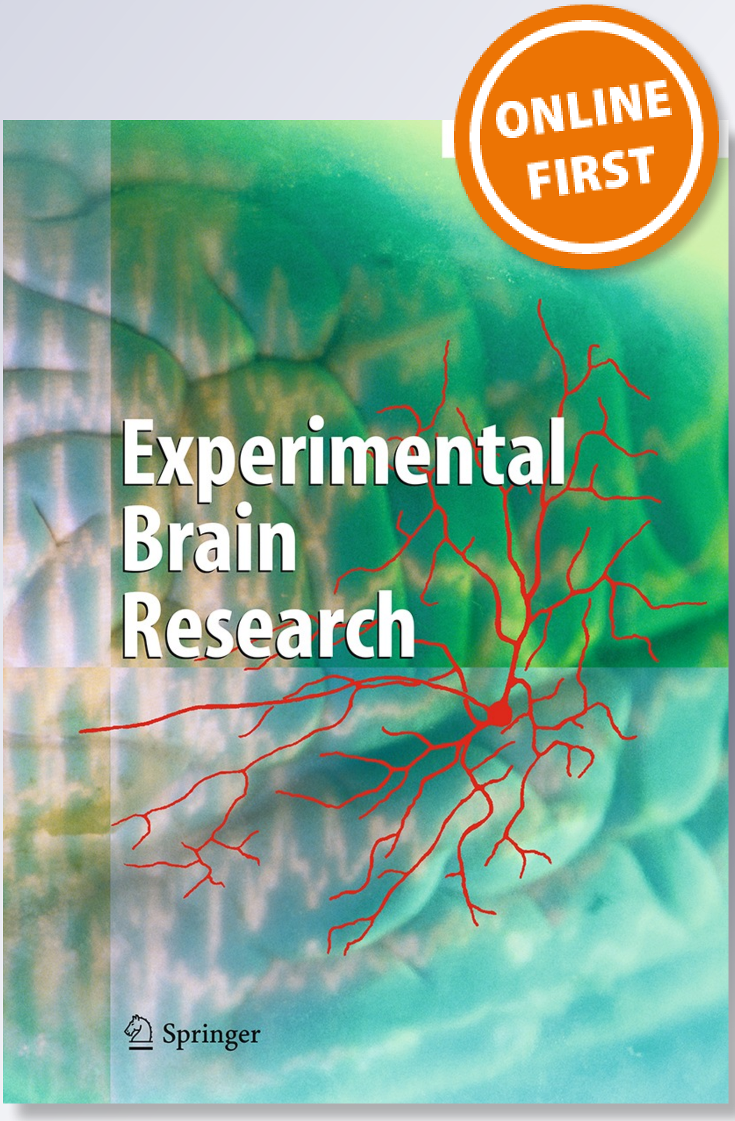
Prior knowledge of spatiotemporal configuration facilitates crossmodal saccadic response

Adele Diederich, Hans Colonius & Farid I. Kandil

Experimental Brain Research

ISSN 0014-4819

Exp Brain Res
DOI 10.1007/s00221-016-4609-5



Exp Brain Res
DOI 10.1007/s00221-016-4609-5

CrossMark

RESEARCH ARTICLE

Prior knowledge of spatiotemporal configuration facilitates crossmodal saccadic response

A TWIN analysis

Adele Diederich[1] · **Hans Colonius**[2] · **Farid I. Kandil**[3,4]

Received: 9 October 2015 / Accepted: 23 February 2016

Abstract Saccadic reaction times from a focused-attention task with a visual target and an acoustic nontarget support the hypothesis that the amount of saccadic facilitation in the presence of a nontarget increases with the prior knowledge of alignment with the target across different blocks of trials. The time-window-of-integration model can account for the size of the effect by having window size depend on the prior knowledge of alignment. Some efforts to identify the neural correlates of the effect are discussed.

Keywords Multisensory integration · Prior knowledge · Focused-attention experiment · Saccadic reaction time

✉ Adele Diederich
a.diederich@jacobs-university.de

Hans Colonius
hans.colonius@uni-oldenburg.de

Farid I. Kandil
f.kandil@uke.de

1 Department of Psychology and Methods, Jacobs University Bremen, Focus Health, Campus Ring 1, 28725 Bremen, Germany

2 Department of Psychology and Cluster of Excellence "Hearing4all", Carl von Ossietzky Universität Oldenburg, 26111 Oldenburg, Germany

3 Department of Psychology and Methods, Jacobs University Bremen, Focus Health, 28725 Bremen, Germany

4 Present Address: Institute of Computational Neuroscience, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany

Introduction

In complex environments, gaze shifts are commonly made in response to sudden cues indicating potentially relevant objects or events. Such an orienting response may be elicited by visual, acoustic, or tactile stimuli, or by some crossmodal combination of them. In a simplified laboratory situation, a common observation is that orienting responses to a visual target stimulus, e.g. a flash, can be facilitated by an auditory stimulus appearing in spatiotemporal proximity even though participants are instructed to ignore the auditory (nontarget) stimulus in this *focused-attention* paradigm (FAP).[1] This has been shown for saccadic eye movements to visual, auditory, or tactile stimuli (e.g., Hughes et al. 1994; Nozawa et al. 1994; Corneil and Munoz 1996; Harrington and Peck 1998; Amlôt et al. 2003; Arndt and Colonius 2003; Diederich and Colonius 2007a; Charbonneau et al. 2013), and also for saccadic head movements in both humans (Goldring et al. 1996) and animals (Whitchurch and Takahashi 2006). Specifically, this crossmodal reaction time (RT) *facilitation* effect is getting stronger when target and nontarget stimuli are presented together in space and time, and it diminishes, or even reverses into *inhibition*, with increasing spatial or temporal distance between the stimuli (Frens et al. 1995; Corneil and Munoz 1996; Hughes and Nelson 1998; Colonius and Arndt 2001; Van Wanrooij et al. 2009). A prominent view of the effect is that high spatiotemporal proximity makes it more likely

[1] This paradigm is also known as "crossmodal exogenous spatial attention" (see, e.g., Van der Stoep et al. 2015; Spence and Driver 1997). Diederich and Colonius (2008a) present a study separating, both empirically and theoretically, multisensory integration from non-spatial cueing effects, via a large range of SOA values.

Published online: 15 March 2016

that participants interpret the visual and acoustic information as originating from one and the same object, thus providing a "common cause" for an event that is speeding up the response (e.g., Körding et al. 2007).

The magnitude of the effect depends on many factors: modality of target and nontarget, relative stimulus intensities, stimulus onset asynchrony (SOA), their absolute position in space, and the spatial configuration between target and nontarget, e.g., whether a visual and an auditory stimuli appear in the same hemifield (*ipsilateral*) or in different hemifields (*contralateral*), and whether they differ with respect to azimuth or elevation (Corneil and Munoz 1996; Spence and Driver 1997; Ten Brink et al. 2014). Intriguingly, it has been demonstrated that a critical factor of the effect of a nontarget acoustic stimulus is its localizability determined by the *perceived* rather than the physical distance between visual target and acoustic nontarget, despite the fact that the participant's task is to only respond to the onset of the visual target and to ignore the acoustic stimulus (Frens et al. 1995; Heuermann and Colonius 2001; Steenken et al. 2008; Colonius et al. 2009).

Location of the target itself plays an important role also in the following way. Since the classic studies by Posner and colleagues (e.g. Posner 1980), it is well known that prior knowledge, or expectation, about the location of a target stimulus speeds up stimulus detection, and the validity of a cue presented in spatiotemporal proximity has been shown to be a modulating factor.[2] However, looking at the role of prior knowledge in saccadic behavior within a *crossmodal* context, with target and nontarget defined in different modalities, has received relatively little attention. In a recent FAP study, Van Wanrooij et al. (2010) measured human head saccades toward a visual target which, in the bimodal condition, was accompanied by a synchronous sound (white noise) of varying vertical disparity to the target. They found reaction times consistently faster in stimulus blocks containing only spatially aligned audiovisual stimuli than in blocks also containing pseudo-randomly presented spatially disparate stimuli.

The present study was set up to corroborate the findings of Van Wanrooij et al. (2010) about the effect of prior knowledge and to extend them to saccadic eye movements with nonsimultaneous presentations of target and nontarget. Furthermore, it generalizes the task to include the horizontal plane which might yield different results given the different sound localization mechanisms for the horizontal vs. vertical plane.

To wrap up, we propose the following empirical hypothesis:

Hypothesis Saccadic facilitation in the presence of a nontarget increases with the level of prior knowledge of spatial alignment of a target with the nontarget.

The hypothesis is tested by varying the relative frequency of alignment of visual target and acoustic nontarget within a block of trials with levels $p = .20$, $p = .50$, and $p = .80$. Note that a weaker version of the hypothesis would be that facilitation is a *nondecreasing* function of p; in any case, faster mean RTs (more facilitation) for $p = .20$ than for $p = .80$ would be considered evidence against any version of the hypothesis.

Another, major goal of the present study was to probe whether the authors' time-window-of-integration model (TWIN) (Colonius and Diederich 2004; Diederich and Colonius 2004) can account for the data. Note that this involves two different aspects: first, whether TWIN is able to quantitatively describe saccadic RTs across the different spatiotemporal conditions and second, how the model incorporates the specific effects of prior knowledge on response speed.

Moreover, it has been shown that prior knowledge of the crossmodal spatial configuration may have an immediate effect on stimulus processing, i.e., from one trial to the next (Wozny and Shams 2011; Mendonça et al. 2014, 2015). Although this short-term effect is not the focus here, we have to consider it as an additional source of crossmodal effects (see also Sarmiento et al. 2015) (see Sect. 3.3).

Material and methods

Participants

Twelve students, aged 19–21, one aged 27, seven females, from Jacobs University Bremen, served as voluntary participants. They were paid or received credit points required for their studies. All had normal or corrected-to-normal vision, normal hearing and were right-handed (self-description, Coren's Lateral Preference Inventory, 1993). They were screened for their ability to follow the experimental instructions (proper fixation, few blinks during trial, saccades toward visual target). They gave their written informed consent prior to their inclusion in the study and the experiment has been conducted according to the principles expressed in the Declaration of Helsinki. Approval for this study was granted by the Academic Integrity Committee of Jacobs University Bremen.

Stimuli and apparatus

Fixation point was a LED (25 mA, 5.95 mcd) located in the medial line at a distance of 90 cm. Auditory stimuli

[2] For a recent modeling approach, see Vossel et al. (2014).

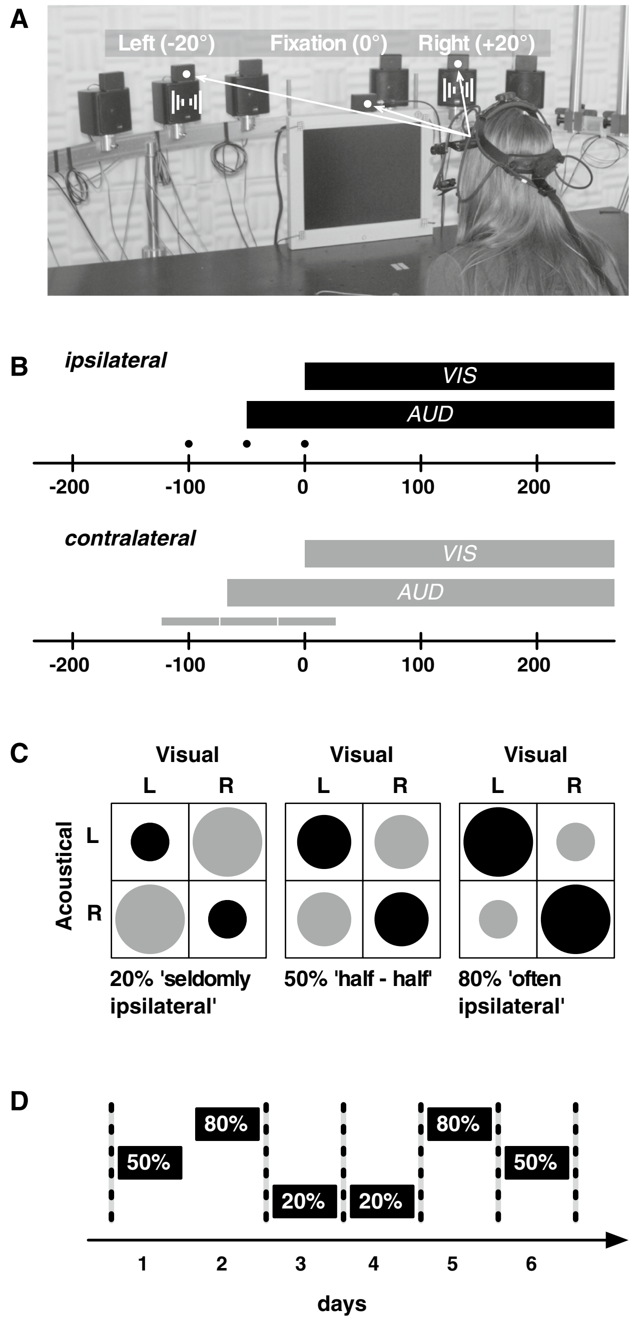


Fig. 1 Setup and experimental design. **a** Experimental setup. Target LED and speakers for the presentation of the visual and the acoustical stimuli, respectively, are installed to the left (−20°) and right (+20°) of the central fixation LED and can be activated independently from each other. In the focused-attention paradigm employed here, the visual stimulus (LED) is always the target stimulus. Participants are instructed to move their eyes from fixation to the visual stimulus as quickly as possible. The acoustical stimulus is only auxiliary and can be presented on the same side as the visual (coincident or ipsilateral condition) or on the opposite side (disparate or contralateral condition). **b** The visual stimulus is by definition presented at $t = 0$ ms. In ipsilateral trials, the acoustical stimulus is presented with an SOA of −100, −50, or 0 ms, whereas in contralateral trials, it can be presented at any time between −125 and −76, −75 and −26, or −25 and +24 ms, respectively. (C) The experimental design comprises three levels of prior knowledge: visual and acoustical stimulus are presented ipsilateral in 20 (*left*), 50 (*middle*), or 80 % (*right*) of the trials. **d** Experiments were run on 6 days. Blocks with different prior knowledge levels were conducted on different days in an A-B-C-C-B-A design that was balanced across participants to minimize sequence effects. Blocks with unimodal stimuli were additionally presented either at the start (day 1) or the end (days 2–6) of each day to control for perceptual or motor learning effects (*black* and *white bars*)

were white noise bursts [59 dB(A), rectangle envelope function], generated by two speakers (Canton Plus XS). The speakers were placed at 20° to the left and right of the fixation LED at the height of the participants' ear level at a distance of 120 cm. The visual stimuli were red LEDs (25 mA, 3.3 mcd) located at 20° to the left and right on top of the speakers at the same viewing distance of 120 cm (Fig. 1a). Note that the distance between stimuli presented at the same side (ipsilateral) was 0°; the distance between two stimuli presented on opposite sides (contralateral) amounted to 40°. One PC controlled the stimulus presentation, and two other interlinked PCs controlled the eye movement registration program (SR-Research EyeLink 1000). The control software for stimulus presentation operated on Realtime Linux (RTLinux), a hard real-time kernel (RTLinux patched kernel) that runs Linux as its idle thread. Signal output was carried out by a computer card (PCIM DDA06/16), equipped with six digital–analog converters and three digital in- and outports, which fed the control electronic with the generated time signals for the LEDs and the loud speakers.

Procedure and experimental design

Participants were seated in a completely darkened, sound attenuated room with the head positioned on a chin rest, elbows, and lower arms resting comfortably on a table. Although the eye tracking equipment takes head movements into account, participants were instructed to keep the head on the chin rest and not to move. The participant began every experimental session with a 10 min of dark adaptation during which the measurement system was adjusted and calibrated. Each trial started with the appearance of the fixation point of random duration (800–1500 ms). When the fixation LED disappeared, the visual target stimulus was turned on for 500 ms without a gap.

For the ipsilateral condition, onset of auditory nontargets was shifted by a stimulus onset asynchrony (SOA) of $-100, -50, 0$ ms (Fig. 1b, upper panel). Negative values mean that the ipsilateral nontarget was presented before the target. For the contralateral condition, onset of the

Table 1 Absolute number of error types (in parentheses: percentage all; percentage within ipsi; percentage within contra) for the three prior knowledge conditions across all participants

Error	Prior knowledge condition		
	20 %	50 %	80 %
No saccade performed	130 (0.95; 1.17; 0.90)	45 (0.38; 0.37; 0.38)	147 (1.12; 1.05; 1.41)
Saccades before any signal	108 (0.79; 0.84; 0.78)	81 (0.68; 0.62; 0.74)	136 (1.04; 1.06; 0.95)
Amplitude not within 3 std	137 (1.00; 0.73; 1.10)	153 (1.29; 0.91; 0.17)	134 (1.02; 0.79; 1.94)
RT < 80	324 (2.37; 4.31; 1.88)	379 (3.20; 4.90; 1.48)	788 (6.00; 6.80; 2.78)
RT < 500	8 (0.01; 0.00; 0.01)	15 (0.13; 0.10; 0.15)	23 (0.18; 0.18; 0.15)
Directional	403 (2.95; 0.95; 3.44)	240 (2.02; 0.25; 3.80)	206 (1.57; 0.50; 5.82)
Total	1110 (8.1; 8.0; 8.1)	913 (7.7; 7.2; 8.2)	1434 (10.9; 10.4; 13.1)

These trials were excluded from further analysis

auditory stimuli was prior to or after the visual target at any SOA between −125 and −76 ms, −75 and −26 ms, or −25 and +24 ms, in steps of 1 ms, respectively, following a uniform distribution (Fig. 1b, lower panel). This was to minimize temporal predictability for contralateral trials. The nontargets were turned off simultaneously with the visual stimulus. Thus, their duration varied between 625 and 500 ms. Stimulus presentation was followed by a break of 2 s in complete darkness before the next trial began, indicated by the onset of the fixation LED. Participants were instructed to gaze at the visual target as quickly and as accurately as possible ignoring any auditory nontargets (focused-attention paradigm). The visual target appeared in combination with the auditory nontarget in either ipsi- or contralateral position. Prior knowledge of ipsi- and contralateral presentations varied in three levels (Fig. 1c). In one condition (central panel), the proportion of ipsi- and contralateral presentation was equal. In the two remaining conditions, ipsilateral presentation occurred in either 80 or 20 % of the trials (Fig. 1c, left and right panel, respectively).

One experimental block consisted of 156 and 180 bimodal trials for equal and unequal frequencies, respectively, randomized over SOA and laterality. Note that the level of prior knowledge was kept invariant within a block during an experimental session. Sequence effects were counteracted by using an A-B-C-C-B-A design that was additionally randomized across participants (Fig. 1d). In addition, unimodal stimuli were presented in separate blocks of 50 visual and 50 auditory stimuli.

Despite training, it cannot be prevented that participants might get faster over the time course of sessions. Therefore, mean response time to unimodal stimuli, as a function of experimental session, served as a measure to gauge this effect. Each participant performed a total of 24 experimental blocks, three bimodal and one unimodal within a single session that lasted about one hour. Each participant was engaged for six hours over the course of two weeks and completed a total of 3, 696 experimental trials.

Results

Data screening and preprocessing

Saccades were screened for anticipation errors (RT < 80 ms), misses (RT > 500 ms), and accuracy: Trials with saccade amplitude deviating more than three standard deviations from the mean amplitude were also excluded from the analysis. Table 1 shows the absolute number and percentages (in parentheses) of different error types for each level of prior knowledge across all participants. The percentage refers to all trials presented in that condition (first entry), to trials presented ipsilateral (second entry), and to trials presented contralateral (third entry). RTs smaller than 80 ms appear more often for ipsilateral trials; direction errors more often for contralateral trials. However, the error rates are rather low (less than 5 %) for most individuals, with some exceptions (see Table 6 for all participants in the "Appendix" section). There is no discernible difference in errors with respect to the level of prior knowledge. There was also no evidence for multiple saccades in the remaining data set.

The design required that prior knowledge was kept invariant within an experimental session of three bimodal blocks of trials. Data inspection revealed that reaction times generally decreased across the six experimental sessions for all participants and the order in which the blocks were presented could bias the results. Therefore, we adjusted the RTs as follows (Fig. 3a shows this for Participant 2). First, in order to estimate the decrease in response times to bimodal stimuli, a regression line was fitted to the six mean unimodal visual response times measured on each of the six sessions/days (black crosses in Fig. 3a). Second, mean response time to bimodal stimuli of a given block was divided by the regression value obtained at the corresponding experimental block value and multiplied by the grand mean response time to unimodal stimuli. Note that the assumption here is that the proportional decrease in unimodal visual RTs is representative for all

Table 2 Repeated measures ANOVA results (output from IBM SPSS) from 12 participants

		df	*F*	Sig.
Prior knowledge	Sphericity assumed	2	3.4	.052
Laterality	Sphericity assumed	1	63.8	.000
SOA	Sphericity assumed	2	111.1	.000
Prior knowledge × Laterality	Greenhouse–Geisser	1.4	14.1	.001
Prior knowledge × SOA	Sphericity assumed	4	0.8	.537
Laterality × SOA	Sphericity assumed	2	7.8	.003
Prior knowledge ×Laterality × SOA	Greenhouse–Geisser	2.4	3.4	.041

bimodal conditions. This seems reasonable given that the speed-up over time should reflect a general habituation effect and we do not have any hypotheses concerning a specific effect of prior knowledge level in the bimodal conditions. Moreover, the latter, if it existed, should be counterbalanced by our randomization across conditions. Only the adjusted response times were included in all further analyses.

Mean saccadic response times

A three-factor repeated measures analysis ($3 \times 2 \times 3$) with factors Prior knowledge (20, 50, 80 %), Laterality (ipsi, contra), and SOA ($-100, -50, 0$) on the means of 12 participants showed strong main effects for Laterality and SOA, and Prior Knowledge missing the 5 % level. Details are given in Table 2.

Figure 2 presents the results averaged across participants. Panel A shows mean RT as a function of SOA, clustered by the level of Prior knowledge; panel B shows the same data clustered by SOA values. Mean RT increases as SOA decreases, i.e., with delaying the acoustic nontarget. Moreover, mean RT to bimodal stimuli presented ipsilateral are shorter than those presented contralateral. This pattern, commonly found in many studies (cf. references in the introduction), holds true under all three levels of Prior knowledge. The significant interactions can be read from the figure as well. Interaction between Prior knowledge and Laterality indicates that the difference between ipsi- and contralateral mean RTs increases with the level of Prior knowledge. This is consistent with the notion that perception of a unitary audiovisual event increases with the Prior knowledge of spatial alignment. The other significant interaction, between Laterality and SOA, is due to the increase in mean RT with SOA approaching zero being more pronounced for ipsi- than for contralateral conditions. This reflects the finding that, for those SOA where it occurs, facilitation is stronger for spatially aligned stimuli than for contralateral stimuli.

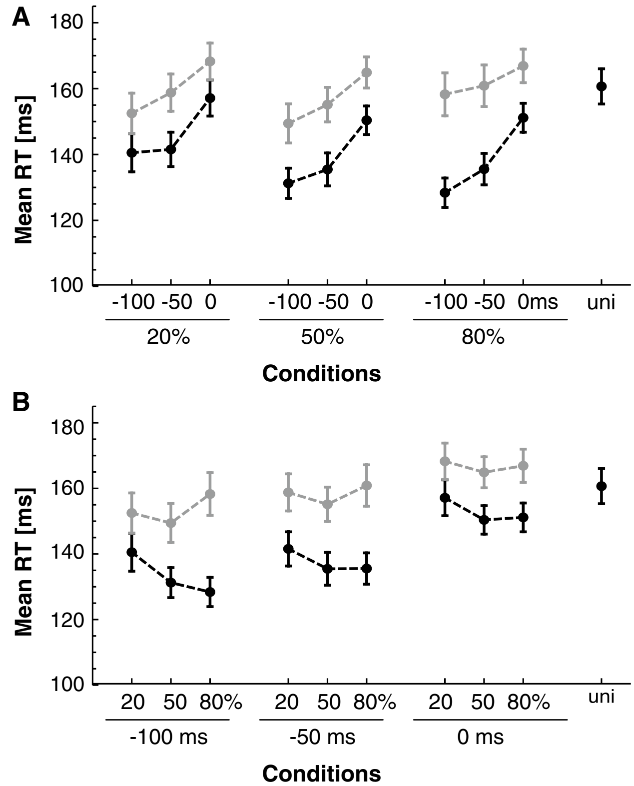


Fig. 2 Reaction time data across all participants. **a** Means and standard deviations for each of the 18 bimodal conditions. RTs are sorted initially by Prior knowledge level (20, 50 and 80 %) and then by SOA. *Black lines* indicate ipsilateral, *gray lines* contralateral conditions. **b** The same results, but sorted initially by SOA and then by Prior knowledge level

Figure 3b, c presents the results for Participant 2. For this participant, effects of prior knowledge of alignment on mean RTs are shown in Panel C. Consistent with the hypothesis, mean RTs in the 80 % Prior Knowledge condition are faster than mean RT in the 20 % Prior Knowledge condition for ipsilateral presentations across all SOAs. For contralateral presentations, there is a tendency for faster responses in the 50 % Prior Knowledge.

A complete list of mean RTs and standard errors for all participants is found in Table 5 in the "Appendix" section. In general, our participants show a rather broad spectrum of results. For individual tests of the hypothesis that increasing the prior knowledge of spatial alignment leads to stronger RT facilitation, we compared each participant's mean RT in the 20 % Prior Knowledge condition with the mean RT in the 80 % condition, under spatial alignment of target and nontarget. Given that most facilitation occurs at SOA $= -100$, this comparison is most informative. In support of the hypothesis, all but two participants (P1 and P3) are significantly faster with the higher prior knowledge about the spatial alignment at SOA $= -100$

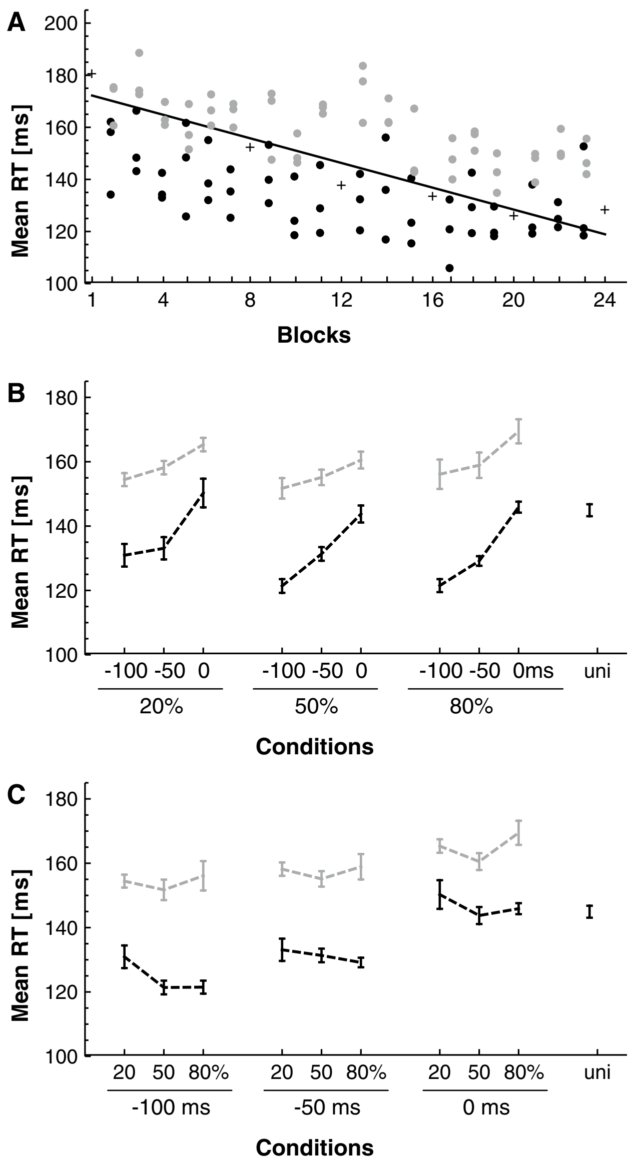


Fig. 3 Reaction time data for Participant 2. **a** Reaction times for the various conditions across all 6 days. The comparison of the RTs for the unimodal visual condition (*black crosses*) reveals a strong speed-up effect of nearly 50 ms for which RTs were corrected. **b** Means and standard deviations for each of the 18 bimodal conditions. RTs are sorted initially by prior knowledge level (20, 50, and 80 %) and then by SOA. *Black lines* indicate ipsilateral and *gray lines* contralateral conditions. **c** The same results but sorted initially by SOA and then by prior knowledge level

($\alpha < .01$ for 6 participants and $\alpha < .05$ for the other 4.); for details, see Table 3. Probing the data of participants P1 and P3 revealed that both had a somewhat noticeable pattern of errors. P3 had a high number of both anticipatory saccades (RT < 80 ms) and directional errors in contralateral conditions, increasing up to 17 in the 80 % Prior Knowledge condition. P1 had an average overall level of errors, the

Table 3 Participants showing a significant facilitation (mean RT) at prior knowledge level 80 % compared to 20 % (Welch's unpaired t-test), with $\alpha < .05$, participant numbers with asterisk: $\alpha < .01$. Only Participants 1 and 3 did not show an effect at SOA = −100 (see text)

SOA (ms)	Participant no.
−100	2, 4, 5*, 6*, 7*, 8*, 9*,10, 11*, 12
−50	4*, 7*, 8*, 11
0	7*, 8, 11, 12

only exception being a high number (7–9 %) of directional errors in the 20 % Prior Knowledge condition . While RTs in both error types were excluded from the analysis, this behavior may have "washed out" any existing effects of prior knowledge.

Trial-to-trial adaptation

As mentioned in the introduction, an alternative account for the RT effects could be that participants, rather than adapting to the prior knowledge based on long-term experience, simply react to the information from the immediately preceding trial(s). For example, they may be better prepared for a saccadic response in the same direction as the response executed in the trial before, i.e., left or right. Such dependencies would require more complex modeling and, possibly, larger sample size than currently available.

However, an analysis of the data from Participant 2, whose data seem sufficiently representative for all participants, revealed that dependencies on the types of the two preceding trials could only account for RT differences of −4.7 ms between the 20 % and the 80 % condition. Alternatively, participants could be more strongly prepared to integrate if the acoustic and the visual stimulus were presented ipsilateral in the preceding trial. Again, it was found that RT differences of only −0.8 ms could be accounted for by this effect. Finally, if stimuli are presented either ipsilateral or contralateral in both the current and the preceding trials, participants might be better prepared for a saccade or, respectively, an antisaccade, to the auxiliary acoustic stimulus. Here, we found an effect of up to +16.7 ms between conditions; however, it points in the direction opposite to this hypothesis. Summarizing, all of these trial-to-trial effects are too small (or point in the opposite direction) to account for the RT differences observed with the prior-probability conditions. Notably, excluding the RTs of the first block of each day from the analysis would have increased both absolute RT reductions and differences between the prior knowledge conditions, indicating that participants actually need the first block to adapt to the prior of that experimental day.

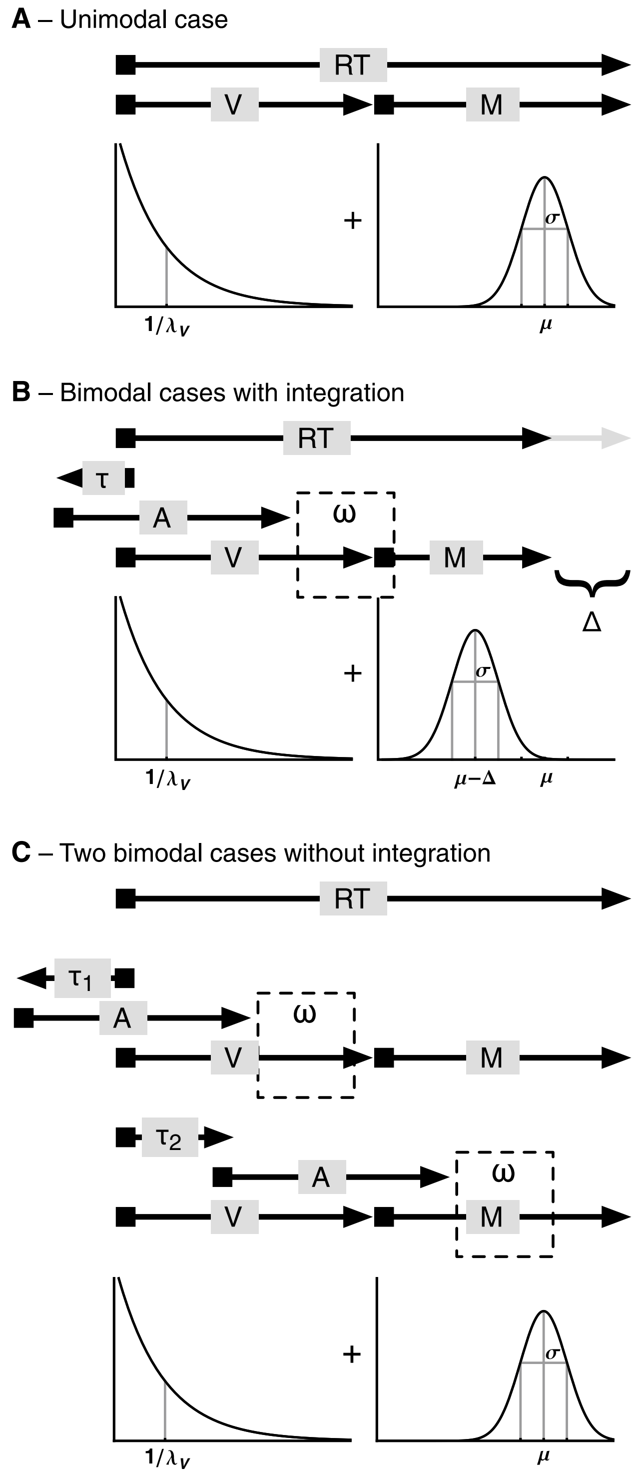


Fig. 4 Schematic presentation of TWIN. **a** The TWIN model ▸ assumes that unimodal RT is the sum of the processing times of a first (sensory) stage (here V for the visual processing time with mean processing time λ_V^{-1}) and a secondary stage (M with mean processing time μ). **b** In a focussed attention task, the nontarget (here A for an acoustical stimulus with mean processing time λ_A^{-1}) is presented τ ms (SOA) before the target (here, the visual stimulus) "opens" the integration window for a certain time (parameter ω). If first-stage processing of the visual stimulus ends within this time window, stimulus integration occurs, influencing the secondary stage (M) processing time. If enhancement occurs, overall RT is reduced by a certain temporal gain (parameter Δ). If inhibition occurs, overall RT is increased by Δ. **c** If the SOA is large enough for the temporal window to close before the visual stimulus processing terminates within the window, or if the acoustical stimulus arrives too late for the visual stimulus to open the window, no integration will occur

Time-window-of-integration (TWIN) model

Because this model and empirical tests have been described in detail in previous publications (Colonius and Diederich 2004; Diederich and Colonius 2007a, b, 2008a, b), we limit presentation to a minimum here. TWIN postulates that a crossmodal stimulus triggers a race mechanism among the activations in very early, peripheral sensory pathways. This first stage is followed by a compound stage of converging subprocesses that comprise neural integration of the input and preparation of a response. The second stage is defined by default: It includes all subsequent processes that are not part of the peripheral processes in the first stage (cf. Fig. 4). The central assumption of TWIN concerns the temporal configuration needed for multisensory integration to occur: *Multisensory integration occurs only if the peripheral processes of the first-stage all terminate within a given temporal interval, the* "time-window-of-integration" (TWIN assumption). Thus, the window acts as a filter determining whether or not the afferent information delivered from different sensory organs is registered close enough in time to trigger multisensory integration. This all-or-none assumption about the filter refers to a single trial only; that is, on average, the amount of crossmodal interaction [measured in ms] is weighed by the probability of interaction occurring, a term that does depend on SOA.

The amount of crossmodal interaction shows up as an increase or decrease in second-stage processing time, but it is assumed not to depend on the stimulus onset asynchrony (SOA) of the stimuli. A basic feature of the TWIN framework is priority of temporal proximity over any other type of proximity. Rather than assuming a joint *spatiotemporal* window of integration, permitting interaction to occur only for both spatially and temporally neighboring stimuli, this framework allows for crossmodal interaction to occur for spatially rather distant stimuli of different modalities, as long as they fall within the time window and some form of perception "unity" is possible. In its simplest version, the TWIN model framework is "agnostic" with respect to the exact nature of the integration mechanism occurring in the second stage. Nevertheless, it affords a number of quantitative predictions without making specific assumption about the probability distributions of the processing times.

Quantitative predictions

The race in the first stage is represented by two statistically independent, nonnegative random variables (V and A) denoting the peripheral processing times for the visual and the auditory stimulus, respectively. With τ as SOA value and ω as integration window width parameter, the requirement for multisensory integration to take place is that the nontarget stimulus (processing time) A wins the race in the first-stage "opening the time window of integration" such that the termination of the target peripheral processing time V falls into the window. Formally,

$$I = \{A + \tau < V < A + \tau + \omega\},$$

A positive τ value indicates that the visual stimulus is presented before the auditory, and a negative τ value indicates the reverse presentation order. An intuitive interpretation of the definition of I is that the winning nontarget will keep the system in a state of heightened reactivity such that the upcoming target stimulus, if it falls into the time window, will trigger crossmodal interaction. If the stimulus from the target modality is the winner of the race in the peripheral channels, second-stage processing is initiated without any multisensory integration mechanism being triggered. The rationale for that is that, if the target is registered first, there is "no reason" for the system to delay preparing the response.[3] In order to derive quantitative predictions some distributional assumptions for the processing stages are required.

We postulate exponential distributions[4] for the peripheral processes, V and A, with expected values $1/\lambda_V$ and $1/\lambda_A$, respectively (λ_V, λ_A are positive-valued parameters characterizing the exponential distributions). Mean RT in the bimodal condition then is

$$\mathrm{E}\,[\mathrm{RT}_{VA}] = \frac{1}{\lambda_V} + \mu - P(I) \cdot \Delta,$$

where μ denotes second-stage mean processing time, Δ the amount of interaction. $P(I)$, the probability of integration to occur, is a function of λ_V, λ_A, window width ω, and the SOA value τ. Note that first-stage duration is defined as the peripheral processing time for the (visual) target stimulus. Mean RT in the unimodal condition is assumed to be

$$\mathrm{E}\,[\mathrm{RT}_V] = \frac{1}{\lambda_V} + \mu.$$

Crossmodal interaction (CI) is defined as the difference between mean RT to the unimodal and crossmodal stimuli, i.e.,

$$\mathrm{CI} \equiv \mathrm{E}\,[\mathrm{RT}_V] - \mathrm{E}\,[\mathrm{RT}_{VA}] = P(I) \cdot \Delta. \qquad (1)$$

In this setting, the term $P(I) \cdot \Delta$ is a measure of the expected amount of intersensory interaction in the second stage, with positive Δ values corresponding to facilitation, negative ones to inhibition. It represents the factorization of CI into a factor $P(I)$ only depending on stimulus parameters λ_V, λ_A, window width ω, and SOA τ on the one hand, and a factor Δ only depending on parameters characterizing the bimodal context, like the spatial separation between the stimuli, on the other. This factorization is the basis of many empirical tests of the TWIN framework (Diederich and Colonius 2007a, b, 2008a, b).

Prior knowledge effects on target positions

We were interested in probing effects of building up prior knowledge (expectation) about the target position across many trials presented over the entire duration of an experimental session. Thus, the question to be addressed next is how the TWIN model could account for facilitatory effects similar to those observed in Van Wanrooij et al. (2010).

Varying prior knowledge about target position should not change effects of stimulus properties like intensity, so peripheral TWIN parameters λ_V, λ_A would not be affected. This leaves the size of the window, parameter ω, and the amount of interaction, parameter Δ, to capture an effect of prior knowledge, expressed as a change in expected crossmodal interaction, that is, the product $\mathrm{CI} = P(I) \cdot \Delta$. Parameter Δ depends on the spatial configuration, in particular the perceived distance between target and nontarget, while $P(I)$ is modulated by parameter ω. In principle, either Δ or ω could be affected by manipulating prior knowledge. A larger window should increase the range of SOAs that show facilitation, while a larger interaction parameter should increase the amount of facilitation at those SOAs where there is any SRT reduction effect.[5] Moreover, it cannot be ruled out that both parameters are affected simultaneously, so it may become quite difficult to disentangle these effects in practice. Given the somewhat restricted range of SOAs in the design, we decided to test only the hypothesis that the size of the window (ω) adapts to prior knowledge, leaving a more general test to future investigations. Evidence for the modulation of window size has previously been found, e.g., for stimulus intensity—low intensity stimuli suggesting a widening of the window

[3] In the TWIN version for a "redundant-signals task" (i.e., signals from any modality are targets), however, any stimulus can open the window (Colonius and Diederich 2012).

[4] The choice of exponentials is made for ease of computation; see Colonius and Diederich (2004).

[5] This observation was made by one of the reviewers.

(Diederich and Colonius 2008b)—and for covariates like age—with elderly participants also featuring a larger window than young adults (Diederich et al. 2008; Colonius and Diederich 2011).

Predicting window width

Specific quantitative predictions about window width can be derived from a decision-theoretic viewpoint developed in Colonius and Diederich (2010, 2011, 2012). It starts from the observation that an individual is faced with a basic decision situation, i.e., either to treat the visual and acoustic information as a unified event or to keep them apart. Whether or not the decision is correct depends on the "state of nature," and both types of error may involve a cost (Körding 2007).

For example, in a predator–prey situation when the potential prey perceives a sudden movement in the dark, it may be vital to recognize whether this is caused by a predator or a harmless wind gust. If visual information is accompanied by some vocalization from a similar direction, it may be adequate to respond to the potential threat by assuming that the visual and auditory information are caused by the same source and to initiate multisensory integration leading to a speeded escape reaction. On the other hand, it may also be disadvantageous, or even hazardous leading to a depletion of resources, to routinely combine information associated with sensory events in a rich dynamic environment which—in reality—are entirely independent and unrelated (Colonius and Diederich 2011, p. 330).

In a Bayesian modeling framework, prior information about the state of nature is combined with the likelihood of the observations under either hypothesis to derive the course of action (Ernst and Banks 2002; Ernst 2006; Shams and Beierholm 2010; Körding et al. 2007; Magnotti et al. 2013; Vossel et al. 2014). In the focused-attention task, prior knowledge amounts to high (or low, or indifferent) probability of a spatial alignment between target and nontarget. How should prior knowledge affect window size in the focused-attention paradigm? According to TWIN, when target and nontarget of different modalities both "fall into the window" they are perceived as a unitary event and subsequent processing is enhanced, with maximal facilitation taking place when the two stimuli are spatially aligned. Thus, whenever prior probability of spatial alignment is high within a block of trials, it would be beneficial for the observer if the time window was relatively large in order to increase the probability of the stimuli being processed as a unitary event.

Given empirical support, the TWIN model will be fitted to the data from all prior knowledge levels while keeping the same parameter values except for window width ω which is adapted to each of the p values.

Assuming that temporal disparity between the "arrival times" of the unimodal signals is the *only* perceptual evidence utilized by the organism and adding an assumption about the corresponding likelihood function under either state of nature, the optimal window width under Bayesian inference, given prior probability p, ω_p, has been shown[6] to be:

$$\omega_p = \left(\frac{1}{\lambda_V} + \frac{1}{\lambda_A}\right) \log\left[\frac{\lambda_V \lambda_A s}{\lambda_V + \lambda_A} \frac{p}{1-p}\right] \tag{2}$$

Obviously, this is an increasing function of prior probability of alignment p; thus, a larger p value should lead to larger window size implying greater response facilitation on average.

Interestingly, it follows from simple algebraic transformations that, for any p with $0 < p < 1$,

$$\omega_{0.5} = \frac{\omega_p + \omega_{1-p}}{2}, \tag{3}$$

which amounts to a theoretical prediction for the optimal window size that is independent of the parameter values. Of course, empirical deviations from this prediction could be due to two different causes: (1) large variability in the parameter estimates that go into estimates of the ω values or (2) the participants not exhibiting optimal behavior in this Bayesian sense.

Parameter estimation for TWIN

Model testing was performed for each participant separately. Eight parameters were estimated from the means of the 18 bimodal conditions (3 different Prior knowledge levels × 2 laterality levels × 3 SOA values): λ_A and λ_V account for the peripheral processing time of the first stage for auditory and visual stimuli, respectively ; μ is the mean time of the second processing stage; ω_{20}, ω_{50}, and ω_{80} indicate the integration window integration width with respect to the three different Prior knowledge levels; and Δ_I and Δ_C stand for the amount of interaction occurring for ipsi- and contralateral conditions, respectively.

Because data from the contralateral condition had to be included for parameter estimation purposes, combined response times to contralateral bimodal stimuli with SOAs between -125 and -76 ms were compared to those with SOA -100 in the ipsilateral condition. Likewise, those with SOA between -75 and -26 ms in the contralateral condition were compared to those with SOA -50 in the ipsilateral condition; finally, those with SOA between -25 and

[6] For details of the derivation, including the inessential constant s, we refer to Colonius and Diederich (2012).

25 ms in the contralateral conditions were compared to those with SOA 0 ms in the ipsilateral condition.

The parameters were estimated by minimizing the χ^2 statistic

$$\chi^2 = \sum_{n=1}^{3} \sum_{j=1}^{2} \left(\frac{\mathrm{RT}_{\mathrm{obs}}(j,n) - \mathrm{RT}_{\mathrm{pred}}(j,n)}{\sigma_{\mathrm{RT}_{\mathrm{obs}}(j,n)}} \right)^2 \tag{4}$$

using the optimization routine *fminsearchbnd*[7] of MATLAB R2014b. Here, $\mathrm{RT}_{\mathrm{obs}}(j,n)$ and $\mathrm{RT}_{\mathrm{pred}}(j,n)$ are, respectively, the observed and predicted values of mean RT to visual–auditory stimuli, presented in spatial positions (ipsilateral, $j = 1$; contralateral, $j = 2$) with SOA (referred to by n from 1 to 3); $\sigma_{\mathrm{RT}_{\mathrm{obs}}(j,n)}$ are the respective standard errors (standard deviation divided by the square root of the number of observations).

For the estimation routine, λ_V and λ_A were restricted to a range consistent with neurophysiological estimates for peripheral processing times (Stein and Meredith 1993; Groh and Sparks 1996), i.e., $5 \leq 1/\lambda \leq 150$. The width of the time window of integration, ω_{20}, ω_{50}, and ω_{80}) was restricted to a positive real number with an upper bound of 600 ms. Mean time for the second stage, μ, was restricted to be positive.

TWIN model fit to the data

Parameters for the model were estimated individually for each participant by minimizing the objective function defined in Eq. 4. As an example illustrating the results, Fig. 5 shows that the model gives a good fit to the data of Participant 2. Considering all participants' data reveals that a few data points deviate from the model predictions in nonsystematic ways, but the overall pattern of the data is captured by the model, and parameter estimates are within a plausible range. The estimates and χ^2 values for all participants are listed in Table 4.

Note that we were not aiming at formal statistical tests of TWIN with this study. Rather, in order to probe the assumption that the effect of prior knowledge can be represented by an appropriate change in the time window width (ω), the focus here is on the ω parameter value as a function of prior probability p. An increase in window size ω implies an increase in the probability of integration, leading to faster responses given all other parameters remain invariant. Thus, in order to be consistent with our hypothesis, ω_{80} should not be smaller than ω_{20}.

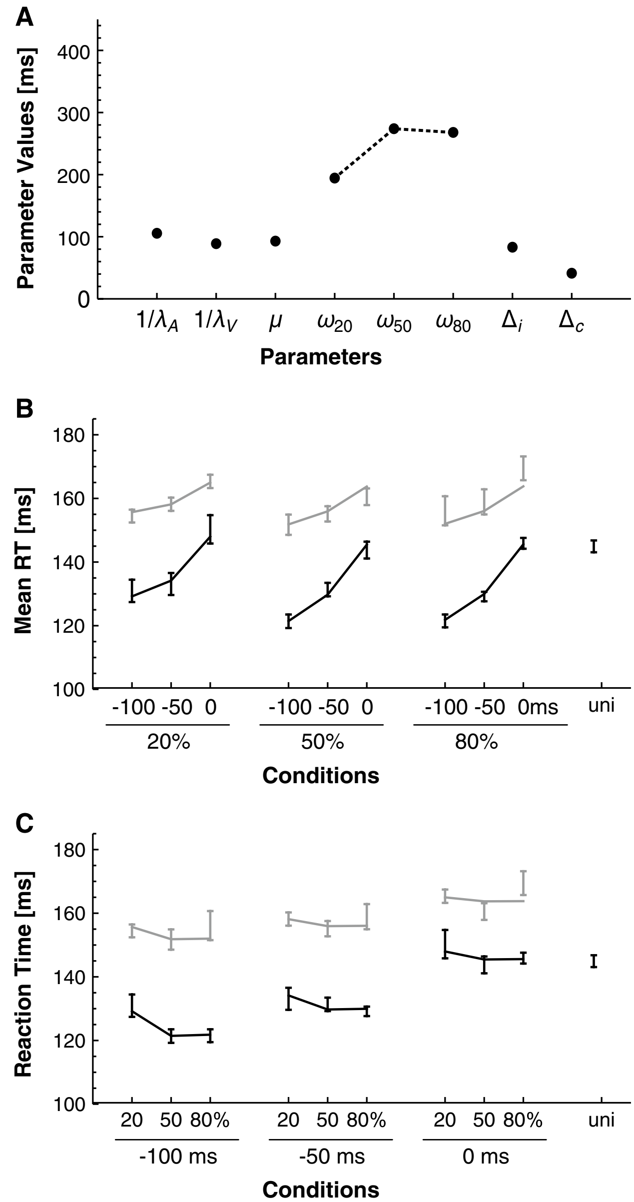


Fig. 5 Parameter estimation results for the same sample participant. **a** Estimated parameters. In this participant, the hypothesized decrease in RTs with increasing Prior knowledge level is reflected in an increase in the ω value between the 20 and the common level of the 50 and the 80 % condition. **b**, **c** Reaction times for that participant (same as in Fig. 3b, c). Here, *solid lines indicate* the estimated mean reaction times predicted by the parameters shown in **a**. Differences between the lines and the centers of the error bars denote the errors of the parameter estimation process and are meant to be minimized

According to Table 4, only Participants 1 and 3 have ω_{20} clearly larger than ω_{80} (by 86 and 240 ms, respectively), that is, values inconsistent with the hypothesis. Note that these two participants were also the only ones not showing a significant mean RT difference between the conditions of prior probability $p = .20$ and $p = .80$ at SOA$= -100$ (see Table 3). For Participants 5 and 9, ω_{20} is

[7] The *fminsearchbnd* routine is similar to the standard *fminsearch* routine except that the range of the parameters of the parameters can be predetermined, for instance, positive real numbers for the residuals or for the time window width. The *fminsearch* uses the Nelder–Mead simplex search method of Lagarias et al. (1998).

Table 4 Estimated parameters individually for each participant. The last column contains the χ^2 values as a measure of goodness of fit

Participant	λ_A^{-1}	λ_V^{-1}	μ	ω_{20}	ω_{50}	ω_{80}	Δ_I	Δ_C	χ^2
1	110.6	101.7	112.4	283.1	247.9	195.7	110.7	84.9	88.4
2	105.5	88.8	92.9	194.4	273.9	267.9	83.1	41.2	7.7
3	47.5	70.6	102.6	472.2	314.6	229.7	76.1	50.5	47.1
4	58.1	92.6	94.5	253.0	296.2	320.0	94.2	75.1	115.2
5	105.7	83.8	90.0	187.2	327.4	183.9	82.4	43.9	99.2
6	149.9	80.2	104.3	141.1	135.3	557.6	57.5	26.1	18.6
7	89.9	95.7	109.3	141.5	283.1	282.4	98.4	47.2	13.9
8	100	83.2	100.6	187.7	295.2	282.6	88.4	63.7	57.6
9	79.2	111.5	110.7	264.8	535.7	257.1	95.6	77.9	15.2
10	80.4	105.1	73.6	211.7	270.6	268.9	60.4	50.0	12.3
11	40.5	88.1	111.4	197.9	206.9	457.5	73.1	58.9	26.9
12	83.5	86.7	136.3	129.5	189.2	191.3	113.6	55.0	18.5

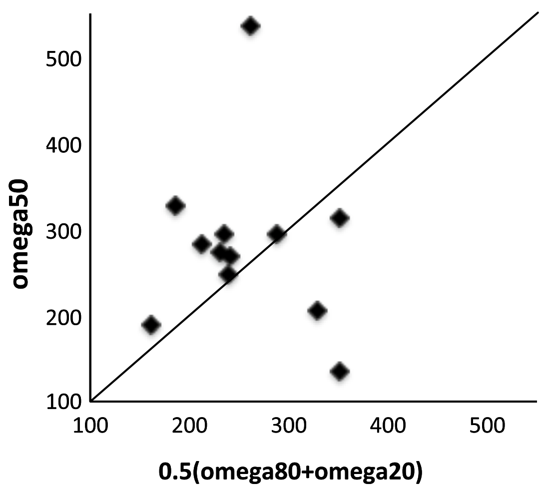


Fig. 6 Prediction of ω_{50} from ω_{20} and ω_{80} According to Eq. (3), all points should ideally lie on the bisecting ($y = x$) line if participants used optimal window sizes under all three conditions

larger than ω_{80} by less than 10 ms, whereas the remaining eight participants feature windows that are consistent with the hypothesis. Given the relatively large variability inherent in estimating the time window parameters (see Kandil et al. 2014), one would not expect the pattern of window sizes to replicate exactly the observed mean RT values, even assuming our assumption that prior knowledge can be represented exclusively by the ω parameter to be exactly true.[8]

Next, we consider the prediction of the relation between time window widths using Eq. (3). Figure 6 shows a scatterplot of the ω_{50} estimates against the values predicted by computing the arithmetic mean of the ω_{20} and ω_{80} estimates. An inspection by eye reveals that for 8 out of 12 participants the prediction holds up within a margin of 100 ms. Again, given the variability inherent in estimating the time window parameters, one would not expect all points to be close to the bisecting ($y = x$) line in Fig. 6. In addition, a further analysis using Eq. (2) shows that the window estimates in Table 4 differ from the optimal values for some participants by up to 200 ms.

[8] Extended simulation studies would be required to probe the statistical significance of the observed differences in ω values.

Discussion and conclusion

The purpose of this focused-attention study, with a visual target and an acoustic nontarget being presented either ipsi- or contralateral, was (1) to investigate the influence of prior knowledge on crossmodal RT facilitation and (2) to test how the time-window-of-integration (TWIN) model (Colonius and Diederich 2004; Diederich and Colonius 2004) can account for this. Specifically, we tested the hypothesis that facilitation increases with the level of prior knowledge of target–nontarget spatial alignment. Prior knowledge was established by presenting spatially aligned vs. nonaligned (ipsi- vs. contralateral) stimuli in a fixed percentage of trials (20, 50, or 80 %) within a block of 156 or 180 trials. Ten out of 12 participants showed significantly faster mean RTs (ipsilateral) in the 80 % compared to the 20 % Prior knowledge condition at SOA = −100 ms. In other words, our results support the idea that the amount of facilitation does depend not only on the specific spatiotemporal crossmodal stimulus conditions in an experimental setting but also on the participants' prior knowledge of alignment contingency within an extended block of trials.

Most of the 10 participants supporting the hypothesis did not show a clear gradation of their amount of facilitation into three levels but only two: There was no difference either between the 20 and 50 % or the 50 and 80 %

conditions, depending on participant (see Table 5). Adopting the idea that spatial alignment of visual and acoustic stimuli increases the crossmodal perception of "unity" or "common cause" of the components (e.g., Körding et al. 2007), our findings suggest that prior knowledge becomes manifest in its effect on perceiving "unity" at two levels only, "often" or "rarely," say.

Further insight into the effect of prior knowledge on response times was gained by presenting the acoustic nontarget at different time points relative to the target (SOAs). First, for both ipsi- and contralateral presentations, we replicated the typical finding that facilitation increases the earlier the nontarget is presented, at least within the critical range of SOAs where facilitation tends to occur (e.g., Frens et al. 1995); see Panel A of Fig. 2. Second, and more interesting, was a statistically significant interaction: The effect of prior knowledge on facilitation was stronger for SOA $= -100$ ms than for SOA $= -50$ ms or SOA $= 0$ ms, but this was observed only for ipsilateral presentations. For example, in panel B of Fig. 2, mean RT for $p = .20$ is clearly longer than for $p = .50$ or $p = .80$ at SOA $= -100$ ms, and the difference levels off with the other two SOA values (ipsilateral presentation). Thus, it seems that it takes some time to process the nontarget (position) in order for prior knowledge to become effective in ipsilateral presentations, whereas responses to contralateral presentations—while still benefiting from early nontarget appearance—do not show this differential effect with respect to the level of Prior knowledge (Fig. 2, contralateral presentation). Moreover, the absence of this effect in the contralateral presentations is probably also due to the randomized SOAs not allowing the buildup of the perception of a unitary event.

In order to better understand this pattern of results, one may speculate about the level at which processing of the nontarget takes place. Note that, in principle, the $p = .80$ and the $p = .20$ are "symmetric" conditions: In either case, localization of the nontarget allows the participant to "predict" the target position with a probability of $p = .80$, alignment in the first case and nonalignment in the second. However, this symmetry was clearly not reflected in our data, and there may be a reason for this: in order to take advantage of prior knowledge in the $p = .20$ ipsilateral condition ("alignment occurs rarely") participants would have to be "prepared" to make an eye movement away from the acoustic nontarget (i.e., an antisaccade). Antisaccades are known to take longer than saccades (Munoz and Everling 2004; Everling and Fischer 1998; Hallett 1978) because a reflexive response must first be suppressed,[9] so even if a prior knowledge advantage had been present, it might have been masked by the slower antisaccade mechanism. In a similar vein, the tendency of several participants to exhibit slower mean RTs to contralateral presentations in the $p = .80$ condition, compared to $p = .50$ or $p = .20$, may be due to the antisaccade effect.

In contrast to the present results, some evidence of an effect of symmetry between conditions $p = .80$ and the $p = .20$ was found in an earlier saccadic eye movement study from our laboratory (Kirchner and Colonius 2005). In a setup quite similar to the present one, we observed clear effects of an auditory nontarget on saccadic responses to a visual target as follows: Whenever Prior knowledge of the target position was either 80 or 20 %, corresponding to conditions $p = .80$ and the $p = .20$ here, saccadic RT was faster compared to 50 % condition ($p = 0.50$). The effect was highly time-dependent, only occurring within a ± 40 ms time interval between target and nontarget presentation. The difference to the present results may arguably be due to the fact that the presentation blocks in the Kirchner study contained 10 % catch trials (unimodal) leading to slower responses (40 ms on average), possibly supporting the occurrence of the "antisaccade effect" mentioned in the previous paragraph. An extension of the TWIN model to account for such data has not yet been investigated.

In sum, identification of the exact eye movement dynamics seems difficult without further systematic variation of the experimental conditions. It has been hypothesized that the effect of prior knowledge becomes more salient under degraded stimulation conditions (Stocker and Simoncelli 2006; Senna et al. 2015). Thus, further insight into the role of prior knowledge might be gained by designing a study that systematically varies, e.g., stimulus intensity, while keeping the level of prior knowledge constant.

On the other hand, without having to specify details of the saccadic mechanism, the TWIN model allows us to describe the effect of prior knowledge on RT facilitation via the multisensory integration time window concept. One major finding of this study comes from fitting the TWIN model to the data (separately for each participant): Keeping all model parameters constant across the three levels of Prior knowledge, with the exception of the window width parameter (ω), it turned out that the observed changes in facilitation due to the level of Prior knowledge could be captured by appropriate changes in the ω parameter. That is, the modeling results suggest that, when Prior knowledge points to a frequent occurrence of spatial alignment, or "unity" of visual and acoustic events ($p = .80$ or $p = .50$), participants adjust their time window width to a larger value than in the case of "rare" alignment. Although this shows that participants were, to some degree, able to adapt to changing

[9] However, empirical evidence for antisaccades in response to an acoustic stimulus is rare (but see Goepel 2010).

conditions, we did not find evidence for optimality of their window sizes, in the sense of a Bayesian modeling approach. One reason may be that participants did not adjust the time window optimally, but it may as well simply be the case that our parameter estimates were too far off the true values. Moreover, whether or not optimal window adaptation would require participants to be aware of their prior knowledge cannot be decided on the basis of our data.

One of the reviewers suggested that rather than perceiving "unity" at two levels only, like "often" or "rarely," there may in reality be a "graded sensitivity" but that the experimental design did not have enough power to detect it. This cannot be dismissed and it is consistent with the TWIN model framework as follows. Within the model, according to Eq. (1) facilitation is the product of amount of facilitation (parameter Δ) and probability of integration $P(I)$, with the latter being a continuous function of window size ω. Thus, prior knowledge affects the probability of integration, i.e., of perceiving "unity" via window size in a graded (i.e., continuous) way.

Prior knowledge, defined by the probability of spatial alignment, is but one example of top-down information with an impact on response behavior under identical stimulus conditions in a crossmodal setting (for review, see Talsma et al. 2010). There is also ample evidence for nonspatial effects, e.g. the difference between focused-attention and redundant-signals paradigms (e.g., Colonius and Diederich 2012), or the effect stop signal probability in a crossmodal stop signal task (Özyurt et al. 2003).

Turning to the issue of neural correlates of the observed behavior, it has been shown in many studies that neural activity in primary motor and in premotor cortex is correlated with changes in reaction time, depending on the amount of information conveyed by a stimulus providing complete, partial, or no information about the direction of a required movement. For example, Dorris and Munoz (1998) had monkeys perform a saccadic task in which the probability of the required saccade being directed into the response field of a neuron varied systematically between blocks of trials (100, 50 or 0 %); they found significant negative correlations between the discharge rate of neurons in the intermediate layers of the superior colliculus and saccadic RT (see also Bell and Munoz 2008 for similar results with 80 % stimulus validity).

While the body of the literature on the neural correlates of prior knowledge or, more generally, of top-down effects on saccadic responses, is vast (for recent review, Corneil and Munoz 2014), most experimental paradigms, often conducted on nonhuman primates and only within the visual modality, differ significantly from the one presented here. A study somewhat closer to our paradigm is by Cohen and colleagues (Cohen et al. 2004) on the lateral intraparietal area (area LIP) of rhesus monkeys. In what they called "predictive-cueing task" they measured saccadic latencies to a visual target presented 200–250 ms after a visual or an auditory cue was presented. Target and cue (nontarget) could be in either of two different locations, and the cue was 80 % predictive. For both of the monkeys tested, they found faster responses when the target was at the "predicted" location, but there was no difference between trials using visual or auditory cues. Notably, in a neutral condition the sensory cues did not systematically affect the monkeys' responses. Recording the activity of 96 LIP neurons during the predictive-cueing task, Cohen et al. found a negative correlation of LIP activity and mean saccade latency. The firing rate of LIP neurons was higher during "predictive" trials than during "nonpredictive" trials. The authors caution, however, that this may stem from a memory trace of the cue that persists following cue offset. Despite no difference in saccadic latency between visual and auditory cues, visual cues elicited higher firing rates than auditory cues, on average. "But, as the task unfolded, these modality-dependent differences were minimized. These observations suggest that the quantity coded by LIP neurons during a behavioral task is dynamic and changes as the sensory, cognitive, or motor demands of a task change." (Cohen et al. 2004, p. 1299).

Given these and similar results from other studies, it seems important to keep in mind the conceptual and technical complexity of claims identifying certain neural structures with cognitive capacities for the domain of saccadic eye movements, as discussed in a critical review by Schall (2004).

Acknowledgments A.D. was supported by German Research Foundation (DFG) Grant DI 506/12-1. H.C. was supported by Cluster of Excellence "HearingForAll" to Oldenburg University and SFB/TRR-31 "Active Listening" (Project B4) funded by the German Research Foundation (DFG). We further thank our student assistants, P. Bogdanova, S. Cheishvili, and A.N. Dinu for their help with collecting the data. Helpful comments from four anonymous reviewers are gratefully acknowledged.

Appendix

See Tables 5 and 6.

Table 5 Mean RTs and standard errors in all conditions (SOA, prior probability) for all 12 participants

Participant	SOA	Prior knowledge condition					
		20 %		50 %		80 %	
		Mean RT	SE	Mean RT	SE	Mean RT	SE
1	−100	144.5	3.2	140.0	2.8	140.7	1.9
	−50	150.8	3.8	147.2	1.9	156.4	2.4
	0	162.6	2.7	166.4	2.1	168.5	1.7
	[−125 to (−76)]	152.2	1.7	156.7	1.9	193.1	5.8
	[−75 to (−26)]	159.2	1.9	161.8	2.2	185.9	4.4
	[−25 to + 24]	172.6	1.7	178.2	2.4	195.4	4.9
2	−100	130.9	3.5	121.4	2.1	121.5	2.0
	−50	133.1	3.5	131.3	2.1	129.2	1.5
	0	150.3	4.5	143.8	2.6	145.9	1.7
	[−125 to (−76)]	154.4	2.0	151.7	3.2	156.1	4.6
	[−75 to (−26)]	158.2	2.1	155.1	2.4	158.9	3.9
	[−25 to + 24]	165.3	2.1	160.5	2.6	169.5	3.9
3	−100	109.9	3.4	108.2	2.2	106.5	2.5
	−50	105.1	2.4	107.4	2.0	112.0	1.6
	0	128.5	3.3	125.5	1.9	127.0	1.5
	[−125 to (−76)]	120.6	1.8	119.4	2.1	138.5	6.5
	[−75 to (−26)]	131.6	1.4	132.4	1.5	135.2	3.1
	[−25 to + 24]	144.8	1.3	143.6	1.6	146.6	2.3
4	−100	114.3	3.6	111.0	2.0	104.9	1.7
	−50	119.7	4.5	110.8	1.5	107.0	1.6
	0	130.0	2.3	130.4	1.7	128.7	1.4
	[−125 to (−76)]	119.7	1.9	117.1	1.7	118.5	2.7
	[−75 to (−26)]	130.4	1.6	129.0	1.6	130.8	2.0
	[−25 to + 24]	143.6	1.3	146.1	1.6	147.6	3.2
5	−100	126.0	2.4	116.7	1.8	118.9	1.3
	−50	128.5	1.8	117.3	1.5	126.5	1.1
	0	143.0	2.9	137.6	1.2	144.1	0.9
	[−125 to (−76)]	144.3	1.2	141.5	1.3	159.9	2.5
	[−75 to (−26)]	149.4	1.3	144.7	1.1	159.2	2.4
	[−25 to + 24]	153.7	1.3	150.6	1.6	157.8	1.8
6	−100	159.6	4.5	156.4	3.6	144.6	2.0
	−50	159.1	3.6	163.6	3.6	155.0	1.7
	0	163.3	3.5	169.0	3.2	166.0	2.1
	[−125 to (−76)]	171.7	2.3	174.1	3.7	171.2	3.3
	[−75 to (−26)]	175.0	2.4	174.6	2.5	167.1	2.9
	[−25 to + 24]	180.9	2.5	176.4	2.1	167.5	2.9
7	−100	149.0	5.8	127.6	3.0	131.1	2.8
	−50	162.3	5.2	144.5	3.8	137.6	1.8
	0	169.6	4.8	157.4	4.0	154.3	2.5
	[−125 to (−76)]	179.2	2.3	168.0	3.9	154.3	4.0
	[−75 to (−26)]	182.0	2.5	172.6	2.7	176.3	5.1
	[−25 to + 24]	188.1	2.8	179.7	3.9	172.9	7.0

Table 5 continued

Participant	SOA	Prior knowledge condition					
		20 %		50 %		80 %	
		Mean RT	SE	Mean RT	SE	Mean RT	SE
8	−100	129.3	2.1	121.0	1.5	115.0	1.2
	−50	136.1	2.4	126.0	1.6	127.2	1.2
	0	149.1	2.4	142.3	1.7	144.3	1.1
	[−125 to (−76)]	140.1	1.0	133.6	1.4	142.5	2.2
	[−75 to (−26)]	148.6	1.0	141.9	1.1	149.1	2.2
	[−25 to + 24]	160.0	1.1	153.1	1.3	157.1	1.5
9	−100	152.5	5.5	143.9	4.7	151.7	3.1
	−50	161.0	5.9	148.8	4.1	155.1	2.9
	0	181.3	8.4	159.3	3.2	173.2	2.8
	[−125 to (−76)]	162.4	2.1	153.4	3.4	160.7	4.7
	[−75 to (−26)]	168.3	2.4	166.0	4.2	169.9	4.7
	[−25 to + 24]	181.5	3.2	177.3	4.0	192.2	5.7
10	−100	139.4	2.7	135.9	1.8	131.8	1.3
	−50	140.0	2.5	134.5	1.6	136.6	1.2
	0	147.2	2.8	146.5	1.5	147.2	1.2
	[−125 to (−76)]	143.7	1.2	138.6	1.4	142.1	2.1
	[−75 to (−26)]	147.5	1.2	144.2	1.4	144.0	1.8
	[−25 to + 24]	154.7	1.3	153.4	1.4	150.1	2.2
11	−100	149.5	5.2	142.1	2.1	128.4	2.1
	−50	144.3	4.3	145.9	2.5	133.0	2.1
	0	173.7	9.2	155.9	2.8	146.6	1.9
	[−125 to (−76)]	152.6	1.7	154.9	3.1	148.6	3.4
	[−75 to (−26)]	158.3	2.7	150.5	1.5	147.5	2.5
	[−25 to + 24]	164.9	1.6	164.1	1.4	155.8	2.4
12	−100	181.7	11.0	151.0	5.3	145.8	6.3
	−50	158.7	8.8	148.5	4.3	151.3	5.0
	0	187.1	6.1	170.8	4.0	168.2	4.6
	[−125 to (−76)]	189.3	3.4	184.4	4.0	196.6	11.7
	[−75 to (−26)]	197.0	3.6	189.1	3.5	206.8	11.8
	[−25 to + 24]	209.3	4.3	196.0	4.0	190.4	8.6

For each participant, the first three rows show the data from the ipsilateral configuration with fixed SOAs, the other three rows those from the contralateral configurations with SOAs varying within the given range

Table 6 Error types (percentages) for each participant under all experimental conditions

Participant	Error	Prior knowledge condition						Overall
		20 %		50 %		80 %		
		ipsi	Contra	ipsi	Contra	ipsi	Contra	
1	Amplitude not within 3 std	0	.1	.4	.6	.2	1.9	5.1
	RT < 80	.5	0	0	0	2.5	1.4	
	Directional	8.8	6.9	0	.2	.8	.5	
2	Amplitude not within 3 std	.5	2.2	.2	1.9	.1	3.2	7.4
	RT < 80	2.8	1.3	7.5	0	9.8	.5	
	Directional	0	1.9	0	3.0	.1	9.3	
3	Amplitude not within 3 std	0	1.6	.2	1.9	1.0	3.2	20.4
	RT < 80	14.4	7.1	17.7	5.8	20.1	8.3	
	Directional	0	8.9	.2	10.0	0	17.1	
4	Amplitude not within 3 std	1.4	2.1	.6	1.9	2.5	3.7	12.8
	RT < 80	6.5	3.8	10.1	6.4	14.5	13.4	
	Directional	0	1.7	0	1.7	0	2.8	
5	Amplitude not within 3 std	.5	.9	0	2.4	.2	1.4	3.8
	RT < 80	2.8	.9	3.4	0	2.7	0	
	Directional	0	1.2	0	2.4	0	2.8	
6	Amplitude not within 3 std	2.8	1.4	1.7	2.1	2.0	2.8	4.2
	RT < 80	.9	.5	1.5	0	1.4	0	
	Directional	0	.2	.2	.4	0	0	
7	Amplitude not within 3 std	4.6	4.2	2.1	2.4	0	0	12.9
	RT < 80	2.3	1.0	4.7	1.9	7.9	4.2	
	Directional	.5	2.2	.9	7.1	2.7	8.3	
8	Amplitude not within 3 std	0	.2	.2	.9	.3	2.8	2.0
	RT < 80	.5	.8	.9	0	2.2	0	
	Directional	0	0	0	.2	0	2.3	
9	Amplitude not within 3 std	2.3	2.4	2.8	3.0	.8	.9	10.8
	RT < 80	6.5	2.3	7.7	3.0	3.7	2.8	
	Directional	0	1.4	.6	6.8	.7	3.2	
10	Amplitude not within 3 std	1.9	.8	0	.2	.1	1.4	1.6
	RT < 80	.5	0	.6	0	1.9	0	
	Directional	0	.2	0	1.1	0	.5	
11	Amplitude not within 3 std	0	1.0	.4	1.9	1.0	2.8	4.8
	RT < 80	.9	.1	.9	.4	2.9	0	
	Directional	.5	.1	0	0	.1	.5	
12	Amplitude not within 3 std	0	0	2.4	1.7	.2	0	21.0
	RT < 80	9.3	2.1	6.2	.4	11.7	1.4	
	Directional	1.4	10.8	.6	12.4	1.7	19.9	

These trials were excluded from further analysis

References

Amlôt R, Walker R, Driver J, Spence C (2003) Multimodal visual-somatosensory integration in saccade generation. Neuropsychologia 41:1–15

Arndt PA, Colonius H (2003) Two stages in crossmodal saccadic integration: evidence from a visual-auditory focused attention task. Exp Brain Res 150:417–426

Bell AH, Munoz DP (2008) Activity in the superior colliculus reflects dynamic interactions between voluntary and involuntary influences on orienting behaviour. Eur J Neurosci 28:1654–1660

Charbonneau G, Vèronneau M, Boudiras-Fournier C, Lepore F, Collignon O (2013) The ventriloquist in periphery: impact of eccentricity-related reliability on audio-visual localization. J Vis 13(12):1–14

Cohen YE, Cohen IS, Gifford GW III (2004) Modulation of lip activity by predictive auditory and visual cues. Cereb Cortex 14:1287–1301

Colonius H, Arndt PA (2001) A two-stage model for visual—auditory interaction in saccadic latencies. Percept Psychophys 63:126–147

Colonius H, Diederich A (2004) Multisensory interaction in saccadic reaction time: a time-window-of-integration model. J Cogn Neurosci 16:1000–1009

Colonius H, Diederich A (2010) The optimal time window of visual-auditory integration: a reaction time analysis. Front Integr Neurosci 4:11

Colonius H, Diederich A (2011) Computing an optimal time window of audiovisual integration in focused attention tasks: illustrated by studies on effect of age and prior knowledge. Exp Brain Res 212(3):327–337

Colonius H, Diederich A (2012) Focused attention vs. crossmodal signals paradigm: deriving predictions from the time-window-of-integration model. Front Integr Neurosci 6:62. doi:10.3389/fnint.2012.00062

Colonius H, Diederich A, Steenken R (2009) Time-window-of-integration (twin) model for saccadic reaction time: effect of auditory masker level on visual-auditory spatial interaction in elevation. Brain Topogr 21:177–184

Corneil BD, Munoz DP (1996) The influence of auditory and visual distractors on human orienting gaze shifts. J Neurosci 16:8193–8207

Corneil BD, Munoz DP (2014) Overt responses during covert orienting. Neuron 82:1230–1243

Diederich A, Colonius H (2004) Modeling the time course of multisensory perception in manual and saccadic reponses. In: Calvert G, Spence C, Stein EB (eds) Handbook of multisensory processes, chap 24. MIT Press, Cambridge, pp 395–408

Diederich A, Colonius H (2007a) Modeling spatial effects in visual-tactile reaction time. Percept Psychophys 69(1):56–67

Diederich A, Colonius H (2007b) Why two "distractors" are better then one: modeling the effect on non-target auditory and tactile stimuli on visual saccadic reaction time. Exp Brain Res 179:43–54

Diederich A, Colonius H (2008a) Crossmodal interaction in saccadic reaction time: separating multisensory from warning effects in the time window of integration model. Exp Brain Res 186(1):1–22

Diederich A, Colonius H (2008b) When a high-intensity "distractor" is better then a low-intensity one: modeling the effect of an auditory or tactile nontarget stimulus on visual saccadic reaction time. Brain Res 1242:219–230

Diederich A, Colonius H, Schomburg A (2008) Assessing age-related multisensory enhancement with the time-window-of-integration model. Neuropsychologia 46(10):2556–2562

Dorris MC, Munoz DP (1998) Behavior probability influences motor preparation signals and time to saccadic initiation. J Neurosci 18(17):7015–7026

Ernst MO (2006) A bayesian view on multimodal cue integration. In: Knoblich G, Thornton IM, Grosjean M, Shiffrar M (eds) Human body perception from the inside out, chap 6. Oxford University Press, New York, pp 105–131

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429–433

Everling S, Fischer B (1998) The antisaccade: a review of basic research and clinical studies. Neuropsychologia 36(9):885–899

Frens MA, van Opstal AJ, van der Willigen RF (1995) Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. Percept Psychophys 57(6):802–816

Goepel J (2010) Antisaccades elicited by visual and acoustic cues—an investigation of children with and without attention deficit hyperactivity disorder. PhD thesis, University of Konstanz, dissertation zur Erlangung des akademischen Grades des Doktors der Naturwissenschaften (Dr. rer. nat.)

Goldring JE, Dorris MC, Corneil BD, Ballantyne PA, Munoz DP (1996) Combined eye-head gaze shifts to visual and auditory targets in humans. Exp Brain Res 111:68–78

Groh JM, Sparks DL (1996) Saccades to somatosensory targets: I. Behavioral characteristics. J Neurophysiol 75:412–427

Hallett PE (1978) Primary and secondary saccades to goals defined by instructions. Vis Res 18:1279–1296

Harrington LK, Peck CK (1998) Spatial disparity affects visual auditory interactions in human sensorimotor processing. Exp Brain Res 122:247–252

Heuermann H, Colonius H (2001) Spatial and temporal factors in visual-auditory interaction. In: Sommerfeld E, Kompass R, Lachmann T (eds) Proceedings of the seventeenth meeting of the international society for psychophysics, Pabst Science Publisher, pp 118–123

Hughes HC, Nelson DM (1998) Spatial characteristics of visual—auditory summation in human saccades. Vis Res 38:3955–3963

Hughes HC, Reuter-Lorenz PA, Nozawa G, Fendrich R (1994) Visual-auditory interactions in sensorimotor processing: saccades versus manual responses. J Exp Psychol Hum Percept Perform 20:131–153

Kandil FI, Diederich A, Colonius H (2014) Parameter recovery for the time-window-of-integration (twin) model of multisensory integration in focused attention. J Vis 14(11):1–20

Kirchner H, Colonius H (2005) Interstimulus contingency facilitates saccadic responses in a bimodal go/no-go task. Cognit Brain Res 25:261–272

Körding KP (2007) Decision theory: What "should" the nervous system do? Science 318(5850):606–610. doi:10.1126/science.1142998

Körding KP, Beierholm U, Ma WJ, Tenenbaum JB, Shams L (2007) Causal inference in multisensory perception. PLoS One 2:e943

Lagarias JC, Reeds JA, Wright MH, Wright PE (1998) Convergence properties of the nelder-mead simplex method in low dimensions. SIAM J Optim 9:112–147

Magnotti JF, Ma WJ, Beauchamp MS (2013) Causal inference of asynchronous audiovisual speech. Front Psychol 4, article 798

Mendonça C, Hiipakka M, van de Par S, Pulkki V, Colonius H (2014) Adaptation to non-individualized spatial sound through audiovisual experience. In: Proceedings AES 55th conference, Helsinki. 27–29 Aug 2014

Mendonça C, Escher A, van de Par S, Colonius H (2015) Predicting auditory space calibration from recent multisensory experience. Exp Brain Res 233:1983–1991

Munoz DP, Everling S (2004) Look away: the antisaccade task and the voluntary control of eye movement. Nat Rev Neurosci 5:218–228

Nozawa G, Reuter-Lorenz PA, Hughes HC (1994) Parallel and serial processes in the human oculomotor system: bimodal integration and express saccades. Biol Cybern 72:19–34

Özyurt J, Colonius H, Arndt PA (2003) Countermanding saccades: evidence against independent processing of go and stop signals. Percept Psychophys 65(3):420–428

Posner MI (1980) Orienting of attention. Q J Exp Psychol 32:3–25

Sarmiento BR, Matusz PJ, Sanabris D, Murray MM (2015) Contextual factors multiplex to control mutlisensory processes. Hum Brain Map. doi:10.1002/hbm.23030

Schall JD (2004) On building a bridge between brain and behavior. Annu Rev Psychol 55:23–50

Senna I, Parise CV, Ernst EO (2015) Hearing in slow-motion: humans underestimate the speed of moving sounds. Sci Rep 5:14054. doi:10.1038/srep14054

Shams L, Beierholm U (2010) Causal inference in perception. Trends Cogn Sci 14:425–432

Spence C, Driver J (1997) Audiovisual links in exogenous covert spatial orienting. Percept Psychophys 59:1–22

Steenken R, Colonius H, Diederich A, Rach S (2008) Visual-auditory interaction in saccadic reaction time: effects of auditory masker level. Brain Res 1220:150–156

Stein BE, Meredith MA (1993) The merging of the senses. MIT Press, Cambridge

Stocker AA, Simoncelli EP (2006) Noise characteristics and prior expectations in human visual speed perception. Nat Neurosci 9(4):578–585

Van der Stoep N, Spence C, Nijboer TW, Van der Stigchel S (2015) On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement. Acta Psychol 162:20–28

Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG (2010) The multifaceted interplay between attention and multisensory integration. Trends Cognit Sci 14(9):400–410

Ten Brink AF, Nijboer TCW, Van der Stoep N, Van der Stigchel S (2014) The influence of vertically and horizontally aligned visual distractors on aurally guided saccadic eye movements. Exp Brain Res 232:1357–1366

Van Wanrooij MM, Bell AH, Munoz DP, Van Opstal AJ (2009) The effect of spatial-temporal audiovisual disparities on saccades in a complex scene. Exp Brain Res 198(2–3):425–437

Van Wanrooij MM, Bremen P, van Opstal AJ (2010) Acquired prior knowledge modulates audiovisual integration. Eur J Neurosci 31(10):1763–1771

Vossel S, Mathys C, Daunizeau J, Bauer M, Driver J, Friston KJ, Stephan KE (2014) Spatial attention, precision, and bayesian inference: a study of saccadic response speed. Cereb Cortex 24:1436–1450

Whitchurch EA, Takahashi TT (2006) Combined auditory and visual stimuli facilitate head saccades in the barn owl (tyto alba). J Neurophysiol 96(2):730–745

Wozny DR, Shams L (2011) Recalibration of auditory space following milliseconds of crossmodal discrepancy. J Neurosci 31(12):4607–4612