

RUNNING HEAD: MULTISENSORY SELECTION AND HIGHER-ORDER COGNITION

Higher-order cognition does not affect multisensory distractor processing

Simon Merz¹, Anne Jensen^{1,3}, Charlotte Bureau¹,
Charles Spence², & Christian Frings¹

¹ University of Trier,

² University of Oxford,

³ Shared first authorship

Words count (3915)

Corresponding author:

Simon Merz

University of Trier

D-54286 Trier

merzs@uni-trier.de

Authors' Note

Anne Jensen, Simon Merz, Charlotte Bureau, and Christian Frings, University of Trier, Department of Psychology, D-54286, Germany; Charles Spence, Crossmodal Research Laboratory, Department of Experimental Psychology, Anna Watts Building, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, England.

The research reported in this article was supported by a grant from the Deutsche Forschungsgemeinschaft to Christian Frings and Charles Spence (FR 2133/5-3).

Correspondence concerning this article should be addressed to Simon Merz, University of Trier, Department of Psychology, D-54286, Germany. Email: merz.s@uni-trier.de.

Abstract

Multisensory processing is required for the perception of the majority of everyday objects and events. In the case of irrelevant stimuli, the multisensory processing of features is widely assumed to be modulated by attention. In the present study, we investigated whether the processing of audiovisual distractors is also modulated by higher-order cognition. Participants fixated a visual distractor viewed via a centrally-placed mirror and responded to a laterally-presented audiovisual target. Critically, a distractor tone was presented from the same location as the mirror, while the visual distractor feature was presented at an occluded location, visible only indirectly via mirror reflection. Consequently, it *appeared* as though the visual and auditory features were presented from the same location though, in fact, they actually *originated* from different locations. Nevertheless, the results still revealed that the visual and auditory distractor features were processed together just as in the control condition, in which the audiovisual distractor features were both actually presented from fixation. Taken together, these results suggest that the processing of irrelevant multisensory information is not influenced by higher-order cognition.

Keywords: Multisensory perception, multisensory selection, distractor processing, higher-order cognition.

Introduction

When you catch sight of yourself in the mirror, you immediately know who it is that is looking at you. This self-recognition ability, commonly observed amongst primates (e.g., Gallup et al., 2002; Gatti, 2016; Mitchell, 1993; Parker et al., 1994; although there are also a few other classes of animals who exhibit this ability), reflects *higher-order cognition* (HOC). HOC typically refers to those top-down cognitive processes that influence perception, attention, and, ultimately, performance (Miller & Wallis, 2009; Palmer, 1975; see Halford, Wilson, & Phillips, 2010, for a review). ‘Top-down’ is typically used to refer to a person’s knowledge, or will, whereas ‘bottom-up’ is determined by the features of the environment, i.e., stimulus-driven (see Evans & Stanovich, 2013, for a recent review). The present study investigates whether presenting part of the distracting information by means of a mirror influences the processing of distractors in a multisensory selection task. In particular, multisensory distractor processing might be modulated when the visual component of an audiovisual distractor is seen indirectly via mirror-reflection and the observer is aware that the auditory and visual features actually originate from different environmental locations.

Using mirrors, previous research has demonstrated that the interpretation of the reflected input determines perception in humans. This influence goes either back to top-down (e.g., Helbig & Ernst, 2007; Maravita et al., 2002; Sambo & Forster, 2011) or bottom-up modulations (e.g., Franz & Packman, 2004; Gallace & Spence, 2005; Holmes & Spence, 2005; Soto-Faraco et al., 2004). For example, in a visuotactile adaptation of the flanker task, Maravita and his colleagues demonstrated that the influence of distracting information is changed if the information is presented via a mirror, presumably because the person knew that the stimuli were actually presented from close to their own body, thus demonstrating a top-down influence. Furthermore, Helbig and Ernst demonstrated that visual and tactile inputs relating to the same object were integrated when the visual input was presented via a mirror placed elsewhere. Even

1 though the visual input appeared to be presented from a different location than the tactile input,
2 the mirror revealed that both inputs actually *originated* from the same spatial location.

3 On the other hand, Soto-Faraco and his colleagues (2004) documented a bottom-up
4 modulation of participants' performance in a tactile spatial flanker task. In particular, a mirror
5 was placed close to the target hand, leading to the perception of a small distance between hand
6 and its mirror reflection. Although the real distractor hand was actually located much further
7 away, the interference effects observed were similar in magnitude to those seen in a condition
8 in which the distractor hand was close without there being a mirror present. Despite the
9 participants' knowledge concerning the presence of the mirror, their performance was mainly
10 driven by the perceived (close) distance induced by the mirror. It can therefore be suggested
11 that the visual reflection was mapped onto (i.e., associated with) the distractor hand, thus
12 creating a strong illusion, that HOC was unable to overcome.

13 Concerning other HOC manipulations, the quality of a barrier, separating both of the
14 participant's hands which received tactile stimulation, has been shown to influence distractor
15 interference in a tactile flanker task (Wesslein, Spence, & Frings, 2015). Specifically, the
16 response flanker effect was reduced when a transparent but impermeable barrier was placed
17 between the participant's hands, as compared to an empty frame. Thus, knowledge concerning
18 the permeability (i.e., HOC), but not spatial separation, was the key factor determining
19 distractor interference. However, another study by Kitagawa and Spence (2005) using a variant
20 of the crossmodal congruency task failed to find any effect on distractor interference with a
21 transparent barrier. Hence, crossmodal interactions might well differ from unisensory ones and
22 may not be influenced by HOC in quite the same way.

23 The present study extends the investigation of the importance of HOC into a
24 multisensory context. In particular, we used our recently-established multisensory flanker task
25 with audiovisual stimuli (Jensen et al., 2019a; see also Merz et al., 2019), in which specific

combinations of visual and auditory features (varying in terms of their color and pitch) were used as stimuli¹. The participants had to respond to specific combinations of audiovisual features that were presented from a target cube², while audiovisual stimuli from a different distractor cube had to be ignored. There was a 35° spatial separation between the two cubes. The general pattern of results observed revealed that reaction times (RTs) were fastest when both features of the distractor matched both features of the target (congruent-congruent condition), while RTs were slowest when neither feature matched (incongruent-incongruent condition). In those conditions in which there was partial feature overlap, that is, where one distractor feature matched the target while the other distractor feature did not (i.e., the congruent-incongruent conditions), RTs were somewhere in-between these two extremes. Critically, the combined processing of the distractor features, that is the signature feature of multisensory distractor processing, was only observed when the participants' gaze was directed at the distractor stimulus.

In our view, distractor processing describes the complete processing of the distractor stimulus, from early perceptual stages all the way through to late response selection stages, as in Pashler's (1994, 1998) account of the three stages of processing (1: perceptual processing; 2: central processing, e.g., response selection; 3: response execution). In our view, the features of the distractor can be processed independently; that is, the two distractor features are processed separately and are never combined during the various stages of information processing. In Jensen et al.'s (2019a) study, this was the case when the target was presented at

¹ To the best of our knowledge, no distractor processing study investigating the influence of HOC has used auditory stimuli (for the discussion of one unpublished experiment in this area, see Spence et al., 2017). In fact, although some distractor processing studies exist which have used auditory stimuli (e.g., Chan et al., 2005; Frings & Spence, 2010; Ulrich, Prislán & Miller, 2020; for a discussion, see Frings, Schneider & Moeller, 2014), far more visual and / or tactile distractor processing studies have been published (e.g., Driver & Grossenbacher, 1996; Eriksen & Eriksen, 1984, for reviews, see Merz et al., 2020; Spence, 2020; Wesslein et al, 2014). Subsequently, studies investigating the influence of HOC on distractor processing have used visual and / or tactile experimental set-ups.

² For experiments investigating multisensory selection, multisensory cubes were specifically developed (7 cm x 7 cm x 7 cm) from which visual, auditory as well as tactile information can be presented so that the information is presented from one physical object. For a detailed description, see Merz et al. (2019).

1 fixation. In contrast, the existence of multisensory distractor processing indicates that the two
2 features are combined and interact somewhere along the various stages of information
3 processing. More precisely, the magnitude of the congruency effect (that is, the difference
4 between congruent and incongruent trials) of one feature is influenced by the congruency of the
5 other feature. In other words, a significant interaction is evidenced. In Jensen et al.'s (2019a)
6 study, this data pattern was observed when the distractor was presented at fixation.

7 *The present study*

8 In the present study, we used an audiovisual flanker task (see also Jensen et al., 2019a)
9 to investigate the influence of HOC on multisensory distractor processing. Therefore, two
10 different conditions were included. In the first (control) condition, the participants gazed at the
11 distractor cube, and the onset of the distractor information preceded the onset of the target
12 information. With this set-up, multisensory distractor processing was expected, as this
13 condition is nearly identical to Experiment 2 of Jensen et al. (2019), which evidenced
14 multisensory distractor processing when the distractors were presented at fixation. In the second
15 (mirror) condition, a mirror reflecting the visual distractor feature, that was actually placed
16 elsewhere, was fixated instead of the distractor itself (see Figure 1 for an illustration of the set-
17 up). The auditory feature, on the other hand, was presented at the location of the mirror.
18 Consequently, the irrelevant visual and auditory inputs *seemed to* originate from the same
19 location, but *actually* originated from different sources. If HOC influences the processing of
20 distractor features, no multisensory distractor processing is expected in the mirror condition as
21 the distractor features were actually separated. By contrast, if HOC does not influence the
22 processing of the distractor features, multisensory distractor processing is expected in both the
23 control as well as the mirror condition.

24 To foreshadow the results, multisensory distractor processing occurred in both
25 conditions and was not further modulated by the presence of the mirror. The fact that the visual

and auditory distractor features did not originate from the same location, despite being perceived as such in the mirror condition, did not qualitatively change the processing of the two features.

Method

Participants. Given a d_z around .5 (based on the published literature and our own prior research), an alpha-level of .05 and a desired power of at least $1-\beta > .80$, we aimed for a minimum of 27 participants (power analyses were run with G-power 3.1.9.2; Erdfelder, Faul, & Buchner, 1996). Thirty-two students from the University of Trier took part in the study in return for course credit. One of the participants was excluded due to an error with the program, as no data was collected. Another participant was excluded due to extremely high error rates (79% errors, sample mean = 9%, SD = 6%), and a third due to extremely high error rates in the catch trials (97% errors, sample mean = 25%, SD = 19%). Consequently, a total of 29 participants (9 male) aged from 19 to 34 years (Mean age = 24 years, SD = 3) were included in the final data analyses.

Apparatus and materials. The participants were seated in a dark soundproofed chamber in front of a 24" CRT computer monitor. The distance to the screen (70 cm) was kept constant by means of a chinrest with added forehead support. The instructions were presented in grey text against a black background in the middle of the screen. Combinations of two different tones (one relatively low in pitch, 440 Hz, and the other relatively higher in pitch, 2794 Hz) and two different colors, namely red (HSV: 0, 100, 12.55) and blue (HSV: 240, 100, 12.55) were used as experimental stimuli. All of the stimuli were presented at a clearly supra-threshold level in order to ensure that the participants were able to successfully execute the task. These were presented via three custom-made multisensory cubes with two LEDs at the front (one on top, and the other at the bottom) and a loudspeaker on the rear (see Merz et al., 2019, for a detailed description of the cubes). The upper LED on each cube was used to present the visual stimuli.

1 In the mirror condition, a mirror was placed in the center of the participant's view, directly
2 below the monitor. Figure 1a depicts a bird's-eye view of the experimental set-up in the mirror
3 condition. In the control condition, the distractor cube was placed at the central location without
4 the mirror and the obstacle, thus matching the visual input between experimental and control
5 condition. The participants responded by means of footpedals using Chronos® (Version 1.0).
6 The experiment was conducted with the E-Prime software (Version 2.0) and the data were
7 analyzed by means of IBM SPSS Statistics (Version 26).

8 *Design.* The experiment consisted of a $2 \times 2 \times 2$ repeated measures design with the
9 variables Condition (control and mirror), Visual Distractor Feature (congruent and incongruent
10 with respect to the visual target feature), and Auditory Distractor Feature (congruent and
11 incongruent with respect to the auditory target feature). Reaction times (RT) as well as error
12 rates (ER) were used as dependent variables.

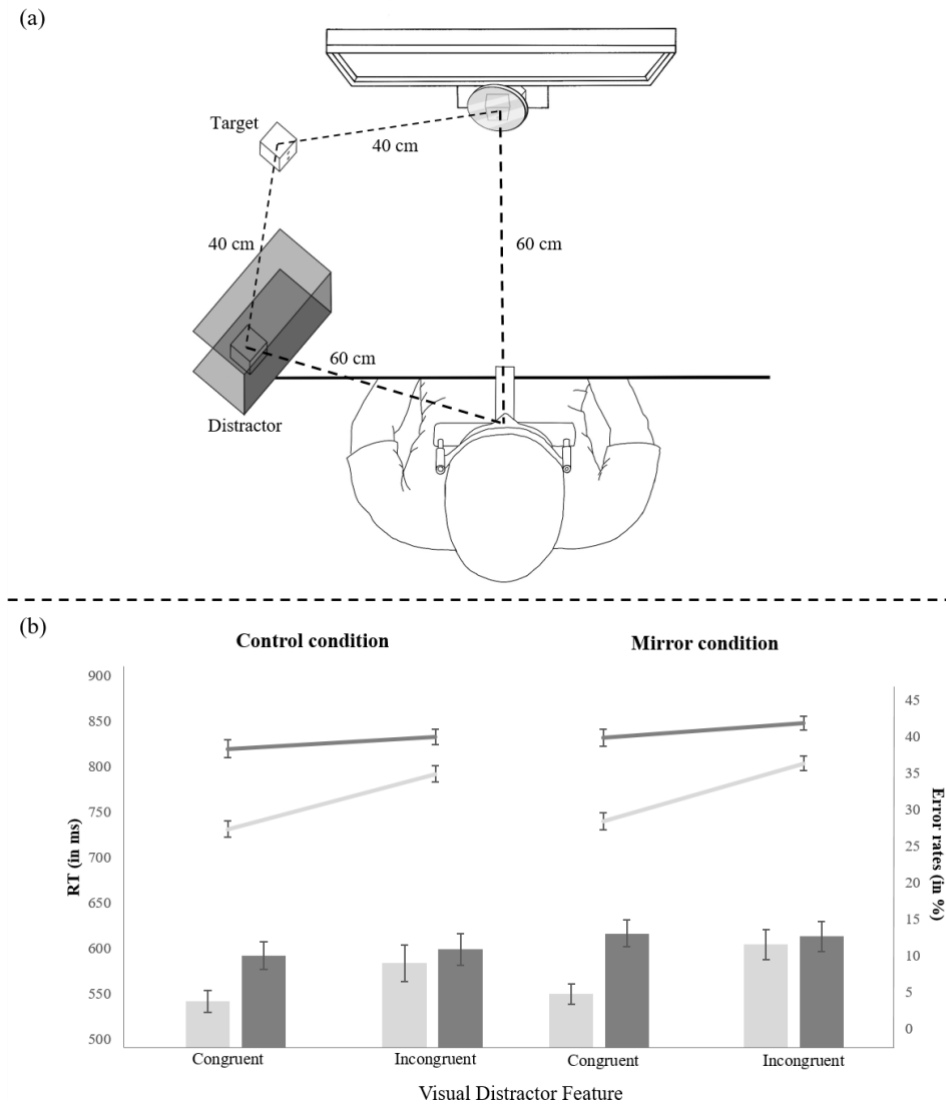


Figure 1. (a) Bird's-eye view of the experimental set-up in the mirror condition. Note that the distractor tone was presented from a third cube behind the mirror that was not visible to the participants. (b) RT (left axis – line diagram) and ER (right axis – bar diagram) as a function of condition, visual and auditory distractor congruency. Visual distractor feature congruency is indicated on the x-axis, auditory distractor congruency is indicated by the color (light gray = congruent; dark gray = incongruent). Error bars indicate within standard errors following Morey (2008).

1 *Procedure.* Before starting the mirror condition, the placement of the mirror was
2 adapted individually, so that the participant was able to see the occluded distractor cube in the
3 mirror. The participants were instructed to fixate the mirror (or on the cube in the control
4 condition) in the center under the monitor throughout the course of the experiment. The task
5 involved discriminating the target via footpedal responses. Specifically, the participant had to
6 respond to two different multisensory target stimuli, a red light combined with a high-pitched
7 tone and a blue light combined with a low-pitched tone. The stimulus-response-mapping was
8 balanced across participants (right vs. left footpedal response). The participants were instructed

1 to respond to the target stimuli presented from the target cube as rapidly and accurately as
2 possible. At the same time, however, they were also told to ignore the distractor stimuli. In
3 order to ensure that the participants processed both target features, and did not just respond to
4 one, occasional catch stimuli with reversed target feature assignment (e.g., red light and low
5 pitch) were also presented from the target cube. The participants were instructed not to respond
6 to these stimuli. These catch trials accounted for 20% of all trials (64 trials). The distractor
7 stimuli were the same as the target and catch stimuli. Thus, the visual and auditory feature of
8 the distractors could either be congruent (target and distractor feature indicate the same
9 response) or incongruent (target and distractor feature indicate different responses) with respect
10 to the corresponding target feature. The mirror manipulation (mirror condition, see Figure 1a,
11 vs. control condition) was conducted within-participants. The conditions alternated within four
12 blocks of 80 trials throughout the experimental phase. Note that this alternation also helped to
13 emphasize the presence of the mirror (cf. Zampini, Guest, Shore, & Spence, 2005). Within each
14 condition, the four different distractor stimuli were presented with the same number of trials
15 (32). This resulted in a total of 320 trials (64 catch trials included). The target side (and set-up
16 side; left, right) was manipulated within-participants and switched in the middle of the
17 experiment. The order of the side (left vs. right first) as well as the order of the conditions
18 (control vs. mirror condition first) were counterbalanced across-participants.

19 The exact sequence of stimuli on each trial was as follows: A cross was presented at the
20 bottom of the screen initiating the start of the trial, followed by the onset of the distractor after
21 250 ms. The target was presented 100 ms after the onset of the distractor. Both stimuli were
22 then presented for 1000 ms and the participants had to respond within 1400 ms of target onset
23 (in catch-trials, if a response was detected during this time, an erroneous response was
24 indicated). Feedback was only provided for false or missing responses and was displayed for
25 500 ms at the bottom of the screen (i.e., just above the mirror). In order to reduce any after-

effects from the previous trial, a 500 ms ‘eraser’ stimulus consisting of white noise and a white flash (HSV: 0, 0, 100) was included in the trial sequence. This stimulus was presented after a 500 ms black screen and was presented from the target cube. The distractor cube in the control condition as well as the cube behind the mirror in the mirror condition additionally presented white noise, while the white flash was only presented from the target. The eraser stimulus was followed by a response-stimulus interval (RSI) of 500 ms.

Instructions were provided on the screen before the experiment. The participants were encouraged to ask the experimenter if anything was unclear. The experiment started with a learning phase for the stimulus-response mapping with 8 trials and a first practice phase with 16 trials. In this practice phase, the target and catch trials both accounted for 50% of all trials and were presented on the target cube while no stimuli were presented from the distractor cube. The second practice phase, with 48 trials, included distractors and was the same as the experimental phase except for the fact that feedback was provided on correct trials and catch trials accounted for 33.3% of all trials. Taken together, the experiment, consisting of 72 practice trials and 320 experimental trials, lasted for around 30 minutes.

Results

We first checked whether the participants completed the multisensory task correctly by examining performance on the catch trials. Comparable to other multisensory flanker studies (e.g., Jensen et al., 2019a, b), the participants responded correctly to 74.8% of these trials, thus indicating that they were able to perform the task adequately.

For the analysis of the RT data, only those trials in which the participants responded correctly to the target were considered. All of the trials in which the RT was lower than 200 ms, as well as those trials with an RT that was 1.5 interquartile ranges above the third quartile of each participant’s individual RT distribution (Tukey, 1977), were excluded from the data

analysis as well. In total, 11.1% of the trials were excluded from the analysis due to these restrictions. Table A1 in the Appendix depicts the mean RTs and ERs.

Reaction time. A 2 (Condition: control vs. mirror) \times 2 (Visual Distractor Feature: congruent vs. incongruent) \times 2 (Auditory Distractor Feature: congruent vs. incongruent) multivariate analysis of variance (MANOVA) was conducted with Pillai's trace as the criterion and mean RT as the dependent variable. The MANOVA revealed a significant main effect of Visual Distractor Feature, $F(1, 28) = 43.52, p < .001, \eta_p^2 = .61$, and Auditory Distractor Feature, $F(1, 28) = 172.59, p < .001, \eta_p^2 = .86$; that is, the participants responded significantly faster when the distractor feature was congruent with the corresponding target feature than when it was incongruent. Furthermore, there was a significant interaction between these two variables, $F(1, 28) = 1.38, p = .001, \eta_p^2 = .55$, indicating multisensory distractor processing, which was not modulated as a function of condition, $F(1, 28) < 0.01, p = .994$. In fact, neither the main effect of condition, $F(1, 28) = 1.38, p = .251$, nor the interactions with the Visual or Auditory Distractor Feature, all $F_s < .26, p_s > .251$, were significant. This clearly indicates that there was no difference between the control and the mirror conditions (see Figure 1b) and suggests that the processing of the distractor features was not modulated by the mirror.

To further support the absence of the critical three-way interaction, Bayesian analysis was conducted with JASP (Wagenmakers et al., 2018a, b), in which the likelihood that the difference of the congruency effects between mirror and control condition is different from zero was tested. This analysis revealed a *Bayes Factor* (BF_{01}) = 5.07, indicating that the null-result concerning the three-way interaction is about five times more likely than the presence of an effect.

Error rates. For the ER, the same MANOVA was conducted as for the RT data. There was an effect of Auditory Distractor Feature, $F(1, 28) = 11.37, p = .002, \eta_p^2 = .29$, and a

significant interaction of Visual and Auditory Distractor Feature, $F(1, 28) = 13.19$, $p = .001$, $\eta_p^2 = .32$. In the ER, there was a marginal effect of Visual Distractor Feature, $F(1, 28) = 3.87$, $p = .059$, $\eta_p^2 = .12$. As for the RT-data, the three way interaction was not significant, $F(1, 28) = 0.64$, $p = .430$, nor was any other effect with condition, all $F_s < 1.61$, $p_s > .216$. Additional Bayesian analysis revealed a $BF_{01} = 3.77$ for the test of the difference of the congruency effects between mirror and control condition against zero, indicating that the null-result concerning the three-way interaction is nearly four times more likely than the presence of an effect..

General Discussion

This study investigated whether multisensory distractor processing is influenced by HOC. While performing a multisensory flanker task with audiovisual stimuli, participants fixated a mirror reflecting the visual distractor feature from a distant cube, while the auditory distractor feature was presented at the mirror location. The question was whether the processing of distractor information is changed when the visual input, which was actually presented from another location, was perceived through the use of a mirror at the comparable location as the auditory information. If knowledge concerning the mirror were to have influenced participants' perception, the combined processing of the two features should have been reduced as the spatial rule is violated (see Spence, 2013, for a critical review). However, compared to a control condition, in which a central cube presented the distractor feature, there were no differences in either main or interaction effects of the distractor features. This indicates that HOC did not influence the combined processing of the distractor features. Consequently, multisensory distractor processing might operate at a rather early stage of information processing, since HOC is typically engaged in later stage information processing (e.g., see Miller & Wallis, 2009).

These results are in line with earlier findings showing no influence of HOC in unisensory distractor interference (e.g., Holmes & Spence, 2005; Soto-Faraco et al., 2004). Though participants were aware of the characteristics of a mirror, their perception was, in some

sense at least, ‘fooled’ by the mirror reflection in these studies. Thus, low-level perceptual features were more important for interference than HOC. Similarly, the illusory spatial co-occurrence of the perceptual input (plus attentional engagement) was sufficient to elicit multisensory distractor processing in the present study. Although the visual and auditory input actually originated from different locations, perceiving them at the same location by means of a mirror resulted in no qualitative change in human information processing.

The results of the present study therefore stand in contrast to Wesslein and colleagues’ (2015) findings suggesting an influence of HOC on distractor interference. In a tactile flanker paradigm, Wesslein and colleagues presented different barriers between the participant’s hand at which tactile stimulation was presented. Their study found a modulation of response interference effects due to the presence of an impermeable barrier, which take place at motor level, whereas the present study focused on distractor processing, seeming to operate on an early perceptual level (Jensen, Merz, Spence, & Frings, 2019b). Please also note that the study by Wesslein and colleagues investigated the influence of HOC on distractor processing with a tactile experimental set-up, whereas the present study used an audiovisual set-up. The present study demonstrated that crossmodal interference effects depend on perceptual links between the features, rather than merely any overlap at the response level. Consequently, it is most likely that the interference and processing of multisensory distractors works at an early perceptual stage of information processing. The results of the present study support this view. Hence, HOC might be able to modulate late response effects but does not necessarily influence earlier perceptual effects.

Conclusions

Multisensory distractor processing was investigated in an audiovisual flanker task by means of manipulating the perceived location of the visual distractor feature by means of a mirror. The results revealed no difference in the interaction of distractor features when the

1 visual feature was seen via mirror reflection. This suggests that knowledge about the spatial
2 separation is not taken into account for the processing of multisensory distractors as long as
3 features seem to originate from the same location. The finding supports the view of an early
4 perceptual processing of multisensory distractors as HOC typically interferes at a later
5 processing stage.

6
7 *Open Practices Statement*

8 The data and code for the experiment is available at PsychArchives
9 (<http://dx.doi.org/10.23668/psycharchives.2688> and
10 <http://dx.doi.org/10.23668/psycharchives.2689>). We thank Stephanie Blasl for the drawings
11 that have been incorporated in Figure 1.

References

- Chan, J. S., Merrifield, K., & Spence, C. (2005). Auditory spatial attention assessed in a flanker interference task. *Acta Acustica united with Acustica*, 91, 554-563.
- Driver, J., & Grossenbacher, P. G. (1996). Multimodal spatial constraints on tactile selective attention. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 209-235). Cambridge, MA: MIT Press.
- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, 28, 1-11.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143-149.
- Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8, 223-241.
- Flowers, J. H., & Wilcox, N. (1982). The effect of flanking context on visual classification: The joint contribution of interactions at different processing levels. *Perception & Psychophysics*, 32, 581-591.
- Franz, E. A., & Packman, T. (2004). Fooling the brain into thinking it sees both hands moving enhances bimanual spatial coupling. *Experimental Brain Research*, 157, 174-180.
- Frings, C., Schneider, K. K., & Moeller, B. (2014). Auditory distractor processing in sequential selection tasks. *Psychological Research*, 78, 411-422.

- 1 Gallace, A., & Spence, C. (2005). Visual capture of apparent limb position influences tactile
2 temporal order judgments. *Neuroscience Letters*, 379, 63-68.
- 3 Gallup, G. G. Jr., Anderson, J. R., & Shillito, D. J. (2002). The mirror test. In M. Bekoff, C.
4 Allen, & G. M. Burghardt (Eds.), *The cognitive animal: Empirical and theoretical*
5 *perspectives on animal cognition* (pp. 325-333). Cambridge, MA: MIT Press.
- 6 Gatti, R. C. (2016). Self-consciousness: Beyond the looking-glass and what dogs found there.
7 *Ethology, Ecology & Evolution*, 28, 232-240.
- 8 Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of
9 higher cognition. *Trends in Cognitive Sciences*, 14, 497-505.
- 10 Helbig, H. B., & Ernst, M. O. (2007). Knowledge about a common source can promote visual-
11 haptic integration. *Perception*, 36, 1523-1533.
- 12 Holmes, N. P., & Spence, C. (2005). Visual bias of unseen hand position with a mirror: Spatial
13 and temporal factors. *Experimental Brain Research*, 166, 489-497.
- 14 Jenkinson, P. M., Haggard, P., Ferreira, N. C., & Fotopoulou, A. (2013). Body ownership and
15 attention in the mirror: Insights from somatoparaphrenia and the rubber hand illusion.
16 *Neuropsychologia*, 51, 1453-1462.
- 17 Jensen, A., Merz, S., Spence, C., & Frings, C. (2019a). Overt spatial attention modulates
18 multisensory selection. *Journal of Experimental Psychology: Human Perception and*
19 *Performance*, 45, 174-188.
- 20 Jensen, A., Merz, S., Spence, C., & Frings, C. (2019b). Perception it is: Processing level in
21 multisensory selection. *Attention, Perception, & Psychophysics*. Advanced online
22 publication. <https://doi.org/10.3758/s13414-019-01830-4>

- 1 Kitagawa, N., & Spence, C. (2005). Investigating the effect of a transparent barrier on the
2 crossmodal congruency effect. *Experimental Brain Research*, 161, 62-71.
- 3 Lajoie, Y., Paillard, J., Teasdale, N., Bard, C., Fleury, M., Forget, R., & Lamarre, Y. (1992).
4 Mirror drawing in a deafferented patient and normal subjects: Visuoproprioceptive
5 conflict. *Neurology*, 42, 1104-1104.
- 6 Ma, X., Jin, Y., Luo, B., Zhang, G., Wei, R., & Liu, D. (2015). Giant pandas failed to show
7 mirror self-recognition. *Animal Cognition*. 18, 713-721.
- 8 Maravita, A., Spence, C., Sergent, C., & Driver, J. (2002). Seeing your own touched hands in
9 a mirror modulates cross-modal interactions. *Psychological Science*, 13, 350-355.
- 10 Merz, S., Jensen, A., Spence, C., & Frings, C. (2019). Multisensory distractor processing is
11 modulated by spatial attention. *Journal of Experimental Psychology: Human Perception*
12 *and Performance*, 45, 1375-1388.
- 13 Merz, S., Frings, C., & Spence, C. (2020). When irrelevant information helps: Extending the
14 Eriksen-flanker task into a multisensory world. *Attention, Perception & Psychophysics*,
15 1-14. doi: 10.3758/s13414-020-02066-3
- 16 Miller, E. K., & Wallis, J. D. (2009). Executive function and higher-order cognition: Definition
17 and neural substrates. *Encyclopedia of Neuroscience*, 4, 99-104.
- 18 Mitchell, R. W. (1993). Mental models of mirror-self-recognition: Two theories. *New Ideas in*
19 *Psychology*, 11, 295-325.
- 20 Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau
21 (2005). *Tutorial in Quantitative Methods for Psychology*, 4, 61-64.

- 1 Palmer, S. E. (1975). Visual perception and world knowledge: Notes on a model of sensory-
2 cognitive interaction. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in*
3 *cognition* (pp. 279-307). San Francisco, CA: Freeman.
- 4 Parker, S. T., Mitchell, R. W., & Boccia, M. L. (1994). *Self-awareness in animals and humans:*
5 *Developmental perspectives*. Cambridge, UK: Cambridge University Press.
- 6 Shore, D. I., Barnes, M. E., & Spence, C. (2006). Temporal aspects of the visuotactile
7 congruency effect. *Neuroscience Letters*, 392, 96-100.
- 8 Soto-Faraco, S., Ronald, A., & Spence, C. (2004). Tactile selective attention and body posture:
9 Assessing the multisensory contributions of vision and proprioception. *Perception &*
10 *Psychophysics*, 66, 1077-1094.
- 11 Spence, C. (2013). Just how important is spatial coincidence to multisensory integration?
12 Evaluating the spatial rule. *Annals of the New York Academy of Sciences*, 1296, 31-49.
- 13 Spence, C. (2020). Extending the study of visual attention to a multisensory world (Charles W.
14 Eriksen Special Issue). *Attention, Perception, & Psychophysics*, 1-13. doi:
15 0.3758/s13414-020-02061-8
- 16 Spence, C., Lee, J., & van der Stoep, N. (2017). Responding to sounds from unseen locations:
17 Crossmodal attentional orienting in response to sounds presented from the rear.
18 *European Journal of Neuroscience*, 51, 1137-1150. doi:10.1111/ejn.13733
- 19 Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison Wesley.
- 20 Ulrich, R., Prislán, L., & Miller, J. (2020). A bimodal extension of the Eriksen flanker task.
21 *Manuscript in revision*.

- 1 Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., et al. (2018a).
2 Bayesian inference for psychology. Part II: Example applications with JASP.
3 *Psychonomic Bulletin & Review*, 25, 58-76.
- 4 Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., et al. (2018b).
5 Bayesian inference for psychology. Part I: Theoretical advantages and practical
6 ramifications. *Psychonomic Bulletin & Review*, 25, 35-57.
- 7 Wesslein, A. K., Spence, C., & Frings, C. (2014). Vision affects tactile target and distractor
8 processing even when space is task-irrelevant. *Frontiers in Psychology*, 5:84.
- 9 Wesslein, A. K., Spence, C., & Frings, C. (2015). You can't ignore what you can't separate:
10 The effect of visually induced target-distractor separation on tactile selection.
11 *Psychonomic Bulletin & Review*, 22, 728-736.
- 12 Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audiovisual simultaneity judgments.
13 *Perception & Psychophysics*, 67, 531-544.
- 14

Appendix

Table A1. Mean RTs (in milliseconds; ER in % in parentheses) as a function of condition, visual and auditory distractor feature congruency.

	Visual distractor feature		
Auditory distractor feature	Congruent	Incongruent	Mean
Mirror condition			
Congruent	742 (5.7)	805 (11.0)	774 (8.4)
Incongruent	834 (12.2)	850 (11.9)	842 (12.0)
Mean	788 (8.9)	827 (11.4)	
Control condition			
Congruent	731 (4.7)	792 (8.6)	762 (6.7)
Incongruent	820 (9.4)	833 (10.0)	826 (9.7)
Mean	776 (7.1)	813 (9.3)	