory and visual stimuli in a multisensory neuron overlap (e.g., King & Palmer, 1985; Meredith et al., 1987). Relatively small differences in temporal onset can occur while still resulting in multisensory integration. The overlap in auditory and visual discharge trains is dependent on stimulus intensity, the distance between the stimuli and the observer, and the time it takes for visual and auditory input to reach a multisensory neuron.