Modality-switching in the Simon task: The clash of reference frames

Manuela Ruzzoli¹ & Salvador Soto-Faraco^{1, 2}

¹ Center for Brain and Cognition, Departament de Tecnologies de la Informació i les
Comunicacions, Universitat Pompeu Fabra, 08018 Barcelona, Spain

² Institució Catalana de Recerca i Estudis Avançats (ICREA)

Address:

Center for Brain and Cognition, Departament de Tecnologies de la Informació i les
Comunicacions, Universitat Pompeu Fabra, Roc Boronat 138, 08018 Barcelona, Spain

Authors' contribution:

M.R. designed the experiments, collected the data, planned and realized data analyses,
interpreted the results, wrote the manuscript.

S.S.F. designed the experiments, interpreted the results, and wrote the manuscript.

Correspondence should be addressed to M. R. (manuela.ruzzoli@upf.edu)

Acknowledgement:

This research was supported by the Ministerio de Economía y Competitividad (PSI2016-75558-P), the AGAUR Generalitat de Catalunya (2014SGR856) and the European Research Council (StG-2010 263145) to SSF. M.R. was supported by the Juan de la Cierva postdoctoral fellowship (JCI-2012-12335) from the Ministerio de Economía y Competitividad. We would like to thank Leonor Castro and Joan Birules for their help in data collection.

Keywords: S-R compatibility, Simon task, spatial representation, vision, touch, audition, sequential compatibility effect.

Word Count: 12.321

Abstract

The representation of spatial information related to an event can influence behaviour even when location is task-irrelevant, as in the case of Stimulus-Response (S-R) compatibility effects on the Simon task. However, unlike single-modality situations, which are often used to study the Simon effect, in real-life scenarios various sensory modalities provide spatial information coded in different coordinate systems. Here, we address the expression of S-R compatibility effects in mixed-modality contexts, where events can occur in one of various sensory modalities (i.e., vision, touch or audition). The results confirm that, in single-modality cases, Simon effects in vision are expressed in an external spatial frame of reference, whilst touch information is coded anatomically. Remarkably, when mixing visual and tactile trials in an unpredictable way, the Simon effect disappeared in vision whereas tactile Simon effects remained expressed in their own (anatomical) frame of reference. Mixing visual and auditory stimuli did not obliterate the visual Simon effect and S-R compatibility effects in an external reference frame were evident for both modalities. The extinction of visual Simon effects as a result of mixing visual and tactile modalities can be interpreted as a consequence of the dynamic reorganization of the weights associated to the different sources of spatial information at play.

In order to plan and execute actions in response to environmental events the brain must estimate, transform and integrate spatial information from a variety of sources into a unified code. Stimulus-Response (S-R) compatibility, or lack of, is a result of how much matching occurs between the spatial information derived from a stimulus and its required response (Fitts & Seeger, 1953). Some spatial representations appear to be quite automatic, happening even when they are not explicitly required to complete the task. For example, if someone unexpectedly shouts 'Look, to the left!' our likely reaction is to first orient towards the person shouting rather than towards the place of interest. This spatial incompatibility is illustrated by the famous Simon effect (Simon & Small, 1969): when a participant is asked to make speeded responses to the colour of a visual stimulus, a benefit is observed if the relative position of the stimulus (i.e., position in the right/left visual hemifield) and the location of the button associated with the manual response (right/left with respect to the body midline) align as compared to when they are misaligned¹.

In vision, spatial S-R compatibility effects occur irrespective of which anatomical effector (i.e., left/right hand) is used to push the corresponding response button thus indicating that S-R compatibility effects are based on the spatial (mis-)alignment between stimulus location and response goal (Riggio, Gawryszewski, & Umiltà, 1986; Wallace, 1971). Indeed, in the classical Simon task, there are at least three sources of spatial information that might bear spatial (in-) compatibility with the location of the stimulus itself: the position of the response button (or the response goal), the position of the responding effector (placed on the right side or on the left side of the body midline) and the anatomical side of the responding effector (left or right hand response). S-R compatibility can arise from the (mis-) match between the position of the stimulus and any of those three possible sources of spatial information (Heister, Schroeder-Heister, & Ehrenstein, 1990; Wallace, 1971). In the classical Simon task, the three possibilities are confounded because the anatomical right (left) hand operates on the right (left) side of the body on the right (left) button. However, crossing the arms over the body midline misaligns the effectors with their position in space and the side of their response goal.

Studies on the Simon effect, testing participants with crossed-arm posture, have shown that visual (Riggio et al., 1986; Wallace, 1971) and auditory (Röder, Kusmirek, Spence, & Schicke, 2007; Roswarski & Proctor, 2000; Simon, Hinrichs, & Craft, 1970) S-R compatibility effects are expressed in terms of their external coordinates. In other words, what matters is the compatibility between the position of the stimulus and the position of the response goal². Conversely, in the tactile Simon effect what matters is the compatibility between the position of the stimulus and the anatomical side of the responding effector (regardless of posture). This means that responses to tactile stimuli are faster when the position of the stimulus matches the anatomical side of the hand (Hasbroucq, 1987) or foot (Medina, McCloskey, Coslett, & Rapp, 2014) used to respond. Collectively, these studies demonstrate that visual, auditory and tactile stimuli abide to distinct S-R compatibility patterns whereby visual and auditory events trigger a spatial code based on external coordinates and tactile events trigger spatial representations based on anatomical coordinates.

All of the studies to date have tested the modalities separately rendering the dominance of a particular S-R compatibility pattern to a trivial issue. However, if one considers real-life situations and the natural environment it is obvious that events can come from different sensory modalities. As a consequence, the interplay amongst the different and potential S-R compatibility patterns is far from trivial. To the best of our knowledge, the problem of switching between multiple S-R compatibility patterns triggered by modality switching has not been studied in the long tradition of research on the Simon effect. Furthermore, although it is known that reference frames play a role in the Simon effect (Klapp, Greim, Mendicino, & Koenig, 1979; Ladavas & Moscovitch, 1984; Lamberts, Tavernier, & D'Ydewalle, 1992; Lu & Proctor, 1995; Roswarski & Proctor, 1996; Stoffer, 1991; Umiltà & Liotti, 1987), their relevance in spatial coding has been investigated mostly in regards to the visual modality.

We measured the Simon effect when participants responded to stimulus properties of one modality (e.g., visual or tactile), on a given block of trials, in both fixed-modality and mixed-modality contexts (e.g., visuo-tactile or visuo-auditory). This was assumed to result in different patterns of spatial compatibility, depending on the given context. One possibility, in mixed-

modality contexts, is that S-R compatibility in external space (triggered by visual and auditory stimuli) will dominate over S-R compatibility in anatomical space (as triggered by touch). Indeed, it has been suggested that it may be more efficient to represent spatial information in external space as a function of required action (e.g., Azañón, Stenner, Cardini, & Haggard, 2015). This hypothesis of visual (hence external space) dominance seems plausible considering that vision has been shown to dominate over the other senses on various tasks (Posner, Nissen, & Klein, 1976; Rock & Victor, 1964), including some multisensory ones, such as the Colavita effect (Colavita, 1974) or the ventriloquism illusion (Bertelson & Aschersleben, 1998). In the specific case of visuo-tactile interactions, the spatial representation, in terms of external space (visually-based reference frame), profoundly affects the subsequent responses to tactile stimuli, despite the fact that their position in external space is completely irrelevant to the task (e.g., Azañón, Camacho, & Soto-Faraco, 2010; Azañón & Soto-Faraco, 2008; Overvliet, Azañón, & Soto-Faraco, 2011). This is similar to what occurs in the crossed-hands effect (Yamamoto & Kitazawa, 2001, see also Heed, Buchholz, Engel, & Röder, 2015; Ossandón, König, & Heed, 2015). This effect reveals a strong tendency to make localization errors when reporting the temporal order of two tactile events presented to each hand while the arms are crossed (as if the responses were based on the arms being uncrossed and thus matching the external frame of reference).

In the present study, our goal was to test S-R compatibility patterns when mixing events in different sensory modalities which operate according to different frames of reference, as in the case of vision and touch. With this aim, Experiment 1 set out to validate our experimental set-up in achieving the desired Simon effect and to replicate previous relevant findings, regarding the dominance of different patterns of S-R compatibility (external vs. anatomical) according to modality (visual and tactile respectively) while manipulating arm posture. In Experiments 2 and 3 we introduced, for the first time, modality-switching on a Simon task by mixing visual and tactile trials in a randomized and unpredictable way. Participants were instructed to provide fast responses to non-spatial features of either visual or tactile stimuli, presented to the left or right side of the body midline. Participants were tested on the same task under uncrossed and crossed

arm conditions. This allowed us to measure possible changes in S-R compatibility patterns in mixed-modality situations as compared to the typical single-modality scenarios seen in Experiment 1. In Experiment 4, we reproduced a mixed-modality scenario but responses were provided through the use of foot pedals rather than manually, as in Experiments 2 and 3, so that the location of the tactile stimulation and the response were misaligned. Finally, in Experiment 5 we tested the Simon effect in a modality-switching context, but this time we mixed visual and auditory trials rather than visual and tactile ones. In such a situation, the cognitive/perceptual load (Braem, Abrahamse, Duthoo, & Notebaert, 2014; Hommel, 1994a) necessary for maintaining more than one active response map is comparable to the visuo-tactile version of the task tested in Experiments 2, 3 and 4; however, unlike vision and touch, which call upon S-R compatibility patterns based on different spatial frames of reference, vision and audition have both been shown to trigger equivalent S-R compatibility patterns, based on external coordinates.

General experimental methods

We report a series of five experiments applying the Simon task on various target modalities in either single (Experiment 1) or mixed-modality contexts (Experiments 2-5). All of the experiments involved visual and tactile targets (Experiment 1, blocked; Experiments 2-4, mixed) except Experiment 5 that involved auditory and tactile targets (mixed). All the experiments were run using a custom-made Matlab code in Psychtoolbox libraries (Brainard, 1997; Pelli, 1997).

Visual stimulus: The visual stimulus was in the form of a disk, with a visual angle diameter of 4.2° , which was presented either to the right or to the left (8.2 visual degrees) of a central fixation cross, for 100 ms. The target disk could be light gray (contrast= 23 cd/m^2) or dark gray (contrast= 3 cd/m^2) presented on a uniform mid-gray background (contrast= 8 cd/m^2).

Tactile stimulus: Vibrotactile stimulation was delivered through Oticon-A bone conduction vibrators (3.8 cm^2 vibrating surface; Oticon) attached to the dorsal side of the index fingers of the left and right hand (i.e., just below the nail). The vibrators were set at a constant and comfortable pressure and secured to the fingers by medical tape. The vibrations lasted for 100

ms and consisted of either 200 Hz sine waves or white noise enveloped with 10 ms on/off ramps. The amplitude of the stimuli was adjusted to a comfortable level and was the same for both types of stimulus.

Auditory stimulus: The auditory stimuli were 200 Hz sine waves or white noise enveloped with 10 ms on/off ramps. The stimuli were played from either the left or the right loudspeaker, placed on respective sides of the monitor, equidistant and at the same height from the fixation cross. The stimulus lasted for 100 ms and had an intensity of 85 dB.

For each stimulus modality, the key stimulus feature (disk lightness: light/dark; vibration type: sine wave/white noise or type of sound: sine wave/white noise), the side of stimulation (left/right) or the combination of the two (stimulus feature x side of stimulation) were both equiprobable and chosen either at random (Experiments 1 and 2) or pseudo-randomly (Experiments 3, 4 and 5, where the number of modality-switch and modality-maintain trials was kept equal).

Procedure: In all of the experiments, the participant's arms and hands were covered from sight with a black cloth, in an attempt to block visual information related to body posture, which might affect tactile stimulus processing, as seen in tactile remapping experiments (e.g., Yamamoto & Kitazawa, 2001). Participants were instructed to maintain visual fixation on a black central cross (0.6° horizontal and vertical dimensions) present for the duration of the entire trial. At the beginning of each trial, the fixation cross flashed (white/black) three times (10 Hz), alerting the participant to the imminent presentation of a stimulus. After a delay, chosen at random between 300 and 500 ms, the target stimulus (visual, tactile or auditory) was presented. Target presentation was jittered in order to avoid possible confounds due to temporal expectation (Mühlberg, Oriolo, & Soto-Faraco, 2014; Pomper, Keil, Foxe, & Senkowski, 2015; van Ede, de Lange, Jensen, & Maris, 2011; van Ede, Jensen, & Maris, 2010). Participants were instructed to respond as quickly as possible to the colour (visual Simon), vibrotactile frequency (tactile Simon) or type of sound (auditory Simon) depending on the experiment and regardless of the spatial position of the stimulus. Responses were provided by pressing the “q” or “p” keys, on a computer keyboard, with the index finger of each hand (or via foot pedals as seen in

Experiment 4). Every subsequent trial started 800-1000 ms post-response. Response mapping was balanced among participants, so that half responded to the dark or light gray stimulus with the right or left finger, and vice versa for the other half. For tactile stimuli, since the modulation of vibrotactile frequency gives rise to the perception of different intensities (as reported by Hollins & Roy, 1996), we referred to the intensity (strong/weak sensation) as opposed to the frequency (sine wave/white noise) when explaining response mapping to the participants. Half of the participants responded to sine waves or white noise with the right or left finger and vice versa for the rest. For auditory stimuli, half responded to sine waves or white noise with the right or left finger or foot, and vice versa for the other half of participants.

With the exception of Experiment 4, in the remaining experiments participants performed the task twice, once while adopting an uncrossed-arm posture and once with a crossed-arm posture (at about the body midline). This ensured that, in the crossed-arm position, the left hand would be used to press the button on the right side of the body, and vice-versa for the right hand (the order of posture conditions was counterbalanced between participants and response mapping). In Experiment 4, participants completed the task (visuo-tactile modality-mixed Simon task with pedal response) adopting the uncrossed-arms position only. In fact, this was a control experiment designed with the unique purpose of excluding the possibility that stimulus-response overlapping was responsible for the vanishing of the visual Simon effect on the modality-mixed Simon task. On the single-modality version of the task, stimulus-response overlapping effects have been previously reported by Medina et al. (2014) (replicated in our S1 experiment as well).

In all of the experiments, the instructions stressed the mapping between the anatomical effector and stimulus feature (colour, vibrotactile intensity, or type of sound), rather than between the stimulus feature and the response button. This was done in order to maintain a unified codification of response-mapping across conditions and experiments, regardless of arm posture. Before the experimental session, participants were familiarized with a training block (16 trials: single-modality, 32 trials: mixed-modality). If participants explicitly reported uncertainty or responses were too slow (>15% of the trials showed RTs >1500 ms), the training session was repeated. Data from the training session(s) was not stored. All of the experiments

reported in this study were run in accordance with the Declaration of Helsinki and approved by the ethics committee, CEIC Parc de Mar (University Pompeu Fabra, Barcelona, Spain). The participants provided written, informed consent before participating in the study.

In the single-modality Simon tasks the sample size was calculated by considering the effect size (difference between compatible and incompatible trials), given a Cohen's d value of 1.5, with 0.8 power and 0.05 alpha. In the mixed-modality Simon task, the sample size was calculated considering the effect size (difference between compatible and incompatible trials), given a Cohen's d value of 1.3, with 0.8 power and 0.05 alpha.

Data Analysis: S-R compatibility was coded on the basis of the relationship between anatomical hand or foot (left/right) and the position of the stimulus (i.e., anatomical S-R compatibility). This is, necessarily, an arbitrary decision to provide a unified scoring and representation of the data, but which has no effect on the final interpretation of the results. We classified the data as compatible if the stimulus position matched the anatomical hand or foot (left/right) required for providing the correct response, regardless the modality of the stimulus or the position of the arms (crossed/uncrossed). Only reaction times (RT) on correct trials, >50 ms and <1500 ms, were included in the analysis. In all five experiments, *response-mapping* (which hand/foot should be used to obtain a correct response) and *posture order* (uncrossed-crossed vs. crossed-uncrossed) were initially included in the analysis as between-participant factors. Since these factors were not significant, nor did they interact with any other relevant factor, we did not consider them in further analyses. Raw data was analyzed using custom-made Matlab codes. We reported the significant results, using repeated measures analyses of variance (ANOVAs), on mean RTs and accuracy, separately. Post-hoc pair-wise comparisons were carried out using Student's t-test. We reported the results of the mean RTs and accuracy; however, the discussion will mainly take into consideration the results of RTs. For complex ANOVAs we show the full set of effects in table format. Only hypotheses-driven interactions will be discussed in the text.

Experiment 1: Single-modality Visual and Tactile Simon tasks

This experiment aimed to replicate previous results of the Simon task using visual and tactile stimuli under a unified protocol and posture manipulation. Two groups of participants were tested (one with visual targets and one with tactile targets) while adopting one of two possible postures (arms-crossed, arms-uncrossed) on separate blocks. We collected responses from a total of 200 trials (50 trials in each compatibility and posture condition), in each modality group.

Participants: Independent samples of 10 right-handed participants were tested on the visual Simon task (5 women; age range 20-34 years) and on the tactile Simon task (6 women; age range 23-34 years). Two additional participants on the tactile Simon task were excluded because of too many errors (performance score of <65% in at least one of the conditions) or too many slow reaction times (RT > 1500 ms for more than 5% of the total number of trials).

Data Analysis: We performed a repeated measures ANOVA with *arm posture* (uncrossed/crossed) and *S-R compatibility* (compatible/incompatible) as within-subjects variables and *modality* (vision/touch) as a between-subjects variable.

Results

Overall, RTs to visual targets were 131 ms faster than responses to tactile targets [$F(1,18)=8.83$; $p=0.008$; $\eta^2_p=0.33$]. There was no significant main effect of *arm posture* on RTs [$F(1,18)=1.48$; $p=0.24$; $\eta^2_p=0.08$], nor was there an interaction between *arm posture* and *modality* [$F(1,18)=1.26$; $p=0.28$; $\eta^2_p=0.07$]. As expected, there was a significant three-way interaction between *arm posture*, *S-R compatibility* and *modality* [$F(1,18)=14.47$; $p=0.001$; $\eta^2_p=0.45$]. In other words, for the visual Simon task, in the uncrossed-arm posture RTs to compatible trials were on average 30 ms faster than on incompatible trials (322 ms vs. 352 ms; $t(18)=-2.30$; $p=0.034$; $d_z=0.73$), whereas in the crossed-arms condition, compatible trials led to RTs that were on average 47 ms slower than on incompatible trials (385 ms vs. 338 ms; $t(18)=3.51$; $p=0.003$; $d_z=1.11$), indicating the presence of a classical Simon effect operating in an external reference frame (Figure 1A). For the tactile Simon task, the magnitude and direction of the S-R compatibility effect was very similar for the uncrossed (compatible = 450 ms vs. incompatible = 509 ms, $t(18)=-4.50$; $p<0.001$; $d_z=1.42$) and crossed postures (452 vs. 510 ms,

$t(18)=-4.30$; $p<0.001$; $d_z=1.34$), indicating the presence of a Simon effect in an anatomical frame of reference (Figure 1B).

The ANOVA on accuracy also confirmed the presence of a significant three-way interaction between *arm posture*, *S-R compatibility* and *modality* [$F(1,18)=9.32$; $p=0.007$; $\eta^2_p=0.34$]. For the visual Simon task with uncrossed-arms spatial compatibility marginally affected accuracy (compatible 96% vs. incompatible 90%, $t(18)=1.95$; $p=0.067$; $d_z=0.62$), whereas in the crossed-arm posture accuracy dropped on compatible trials in comparison to incompatible ones (90% vs. 98%, $t(18)=-3.46$; $p=0.003$; $d_z=1.10$). For the tactile Simon task, accuracy was higher for the compatible trials than the incompatible trials for both the uncrossed (97% vs. 93%; $t(18)=1.39$; $p=0.182$; $d_z=0.44$) and crossed-arm conditions (99% vs. 93%; $t(18)=2.37$; $p=0.029$; $d_z=0.75$), however the effect was only significant for the crossed-arms condition.

Discussion

The results from the Experiment 1 are straightforward and confirm previous evidence on visual and tactile Simon effects, respectively. The crossed and uncrossed-arm manipulations allowed us to disambiguate the relevant pattern of S-R compatibility effects in the visual and tactile domains. Indeed, the coding of spatial information coming from body posture influences the RTs to visual and tactile events according to different spatial coordinates (i.e., frames of reference) (Wallace, 1971). Namely, visual Simon effects are determined by spatial compatibility between stimulus location and the location of the response goal, or, in other words, visual Simon effects are expressed in terms of external space, as previously reported by others (Riggio et al., 1986; Wallace, 1971). In contrast, tactile Simon effects are determined by spatial compatibility between stimulus location and the anatomical side of the responding hand. This result is in line with work by Hasbroucq (1987) and a recent study by Medina et al. (2014; Experiment 2) on tactile Simon effects in uncrossed or crossed-arm positions. Interestingly, in Medina et al.'s study (2014), responses were delivered via foot pedals rather than by manual key-press. Their result therefore reinforced the claim that, even in the absence of overlap between the stimulus location on the skin and the location of the response, tactile S-R compatibility depends on the spatial match between stimulus position and the anatomical side of

the responding effector. (We have replicated this particular finding in six additional participants, with the tactile stimuli described in Experiment 1 but using foot pedal responses; see Supplemental materials).

The results of Experiment 1 validate the experimental setting for attaining different S-R compatibility patterns as a function of stimulus modality. In the following experiments, we will capitalize on postural and modality manipulations in order to disentangle the interplay of S-R compatibility patterns when the spatial frames of reference associated to each sensory modality alternate unpredictably from one trial to the next (through switches in target modalities on the Simon task).

Experiments 2 and 3: Visuo-tactile mixed-modality Simon task

In real-life, events requiring a response can be unpredictable and originate from different locations and sensory modalities. To address compatibility effects in these situations, we designed a mixed-modality (visuo-tactile) Simon task with the aim of testing the consequences of switching between S-R compatibility patterns (here, grounded on external vs. anatomical frames of reference). The importance of defining a coordinate system (or frame of reference) for unequivocally interpreting spatial relationships is imperative to action planning. It has already been shown that the misalignment of different spatial reference frames in the visual modality has a direct influence on the Simon effect (Ladavas & Moscovitch, 1984; Nicoletti & Umiltà, 1994; Umiltà & Liotti, 1987), however, to the best of our knowledge, no study as of yet has considered modality-switching as a means to predispose the activation of different S-R compatibility patterns derived from the implicit use of alternative frames of reference.

The logic of these experiments is based on the assumption that changes in stimulus modality will engender switches in the respective spatial frame of reference used to encode the location of the stimulus, and hence, the pattern of S-R compatibility. One might assume that the frame of reference used to refer a stimulus to a spatial location is chosen after stimulus presentation, but not before. In such a case, the Simon effect should not differ between the mixed-modality and the single-modality situations, where the sensory modality of the stimulus is fixed beforehand.

Alternatively, one could assume that the single-modality version of the task allows for expectancy effects in one sensory modality and hence for anticipation effects in regards to the frame of reference. Because this anticipation of the relevant spatial code is ineffective when modalities are randomly mixed, the Simon effect is expected to change between the single and mixed-modality versions of the task. Specifically, we hypothesized that if events in different modalities are presented in an unpredictable way on a Simon task (ensuing switches in spatial reference frame), then the preferential pattern of S-R compatibility will be the one corresponding to the visual modality, that is, the one coded in external coordinates. Our hypothesis is based on the evidence of visual dominance over the other senses (Hecht & Reiner, 2009; Rock & Victor, 1964; Sinnett, Spence, & Soto-Faraco, 2007) and especially over touch in the coding of space (as the case of visuo-tactile remapping) (Heed et al., 2015; Yamamoto & Kitazawa, 2001). On a visuo-tactile Simon task, we therefore expect that visual Simon effects will manifest in their own external reference frames, whilst S-R compatibility effects for tactile events will be eliminated or even reversed, leading to a dominance of the external frame of reference over the anatomical one.

Method

Participants: We analyzed data from two independent samples of 16 right-handed participants each (Experiment 2: 8 women; age range 18-35 years; Experiment 3: 12 women; age range 18-34 years). Data from four additional participants (two from each experiment) were excluded from the analysis due to low accuracy (<65%, in at least one of the conditions) or slow RTs (> 1500 ms, on more than 5% of trials).

Procedure: A very similar procedure, as in Experiment 1, was applied to Experiments 2 and 3. The only difference was that the modality of the stimulus varied unpredictably between trials. Participants were instructed to respond as fast as possible to either colour (in the case of a visual stimulus) or frequency of vibration (in the case of a tactile stimulus), regardless of spatial location (left or right). Arm posture was manipulated, within-participants, in separate blocks (order was counterbalanced). The only difference between Experiments 2 and 3 were in the number of trials (200 and 416, respectively) and trial distribution. Experiment 3 could be

considered as an extended version of Experiment 2, however due to the value of reproducibility in psychology (Ioannidis, 2005; Open Science Collaboration, 2015), we decided to fully report both experiments. In Experiment 2, 200 trials were collected for each possible arm position (crossed/uncrossed) (50 trials per condition: modality x S-R compatibility). The experiment had a duration of approximately 45 minutes (breaks included). Experiment 3 was exactly the same as Experiment 2 with one small difference in that the number of modality-maintain trials (i.e., two consecutive trials of the same modality) was the same as the number of modality-switch trials (104 trials each). For Experiment 3, 416 trials were collected for each arm posture (104 trials for each condition: modality x compatible/incompatible stimulus-response postures) and took about 80 minutes (breaks included).

Results

We conducted a repeated measures ANOVA on correct RTs with *modality* (vision/touch), *arm posture* (uncrossed/crossed) and *S-R compatibility* (compatible/incompatible) as within-subjects variables on data from Experiments 2 and 3, separately. Results from Experiment 2 (see Figure 2) revealed a significant main effect of *modality* ($F(1,15)=108.38$; $p<0.001$; $\eta^2_p=0.88$) indicating that responses to visual stimuli were on average faster than responses to tactile stimuli (532 ms vs. 672 ms). The main effect of *arm posture* approached significance ($F(1,15)=4.1$; $p=0.061$; $\eta^2_p=0.21$), with responses in the uncrossed-arm condition being slightly faster than those in the crossed-arm condition (577 vs. 627 ms). The main effect of *S-R compatibility* was also significant ($F(1,15)=4.63$; $p=0.048$; $\eta^2_p=0.24$) with compatible trials receiving on average faster responses than incompatible ones (589 ms vs. 615 ms). The interaction between *modality* and *S-R compatibility* also reached significance ($F(1,15)=24.37$; $p<0.001$; $\eta^2_p=0.62$). While there was no difference in RT between visuo-compatible and visuo-incompatible trials (534 ms and 530 ms, respectively; $t(15)=0.50$; $p=0.730$; $d_z=0.13$), the effect of S-R compatibility on tactile stimuli was significant (643 ms vs. 701 ms, $t(15)=-3.57$; $p=0.003$; $d_z=0.89$).

Critically, there was a significant three-way interaction of *modality* x *arm posture* x *S-R compatibility* ($F(1,15)=6.5$; $p=0.022$; $\eta^2_p=0.30$). As seen in Experiment 1, the disparate effects of arm crossing in the S-R compatibility effect of each modality can potentially dissociate the

dominant frame of reference. In follow-up planned t-tests, the effect of *S-R compatibility* on the tactile modality was evident for both uncrossed (616 ms vs. 665 ms, $t(15)=-2.83$; $p=0.013$; $d_z=0.71$) and crossed (671 ms vs. 737 ms, $t(15)=-3.56$; $p=0.003$; $d_z=0.89$) arm postures, thus replicating the findings of Experiment 1 (tactile Simon effect). For visual stimuli, there was no significant main effect of arm posture. This result differs from those found under single-modality conditions (Experiment 1). A closer inspection of the direction of the effects on visual stimuli revealed a weak trend for S-R compatibility in external space, with a 13 ms S-R compatibility effect when the arms were uncrossed (compatible = 508 ms vs. incompatible = 521 ms, $t(15)=-1.27$; $p=0.224$; $d_z=0.32$) and a -21ms reversed compatibility trend when the arms were crossed (560 vs. 539 ms, respectively, $t(15)=1.27$; $p=0.223$; $d_z=0.32$). Thus, when two modalities were mixed, unpredictably, on a Simon task, the visual S-R compatibility effect disappeared (statistically) or was, at best, strongly attenuated. In contrast, the S-R compatibility effect remained full strength for tactile stimuli. In Experiment 2, accuracy only yielded a significant main effect of *modality* ($F(1,15)=25.72$; $p<0.001$; $\eta^2_p=0.63$), whereby visual performance was superior to tactile performance (96% vs. 88%).

Data from Experiment 3 (Figure 4) was analyzed in the same way as data from Experiment 2, and the results confirmed the somehow unexpected picture emerging from Experiment 2. We found a significant main effect of *modality* ($F(1,15)=54.66$; $p<0.001$; $\eta^2_p=0.78$) because RTs to visual stimuli were on average 97 ms faster than RTs to tactile stimuli (496 ms vs. 593 ms). The main effect of *arm posture* in Experiment 3 was also significant ($F(1,15)=15.90$; $p=0.001$; $\eta^2_p=0.51$), with RTs in the uncrossed condition being faster than those in the crossed condition (511 ms vs. 578 ms). Furthermore, the main effect of *S-R compatibility* was significant ($F(1,15)=10.29$; $p=0.006$; $\eta^2_p=0.41$) with compatible RTs being on average 19 ms faster than incompatible ones (535 ms vs. 554 ms). The two-way interaction between *modality* and *S-R compatibility* reached significance ($F(1,15)=33.02$; $p<0.001$; $\eta^2_p=0.69$), with S-R compatibility effects causing significantly faster RTs on tactile trials (compatible = 570 ms vs. incompatible = 615 ms, $t(15)=-4.75$; $p<0.001$; $d_z=1.19$) but not visual ones (compatible = 499 ms vs. incompatible = 493 ms, $t(15)=1.30$; $p=0.214$; $d_z=0.33$). Lastly, there was no three-way

interaction ($F(1,15)=1.21$; $p=0.287$; $\eta^2_p=0.08$). Seeing as in Experiment 3, S-R compatibility effects did not vary as a function of arm posture in either modality (tactile: strong and positive for crossed/uncrossed; visual: null in both postures), this result was trivial.

The analysis on accuracy data from Experiment 3 only revealed a significant two-way interaction between *modality* and *S-R compatibility* ($F(1,15)=6.28$; $p=0.024$; $\eta^2_p=0.30$). For visual trials, accuracy did not differ between compatible and incompatible trials (97%, both; $t(15)=-1.37$; $p=0.192$; $d_z=0.34$). For tactile trials, we observed a slight advantage for the compatible trials (97% vs. 96%; $t(15)=2.17$; $p=0.047$; $d_z=0.54$). In sum, the results of Experiment 3 closely replicated those found in Experiment 2, only this time the Simon effect in vision collapsed completely, as the non-significant trend for visual Simon effects seen in Experiment 2, also disappeared.

In Experiments 2 and 3, mixed-modalities on the Simon task and thus different sources of spatial information competed, in such a way that, one spatial code prevailed over the other setting the corresponding S-R compatibility pattern. This may have occurred due to the influence of the frame of reference used in the prior trial, generating costs for trials involving modality switching, and/or benefits when a modality is maintained across consecutive trials. Therefore, we conducted an analysis of the RTs as a function of the responses to the prior trial (N-1), in order to investigate cognitive control mechanisms for modality switching, as seen on the Simon task. It has been previously observed that on tasks (e.g., discrimination tasks), involving stimuli presented to different sensory modalities, mixed, trial-by-trial, RTs are slower for modality-switch trials than for modality-maintain trials (Hunt & Kingstone, 2004; Murray, Santis, Thut, & Wylie, 2009) thus accounting for modality-specific control mechanisms as opposed to an amodal executive system.

The analysis of the RTs as a function of prior trial (N-1) in Experiments 2 and 3 aimed at answering two specific questions. First, it was important to confirm whether or not there were modality-switching costs, to see whether our results conform to prior cross-modal literature. Second, and most importantly, we wanted to measure whether visual Simon effects were present

on visual trials preceded by another visual trial but absent on visual trials preceded by a tactile trial. The answer to these questions will allow us to explore possible mechanisms responsible for the disappearance of the visual Simon effect in Experiments 2 and 3.

We use the terms *maintain* and *switch* to refer to the modality of the current trial (N) as a function of the modality of the previous trial (N-1). Hence, maintain trials were those in which, the preceding trial was the same modality and switch trials were those in which, the modality of the previous trial was different. So, a *visual switch* trial was a visual trial presented after a tactile trial and vice versa. We performed a repeated measures ANOVA with *modality* (vision/touch), *arm posture* (uncrossed/crossed), *modality transition* (maintain/switch) and *S-R compatibility* (compatible/incompatible) as within-subjects variables. A complete list of ANOVA terms along with their interactions and reported significance values can be found in Tables 1 and 2 for Experiments 2 and 3, respectively. Below, we report the analyses that respond specifically to the claims posited by our hypotheses.

a) *Are there modality-switching costs?*

For both Experiments 2 and 3, the results of trial N as a function of the modality presented in trial N-1 were very similar. For Experiment 2 (Figure 3), the term *modality transition* was significant [$F(1,15)=47.27$; $p<0.001$; $\eta^2_p=0.76$] demonstrating that switching between modalities (visual or tactile) on a Simon task induced longer RTs than when the modality was maintained (548 ms vs. 665 ms, respectively). This *modality transition* effect interacted with *modality* [$F(1,15)=6.45$; $p=0.023$; $\eta^2_p=0.30$] in the following way whereby switching from touch to vision produced RTs that were on average 127 ms longer than when the visual modality was maintained for both N and N-1 trials [$t(15)=-6.76$; $p<0.001$; $d_z=1.69$]. Analogously, switching from vision to touch induced a cost of 85 ms when compared to responding to maintained tactile trials [$t(15)=-5.27$; $p<0.001$; $d_z=1.32$]. Therefore, switch costs were present in both directions but steeper in vision than in touch. *Arm posture* approached significance [$F(1,15)=3.65$; $p=0.076$; $\eta^2_p=0.20$] but did not interact with *modality transition*, indicating that switch cost was unaffected by arm posture.

Data from Experiment 3 (see Table 2 and Figure 5) confirmed the patterns observed in Experiment 2. Switching between modalities led to RT costs and on average maintain trials exhibited faster RTs than switch trials (80ms) as seen with the significant main effect of *modality transition* [$F(1,15)=46.66$; $p<0.001$; $\eta^2_p=0.76$]. *Modality transition* interacted with *modality* [$F(1,15)=15.23$; $p=0.001$; $\eta^2_p=0.50$], similarly to Experiment 2, and once again switch costs were steeper in vision than in touch. More specifically, switching from touch to vision slowed RTs by approximately 106 ms (visual maintain = 442 ms vs. visual switch = 548 ms; $t(15)=-7.08$; $p<0.001$; $d_z=1.77$), whereas switching from vision to touch slowed reaction times by an average of 55ms (tactile maintain = 563 ms vs. tactile switch = 618 ms $t(15)=-4.71$; $p<0.001$; $d_z=1.18$). Switching costs were not modulated by *arm posture* (*arm posture* x *modality transition* $F(1,15)=1.28$; $p=0.276$; $\eta^2_p=0.08$).

- b) *Is there a visual Simon effect on visual trials following another visual trial, but washed out when there is a modality switch from N-1 to N?*

For Experiment 2 (Figure 3), the three-way interaction between *modality*, *modality transition* and *S-R compatibility* was far from significant [$F(1,15)=2.21$; $p=0.158$; $\eta^2_p=0.13$], as was the four-way interaction between *modality*, *arm posture*, *modality transition* and *S-R compatibility* [$F(1,15)=0.34$; $p=0.566$; $\eta^2_p=0.02$], thus indicating that, overall, the Simon effect for both the visual and tactile trials did not differ significantly based on the modality of the N-1 trial.

However, due to the specific question we wanted to address, we looked in more detail at the three-way interaction (*modality*, *modality transition* and *S-R compatibility*) and computed a series of planned t-test comparisons: the compatibility effect was of -11 ms and 4 ms for visual maintain and visual switch trials, respectively [$t_{\text{maintain}}(15)=0.94$; $p=0.363$; $d_z=0.24$; $t_{\text{switch}}(15)=-0.27$; $p=0.794$; $d_z=0.07$], thus indicating the absence of visual Simon effects in visual trials following other visual trials. In the case of touch, S-R compatibility was 62 ms and 42 ms for maintain and switch trials respectively [$t_{\text{maintain}}(15)=-3.42$; $p=0.004$; $d_z=0.86$; $t_{\text{switch}}(15)=-2.37$; $p=0.031$; $d_z=0.59$], thus confirming the presence of tactile Simon effects for tactile trials regardless of prior trial modality.

For Experiment 3 (Figure 5), the three way interaction between *modality*, *modality transition* and *S-R compatibility* approached significance [$F(1,15)=3.90$; $p=0.067$; $\eta_p^2=0.21$]. Meanwhile, in Experiment 2, where we explored this interaction through planned t-test comparisons: S-R compatibility effects were on average -12 ms for visual maintain trials and -3 ms for visual switch trials [$t_{\text{maintain}}(15)=2.32$; $p=0.035$; $d_z=0.58$; $t_{\text{switch}}(15)=-0.45$; $p=0.661$; $d_z=0.11$], indicating the absence of visual Simon effects when the visual modality was maintained. In contrast, tactile Simon effects showed on average 49 ms and 40 ms for maintain and switch trials respectively with both resulting in significant values on separate analyses [$t_{\text{maintain}}(15)=-5.56$; $p<0.001$; $d_z=1.39$; $t_{\text{switch}}(15)=-3.46$; $p=0.004$; $d_z=0.87$].

Discussion

Contrary to our initial hypothesis, results indicated that when participants responded to targets in a modality-switching Simon task, the visual S-R compatibility faltered, even under the canonical, uncrossed-arms condition, whilst the tactile S-R compatibility persisted and was expressed in its ‘native’, anatomical and spatial frame of reference. This pattern, initially found in Experiment 2, was confirmed in a subsequent independent experiment (Experiment 3). We reported here two independent replications of the same finding in order to accentuate the consistency of the main result (i.e., the disappearance of visual Simon effects on visuo-tactile switching tasks), which was contrary to our initial hypothesis. Given the consistence of evidence across Experiments 2 and 3 and given that the Simon effect (in vision) is one of the most reproducible effects in experimental psychology (J.R. Simon, 2011), one must wonder at what caused such a dramatic collapse.

One possible explanation may be related to the overall slower RTs in these experiments as compared to single-modality Simon tasks (Experiment 1). Some studies have reported that S-R compatibility effects are strongly reduced (or even reversed) for slow responses (De Jong, Liang, & Lauber, 1994; Hommel, 1994b). In order to test for any possible influence of overall RTs on the Simon effects reported in our previous experiments, we decided to run a distributional RT analysis (see Supplemental material). For Experiment 2, the magnitude of the visual Simon effect for crossed and uncrossed-arms separately did not change as a function of

the speed of participants' RTs, thus excluding the possibility that slower RTs could account for the attenuation of the Simon effect in the visual modality. The RT distributional analysis on data from Experiment 3 (see Supplemental material) revealed that for the fastest, visual RT quartile, there was, in fact, some evidence for S-R compatibility expressed in an external reference frame.

Another, theoretically more interesting possibility for explaining the disappearance or at least the strong attenuation of visual Simon effects on the modality-switching version of the task, is to consider the influence of carryover effects from trial-to-trial. Note that, switching from one modality to another also implicitly involved a change in the spatial frame of reference (from/to anatomical to/from external). Hence, it is possible that switching is asymmetrical, in the sense that, it is harder to switch from an (tactile) anatomical frame of reference to an (visual) external one, than the other way around. Following this prediction, the visual Simon effect was attenuated because the anchor to the anatomical frame of reference prevented visual S-R compatibility from manifesting itself. When the modality of the previous trial was taken into account in statistical analyses, we observed modality-switching costs for both visual and tactile trials, which were stronger when going from touch to vision than vision to touch. The presence of modality-switching costs suggests that the control mechanism (i.e., executive functions) activated by the modality-switching Simon task is, in fact, modality-dependent, in agreement with previous studies (Hunt & Kingstone, 2004; Murray et al., 2009). This asymmetry in switching cost may partially explain, why, contrary to our initial expectations, the tactile S-R compatibility pattern remained whilst the visual Simon effect vanished on mixed-modality trials. Indeed, being harder to switch away from touch, no cognitive resources are left for the processing of the irrelevant spatial information in the visual trials, which stops to affect the responses, hence no S-R compatibility effects are observed (Kiesel et al., 2010; Monsell, 2003; Rogers & Monsell, 1995). Nonetheless, the asymmetrical modality switch cost alone cannot account for the collapse of visual Simon effects, considering that S-R compatibility effects also disappeared on modality-maintain visual trials, or when no switching was involved. Our study, to the best of our knowledge, is the first time that a modality switch manipulation is introduced

to the Simon task; however, cross-modal spatial compatibility effects have, in fact, been previously investigated (Lukas, Philipp, & Koch, 2010a, 2010b, 2014; Van der Lubbe & Van der Helden, 2006). In one study, by Lukas and collaborators (2010a), participants were instructed to switch their attention between modalities (auditory and visual) according to a cue presented at the beginning of each trial. Participants had to respond to the spatial position of either the simultaneously presented auditory or visual stimulus. There was a consistent and asymmetrical modality-switching cost, so that, RTs were shorter for visual stimuli as compared to auditory stimuli and visual distracters induced larger interference when processing auditory targets than the other way around (Lukas et al., 2010a). In contrast to these findings, spatial information in our study was completely irrelevant (as in the Simon task). Furthermore, on our version of the modality-switch Simon task, attention was not directly manipulated by a cue, so voluntary pre-selection of stimulus modality (and necessarily its frame of reference) was impossible.

We hypothesized that in Experiments 2 and 3, the unpredictable changes in stimulus modality involved uncertainty about the relevant S-R compatibility pattern (based on either an external or an anatomical frame of reference). By switching modalities on a Simon task, we introduced an important source of conflict, which had an impact on the management of spatial information. In fact, if on the single-modality Simon tasks (visual or tactile) the possible (mis)match of spatial information was resolved by abiding to one (external) frame of reference or another (anatomical), on the modality-switching version, it was impossible to select an appropriate frame of reference prior to the appearance of the stimulus (incidentally, it would appear that it was not possible to prepare for both frames of reference at the same time either).

However, before adopting the above conclusion, it is important to evaluate two possible alternative explanations for the vanishing of the visual Simon effect, which do not involve direct competition between S-R compatibility patterns or differential frames of reference. The first alternative explanation refers to the proximity between location of stimulation and location of response in the case of tactile events, which could have simply magnified the tactile S-R compatibility effects to the detriment of the visual S-R compatibility effects (we test this

explanation in Experiment 4). The second alternative account refers to the fact that any modality switch in the Simon task can cancel out its resulting S-R compatibility effect due to the higher cognitive cost of maintaining two possible responses. The latter could provide a simpler account than the more nuanced interpretation we hypothesized which, related to the implicit switch of frames of reference in coding stimuli for both vision and touch (tested in Experiment 5).

Experiment 4: Visuo-Tactile mixed-modality Simon Task with foot response

Contrary to visual stimuli, which occur at a fixed position in space and are independent of bodily manipulations, the locations of tactile stimuli (presented to the index finger) necessarily move along with their response effector, which changes according to the position of the arm(s). Therefore, reactions to congruent, tactile stimuli may be faster than reactions to incongruent ones simply because stimulus location physically overlaps with the position of the responding finger. This alternative could invalidate the classical account of the Simon effect for the tactile modality, not only on the mixed-modality version of the task (Experiments 2 and 3), but also on the tactile Simon task in general (Experiment 1). In a recent study, Medina and collaborators (2014) tested participants using a tactile Simon task, where the stimulation was presented to the fingers but responses were provided by foot pedals³. Notwithstanding, the novel experimental manipulation, which we introduced in Experiments 2 and 3 (i.e., mixed-modality Simon task) and the ensuing results, a surprising obliteration of the visual Simon effect and a strong maintenance of the tactile Simon effect, steers doubts. In particular, one could imagine that any S-R compatibility Simon effect would simply vanish under mixed-modality conditions with a potential effect of stimulus-effector overlap (only present in the tactile modality) remaining (accounting for faster responses to congruent targets than incongruent ones). This peculiarity of tactile events could have favoured S-R compatibility for the tactile modality to the detriment of the visual modality (supported by our results in Experiments 2 and 3). To account for this possibility, we ran a new experiment (Experiment 4), which was a replication of Experiments 2 and 3 (visuo-tactile mixed-modality Simon task), but in which participants were asked to respond by using foot pedals instead of finger presses, so that the spatial positions of the stimuli

(on the finger) and those of the response effector (the foot) were misaligned in terms of their anatomical location. Given the focus of this particular experiment, only the canonical, uncrossed-arm posture was used.

Methods

Participants: We tested 19 right-handed participants. Three participants were excluded because of low performance accuracy or slow RTs (<65% in at least one of the conditions or RTs > 1500 ms for more than 5% of the total number of trials). Data from the remaining 16 participants (11 women; age range 18-28 years) underwent statistical analysis.

Procedure: The procedure was similar to the one used in Experiment 3. Participants were instructed to respond as fast as possible to the colour (in the case of a visual stimulus) or frequency (in the case of a tactile stimulus) of the stimulus, regardless of its spatial position (left or right). A total of 416 trials (104 trials per condition: modality-compatible and modality-incompatible stimulus-response positions) were collected in about 45 minutes (breaks included).

Results

We conducted a repeated measures ANOVA on RTs with *modality* (vision/touch) and *S-R compatibility* (compatible/incompatible) as within-participants variables. The main effect of *modality* was significant ($F(1,15)=20.23$; $p<0.001$; $\eta^2_p=0.57$). Visual stimuli were responded to significantly faster than tactile stimuli (515 ms vs. 617 ms, respectively). Critically, the interaction between *modality* and *S-R compatibility* was also significant ($F(1,15)=14.82$; $p=0.002$; $\eta^2_p=0.50$). In the visual modality (Figure 6A), RTs to compatible trials were 20 ms slower than RTs to incompatible trials (525 ms vs. 505 ms), a difference which, had a tendency toward significance in the opposite direction with respect to the typical Simon effect [$t(15)=2.03$; $p=0.06$; $d_z=0.51$]. On the contrary, in the tactile modality (Figure 6B), the Simon effect was strong and in the expected direction [593 ms vs. 640 ms; $t(15)=-2.70$; $p=0.016$; $d_z=0.68$].

The repeated measures ANOVA on accuracy, with *modality* (vision/touch) and *S-R compatibility* (compatible/incompatible) as within-subjects variables, only revealed a significant

main effect of *modality* ($F(1,15)=17.29$; $p<0.001$; $\eta^2_p=0.54$), with responses to visual stimuli being more accurate than responses to tactile stimuli (97 % vs. 92 %).

Discussion

Tactile events presented on the index fingers may trigger faster responses when the location of stimulus presentation aligns with the necessary responding effector (in the case of compatible trials). In contrast, RTs for tactile events occurring on the opposite response effector (or incompatible trials) may be significantly slower.. This may lead tactile events, presented on the fingers, to amplify tactile Simon effects to the detriment of visual Simon effects thus offering a simplified explanation for the pattern of results seen in Experiments 2 and 3 and therefore invalidating the proposed hypothesis based on a competition between different frames of reference. Yet, Experiment 4 showed that tactile Simon effects remained strong, while visual Simon effects disappeared (and even tended to reverse), when responses were provided with foot pedals, rather than key-presses, in line with results from Medina et al. (2014). Furthermore, the fact that S-R compatibility in the visual modality had a tendency to reverse (although non-significant) cannot be explained by carry-over effects due to arm posture manipulation; since this experiment did not include crossed-arm conditions we will refrain from discussing this non-significant pattern any further.

Overall, the pattern of results from the visuo-tactile mixed-modality Simon task with pedal responses fails to provide evidence for faster responses due to anatomical overlap as the reason for the strong, tactile Simon effects and reduced, visual Simon effects. Therefore, the interpretation of a competition between frames of reference in the modality-switching Simon task appears to stand, as a potential explanation, for the dampening of visual Simon effects.

Experiment 5 Audio-Visual mixed-modality Simon Task

It has been shown that an increase in task difficulty is associated with an increase in cognitive/perceptual load which, in turn decreases the magnitude of the Simon effect (Hommel, 1994a). Cognitive/perceptual load may partially explain the modulations of the Simon effect seen in our experiments, because during the modality-switching experiments (Experiments 2-4)

participants had to maintain two possible response mappings (one for each modality) while in the single-modality version of the task (Experiment 1) participants only had to maintain one. Switching between modalities (Experiments 2 and 3) could have slowed RTs to the point that the influence of irrelevant, spatial information ceased to be effective (see also Hommel, 1993a). Although this alternative explanation fails to provide an account for the asymmetrical modulation of visual and tactile Simon effects, the nature of cognitive load effects proposes a very different rationale compared to the cost of switching frames of reference, which we propose here. To test the cognitive/perceptual load hypothesis, we decided to run a fifth experiment in which there was modality switching, so that the cognitive/perceptual load would be equivalent to Experiments 2-4, but that did not induce switches between frames of reference (and hence, S-R compatibility patterns). To do this we used auditory and visual stimuli, because audition and vision are distal modalities and their dominant S-R compatibility patterns overlap. Both audition and vision use an external frame of reference to encode space. In other words, in both modalities RTs are faster when the stimulus location matches that of the response goal, regardless of anatomical hand used to make the required response (for vision: Riggio et al., 1986; Wallace, 1971; for audition: Röder et al., 2007; Roswarski & Proctor, 1996; Simon et al., 1970). Thus, in Experiment 5 we used an audio-visual mixed-modality Simon task to maintain the cognitive demands of a mixed-modality Simon task whilst eliminating possible switches in frames of reference.

According to our hypothesis, that modality switching involves changing between spatial frames of reference, this experiment should induce significant Simon effects (deployed in external space) for both modalities, that is, a pattern of S-R compatibility, which should reverse under crossed-arm posture.

Methods

Participants: We tested 18 right-handed participants. Two participants were excluded based on low performance accuracy ($< 65\%$ in at least one of the conditions) or RTs > 1500 ms for more than 5% of the total number of trials. Data from the remaining 16 participants (8 women; age range 19-25 years) were statistically analyzed.

Procedure: The procedure was similar to the one used in Experiment 3. Participants were instructed to respond as fast as possible to the type of sound (in the case of auditory stimuli) or colour (in the case of visual stimuli) of the target, regardless of spatial position.

Results

We conducted a repeated measures ANOVA on the RTs with *modality* (vision/audition), *arm posture* (uncrossed/crossed) and *S-R compatibility* (compatible/incompatible) as within-participants variables. A significant main effect of *modality* was present ($F(1,15)=17.47$; $p<0.001$; $\eta^2_p=0.54$), with RTs to visual stimuli being significantly faster than those to auditory stimuli, 60 ms on average (450 ms (V) vs. 510 ms (A)). The main effect of *arm posture* was also significant ($F(1,15)=22.44$; $p<0.001$; $\eta^2_p=0.460$), with RTs under crossed-arm conditions being significantly slower, on average by 53 ms, than those under uncrossed-arm conditions. The two-way interactions between *modality* and *S-R compatibility* ($F(1,15)=11.61$; $p=0.004$; $\eta^2_p=0.44$) and *arm posture* and *S-R compatibility* ($F(1,15)=51.86$; $p<0.001$; $\eta^2_p=0.78$) both reached significance; however, the three-way interaction, between *modality*, *S-R compatibility* and *arm posture*, did not ($F(1,15)=3.26$; $p=0.091$; $\eta^2_p=0.18$). Indeed, contrary to Experiments 2 and 3 (mixed-modality visuo-tactile Simon tasks) for which we expected the opposite pattern of results, in Experiment 5 we did not expect the same. Audition and vision share patterns of S-R compatibility therefore, crossing the arms would lead to a reversal of S-R compatibility effects for both modalities (as per our coding of S-R locations). In order to verify this hypothesis, we executed planned t-test comparisons between visual and auditory trials, in order to confirm the significant two-way interaction (*arm posture* x *S-R compatibility*) within each modality. For visual trials (Figure 7A), the compatibility effect was significant in the uncrossed-arm posture (19 ms; $t(15)=-2.86$; $p=0.012$; $d_z=0.72$) as well as in the crossed-arm posture (-38 ms; $t(15)=5.42$; $p<0.001$; $d_z=1.36$) and results were in opposite directions, as expected. For auditory trials (Figure 7B), the compatibility effects were of approximately 26 ms ($t(15)=-3.46$; $p=0.003$; $d_z=0.87$) and -5 ms ($t(15)=1.01$; $p=0.327$; $d_z=0.25$) for the uncrossed and crossed-arm postures.

An analysis of accuracy data revealed a significant two-way interaction between *modality* and *S-R compatibility* ($F(1,15)=6.27$; $p=0.024$; $\eta^2_p=0.29$), where accuracy was equivalent for visual

compatible and incompatible trials (94 vs. 95 %; $t(15)=-1.04$; $p=0.320$; $d_z=0.26$) and slightly higher in the compatible than incompatible trials for audition (95 vs. 93 %; $t(15)=-2.19$; $p=0.04$; $d_z=0.55$). Nonetheless, these accuracy values are very high overall and their variability could potentially be lowered by saturation. The two-way interaction between *arm posture* and *S-R compatibility* ($F(1,15)=13.08$; $p=0.003$; $\eta^2_p=0.47$) was significant: with accuracy being higher on uncrossed-arm compatible trials than uncrossed-arm incompatible trials (96 % vs. 93 %; $t(15)=3.76$; $p=0.002$; $d_z=0.94$). In the crossed-arms condition, the pattern of results tended to reverse marginally, with heightened accuracy on incompatible crossed-arm trials than compatible crossed-arm trials (93 % vs. 95 %; $t(15)=-1.93$; $p=0.07$; $d_z=0.48$).

In summary, both visual and auditory Simon effects survived in an audio-visual mixed-modality Simon task. These effects were modest but significant in both modalities, consistent with the dominance of an external frame of reference; although the Simon effect in the crossed-arms posture was not evident in the auditory modality. Analogously to Experiments 2 and 3, we ran the N-1 analysis for Experiment 5 (see Supplemental materials).

Discussion

As advanced by our hypothesis, the results from Experiment 5 revealed that the visual Simon effect remained, despite mixing modalities. This is in stark contrast to the visuo-tactile mixed-modality experiments (Experiments 2 and 3) and supports the hypothesis that switching between S-R compatibility patterns, via spatial reference frames associated to each modality, obliterates the visual Simon effect. Indeed, here, audio-visual switching did not involve a switch in S-R compatibility patterns, and hence, there was no effect on frames of reference (as all modalities were expressed in an external frame of reference). As previously shown, using crossed-arm manipulations, the conflict introduced by the processing of irrelevant spatial information for visual (Riggio et al., 1986; Wallace, 1971) and auditory (Röder et al., 2007; Roswarski & Proctor, 2000; Simon et al., 1970) events is resolved by considering external space as the dominant frame of reference. The results from this experiment therefore negate the alternative explanation that the collapse of the visual Simon effect in Experiments 2-4 could be attributed to cognitive/perceptual overload (Hommel, 1994a).

Analysis of the visual Simon effect across experiments

In order to appreciate higher-order effects across experiments, we ran a mixed ANOVA on the RTs of visual trials in the uncrossed and crossed-arm postures separately, with *S-R compatibility* (compatible/incompatible) as within-participants variables and *experiment* (5 for the uncrossed-arm posture and 4 for the crossed-arm posture) as a between-participants variable. Overall, when the participants' arms were uncrossed, RTs on the visual single-modality Simon task (Experiment 1) were faster than RTs on the mixed-modality experiments (all p s < 0.025), with the exception of Experiment 4 (foot response) where there was a main effect of *experiment* ($F(4,69)=7.43$; $p<0.001$; $\eta^2_p=0.30$). Interestingly, the two-way interaction between *S-R compatibility* and *experiment* was significant ($F(4,69)=4.29$; $p=0.004$; $\eta^2_p=0.20$). The visual Simon effect disappeared when visual trials were intermixed with tactile ones ($t_2(69)=-1.53$; $p=0.131$; $d_z=0.38$; and $t_3(69)=0.17$; $p=0.862$; $d_z=0.04$, for Experiments 2 and 3) and even reversed, when participants made pedal responses such as those required on Experiment 4 ($t_4(69)=2.27$; $p=0.026$; $d_z=0.57$). Conversely, the visual Simon effect was present either when visual trials were presented alone (Experiment 1; $t_1(69)=-2.71$; $p=0.009$; $d_z=0.68$) or intermixed with auditory trials (Experiment 5; $t_5(69)=-2.19$; $p=0.032$; $d_z=0.55$).

Crossing the arms over the body midline also revealed the external coding of visual S-R comparability, either when the visual target was presented alone (Experiment 1; $t_1(69)=3.31$; $p=0.002$; $d_z=0.83$) or intermixed with auditory trials (Experiment 5; $t_5(69)=3.34$; $p=0.002$; $d_z=0.84$). When visual trials were intermixed with tactile trials, the visual Simon effect was also strongly reduced when a crossed-arm posture was adopted ($t_2(69)=1.89$; $p=0.064$; $d_z=0.47$; and $t_3(69)=0.97$; $p=0.339$; $d_z=0.24$).

Summary of the results

The main findings related to S-R spatial compatibility (Simon effect) to emerge from the five experiments above are summarized in Figure 8. We capitalized on the arm crossing manipulation to disambiguate the dominant spatial reference frame in S-R compatibility effects

(external vs. anatomical). In single-modality situations, the pattern of RTs indicated the prevalence of an external frame of reference for spatial representation in the visual modality. In contrast, RTs to tactile targets indicated the prevalence of an anatomical frame of reference. The present study introduced, for the first time, mixed-modality blocks requiring switches from trial to trial, so that the spatial representations associated with different modalities must be necessarily adjusted. This context is more akin to what the cognitive system must deal with in real-life situations, where stimuli coming from different modalities require motor responses. Remarkably, we found that when visual and tactile events were mixed on a Simon task, the visual Simon effect disappeared. This was shown in three independent experiments (Experiments 2, 3 and 4). The collapse of visual Simon effects was corroborated even when the analysis was restricted to visual trials following other visual trials (Experiments 2, 3), and when the location of the tactile stimulation and the location of the response effector were misaligned (Experiment 4). In all of these cases, visual Simon effects disappeared and tactile Simon effects remained. Critically, when mixed-modality blocks involved a switch from audition to vision or vice versa (Experiment 5), the visual Simon effect remained and its consequent S-R compatibility effect was expressed in the characteristic, external frame of reference.

General Discussion

Irrelevant spatial representations in a cross-modal scenario

The abundant literature on the Simon effect is accompanied by a stimulating debate on its underlying cognitive mechanisms (for a recent discussion see Hommel, 2011; Van der Lubbe & Abrahamse, 2011). In his fundamental paper, Wallace (1971) first proposed that the body (i.e., the responding hand) represents a source of spatial information that should be treated as a stimulus itself. According to the *coding hypothesis* proposed by Wallace (1971), the hand is represented in the form of a “body code” and, as such, can be associated with a location in space, which is updated as it moves from one position to another (Wallace, 1971, p. 359). The *coding hypothesis* (Wallace, 1971) or the more recent *referential-coding hypothesis* (Hommel, 1993b) provides a parsimonious framework to interpret the S-R compatibility pattern arising

from switches in modality. The emergence of S-R compatibility patterns implicates the coding of the stimulus (including the hand) in reference to an “...unintentional defined object[s] or frame of reference (or several of them) [...]” (Hommel, 1993b, p. 209). Thus, the necessity to manipulate different, spatial frames of reference is somehow contemplated in these frameworks, albeit neither proposal, by Wallace or Hommel, specified the conditions, which determine the relevance of one frame of reference to another.

Crossing the arms over the body midline can reveal the management of spatial representations as they correspond to the “body code” (Nicoletti & Umiltà, 1994; Riggio et al., 1986; Wallace, 1971). The coding of the spatial position of the responding hand, as postulated by Wallace (1971), even if not directly responsible for S-R compatibility on the visual Simon task (dependent on the position of the response button), affects the results as shown by crossing the arms on the modality-switching tasks. We argue that the postural manipulation (i.e., crossing the arms) as well as the uncertainty triggered by switches in modality leads to the emergence of different, yet coexisting, spatial representations of the stimuli. On one hand, crossing the arms leads to a mismatch between sources of spatial information even within a single modality. On the other, many studies in the past have demonstrated that the expression of S-R compatibility patterns depends on the modality of the sensory stimulus (Hasbroucq & Gliard, 1991; Medina et al., 2014; Riggio et al., 1986; Röder et al., 2007; Roswarski & Proctor, 2000; Wallace, 1971). Hence, switching between stimulus modalities forces one to select the relevant frame of reference according to the spatial representation that has to be formed, as a function of sensory modality. The unpredictable task of modality-switching prevents both preparation and/or pre-activation of a particular frame of reference and, implies that the relevance (or weight) assigned to the different sources of spatial information must be actively re-calculated to build up the necessary S-R compatibility pattern.

Manipulations of the Simon task in the past have introduced switching, or competition, between multiple frames of reference within the visual modality (Hommel, 1993b; Hommel, Musseler, Aschersleben, & Prinz, 1994; Klapp et al., 1979; Ladavas & Moscovitch, 1984; Lamberts et al., 1992; Roswarski & Proctor, 1996; Stoffer, 1991; Umiltà & Liotti, 1987). These

studies provided an intuition that S-R compatibility effects may collapse in such situations, perhaps because reacting quickly, in more than one frame of reference, is an excessively demanding task (although please note, once more, that spatial representations about stimulus location are irrelevant on the Simon task). Results from Experiments 2-4 in this study, by inducing modality switches, provide direct empirical evidence for the cognitive challenge of forming and maintaining multiple reference frames concurrently. Hence, a comprehensive theory on spatial representations, as the numerous ones proposed since the initial discovery of the Simon effect by Simon & Small (1969), should consider situations in which spatial codes arise from more than one sensory modality, as seen in naturalistic settings. To arrive at such an account, hypotheses should generalize to cross-modal situations. For example, in the context of spatial coding, the influence of response modality has been given a lot of weight (see Lu & Proctor, 1995) and has been shown to cause asymmetries on the spatial Stroop effect for example (using key-press or vocal responses; Virzi & Egeth, 1985). This would suggest that some stimulus attributes automatically facilitate responses of a particular kind. More recently, the implication of S-R compatibility patterns has been considered in relation to human-computer interface designs (Proctor & Vu, 2016). This study adds to this line of research on how spatial representations are formed in real-life contexts by considering both sensory modalities (vision and audition; vision and touch) and the intimate relationship between stimulus and response as seen on the Simon task.

It is interesting to note that when mixing modalities (visual and tactile), it is the visual (irrelevant) spatial dimension that ceases to affect the responses to the stimuli, rather than the tactile dimension. Despite the fact that this could suggest that the anatomical (tactile) frame of reference dominates over the external (visual) one when there is competition between the two, we do not believe that this is the case. Instead, the lack of visual Simon effects seems to imply that irrelevant spatial representation of the visual stimulus simply stops affecting the responses, rather than being re-coded in an anatomical frame of reference altogether, as anatomical dominance would imply. To explain this pattern, we argue that when more than one frame of

reference can be formed and modality (and subsequent frame of reference) is uncertain, the weight of different sources of spatial information must be re-assigned on a trial-by-trial basis following stimulus presentation, but not before. Here, the concept of ‘weight’ refers to the relevance of a given source of spatial information, as a function of its spatial representation of a stimulus, at a given moment in time. We think that the management of spatial information happens on-line so that selection of a relevant reference frame occurs accordingly. We suggest the idea of active re-weighing of spatial representations rather than the concept of a competition between frames of reference because if these (external and anatomical) frames of reference engaged in a competition, and the anatomical frame of reference dominated then we would expect the anatomical frame of reference to take over in all instances of S-R compatibility. Instead, we observe evidence for the existence of some form of visual stimuli spatial information processing even under modality-switch conditions where the Simon effect disappeared (see: residual, visual Simon effects for the fastest RTs in the distributional analysis of Experiment 3). Please note that this ‘on the fly re-weighing’ argument does not yet explain the asymmetry between vision and touch. This will be discussed later.

The role of task difficulty

Cognitive and/or perceptual cognitive load has been shown to diminish the magnitude of the Simon effect (Hommel, 1994a). Therefore, an alternative explanation of the lack of Simon effects on visual trials in the visuo-tactile mixed-modality conditions of this study could be that mixing modalities increases cognitive load (because of the need to maintain two different response mappings). Longer RTs made the relevance of the spatial dimension decay (De Jong et al., 1994). Yet, this does not appear to be the reason for the elimination of the Simon effects seen in this particular case. First, if task difficulty caused the present pattern of results, then one would expect S-R compatibility effects to vanish on (more difficult, hence slower to respond) tactile trials, and remain for the (easier, hence faster to respond) visual trials. Furthermore, accuracy was also higher on visual than tactile trials, thus indicating that overall, the tactile task was more difficult than the visual one. Second, if task difficulty were to explain the

disappearance of the Simon effect on visual trials, then one would expect compatibility effects at least on modality-maintain trials where RTs are significantly faster than modality-switch trials on the modality-mixed version of the experiment. However, visual Simon effects were not evident even on modality-maintain trials, where modality switching was absent (see Experiments 2 and 3). Third, the results of Experiment 5 show that the added difficulty of modality switching per se does not obliterate the Simon effect, when the frame of reference is unambiguous. Auditory stimuli indeed received responses more slowly than visual ones, yet we were able to appreciate S-R compatibility in an external reference frame.

The attentional selection account on the modality-switching Simon task

Besides the *referential coding hypothesis*, another relevant theoretical framework regarding the cognitive mechanisms behind the Simon effect puts emphasis on the role of attention (Hommel, 2011; Nicoletti & Umiltà, 1994; Sheliga, Craighero, Riggio, & Rizzolatti, 1997; Van der Lubbe & Abrahamse, 2011). Briefly, according to the *early attentional selection hypothesis*, it is the programming of an attention shift toward stimulus location rather than stimulus position itself (as suggested by the *coding hypothesis* Hommel, 1993b; Wallace, 1971), which causes the S-R compatibility pattern to emerge on the Simon task (Nicoletti & Umiltà, 1994; Stoffer, 1991). According to the *late attentional selection hypothesis*, selecting a location in space primes action towards that location and eventually shifts attention in that direction (Sheliga et al., 1997).

In order to explain the results from Experiments 2-4, one does not need to appeal to the attentional selection (early or late) account of the Simon effect. In our experiments we did not directly modulate the direction of exogenous (e.g., through stimulus salience) or endogenous attention (e.g., cueing for one or the other modality). Whereas the appearance of a lateralized stimulus is likely to induce an exogenous shift of attention to that side of space, this transient shift in attention should occur in the same way on single (Experiment 1) and mixed-modality versions of the Simon task (Experiments 2-5). Yet, the Simon effect occurred in some cases (Experiments 1 and 5 and Experiments 2, 3 and 4 for tactile events only), and not in others

(Experiments 2, 3 and 4 for visual events). Therefore, spatial orienting is presumably orthogonal to the effect of modality-switching in our study.

Nevertheless, there is a similarity between the modality-switching paradigm presented here and previous experimental situations based on the attentional selection account where the visual Simon effects disappeared (Hommel, 1993b, 1994a; Klapp et al., 1979; Ladavas & Moscovitch, 1984; Lamberts et al., 1992; Roswarski & Proctor, 1996; Stoffer, 1991; Umiltà & Liotti, 1987). Specifically, in the no-delay condition in a study by Umiltà and Liotti (1987; Experiment 3), similarly to our modality-switching version of the Simon task, the critical factor is that more than one spatial code can be formed prior to stimulus appearance. We think that under conditions of uncertainty, when spatial information can be framed according to multiple (external, anatomical) anchors, the weight given to each source of information must be actively re-calculated eventually leading to the disappearance of the visual Simon effect.

Visual (external) reference vs. tactile (anatomical) reference frames

The present results also reveal an interesting asymmetry between how visual and tactile spatial representations prevail when taking into account multiple frames of reference. Indeed, Experiments 2, 3 and 4 show that, while visual S-R compatibility disappears, tactile events continue to be encoded in an anatomical frame of reference, in cross-modal contexts. Furthermore, the magnitude of the switching cost was the same when switching from vision to audition and audition to vision (Experiment 5). In fact, the cost of switching from touch to vision was larger than switching from vision to touch (Experiments 2 and 3). This may suggest that tactile S-R compatibility patterns (anatomical frames of reference) are more resilient to the conflict imposed by switching between modalities as compared to visual, externally framed S-R compatibility patterns. As a consequence, the visual Simon effect vanished, whilst the tactile Simon effect remained. This result is unexpected because, in many other contexts, vision dominates over the other modalities (Colavita, 1974; Rock & Victor, 1964; Sinnett et al., 2007), especially when it comes to the spatial domain (Lukas et al., 2010a). Indeed, in tactile remapping for example, spatial information in regards to touch seems to automatically align to

external coordinates within a visual and spatial frame of reference, rather than remaining in an anatomical one (Röder, Rösler, & Spence, 2004). According to several studies, tactile information is re-mapped automatically into external coordinates for action and this spatial code then competes with the original, tactile anatomical code, often leading to erroneous spatial localization of touch when the hands are crossed (Yamamoto & Kitazawa, 2001), producing fleeting orientation of attention (Azañón et al., 2010; Azañón & Soto-Faraco, 2008) and even misdirected saccades to tactile events (Overvliet et al., 2011). Furthermore, this remapping of tactile stimuli into external coordinates has been hypothesized to be mandatory for action (Heed et al., 2015), and to be correlated with tactile awareness (Kitazawa, 2002). The apparent anomaly, in Experiments 2, 3 and 4 is therefore that the visual reference frame (and hence Simon effect) vanished, while the tactile reference frame (and its associated Simon effect) remained, suggesting that in the latter the anatomical S-R compatibility pattern is harder to shut down.

Related to these results, some studies have shown that in the visual modality, when visual location cues (i.e., external) are rendered ineffective by adopting an unusual body posture (i.e., a head tilt), visual S-R compatibility relies on an egocentric frame of reference (Ladavas & Moscovitch, 1984). Furthermore, on tasks involving implicit modality expectation, touch appears to be more difficult to ignore as compared to vision and audition (Spence, Nicholls, & Driver, 2001). Indeed, contrary to vision, which is a distal modality; tactile sensations originate from stimuli directly impinging on the bodily surface. For these reasons, according to some authors, the processing of tactile events intrinsically gives a stronger weight to an anatomical frame of reference (“body code”) (Berlucchi, Crea, Di Stefano, & Tassinari, 1977; Gallace & Spence, 2009), which seems to be more difficult to silence.

It is important to note however, that touch has often been overlooked in the literature regarding spatial S-R compatibility, when compared to both vision and audition. In contrast, touch may be used to enhance human-computer interactions. Especially considering the implication that basic research has for the design of human-computer interfaces (see Proctor &

Vu, 2016), the present results can add to our knowledge in this field, showing that switches in modality engender switches in frames of reference necessary for spatial coding.

Conclusion

Because the processing of irrelevant spatial information is (often) triggered in an automatic fashion (i.e. not under volitional control) it may induce performance costs, as shown by the Simon effect. It is well known that the single-modality Simon effect involves competition between relevant and irrelevant response codes according to a modality-relevant frame of reference. However, when multiple frames of reference come into play, as in the case of the mixed-modality Simon task, the orchestration of spatial information necessarily relies on the re-assignment of the weight given to different (possible) sources of spatial information in an active, on-line manner. According to the present results, this re-assignment can alter the typical (i.e., the most frequently tested, unimodal) S-R compatibility pattern in an important way, so that certain compatibility effects are eliminated or at least, strongly reduced (visual Simon effect in visuo-tactile mixed-modality conditions), whilst others remain (tactile Simon effect).

Bibliography

- Azan, E., Cardini, F., Haggard, P., Stenner, M., & Cardini, F. (2015). Report Dynamic Tuning of Tactile Localization to Body Posture Dynamic Tuning of Tactile Localization to Body Posture. *Current Biology*, 25(4), 1–6. <http://doi.org/10.1016/j.cub.2014.12.038>
- Azañón, E., Camacho, K., & Soto-Faraco, S. (2010). Tactile remapping beyond space. *European Journal of Neuroscience*, 31(10), 1858–1867. <http://doi.org/10.1111/j.1460-9568.2010.07233.x>
- Azañón, E., & Soto-Faraco, S. (2008). Changing reference frames during the encoding of tactile events. *Current Biology : CB*, 18(14), 1044–9. <http://doi.org/10.1016/j.cub.2008.06.045>
- Berlucchi, G., Crea, F., Di Stefano, M., & Tassinari, G. (1977). Influence of spatial stimulus-response compatibility on reaction time of ipsilateral and contralateral hand to lateralized light stimuli. *Journal of Experimental Psychology. Human Perception and Performance*, 3(3), 505–17. <http://doi.org/10.1037/0096-1523.3.3.505>
- Bertelson, P., & Aschersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin & Review*, 5(3), 482–489. <http://doi.org/10.3758/BF03208826>
- Braem, S., Abrahamse, E. L., Duthoo, W., & Notebaert, W. (2014). What determines the specificity of conflict adaptation? A review, critical analysis, and proposed synthesis. *Frontiers in Psychology*, 5(SEP), 1–13. <http://doi.org/10.3389/fpsyg.2014.01134>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–6.
- Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, 16(2), 409–412. <http://doi.org/10.3758/BF03203962>
- De Jong, R., Liang, C.-C., & Lauber, E. (1994). Conditional and Unconditional Automaticity: A Dual-Process Model of Effects of Spatial Stimulus-Response Correspondence. *Journal of Experimental Psychology: Human Perception and Performance*, 20(4), 731–750. <http://doi.org/10.1037/0096-1523.20.4.731>
- 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(1), 143–149. <http://doi.org/10.3758/BF03203267>

- Fitts, P., & Seeger, C. (1953). S-R compatibility: spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 46(3), 199–210.
<http://doi.org/10.1037/H0062827>
- Gallace, A., & Spence, C. (2009). The cognitive and neural correlates of tactile memory. *Psychological Bulletin*, 135(3), 380–406. <http://doi.org/10.1037/a0015325>
- Hasbroucq, T. (1987). Preliminary evidence for body-centered coding of tactile motor events. *Perceptual and Motor Skills*, 64, 631–634.
- Hasbroucq, T., & Gliard, Y. (1991). Stimulus-response compatibility and the Simon effect: toward a conceptual clarification. *Journal of Experimental Psychology. Human Perception and Performance*, 17(1), 246–266. <http://doi.org/10.1037/0096-1523.17.1.246>
- Hecht, D., & Reiner, M. (2009). Sensory dominance in combinations of audio, visual and haptic stimuli. *Experimental Brain Research*, 193(2), 307–14. <http://doi.org/10.1007/s00221-008-1626-z>
- Heed, T., Buchholz, V. N., Engel, A. K., & Röder, B. (2015). Tactile remapping: From coordinate transformation to integration in sensorimotor processing. *Trends in Cognitive Sciences*, 19(5), 251–258. <http://doi.org/10.1016/j.tics.2015.03.001>
- Heister, G., Schroeder-Heister, P., & Ehrenstein, W. H. (1990). Spatial coding and spatio-anatomical mapping: evidence for a hierarchical model of spatial stimulus-response compatibility. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-Response compatibility* (pp. 117–143). Elsevier B.V.
- Hollins, M., & Roy, E. A. (1996). Perceived intensity of vibrotactile stimuli: the role of mechanoreceptive channels. *Somatosensory & Motor Research*, 13(3–4), 273–86.
<http://doi.org/10.3109/08990229609052583>
- Hommel, B. (1993a). The relationship between stimulus processing and response selection in the Simon task: Evidence for a temporal overlap. *Psychological Research*, 55(4), 280–290.
<http://doi.org/10.1007/BF00419688>
- Hommel, B. (1993b). The role of attention for the Simon effect. *Psychological Research*, 55(3), 208–222. <http://doi.org/10.1007/BF00419608>

- Hommel, B. (1994a). Effects of irrelevant spatial S-R compatibility depend on stimulus complexity. *Psychological Research*, 56(3), 179–184. <http://doi.org/10.1007/BF00419705>
- Hommel, B. (1994b). Spontaneous decay of response-code activation. *Psychological Research*, 56(4), 261–8. <http://doi.org/10.1007/BF00419656>
- Hommel, B. (2011). Attention and spatial stimulus coding in the Simon task: A rejoinder to van der Lubbe and Abrahamse (2010). *Acta Psychologica*, 136(2), 265–268. <http://doi.org/10.1016/j.actpsy.2010.10.002>
- Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (1994). The theory of event-coding (TEC): A framework for perception and action planning. *Behavioural and Brain Sciences*, 24, 849–937. <http://doi.org/10.1017/S0140525X01000103>
- Hunt, A. R., & Kingstone, A. (2004). Multisensory executive functioning. *Brain and Cognition*, 55(2), 325–327. <http://doi.org/10.1016/j.bandc.2004.02.072>
- Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8), e124. <http://doi.org/10.1371/journal.pmed.0020124>
- Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching--a review. *Psychological Bulletin*, 136(5), 849–74. <http://doi.org/10.1037/a0019842>
- Kitazawa, S. (2002). Where conscious sensation takes place. *Consciousness and Cognition*. [http://doi.org/10.1016/S1053-8100\(02\)00031-4](http://doi.org/10.1016/S1053-8100(02)00031-4)
- Klapp, S. T., Greim, D. M., Mendicino, C. M., & Koenig, R. S. (1979). Anatomic and environmental dimensions of stimulus-response compatibility: implication for theories of memory coding. *Acta Psychologica*, 43(5), 367–379. [http://doi.org/10.1016/0001-6918\(79\)90031-3](http://doi.org/10.1016/0001-6918(79)90031-3)
- Kornblum, S., Hasbroucq, T., & Osman, a. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility--a model and taxonomy. *Psychological Review*, 97(2), 253–270. <http://doi.org/10.1037/0033-295X.97.2.253>
- Ladavas, E., & Moscovitch, M. (1984). Must egocentric and environmental frames of reference be aligned to produce spatial S-R compatibility effects? *Journal of Experimental*

- Psychology. Human Perception and Performance*, 10(2), 205–215.
<http://doi.org/10.1037/0096-1523.10.2.205>
- Lamberts, K., Tavernier, G., & D’Ydewalle, G. (1992). Effects of multiple reference points in spatial stimulus-response compatibility. *Acta Psychologica*, 79(2), 115–130.
[http://doi.org/10.1016/0001-6918\(92\)90028-C](http://doi.org/10.1016/0001-6918(92)90028-C)
- Lu, C.-H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, 2(2), 174–207. <http://doi.org/10.3758/BF03210959>
- Lukas, S., Philipp, A. M., & Koch, I. (2010a). Switching attention between modalities: Further evidence for visual dominance. *Psychological Research*, 74(3), 255–267.
<http://doi.org/10.1007/s00426-009-0246-y>
- Lukas, S., Philipp, A. M., & Koch, I. (2010b). The role of preparation and cue-modality in crossmodal task switching. *Acta Psychologica*, 134(3), 318–22.
<http://doi.org/10.1016/j.actpsy.2010.03.004>
- Lukas, S., Philipp, A. M., & Koch, I. (2014). Crossmodal attention switching: Auditory dominance in temporal discrimination tasks. *Acta Psychologica*, 153, 139–146.
<http://doi.org/10.1016/j.actpsy.2014.10.003>
- Medina, J., McCloskey, M., Coslett, B., & Rapp, B. (2014). Somatotopic representation of location: evidence from the Simon effect. *Journal of Experimental Psychology. Human Perception and Performance*, 6, 2131–2142. <http://doi.org/10.1037/a0037975>
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140.
[http://doi.org/10.1016/S1364-6613\(03\)00028-7](http://doi.org/10.1016/S1364-6613(03)00028-7)
- Mühlberg, S., Oriolo, G., & Soto-Faraco, S. (2014). Cross-modal decoupling in temporal attention. *European Journal of Neuroscience*, 39(12), 2089–2097.
<http://doi.org/10.1111/ejn.12563>
- Murray, M. M., Santis, L. De, Thut, G., & Wylie, G. R. (2009). The costs of crossing paths and switching tasks between audition and vision. *Brain and Cognition*, 69(1), 47–55.
<http://doi.org/10.1016/j.bandc.2008.05.004>

- Nicoletti, R., & Umiltà, C. (1994). Attention shifts produce spatial stimulus codes. *Psychological Research*, 56(3), 144–150. <http://doi.org/10.1007/BF00419701>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. <http://doi.org/10.1126/science.aac4716>
- Ossandón, J. P., König, P., & Heed, T. (2015). Irrelevant tactile stimulation biases visual exploration in external coordinates. *Scientific Reports*, 5, 10664. <http://doi.org/10.1038/srep10664>
- Overvliet, K. E., Azañón, E., & Soto-Faraco, S. (2011). Somatosensory saccades reveal the timing of tactile spatial remapping. *Neuropsychologia*, 49, 3046–3052. <http://doi.org/10.1016/j.neuropsychologia.2011.07.005>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4), 437–42.
- Pomper, U., Keil, J., Foxe, J. J., & Senkowski, D. (2015). Intersensory selective attention and temporal orienting operate in parallel and are instantiated in spatially distinct sensory and motor cortices. *Human Brain Mapping*, 36(8), 3246–3259. <http://doi.org/10.1002/hbm.22845>
- Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: an information-processing account of its origins and significance. *Psychological Review*, 83(2), 157–71. <http://doi.org/10.1037/0033-295X.83.2.157>
- Proctor, R. W., & Vu, K.-P. L. (2016). Principles for Designing Interfaces Compatible With Human Information Processing. *International Journal of Human-Computer Interaction*, 32(1), 2–22. <http://doi.org/10.1080/10447318.2016.1105009>
- Riggio, L., Gawryszewski, L. de G., & Umiltà, C. A. (1986). What is crossed in crossed-hand effects? *Acta Psychologica*, 62, 89–100. [http://doi.org/10.1016/0001-6918\(86\)90006-5](http://doi.org/10.1016/0001-6918(86)90006-5)
- Rock, I., & Victor, J. (1964). Vision and Touch: an experimentally created conflict between the two senses. *Science*, 143(3606), 594–6. <http://doi.org/10.1126/science.143.3606.594>
- Röder, B., Kusmierek, A., Spence, C., & Schicke, T. (2007). Developmental vision determines the reference frame for the multisensory control of action. *Proceedings of the National*

- Academy of Sciences of the United States of America*, 104(11), 4753–8.
<http://doi.org/10.1073/pnas.0607158104>
- Röder, B., Rösler, F., & Spence, C. (2004). Early vision impairs tactile perception in the blind.
Current Biology : CB, 14(2), 121–4. <http://doi.org/10.1016/j.cub.2003.12.054>
- Rogers, R., & Monsell, S. (1995). Cost of a predictable switch between simple cognitive tasks.
Journal of Experimental Psychology. General, 124(2), 207–231.
<http://doi.org/10.1037/0096-3445.124.2.207>
- Roswarski, T. E., & Proctor, R. W. (1996). Multiple spatial codes and temporal overlap in choice-reaction tasks. *Psychological Research*, 59(3), 196–211.
<http://doi.org/10.1007/BF00425834>
- Roswarski, & Proctor, R. W. (2000). Auditory stimulus-response compatibility: is there a contribution of stimulus-hand correspondence? *Psychological Research*, 63(2), 148–58.
<http://doi.org/10.1007/PL00008173>
- Salzer, Y., Aisenberg, D., Oron-Gilad, T., & Henik, A. (2014). In touch with the Simon effect.
Experimental Psychology, 61(3), 165–179. <http://doi.org/10.1027/1618-3169/a000236>
- Sheliga, B. M., Craighero, L., Riggio, L., & Rizzolatti, G. (1997). Effects of spatial attention on directional manual and ocular responses. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, 114(2), 339–351.
<http://doi.org/10.1007/PL00005642>
- Simon, J. R. (2011). “The Simon effect”: A potent behavioral mechanism. *Acta Psychologica*.
<http://doi.org/10.1016/j.actpsy.2010.04.007>
- Simon, J. R., Hinrichs, J. V, & Craft, J. L. (1970). Auditory S-R compatibility: reaction time as a function of ear-hand correspondence and ear-response-location correspondence. *Journal of Experimental Psychology*, 86(1), 97–102. <http://doi.org/10.1037/h0029783>
- Simon, J. R., & Small, A. M. (1969). Processing Auditory Information : Interference from an irrelevant cue. *Journal of Applied Psychology*, 53(5), 433–435.
<http://doi.org/10.1037/h0028034>

- Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: the Colavita effect revisited. *Perception & Psychophysics*, 69(5), 673–86.
- Spence, C., Nicholls, M. E., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, 63(2), 330–6.
<http://doi.org/10.3758/BF03194473>
- Stoffer, T. H. (1991). Attentional focussing and spatial stimulus-response compatibility. *Psychological Research*, 53(2), 127–135. <http://doi.org/10.1007/BF01371820>
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662. <http://doi.org/10.1037/h0054651>
- Umiltà, C., & Liotti, M. (1987). Egocentric and relative spatial codes in S-R compatibility. *Psychological Research*, 49, 81–90. <http://doi.org/10.1007/BF00308672>
- Van der Lubbe, R. H. J., & Abrahamse, E. L. (2011). The premotor theory of attention and the Simon effect. *Acta Psychologica*, 136(2), 259–264.
<http://doi.org/10.1016/j.actpsy.2010.09.007>
- Van der Lubbe, R. H. J., & Van der Helden, J. (2006). Failure of the extended contingent attentional capture account in multimodal settings. *Advances in Cognitive Psychology*, 2(4), 255–267. <http://doi.org/10.2478/v10053-008-0060-x>
- van Ede, F., de Lange, F., Jensen, O., & Maris, E. (2011). Orienting attention to an upcoming tactile event involves a spatially and temporally specific modulation of sensorimotor alpha- and beta-band oscillations. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 31(6), 2016–24. <http://doi.org/10.1523/JNEUROSCI.5630-10.2011>
- van Ede, F., Jensen, O., & Maris, E. (2010). Tactile expectation modulates pre-stimulus beta-band oscillations in human sensorimotor cortex. *NeuroImage*, 51(2), 867–76.
<http://doi.org/10.1016/j.neuroimage.2010.02.053>
- Virzi, R. A., & Egeth, H. E. (1985). Toward a translational model of Stroop interference. *Memory & Cognition*, 13(4), 304–19.
- Wallace, R. J. (1971). S-R compatibility and the idea of a response code. *Journal of*

Experimental Psychology, 88(3), 354–360. <http://doi.org/10.1037/h0030892>

Yamamoto, S., & Kitazawa, S. (2001). Reversal of subjective temporal order due to arm crossing. *Nature Neuroscience*, 4(7), 759–765. <http://doi.org/10.1038/89559>

Footnotes

¹ There are other forms of S-R compatibility which do not involve spatial information, such as the Stroop effect (Stroop, 1935) or the Flanker effect (Eriksen & Eriksen, 1974) (see also Kornblum, Hasbroucq, & Osman, 1990).

² In many studies, even when posture is manipulated, the location of the response goal (button) and position of the effector in space are confounded (unless tools are used, for example). Yet, Riggio et al. (1986) in their study disentangled these two sources of spatial information. In two experiments, participants responded to visual stimuli by adopting an uncrossed-arm posture but (un)crossing the index fingers (Experiment 1) or by operating with (un)crossed sticks (Experiment 2). The results were consistent in indicating that visual Simon effects depend on the matching of spatial information between the position of the stimulus and the position of the response goal (external frame of reference).

³ As mentioned in the discussion of Experiment 1, we have also replicated the results shown by Medina and collaborators (2014), under single modality conditions; see Supplemental materials.

Figures

A. Visual Simon

B. Tactile Simon

Figure 1. Results from Experiment 1. **A.** Single modality Visual Simon task. **B.** Single modality Tactile Simon task. The reaction times (in ms) for uncrossed (II) and crossed (X) arm postures are depicted leftmost in both inserts. Rightmost inserts represent accuracy (proportion of correct responses) and the congruency effect (Incompatible RTs – Compatible RTs, in ms) for both experiments. Error bars represent the Standard Error of the Mean (S.E.M.).

A. Visual Trials

B. Tactile Trials

Figure 2. Results from Experiment 2: Visuo-tactile mixed-modality Simon task. **A.** Visual trials; **B.** Tactile trials. The reaction times (in ms) for uncrossed (II) and crossed (X) arm postures are depicted leftmost in both inserts. Rightmost inserts represent accuracy (proportion of correct responses) and the congruency effect (Incompatible RTs – Compatible Rts, in ms) for both trial modalities. Error bars represent the Standard Error of the Mean (S.E.M.).

A. Visual Trials

B. Tactile Trials

Figure 3. Results of the N-1 Analysis from Experiment 2: Visuo-tactile mixed-modality Simon task. A. Visual trials; B. Tactile trials. The reaction times (in ms) for compatible (dark gray) and incompatible (light gray) conditions are plotted as a function of the modality of the previous trials (x-axis; maintain vs. switch) for the uncrossed (II) and crossed (X) arm posture. Error bars represent the Standard Error of the Mean (S.E.M.).

A. Visual Trials

B. Tactile Trials

Figure 4. Results from Experiment 3: Visuo-tactile mixed-modality Simon task. **A.** Visual trials; **B.** Tactile trials. The reaction times (in ms) for uncrossed (II) and crossed (X) arm postures are depicted leftmost in both inserts. Rightmost inserts represent accuracy (proportion of correct responses) and the congruency effect (Incompatible RTs – Compatible RTs, in ms) for both trial modalities. Error bars represent the Standard Error of the Mean (S.E.M.).

A. Visual Trials

B. Tactile Trials

Figure 5. Results of the N-1 Analysis from Experiment 3: Visuo-tactile mixed-modality Simon task. **A.** Visual trials; **B.** Tactile trials. The reaction times (in ms) for compatible (dark gray) and incompatible (light gray) conditions are plotted as a function of the modality of the previous trials (x-axis; maintain vs. switch) for the uncrossed (II) and crossed (X) arm posture. Error bars represent the Standard Error of the Mean (S.E.M.).

A. Visual Trials

B. Tactile Trials

Figure 6. Results from Experiment 4: Visuo-tactile mixed-modality Simon task with foot response. **A.** Visual trials; **B.** Tactile trials. The reaction times (in ms) for uncrossed arm posture are depicted leftmost in both inserts. Rightmost inserts represent accuracy (proportion of correct responses) and the congruency effect (Incompatible RTs – Compatible RTs, in ms) for both trial modalities. Error bars represent the Standard Error of the Mean (S.E.M.).

A. Visual Trials

B. Auditory Trials

Figure 7. Results from Experiment 5: Audio-Visual mixed-modality Simon task. **A.** Visual trials; **B.** Auditory trials. The reaction times (in ms) for uncrossed (II) and crossed (X) arm postures are depicted leftmost in both inserts. Rightmost inserts represent accuracy (proportion of correct responses) and the congruency effect (Incompatible RTs – Compatible RTs, in ms) for both trial modalities. Error bars represent the Standard Error of the Mean (S.E.M.).

Figure 8. Summary of the Simon effect from 5 Experiments. Experiments 1 are single modality Simon tasks (visual and tactile, respectively). Experiments 2 and 3 are Visuo-tactile mixed-modality Simon task. Experiment 4 is Visuo-tactile mixed-modality Simon task with foot response. Experiment 5 is Audio-Visual mixed-modality Simon task. The reaction times (in ms) for uncrossed (II - top graph) and crossed (X – bottom graph) arm postures are plotted for compatible (dark gray) and incompatible (light gray) conditions. For Experiment 4 we only ran the uncrossed arm posture condition. Shadow areas highlight the visual Simon effect across experiments. Error bars represent the Standard Error of the Mean (S.E.M.).

Tables

Table 1. Statistics from Experiment 2 N-1 analysis. Repeated measures ANOVA with *modality* (Vision/Touch), *arm posture* (Uncrossed/Crossed), *modality transition* (Maintain/Switch) and *S-R compatibility* (Compatible/Incompatible) as within participant variables. [Abbreviations: MOD.=modality; ARM POST.=arm posture; MOD. TRANS.= modality transition]

	SS	D. of F.	MS	F	p	Partial eta-squared
Intercept	92,63	1,15	92,629	484,94	0,000	0,97
MODALITY	1,26	1,15	1,259	102,72	0,000	0,87
ARM POSTURE	0,15	1,15	0,152	3,65	0,076	0,20
MODALITY TRANSITION	0,72	1,15	0,723	47,27	0,000	0,76
S-R	0,04	1,15	0,040	3,85	0,069	0,20
MODALITY*ARM POSTURE	0,01	1,15	0,007	1,73	0,209	0,10
MODALITY*MODALITY TRANSITION	0,03	1,15	0,029	6,45	0,023	0,30
ARM POSTURE*MODALITY TRANSITION	0,01	1,15	0,011	2,14	0,164	0,12
MODALITY*S-R	0,05	1,15	0,053	18,01	0,001	0,55
ARM POSTURE*S-R	0,00	1,15	0,000	0,01	0,922	0,00
MODALITY TRANSITION*S-R	0,00	1,15	0,000	0,00	0,983	0,00
MODALITY*ARM POSTURE*MODALITY TRANSITION	0,00	1,15	0,001	0,99	0,335	0,06
MODALITY*ARM POSTURE*S-R	0,01	1,15	0,010	8,20	0,012	0,35
MODALITY*MODALITY TRANSITION*S-R	0,00	1,15	0,004	2,21	0,158	0,13
ARM POSTURE*MODALITY TRANSITION*S-R	0,00	1,15	0,004	2,69	0,122	0,15
MOD. * ARM POST. * MOD. TRANS. * S-R	0,00	1,15	0,001	0,34	0,566	0,02

Table 2. Statistics from Experiment 3 N-1 analysis. Repeated measures ANOVA with *modality* (Vision/Touch), *arm posture* (Uncrossed/Crossed), *modality transition* (Maintain/Switch) and *S-R compatibility* (Compatible/Incompatible) as within participant variables. [Abbreviations: MOD.=modality; ARM POST.=arm posture; MOD. TRANS.= modality transition]

	SS	D. of F.	MS	F	p	Partial eta-squared
Intercept	75,42	1,15	75,421	1069,37	0,000	0,99
MODALITY	0,59	1,15	0,586	51,79	0,000	0,78
ARM POSTURE	0,29	1,15	0,289	15,76	0,001	0,51
MODALITY TRANSITION	0,42	1,15	0,415	46,66	0,000	0,76
S-R	0,02	1,15	0,023	9,37	0,008	0,38
MODALITY*ARM POSTURE	0,00	1,15	0,001	0,53	0,476	0,03
MODALITY*MODALITY TRANSITION	0,04	1,15	0,041	15,23	0,001	0,50
ARM POSTURE*MODALITY TRANSITION	0,00	1,15	0,003	1,28	0,276	0,08
MODALITY*S-R	0,04	1,15	0,044	31,96	0,000	0,68
ARM POSTURE*S-R	0,00	1,15	0,000	0,01	0,910	0,00
MODALITY TRANSITION*S-R	0,00	1,15	0,000	0,01	0,942	0,00
MODALITY*ARM POSTURE*MODALITY TRANSITION	0,00	1,15	0,001	0,99	0,336	0,06
MODALITY*ARM POSTURE*S-R	0,00	1,15	0,001	1,10	0,311	0,07
MODALITY*MODALITY TRANSITION*S-R	0,00	1,15	0,001	3,90	0,067	0,21
ARM POSTURE*MODALITY TRANSITION*S-R	0,00	1,15	0,003	2,24	0,155	0,13
MOD. * ARM POST. * MOD. TRANS. * S-R	0,00	1,15	0,000	0,05	0,821	0,00

Table 3. Statistics from Experiment 5 N-1 analysis. Repeated measures ANOVA with *modality* (Vision/Audition), *arm posture* (Uncrossed/Crossed), *modality transition* (Maintain/Switch) and *S-R compatibility* (Compatible/Incompatible) as within participant variables. [Abbreviations: MOD.=modality; ARM POST.=arm posture; MOD. TRANS.= modality transition]