Consciousness and Cognition 80 (2020) 102900

Contents lists available at ScienceDirect

Consciousness and Cognition

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

Seeing an image of the hand affects performance on a crossmodal congruency task for sequences of events

Alan O' Dowd[a,*], Francesca Sorgini[a,b], Fiona N. Newell[a]

[a] School of Psychology and Institute of Neuroscience, Trinity College Dublin, Ireland
[b] The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy

ARTICLE INFO

Keywords:
Crossmodal congruency task
Hand image
Multisensory
Touch
Vision
Visual hand identity

ABSTRACT

The crossmodal congruency effect (CCE) is augmented when viewing an image of a hand compared to an object. It is unclear if this contextual effect extends to a non-spatial CCE. Here, participants discriminated the number of tactile vibrations delivered to the hand whilst ignoring visual distractors on images of their own or another's hand or an object. The CCE was not modulated by stimulus context. Viewing one's hand from a third person perspective increased errors relative to viewing an object (Experiment 1). Errors were reduced when viewing hands, from first or third person perspectives, with additional identity markers (Experiments 2 and 3). Our results suggest no effect of context on the non-spatial CCE and that differences in task performance between hand and object images depend on their visual properties. These findings are discussed in light of the relationship between body representation and perception of body-centred stimuli in the temporal domain.

1. Introduction

Our perception of, and interaction with, the external world is strongly influenced by the experience of recognising one's own body as distinct from another's. Over the past few years, a number of studies have demonstrated the multisensory nature of the body representation, particularly in the ability to localise stimulation on the body. This has been highlighted robustly with the crossmodal congruency task (CCT) (Spence, Pavani, & Driver, 2004; Spence, Pavani, Maravita, & Holmes, 2008). The CCT typically involves the presentation of a tactile stimulus to different locations on the body, most commonly the index finger and thumb, and visual distractors in external space which are also arranged in corresponding spatial locations on or near the physical body. The task for the participant is to indicate the elevation of the tactile stimulation whilst ignoring the distractors. Across trials, these visuotactile stimuli are presented as spatially congruent or incongruent and the difference in performance, measured in terms of speed or accuracy, between these trial types is known as the crossmodal congruency effect (CCE).

Several studies have provided evidence that the CCE occurs when visual distractors are superimposed on images of the body and have reported that viewing a body part, or indeed a full body, significantly increases the magnitude of the CCE relative to viewing an object (Aspell, Lenggenhager, & Blanke, 2009; Igarashi, Kimura, Spence, & Ichihara, 2008; Igarashi, Kitagawa, & Ichihara, 2004; Igarashi, Kitagawa, Spence, & Ichihara, 2007; Salomon, Van Elk, Aspell, & Blanke, 2012; Thomas, Press, & Haggard, 2006). The occurrence of a large CCE when viewing an image of a body part may reflect the role of 'context' in multisensory integration in which cross-sensory events arising in a shared context are more likely to be integrated than those arising in separate contexts (Chen & Spence, 2017; Welch & Warren, 1980). The context can be influenced by prior learning from everyday experiences; for example, a

* Corresponding author.
E-mail address: odowda1@tcd.ie (A. O' Dowd).

https://doi.org/10.1016/j.concog.2020.102900
Received 8 November 2019; Received in revised form 25 January 2020; Accepted 12 February 2020
Available online 28 February 2020

visual event appearing briefly on the body is typically accompanied by a tactile sensation at the same location (Johnson, Burton, & Ro, 2006; Mirams, Poliakoff, Brown, & Lloyd, 2010). Even when viewing a body part is non-informative to the tactile task, the visual percept can nevertheless influence tactile perception by increasing tactile sensitivity (Kennett, Taylor-Clarke, & Haggard, 2001; Taylor-Clarke, Kennett, & Haggard, 2002; Tipper et al., 1998).

Despite clear evidence for a role of viewing the body on tactile perception, supporting evidence that this crossmodal interaction is further influenced by the perception of body identity is equivocal. The processing of stimuli associated with the self can be prioritized compared to stimuli associated with another (e.g. Frassinetti, Ferri, Maini, Benassi, & Gallese, 2011; Schäfer, Wesslein, Spence, Wentura, & Frings, 2016; Sui & Rotshtein, 2019). There is evidence that the spatial CCE is enhanced when viewing an image of one's own body part compared to body parts not attributed to the self, through self-recognition or perceived ownership (Aspell et al., 2009; Maravita, Spence, Sergent, & Driver, 2002; Salomon et al., 2012; Zopf, Savage, & Williams, 2010). However, viewing another's body or body part can also produce an enhancement of visual, tactile and crossmodal spatial discrimination compared to viewing an object (Haggard, 2006; Salomon et al., 2012; Thomas et al., 2006; Whiteley, Spence, & Haggard, 2008). As such, the extent to which crossmodal interactions are uniquely modulated by the compatibility between a viewed body part and oneself is unclear.

The processing of multisensory information relating to the body is thought to occur within fronto-parietal regions of the brain (Blanke, 2012; Tsakiris, Hesse, Boy, Haggard, & Fink, 2007). Neurophysiological work with non-human primates established the presence of bimodal neurons within this network that represent a somatosensory map of the body surface as well as the visual space on, and within reachable distance from, the body (termed peripersonal space, PPS) (Rizzolatti, Fadiga, Fogassi, & Gallese, 1997). Evidence from patient and neuroimaging studies supports the existence of a comparable visuotactile PPS representation in the human brain (Làdavas, Pellegrino, Farnè, & Zeloni, 1998; Makin, Holmes, & Zohary, 2007). Some suggest that the visual representation of one's own body can change visuotactile interactions via the projection of PPS onto the image of the self-identified body part (Maravita et al., 2002; Salomon et al., 2012). Indeed, the multisensory, body-centred representation of PPS is thought to support one's sense of bodily self-consciousness (e.g. body ownership; Serino, 2019). Such a projection of PPS could contribute to the reported contextual influence of viewing a body identified as one's own on visuotactile interactions (Aspell et al., 2009; Maravita et al., 2002; Salomon et al., 2012).

Most studies based on the CCT have investigated the role of spatial proximity between visual and tactile stimuli on the spatial perception of touch. However, in everyday life, multisensory events are characterised not only by where they arise in space but also when or how often these events occur across each modality. Indeed, both spatial and temporal coincidence can be important precursors to multisensory integration in the brain (Holmes & Spence, 2005; Shore, Barnes, & Spence, 2006; Stein & Meredith, 1990; Wallace et al., 2004) often in a task-dependent manner (see McGovern, Astle, Clavin, & Newell, 2016; Spence, 2019). In the context of body representation, evidence for an effect of viewing the body on the integration of visual and tactile inputs in the temporal domain is mixed. For example, Ide and Hidaka (2013) reported that viewing a hand arranged in a plausible orientation with respect to the body widens the temporal binding window for vision and touch compared to viewing an incompatible hand or an image of an arrow. Maselli, Kilteni, López-Moliner, and Slater (2016) also reported a dilated temporal binding window for tactile stimuli presented on a participant's hand and visual stimuli presented close to a virtual hand, the size of which correlated with the strength of virtual hand ownership. Other studies involving visual and tactile asynchrony detection (Keys, Rich, & Zopf, 2018) and temporal order judgements (Smit, Rich, & Zopf, 2019) have reported no effect of the body context on the visuotactile temporal binding window.

The following experiments were designed to assess whether visuotactile interactions for discriminating sequences of events are modulated when viewing a hand or an object. There were several advantages to using the CCT, not least as it allowed us to maintain important aspects of the paradigm to measure the effect of task-irrelevant visual information on tactile perception of events on the body. We also investigated a role for visual body identity in crossmodal temporal discrimination by assessing differences in the extent of the CCE when viewing images of one's own over an unfamiliar other's hand. While there is evidence that the CCE is larger when viewing images of one's own body than another's (Salomon et al., 2012) and when viewing an image of a hand compared to an object (Igarashi et al., 2007, 2008) it is not yet well established if visual self-recognition specifically influences crossmodal interactions. Recognition of one's own hand in the absence of motor cues can be difficult (Van Den Bos & Jeannerod, 2002; Tsakiris, Haggard, Franck, Mainy, & Sirigu, 2005). However, visuotactile interactions can be influenced by the perceived compatibility of a hand with one's own body (Igarashi et al., 2004; Ide & Hidaka, 2013; Maselli et al., 2016; Pavani, Spence, & Driver, 2000; Zopf et al., 2010) and viewing a mirrored image of one's own hand compared to a rubber hand enhances the CCE (Maravita et al., 2002). Moreover, a self-advantage effect has been found when viewing one's own hand compared to that of another in an implicit task context (Frassinetti et al., 2011). In the following study we conducted three experiments using a non-spatial CCT (Forsberg, O'Dowd, & Gherri, 2019; Holmes, Sanabria, Calvert, & Spence, 2006, 2007; Poole, Couth, Gowen, Warren, & Poliakoff, 2015). Similar to the spatial CCT (e.g. Spence et al., 2004, 2008), the non-spatial CCT can be sensitive to the temporal synchronicity and spatial compatibility of visual and tactile signals, with larger CCEs for synchronous (or near-synchronous) and/or co-located signals (Holmes, Sanabria, Calvert, & Spence, 2006; Poole et al., 2015). In our experiments, participants were instructed to discriminate tactile sequences of events whilst ignoring visual distractors that were superimposed on an image of either the participant's own hand, an unfamiliar 'other' hand or on an object. We manipulated the specific viewing conditions of the images of the hands and object across experiments in order to determine whether the perception of tactile sequences of events is affected by particular contexts of viewing the body.

2. Experiment 1

To examine whether tactile perception for sequences of events is modulated by visual stimuli as well as context of the body, we presented distractors on an image of one's own hand, another's hand and an object. All images of these context stimuli (hand and object) were presented from an allocentric view (i.e. third person, or 'mirror-image' view in which the hand is pointed towards the participant), consistent with several previous spatial CCT studies which presented body parts from this perspective (Aspell et al., 2009; Igarashi et al., 2004; Salomon et al., 2012; Thomas et al., 2006). Our hypotheses were as follows:

1. A significant non-spatial CCE would arise in response to the visual distractors (e.g. see Holmes et al., 2006, 2007; Poole et al., 2015). That is, that performance would be worse on incongruent than congruent trials.
2. A larger CCE would arise when viewing an image of a hand compared to an object, provided that context influences non-spatial visuotactile interactions in a manner compatible with the spatial CCE (e.g. Igarashi et al., 2007, 2008).
3. A larger CCE would arise when viewing identified images of one's own hand, if self-identification further influences temporal visuotactile interactions (e.g. Aspell et al., 2009; Maravita et al., 2002; Salomon et al., 2012).

2.1. Materials and methods

2.1.1. Participants

Twenty-eight participants (3 males) with a mean age of 19.3 years (and an age range of 18–21 years) were initially recruited for Experiment 1. All were students of the School of Psychology, Trinity College Dublin and took part in exchange for course credits or on a voluntary basis. Participants were required to be right-handed, have normal or corrected-to-normal vision, no tactile impairments and be neurologically healthy to be included in the study. The recruitment procedure excluded anyone with distinctive markings on their right hand, such as obvious scars, birth marks or tattoos. All experiments were approved by the School of Psychology Research Ethics Board (Trinity College Dublin) and conducted in accordance with the Declaration of Helsinki. Data were obtained in accordance with EU General Data Protection Regulation (GDPR). Written consent was obtained prior to each experiment.

2.1.2. Stimuli and apparatus

An image of each participant's right hand (with their index finger outstretched, remaining fingers and thumb tucked underneath the palm and palm facing downwards) was taken from an allocentric view, at a constant distance (15 cm) and under constant luminance conditions across participants. The images of the hand therefore appeared as mirror-images of each participant's physical right hand during the experiment. Prior to taking the image, participants were instructed to remove any hand jewellery and to have trim nails without nail polish. A random selection of 5 images of hands from different participants was morphed with an image of the experimenter's hand to create one image representing an 'other' hand and to ensure that the 'other' hand was unfamiliar and not identifiable. All participants were shown the same image of the morphed 'other' hand. The object stimulus comprised an image of an unfamiliar, three-dimensional object constructed from LEGO that was created to closely match the shape, size and colour of a hand. Thus we made every effort to reduce potential confounds based on image differences in scale, visual complexity or saliency between images of the hand and object. All images were edited to remove irrelevant background features and superimposed (using GIMP 2.8 image manipulation software) onto plain black backgrounds. The scale of each image was adjusted to an average size which was determined based on size measurements taken from four original images of hands (2 female, 2 male).

Prior to the experiment, each participant conducted a hand recognition pre-study in order to ensure that our self- and other-hand conditions were valid. To that end, an image of the participant's own hand was presented in a printed 2 × 4 array together with a random selection of images of 7 other hands. All images of hands were scaled to approximately match in size. The position of the image of the participants' own hand and the distractor hand images in the array was randomised across participants.

The main experiment (i.e. the CCT) included visual and tactile stimuli. An illustration of the experimental apparatus is shown in Fig. 1a. The experiment was programmed on Psychopy (version 1.84.2; Peirce et al., 2019) and run on a desktop computer (Alienware X51 R3, 64-bit, Intel Core i5-6400 @ 2.70 GHz, Nvidia GeForce GTX 970 graphics). The screen (59.8 cm by 33.6 cm, 2560 by 1440 resolution, 60 Hz refresh rate) was positioned approximately 15 cm away from the tip of the participant's hand and at a distance of 45 cm from the centre of the participant's body. This ensured that the screen was well within the boundaries of PPS for the hand and as close as possible to the boundaries of the body-PPS (given the constraints associated with the positioning of the apparatus).

Fig. 1b provides an example of visual displays used in the experiment. The distractor stimulus was positioned in the middle of the index finger thus ensuring that most of the information of the hand or object image was always visible. The display, including the distractor, therefore comprised of an image of a hand (self or other) or object, which was positioned off-centre, 1.3 degrees of visual angle to the right of a central fixation point. This spatial position of the distractor ensured that it was presented in central vision as well as in the same hemifield (i.e. the right visual field) as the tactile events (on the right hand), such that attention was directed to a single hemifield. The size of the hand or object image subtended a visual angle of approximately 8 by 8 degrees in the horizontal and vertical dimensions respectively.

The tactile stimulus comprised of vibrations which were delivered to the fingertip via the single black coin vibration motor (10 mm × 3 mm; 200 Hz; 500 mV). The motor was mounted on to a plastic pinhead that was secured to a black support column positioned 15 cm above the table surface. The motor was connected to a custom-built electrical circuit which was located to the right of the apparatus and connected to the computer via USB. The circuit was activated using an Arduino R3 Uno controller board and an Arduino script (version 1.8.9). The circuit and the computer were occluded from view with an opaque black cloth. All participants wore sound-attenuating headphones and white noise was played at 50 dB for the duration of the experiment to mask the sound of the vibrations of the coin motor.

A single visual event involved the distractor stimulus presented for a duration of 200 ms. When two visual events were presented

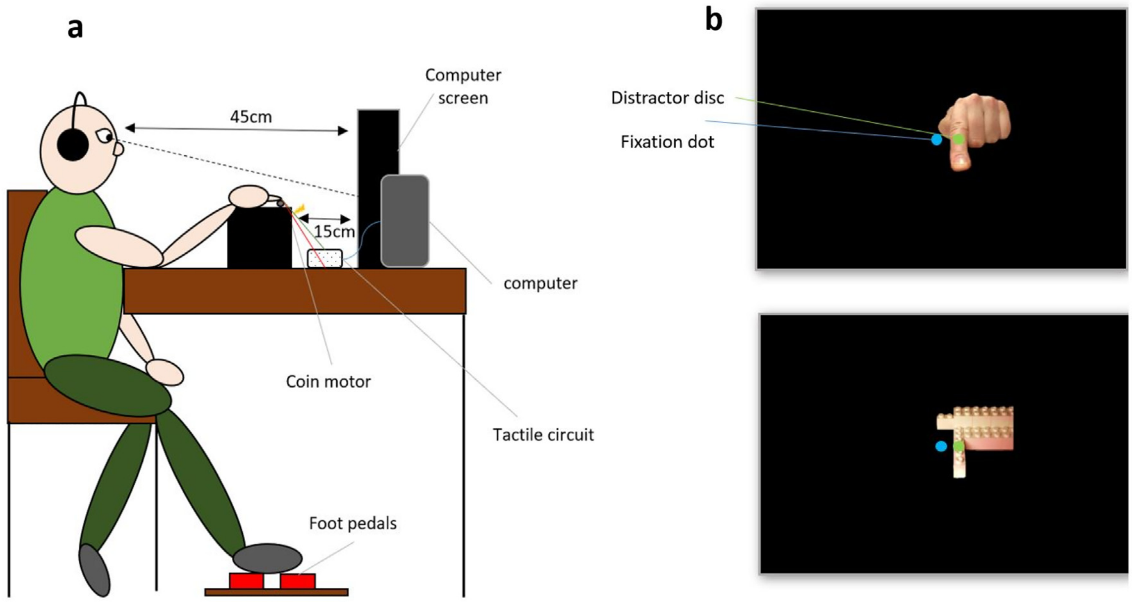

Fig. 1. An illustration of the experimental set-up of Experiment 1. (a) This schematic shows how a typical participant positioned their right hand, which was always visible, over a coin vibration motor which delivered tactile vibrations to the index finger. (b) An illustration of a stimulus scene in which a distractor (green disc) was superimposed on an image of a hand (upper image) or object (lower image). In this experiment, context stimuli (hand or object) were presented from an allocentric view. A central fixation dot (blue) was present on the screen throughout the experiment.

these consisted of two presentations of the distractor for 100 ms each separated by an interval of 50 ms. A single tactile event consisted of a vibration on the fingertip for a duration of 200 ms and double tactile event consisted of two vibrations for 100 ms each separated by a period of 50 ms. The timing of the visual and tactile events was verified with a photodiode and oscilloscope respectively.

Participants responded to each trial using one of two pedals (9 cm by 6.5 cm), each positioned directly under the toes and the heel of the right foot respectively, and connected to the computer via a modified serial response box. The foot pedals were mounted on a wooden support and their position was directly aligned with the location of the right hand on the table. Individual non-slip mats were used to prevent movement of the foot pedals and wooden support. Each toe or heel press indicated the participant's response to the tactile events as 'single' or 'double' respectively.

During the experiment, trials were blocked according to each of the context stimuli (self-, other-hand, object) and each block was repeated twice, yielding a total of six blocks. The order of these blocks was counterbalanced across participants. Each block consisted of 48 trials (24 congruent and 24 incongruent visuotactile presentations) and trial order was random across participants. This design yielded a total of 96 trials per stimulus context and 288 trials in total.

2.1.3. Procedure

All experiments took place in a dimly-lit, sound-attenuated testing space in the basement of the Institute of Neuroscience, Trinity College Dublin. Each participant first performed a hand identification task in which they had to identify an image of their own hand from an array of images including the target (own hand) and 7 images of hands from other individuals. All hand images in the array were presented from an allocentric view. There was no time limit imposed on this task. Participants indicated their response verbally which was manually recorded by the experimenter onto a computer. Participants were provided with feedback on their performance on this task.

Following the hand identity task, the participant was then taken to the testing space and placed in a seated position at a table for the main experiment. They were instructed to rest their right elbow on the table and to keep their left hand by their side so that it was occluded from view for the duration of the experiment. The participant then placed the tip of their right index finger onto the coin vibration motor and were instructed to maintain that position and posture during the experimental blocks. Prior to the start and throughout the experiment, the experimenter ensured that the participant's index finger and hand was aligned with the image of the hand displayed on the computer screen and adjusted the position of the apparatus accordingly. The participant's right hand was visible throughout the experiment to enhance the effect of viewing a mirror image on the screen.

A brief training period was conducted before the main experiment to familiarize the participants with the task and the order and timing of the visual and tactile events. As shown in Fig. 2 below, each trial began with an image of a hand or object stimulus along with a central fixation dot, presented for a duration of 250 ms. The distractors were then presented on the image of the hand or object along with tactile stimulation on the fingertip. Depending on the nature of the trial, the distractor event comprised of a single or a double presentation of a green disc whilst the tactile stimulus comprised of a single or double tactile vibration on the fingertip.

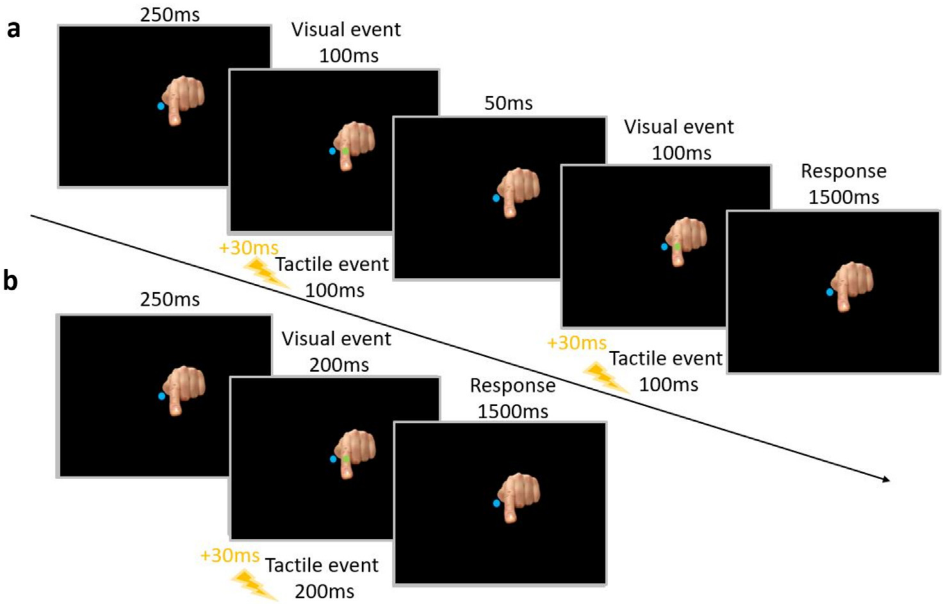

Fig. 2. An illustration of the sequence of stimuli presented during a typical (a) single and (b) double visuotactile congruent trial in the CCT of Experiment 1. In this experiment, context stimuli (hand or object) were presented from an allocentric view. See main text for further details.

Consistent with previous CCT studies, the onset of the distractor stimulus preceded that of the tactile vibration by a short interval of 30 ms. This timing is considered within the temporal binding window of visuo-tactile integration (Meredith, Nemitz, & Stein, 1987; Spence & Squire, 2003) and is known to maximise the size of both the spatial (Spence et al., 2004, 2008) and non-spatial (Poole et al., 2015) CCE. The relative delay in the onset of stimulation across modalities may also help compensate for the differences in transduction latency across vision and touch (Spence, Shore, & Klein, 2001).

Both the fixation dot and the image of the hand or object remained on the screen for the duration of each trial in the experiment (see Fig. 2). The eye movements of the participants were covertly monitored by the experimenter to ensure that the participant was looking at the screen and maintaining fixation throughout the experiment. Participants were instructed to fixate on the centre of the visual scene throughout the experiment and they were also explicitly instructed to ignore the distractors and to perform the task based on tactile information only. Specifically, their task was to indicate if they felt a single or double vibration on their index finger for every trial via the foot pedals and to respond as quickly and as accurately as possible. Participants were given 1500 ms to respond and either a response or the end of the trial triggered the onset of the subsequent trial. The experiment took approximately 40 min for each participant to complete in one experimental session.

2.1.4. Design

The experiment was based on a 2 (congruency) × 3 (stimulus context) within-subjects design.

2.1.5. Data analysis

Trials with missing data were removed (i.e. where participants either failed to make a response, made a response outside of the response time allocated, or did not press the pedal with enough force to register a response), as were statistical outliers defined as responses which were 2.5 standard deviations beyond the population mean. The mean correct RTs and percentage errors were calculated per participant for each of the temporal sequence and stimulus context conditions. All analyses were conducted using R (Team R, 2017) via Rstudio (version 3.5.0) (Team R, 2015). For the ANOVAs, the ez package (Lawrence, 2016) was used with type 3 sum of squares to test for main effects and interactions. All assumptions of ANOVA were satisfied. Significance values are reported with Greenhouse-Geisser corrections in cases where sphericity was violated and we used Bonferroni correction in cases of multiple, post-hoc comparisons.

The data were analysed with a within-subjects 2 (distractor number) × 2 (congruency) 3 (stimulus context) ANOVA. We included distractor number as an independent variable to better explore the effect of different visuotactile combinations on task performance. For example, on incongruent trials, a response of ‘2’ may occur in response to 2 distractors and 1 target and a response of ‘1’ may occur in response to 1 distractor and 2 targets. Following on from audiovisual research (Andersen, Tiippana, & Sams, 2004; Shams, Kamitani, & Shimojo, 2000, 2002), we have labelled these erroneous responses as ‘fission-like’ and ‘fusion-like’ responses respectively (note, we recognise that these effects are not identical to those reported in multisensory illusions which typically involve shorter stimulus durations). The dependent variables were speed and accuracy of responses to the tactile events.

The ANOVA was conducted on the percentage of errors only. Response time data per congruency condition (congruent, incongruent) and stimulus context (self-, other-hand, or object) are reported in all main tables and analyses of RT data can be found in

Table 1

The means and standard deviations (in parentheses) of errors (%), RTs (ms) and IE scores (ms) across congruency conditions (Cong = Congruent, Incong = Incongruent) for each stimulus context (Self-hand, Other-hand, Object) presented from an allocentric view in Experiment 1.

	Overall		Self		Other		Object	
	Cong	Incong	Cong	Incong	Cong	Incong	Cong	Incong
Error (%)	8.9 (9.9)	29.3 (16.9)	12.1 (14.4)	33.9 (17.3)	7.4 (6.4)	27.6 (15.7)	7.1 (6.4)	26.3 (17.4)
			22.99 (26.63)		18.18 (23.57)		16.97 (24.92)	
RT (ms)	819 (80)	836 (83)	809 (81)	828 (84)	815 (74)	830 (82)	832 (88)	849 (88)
IE (ms)	910 (1 4 4)	1264 (3 8 4)	949 (2 0 5)	1345 (4 2 1)	884 (1 0 4)	1210 (3 3 9)	898 (98)	1236 (7 9 1)

Supplementary materials. Inverse efficiency scores (IE; the mean correct RTs divided by the proportion of correct responses; Townsend & Ashby, 2014) are reported in Tables 1–3 across experiments. There was no evidence of a speed-accuracy trade-off in the data and given the recommendations that IE is not analysed if accuracy is below 90% in an experimental condition (Bruyer & Brysbaert, 2011), these scores were not subjected to further analysis.

Bayesian within-subjects ANOVAs and paired t-tests were conducted to quantify Bayes Factors (BF_{01}) for non-significant interactions using the BayesFactor package, with default prior settings (Morey, Rouder, & Jamil, 2018; see Rouder, Speckman, Sun, Morey, & Iverson, 2009, and Rouder, Morey, Speckman, & Province, 2012, for information on models, priors, and methods of computation for Bayesian ANOVAs and t-tests using the BayesFactor package). This allowed us to assess the strength of support for the null hypotheses (Dienes, 2014). A cut-off of $BF_{01} > 3$ was taken as indication of robust support for the null hypothesis (Jeffreys, 1961). Finally, exploratory signal detection analyses (McGovern, Roudaia, Stapleton, McGinnity, & Newell, 2014; Rosenthal, Shimojo, & Shams, 2009) were conducted to examine whether the influence of the different numbers of distractors (1, 2) on the discrimination of tactile sequences of events may be due to perceptual or response-based factors. These analyses were conducted on the data for each experiment and are presented in Supplementary materials.

2.2. Results

We first analysed the participants' responses to the hand recognition test. Eight participants (28.6%) failed to recognize an image of their own hand in the array. Given that self-hand recognition was a necessary requirement for this study to ensure a distinction between self- and other-hands, the data from these participants were not included in subsequent analyses. Data from a total sample of 20 participants were deemed eligible for further analysis. Of these participants, data from one participant (female, 20 years) were removed due to consistently low levels of accuracy (< 50%) throughout the experiment, suggesting a failure to follow the instructions of the task. Responses to trials with missing values and statistical outliers (RT) were removed which resulted in a further removal of 4% and 3.6% of overall trials respectively. Table 1 shows the descriptive statistics for percentage error, reaction times and inverse efficiency scores for Experiment 1.

The mean error rate across all 19 participants was 19.08% ($SD = 17.16$). The within-subjects 2 (distractor number) × 2 (congruency) × 3 (stimulus context) ANOVA on mean error rates revealed a main effect of distractor number ($F(1,18) = 27.82$, $p < 0.001$, $\eta^2_G = 0.14$) and congruency ($F(1,18) = 42.80$, $p < 0.001$, $\eta^2_G = 0.24$). There was also a main effect of stimulus context ($F(2,36) = 5.44$, $p = 0.009$, $\eta^2_G = 0.02$) (Fig. 3a). The two-way interaction between congruency and stimulus context failed to reach significance ($p = 0.81$) and the Bayes Factor indicated support for the null hypothesis ($BF_{01} = 12.5$).

Post-hoc pairwise-comparisons (Bonferroni corrected) confirmed more errors were made in response to the self-hand ($M = 23.01\%$, $SD = 19.18$) than the object ($M = 16.72\%$, $SD = 16.20$) ($p = 0.04$). No other pairwise comparisons reached significance ($p > 0.05$). The Bayes Factors for the comparisons of other-hand and object and self-hand and other-hand were inconclusive ($BF_{01} = 2.5$; $BF_{01} = 0.66$). The two-way interaction between distractor number and congruency was significant ($F(1,18) = 18.22$, $p < 0.001$, $\eta^2_G = 0.24$), as shown in Fig. 3b. Post-hoc pairwise comparisons confirmed a significant congruency effect only in response to 2 distractors ($M\ diff = 41.60\%$, $SD = 31.36$). All other interactions were non-significant (all $ps > 0.05$).

2.3. Discussion

The results confirmed a large CCE on the error responses across all stimuli, with a significant increase in errors to judgements of tactile events in the incongruent than congruent trials. Further analysis of the performance across the number of visual distractors shown suggested that tactile discrimination was more difficult when 2 distractors rather than 1 distractor was presented on incongruent trials. This finding is consistent with previous reports (Holmes et al., 2006) and may be due to a combination of perceptual sensitivity as well as a response bias (see Supplementary materials).

Although the effect of congruency was present for RTs also (see Supplementary materials), it was weaker than that found for percentage error. The results suggest more errors were made in the context of viewing one's own hand relative to an object. Moreover, there was no difference in the size of the CCE on errors made (nor on RTs) across stimulus context, with strong support for the null hypothesis (Dienes, 2014). As such, there was no strong evidence that the visuotactile CCE for sequences of events was modulated by stimulus corporeality or body identity, unlike previous reports of the CCE involving spatially congruent or incongruent visuo-tactile information (e.g. Aspell et al., 2009; Igarashi et al., 2004, 2007; Igarashi et al., 2008; Maravita et al., 2002; Salomon

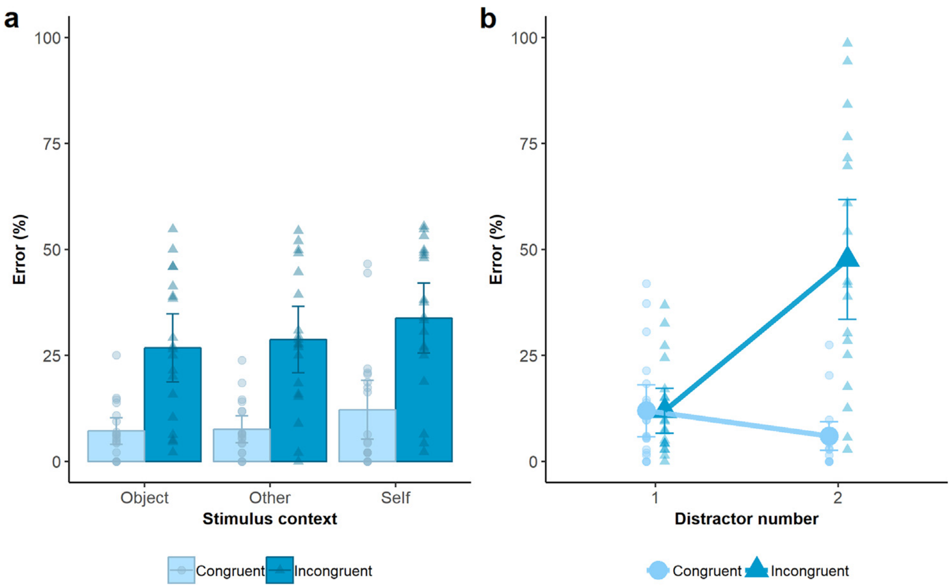

Fig. 3. The results (mean % error) of Experiment 1. (a) Plot shows % errors made to each of the stimulus contexts presented from an allocentric view as a consequence of visuotactile congruency. (b) Plot shows the interaction between distractor number and congruency. Individual mean error rates are shown. Error bars represent 95% confidence intervals.

et al., 2012; Thomas et al., 2006). Given this discrepancy with previous findings, and that our results suggested a greater error rate for the self-hand than object stimulus context, we conducted additional experiments to further explore the role of visual body identity and context on visuotactile interactions for sequences of events.

3. Experiment 2

To examine visuotactile integration of sequences of events across the contexts of viewing a hand or object further, in the following experiment we presented images of the hand from a more naturalistic view, that is, aligned with the position of the participant's own hand. As such, all context stimuli, including the object, were presented from an egocentric (i.e. first person) view in which one's own body is typically experienced, with the hand extending away from the body. This view is also consistent with that usually presented in the spatial CCT (Spence et al., 2004, 2008) in which both visual and tactile body information is typically aligned. To ensure the images of the hand were aligned with that of the participant's real hand, we also used an apparatus that allowed us to present the images on or as near as possible to the actual spatial location of the participant's real hand. We predicted that visual body identity may affect the CCE such that a difference between 'self-' and 'other-hand' might emerge by presenting the image of one's own hand from a perspective that is compatible with that usually experienced. Indeed, there is evidence that the visual representation of one's own dominant hand is distinct from the representation of another's hand, particularly when shown from an egocentric view, facilitating self-other distinction (Brady, Maguinness, & Choisdealbha, 2011; Conson, Aromino, & Trojano, 2010; Saxe, Jamal, & Powell, 2005).

3.1. Methods

3.1.1. Participants

Nineteen participants (6 males) with a mean age of 19.7 years (and a range of 18–29 years) were recruited for Experiment 2. All were students of the School of Psychology, Trinity College Dublin, and some took part in the experiment in exchange for course credits. All participants were naïve to the task and none took part in Experiment 1. All participants reported to be right-handed, had normal or corrected-to-normal vision, no tactile impairments and neurologically healthy. As in Experiment 1, anyone with distinctive markings on their right hand, such as obvious scars, birth marks or tattoos, were excluded from taking part.

3.1.2. Design and procedure

The experimental set-up and procedure were almost identical to that of Experiment 1 with the exception of the viewpoint of the hand and object images shown. An image of each participant's right hand was taken under constant luminance conditions with the hand in the same pose as that of Experiment 1. The camera was now positioned a fixed distance above the hand so as to achieve an egocentric view. To maintain this view during the CCT, a computer screen was placed flat on a table using a raised wooden column. The participant completed the experiment from a standing position in which they looked down at the image displayed on the screen (see Fig. 4a). The participant's physical hand rested on a flat surface positioned 20 cm directly under the screen and was occluded

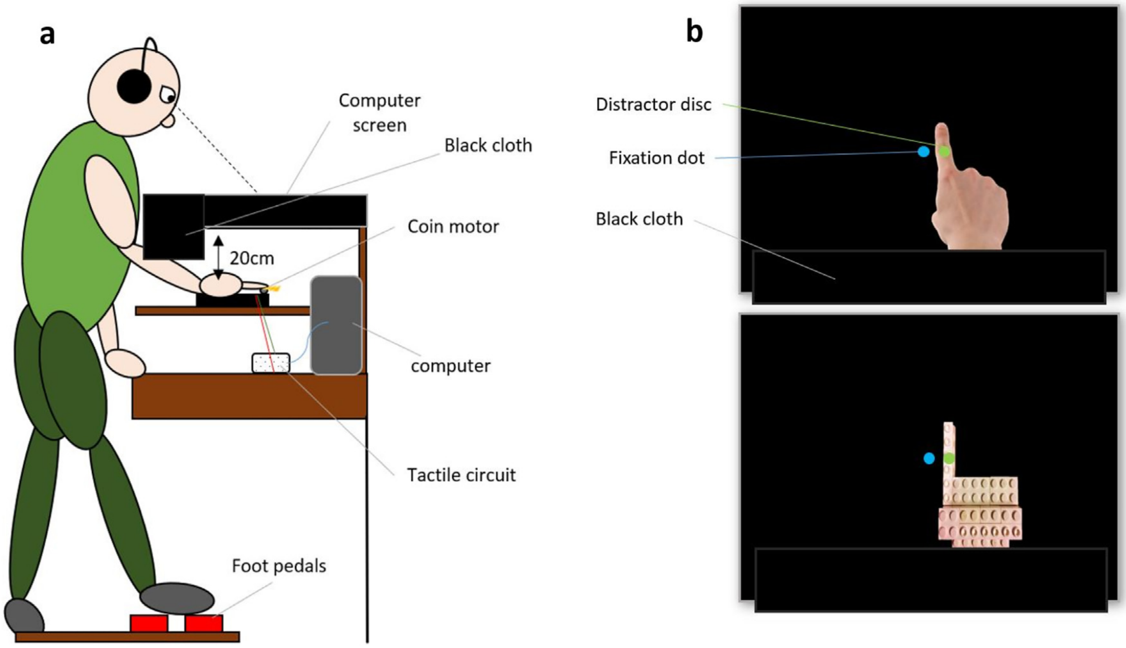

Fig. 4. An illustration of the experimental set-up of Experiment 2. (a) This schematic shows how the participant stood at the apparatus and looked down onto a flat computer screen to view a visual image of a hand or object stimulus. (b) An example of the images of the hand (upper panel) and object (lower panel) used in the experiment. In this experiment, context stimuli (hand or object) were presented from an egocentric view. See main text for details.

from view throughout the experiment by the screen and a black cloth (Fig. 4b). The black cloth also covered the bottom portion of the screen that corresponded to the location of the participant's arm. This set-up allowed for all visual context stimuli (self-, other- hand and object) to be displayed from an egocentric view and spatially aligned with the position of the participant's actual hand. These images of the hands or objects subtended a visual angle of approximately 10 by 16 degrees in the horizontal and vertical dimensions respectively.

The experiment was based on a 2 (congruency) × 3 (stimulus context) within-subjects design. Prior to the experiment, the participants were first required to explicitly identify their hand from a 2 × 4 array of hand images shown from an egocentric view. However, based on the results of Experiment 1, here we also added a self-report measure by asking participants to rate the degree to which they self-identified with the images of each hand or object presented during the completed block of CCT trials. A 7 point Likert scale was used to address the question "I identified the stimulus on screen as my hand" on a range from −3 ("Definitely not") to +3 ("Definitely"). The analysis of these data can be found in Supplementary materials. The entire experiment was completed in one session, lasting approximately 40 min.

3.2. Results

All participants correctly identified the image of their own hand in the array of hand images, all presented from an egocentric view. Trials with missing values or statistical outliers (RT) amounted to a removal of 4.9% and 3% of the data from the final dataset, respectively. Table 2 shows the descriptive statistics for percentage error, reaction times and inverse efficiency scores for Experiment 2. The mean error rate across all participants was 24.46% (SD = 21.11). The within-subjects 2 (distractor number) × 2 (congruency) × 3 (stimulus context) ANOVA on mean error rates yielded a main effect of distractor number ($F(1,18)$ = 26.71,

Table 2
The means and standard deviations of error rates (%), RTs (ms) and IE scores (ms) across congruency conditions (Cong = Congruent, Incong = Incongruent) for each stimulus context (Self-hand, Other-hand, Object) presented from an egocentric view in Experiment 2.

	Overall		Self		Other		Object	
	Cong	Incong	Cong	Incong	Cong	Incong	Cong	Incong
Error (%)	12.7 (12.3)	36.2 (21.6)	5.2 (6.0)	31.7 (20.3)	5.3 (7.1)	24.4 (20.1)	27.5 (6.1)	52.6 (13.4)
			18.53 (28.11)		14.89 (24.99)		39.93 (21.99)	
RT (ms)	805 (125)	832 (118)	796 (130)	829 (122)	817 (125)	843 (121)	803 (127)	826 (117)
IE (ms)	943 (212)	1479 (654)	842 (135)	1308 (393)	866 (135)	1211 (487)	1121 (231)	1919 (798)

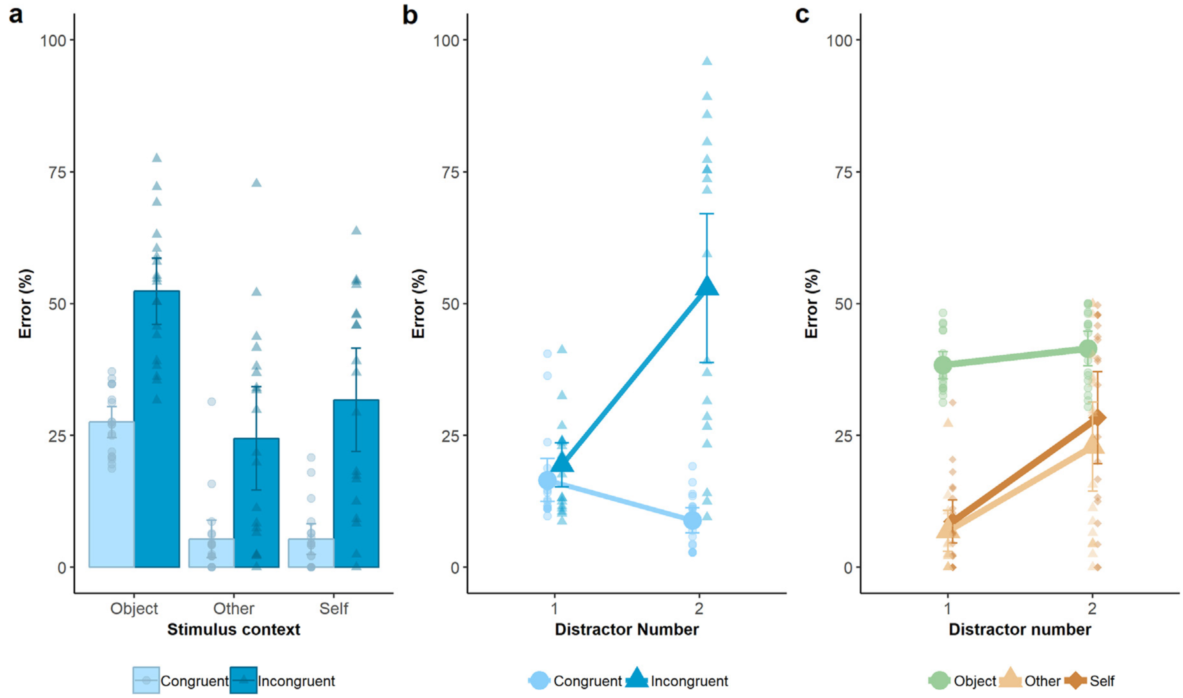

Fig. 5. The results (mean % error) of Experiment 2. (a) Plot shows % errors made to each of the stimulus contexts presented from an egocentric view as a consequence of visuotactile congruency. (b) Plot shows the interaction between congruency and distractor number**.** (c) Plot shows the interaction between distractor number and stimulus context. Individual mean % errors are shown. Error bars represent 95% confidence intervals.

$p < 0.001$, $\eta^2_G = 0.11$) and congruency ($F(1,18) = 39.69$, $p < 0.001$, $\eta^2_G = 0.30$). There was also a main effect of stimulus context ($F(2,36) = 62.33$, $p < 0.001$, $\eta^2_G = 0.27$). The two-way interaction between congruency and stimulus context failed to reach significance ($p = 0.14$) and the Bayes Factor indicated support for the null hypothesis ($BF_{01} = 9.1$)

Post- hoc pairwise comparisons (Bonferroni corrected) revealed more errors were made to the object ($M = 39.93\%$, $SD = 21.99$) than either the self- ($M = 18.53\%$, $SD = 28.11$; $p < 0.001$) or other-hand ($M = 14.89\%$, $SD = 24.99$) (both $p < 0.001$) with no difference between the self- and other-hand ($p = 0.13$) (see Fig. 5a). The Bayes Factor for the comparison of the self- and other-hand was inconclusive ($BF_{01} = 0.64$).

The two-way interaction between distractor number and congruency was significant ($F(1,18) = 20.43$, $p = 0.003$, $\eta^2_G = 0.25$) as shown in Fig. 5b. Post-hoc pairwise comparisons confirmed a significant congruency effect emerged only in response to 2 distractors (*M diff* = 44.03%, $SD = 33.05$) ($p < 0.001$). There was a large difference in error rates between the object and the self- (*M diff* = 29.67%, $SD = 10.33$) and other-hand (*M diff* = 31.47%, $SD = 10.18$) in response to 1 distractor ($p < 0.001$) and no difference between the self- and other- hand (*M diff* = 1.81%, $SD = 7.41$) ($p = 0.93$). The Bayes factor the self- and other-hand comparison was inconclusive ($BF_{01} = 2.65$). There were comparatively smaller differences between the object and self- (*M diff* = 13.12%, $SD = 15.20$) ($p = 0.005$) and other-hand (*M diff* = 18.60%, $SD = 15.72$) ($p < 0.001$) in response to 2 distractors and no difference between the self- and other-hand (*M diff* = 5.48%, $SD = 9.44$) ($p = 0.07$). The Bayes factor the self- and other-hand comparison was inconclusive ($BF_{01} = 0.38$). All other interactions failed to reach significance (all *ps* > 0.05).

3.3. Discussion

Consistent with the findings from Experiment 1 we found a large non-spatial CCE on error responses when participants viewed visual distractors during a task in which they reported the number of tactile events in a sequence. As in Experiment 1, this was driven primarily by responses to 2 distractors, due to a combination of perceptual sensitivity and response bias (see Supplementary materials). Also consistent with Experiment 1 was the presence of a weaker effect of congruency for RTs and no modulation of RT by stimulus context (see Supplementary materials). There was no evidence that stimulus context influenced the magnitude of the non-spatial CCE on error rate (or RTs): again, the effect was present and comparable for both the self- and other-hand conditions and the object, with strong support for the null hypothesis (Dienes, 2014). However, in contrast to Experiment 1, a relative increase in errors to the object was found across both the congruent and incongruent sequence conditions, suggesting a cost in tactile discrimination and larger interference from the distractors when viewing the object. This was observed across distractor conditions but largely driven by differences in error rates across stimulus contexts in response to 1 distractor. A signal detection analysis (see Supplementary materials) indicated that poorer performance in response to the object relative to hand contexts was due to reduced perceptual sensitivity to the tactile vibrations. We found no evidence for an influence of visual body identity on visuotactile interactions.

The principle focus of Experiments 1 and 2 was to investigate the influence of viewing images of a hand versus object stimuli on the CCE for sequences of events, and specifically, whether self-recognition modulated the size of the CCE. We found no robust evidence for a role of visual body identity on the CCE in this task. One possibility why visual body identity did not affect the CCE in Experiment 1 was that recognition of one's own hand was not sustained during the main experiment. Here we assessed self-recognition throughout the experiment using a rating task and the results suggested that there was no loss in identifying one's own hand and no evidence that the other's hand was identified as one's own during the experiment. Nevertheless, the finding that self-identification of a body part did not influence the visuotactile CCE is inconsistent with some previous studies examining spatial (Aspell et al., 2009; Maravita et al., 2002; Salomon et al., 2012; Zopf et al., 2010) and temporal (Maselli et al., 2016) visuotactile interactions and evidence for the prioritization of self-relevant information (Frassinetti et al., 2011; Humphreys & Sui, 2016; Schäfer et al., 2016; Sui & Rotshtein, 2019). Therefore, this finding deserves further scrutiny. To that end, we attempted to further enhance self-recognition of the hand in the following experiment to measure the subsequent effect of visual body identity on the CCE for sequences of events.

4. Experiment 3

The following experiment was designed to address the possibility that presenting images of hands alone was not sufficiently effective in evoking a perception of visual body identity. Indeed, previous work has implied that differences in the spatial CCE across self versus other stimuli may be driven by more reliable visual cues to identity, such as images of faces (Salomon et al., 2012). Thus, in the following experiment we added extra cues to identity, in the form of images of faces or names, to each of the visual displays. We used images of faces and names in order to assess the role of salient corporeal and non-corporeal social cues to identity on crossmodal integration and in response to evidence for self-prioritization effects in response to these identity cues (Humphreys & Sui, 2016). We expected that, by reinforcing visual body identity, images of one's own hand would be associated with a greater CCE for sequences of events than images of another unfamiliar person's hand or a control object.

4.1. Methods

4.1.1. Participants

The same participants who took part in Experiment 2 also conducted this experiment.

4.1.2. Design and procedure

The experimental set-up and procedure were near identical to that of Experiment 1 with the exception of the content of the visual displays. We adapted the visual displays used in Experiment 1 as follows: first we used images depicting an allocentric view of the hand, and included an image of a face or name to each display, across the two body contexts (self- and other-hand). An image of each participant's face was taken under constant luminance conditions while the participant stood against a plain white wall and adopted a neutral expression. Images were cropped so that only the face was presented. Fig. 6 provides an example of the stimuli used in this experiment. In any one trial, the face images were positioned off-centre, to the left of fixation as well as the image of the hand (the hand was positioned to the right of fixation as in Experiments 1 and 2). In this way, both the image of the face and the hand were aligned with the physical position of the participant, and were convincingly presented as a 'mirror' image of the participant's face and right hand during the experiment. The image of the hand slightly overlapped the bottom right of the face image to ensure that relative image sizes were maintained and that all information contained in the scene was visually accessible from the central fixation point. The image of the participant's own hand was always paired with an image of their face (self), whereas the image of another person's hand was paired with an unfamiliar face selected from the Chicago-Face Database (Ma, Correll, & Wittenbrink, 2015). Faces used for the other-hand condition were selected to be sex-matched to the participant's own face. All face images (self and other) were presented with a neutral expression. Finally, the image of the control object (used in previous experiments) was paired with a face-like configuration made from LEGO (Fig. 6). The hand or object image occupied an area of approximately 14 by 15 degrees in visual angle.

We also decided to use names as an alternative to face images in order to investigate whether visuotactile integration was specifically affected by visual body identity or semantic knowledge of identity more generally (see Fig. 6). For the self-hand condition, an image of the participant's own hand was paired with an image of their name. For the other-hand condition, the images of the hand were paired with a name that was sex-matched to the participant and was composed of the same number of letters as the participant's name. Finally, in the object condition the object image was paired with a nonsense word that was also matched in letter length with each participant's name. The letters were written in upper case and in a white, sans-serif font to contrast with the black background and were positioned to the left of the scene thus matching the location of the face image in the face conditions. The visual angle subtended by the name stimuli varied from approximately 9 to 14 degrees horizontally, depending on word length, and was approximately 9 degrees vertically.

Experiment 3 was based on 2 (congruency) × 3 (stimulus context) × 2 (identity marker) within-subjects design. As in Experiments 1 and 2, the participants were required to first identify their hand and face in separate 2 × 4 arrays of hand and face images. The same self-identity ratings were provided at the end of each experimental block, as described in Experiment 2. The analysis of these data can be found as Supplementary material. To avoid carry-over or learning effects in the data, the order of the experimental sessions of Experiment 3 and Experiment 2 were counterbalanced across participants and separated by a duration of at least two weeks. Moreover, the order of the face and name blocks was counterbalanced across participants and the presentation of

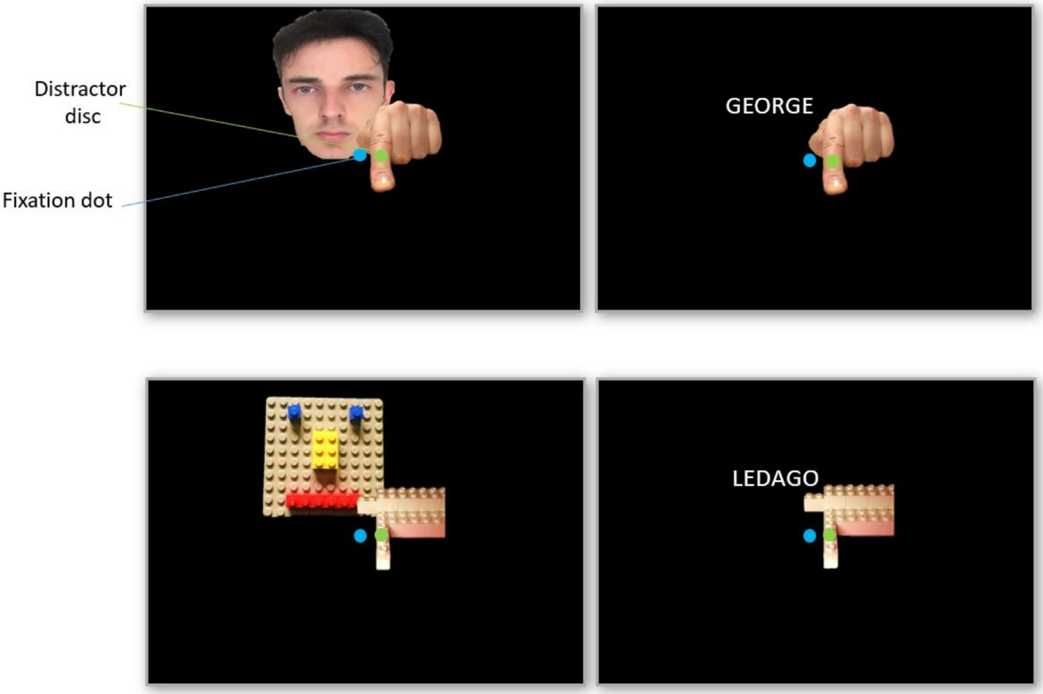

Fig. 6. An example of the visual stimuli used in Experiment 3 in which an image of participant's own face (self), an 'other' face or face-like object (panels on the left), or an image of participant's own name (self), 'other' name or nonsense word (panels on the right) were included. In this experiment, context stimuli (hand or object) were presented from an allocentric view. See main text for further details. Note, the face shown is for illustrative purposes and was not included in the experiment.

these blocks was separated by a duration of at least two weeks. Each condition (face or name) was completed in one experimental session, lasting approximately 40 min.

4.2. Results

All participants correctly identified the image of their own hand in the array of hand images, both from an allocentric or egocentric view. Trials with missing values or statistical outliers amounted to a removal of 3.7% and 3.2% of the data from the final dataset, respectively. Table 3 shows the descriptive statistics for percentage error, reaction times and inverse efficiency scores for Experiment 3. Overall, participants' error rate was on average 27.50% ($SD = 20.84$). The within-subjects 2 (distractor number) × 2 (congruency) × 3 (stimulus context) × 2 (identity marker) ANOVA on mean error rates revealed a main effect of distractor number ($F(1,18) = 50.91, p < 0.001, \eta^2_G = 0.16$) and congruency ($F(1,18) = 62.65, p < 0.001, \eta^2_G = 0.32$). There was also a main effect of stimulus context ($F(2,36) = 94.41, p < 0.001, \eta^2_G = 0.18$) (Fig. 7a). There was no two-way interaction between congruency and stimulus context ($p = 0.78$) and the Bayes Factor indicated support for the null hypothesis ($BF_{01} = 20$).

Post-hoc pairwise comparisons (Bonferroni corrected) confirmed higher error rates in response to the object ($M = 40.22\%$, $SD = 28.34$) than either the self- ($M = 20.75\%$, $SD = 30.48$) ($p < 0.001$) or other-hand ($M = 22.26\%$, $SD = 32.30$) ($p < 0.001$) with no difference between the self- and other-hand ($p = 0.46$). The Bayes Factor for the comparison of self- and other-hands was inconclusive ($BF_{01} = 1.57$). The main effect of identity marker failed to reach significance ($p = 0.57$).

A two-way interaction between distractor number and congruency ($F(1,18) = 39.36, p < 0.001, \eta^2_G = 0.45$) was found, as shown

Table 3
The means and standard deviations of error rates (%), RTs (ms) and IE scores (ms) across congruency conditions (Cong = Congruent, Incong = Incongruent) and for each stimulus contexts (Self-hand, Other-hand, Object) shown from an allocentric view in Experiment 3.

	Overall		Self		Other		Object	
	Cong	Incong	Cong	Incong	Cong	Incong	Cong	Incong
Error (%)	14.3 (12.9)	40.8 (18.8)	6.7 (7.8)	33.7 (18.0)	8.9 (10.8)	34.9 (18.5)	27.1 (8.5)	53.7 (12.3)
			20.75 (30.48)		22.26 (32.30)		40.22 (28.34)	
RT (ms)	768 (130)	782 (136)	778 (138)	800 (136)	761 (119)	772 (130)	764 (136)	776 (144)
IE (ms)	915 (205)	1426 (412)	839 (162)	1267 (284)	844 (141)	1249 (278)	1061 (222)	1762 (432)

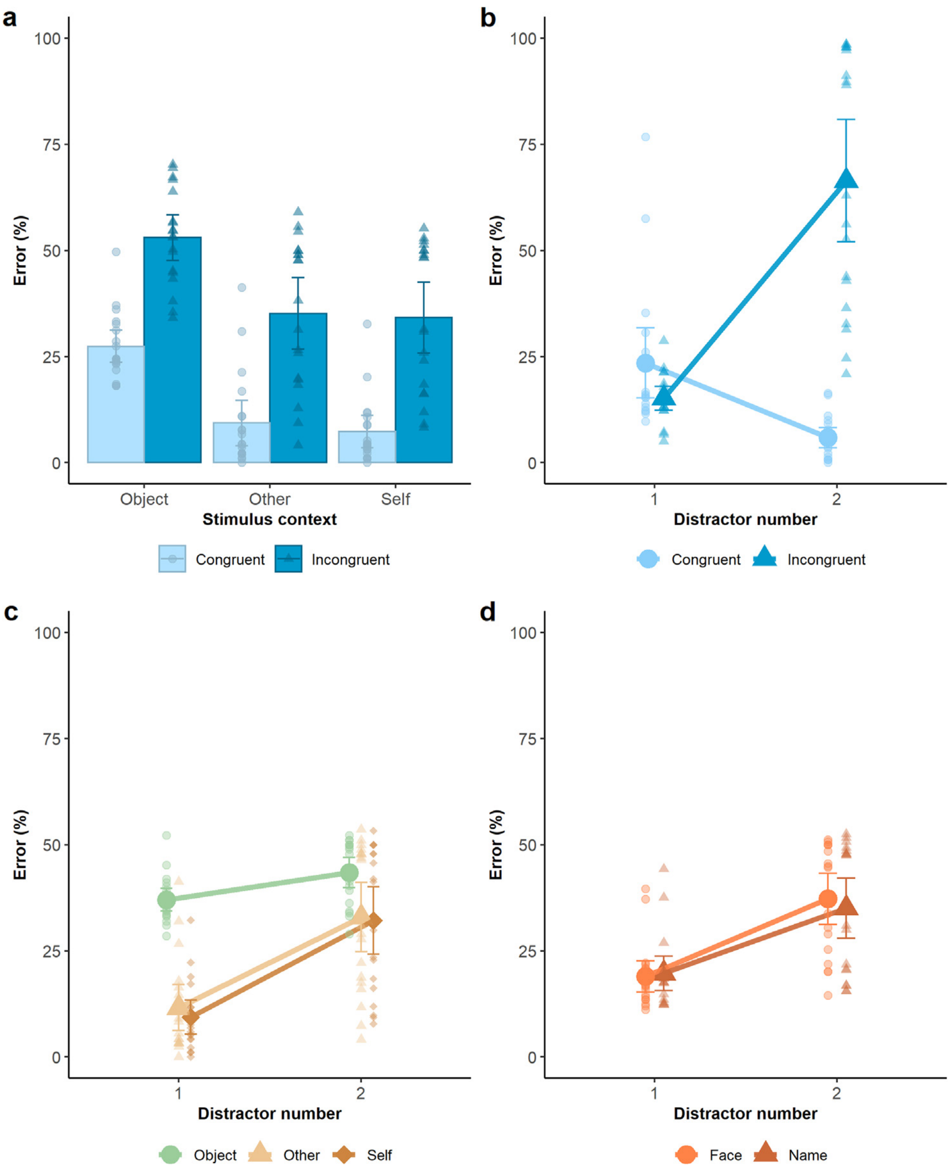

Fig. 7. Plots showing the results (mean % error) from Experiment 3. (a) Plot shows % errors made to each of the stimulus contexts shown from an allocentric view as a consequence of visuotactile congruency. (b) Plot shows the interaction between distractor number and congruency. (c) Plot shows the interaction between distractor number and stimulus context. (d) Plot shows the interaction between distractor number and identity marker. Individual mean errors are shown. Error bars represent 95% confidence intervals.

in Fig. 7b. Post-hoc pairwise comparisons confirmed a congruency effect in response to 2 distractors only (*M diff* = 60.53%, *SD* = 33.10) ($p < 0.001$). The two-way interaction between distractor number and stimulus context also reached significance ($F(2,36) = 35.68$, $p < 0.001$, $\eta^2_G = 0.04$) as shown in Fig. 7c. Post-hoc pairwise comparisons confirmed large differences in error rates between the object and the self- (*M diff* = 27.64%, *SD* = 6.71) ($p < 0.001$) and other-hand (*M diff* = 25.42%, *SD* = 8.33) ($p < 0.001$) in response to 1 distractor with no difference between the self- and other-hand (*M diff* = 2.21%, *SD* = 6.50) ($p = 0.49$). The Bayes Factor for the comparison of self- and other-hand was inconclusive ($BF_{01} = 1.69$). There were smaller differences in error rates between the object and the self- (*M diff* = 11.32%, *SD* = 9.68) and other-hand (*M diff* = 10.51%, *SD* = 10.55) ($p < 0.001$) in response to 2 distractors, with no difference between the self- and other- hand (*M diff* = 0.81, *SD* = 6.16) ($p = 1.00$), as supported by the Bayes Factor ($BF_{01} = 3.70$). Finally, a significant two-way interaction was found between distractor number and identity marker ($F(1,18) = 5.98$, $p = 0.025$, $\eta^2_G = 0.001$) as shown in Fig. 7d. However, post-hoc pairwise comparisons indicated no differences between the face or name identity markers in response to either 1 or 2 distractors (*ps* > 0.05). All other interactions failed to reach significance (all *ps* > 0.05).

4.3. Discussion

The results of Experiment 3 provide further evidence that the CCE extends beyond the spatial domain, in that performance in discriminating tactile sequences of events was significantly worse in response to the presentation of incongruent visual events, particularly in response to 2 distractors. Unlike Experiments 1 and 2, an effect of congruency on RTs was not found (see Supplementary materials). However, consistent with previous results, the results of Experiment 3 also suggest that the magnitude of the CCE on error rates or RTs was not affected by stimulus context or visual self-recognition, even when the display included markers (face or name) to the identity of one's own hand over another's, with strong support for the null hypothesis (Dienes, 2014). Again, the self-report data add supporting evidence that participants identified their own hands throughout the CCT (see Supplementary materials). Overall performance in discriminating tactile sequences of events was significantly worse on trials in which an object context was presented, a result that is consistent with that of Experiment 2. A signal detection analysis (see Supplementary materials) indicated that this was due to overall reduced perceptual sensitivity to the vibrations in response to the object than the hand, as in Experiment 2.

5. General discussion

Across three separate experiments we investigated whether the CCE also generalises to tasks beyond the spatial domain, and assessed the influence of viewing images of the hand or an object on the discrimination of tactile sequences of events and the role of body identity in this process. Participants were required to discriminate the presence of 1 or 2 tactile sequences of events on their hand whilst ignoring numerically congruent or incongruent visual distractors superimposed onto images of a self-hand, other-hand or object. The results of all experiments suggest that the CCE also applies to the temporal domain, specifically in tactile perception of sequences of events, but mainly when attempting to ignore 2 distractors. The magnitude of the CCE for error rates is compatible with that of previous studies using a similar non-spatial set-up (e.g. see Forsberg et al., 2019; Holmes et al. 2006, 2007; Poole et al., 2015), suggesting that the non-spatial CCT yields largely consistent results. The RTs appeared less sensitive to non-spatial visuotactile congruency than error rates, and the CCEs for RTs were larger in magnitude than the CCE reported by Poole et al. (2015) but smaller than the CCE reported by Holmes et al. (2006, 2007). Unlike error rates, the RTs were consistently unaffected by stimulus context, suggesting that RT may be a less sensitive measure of stimulus induced changes to the discrimination of tactile sequences of events than accuracy.

We found no difference in the magnitude of the CCE (i.e. a greater number of errors to the crossmodal incongruent than congruent trials) between the hand and object conditions across all experiments. This latter result is not consistent with previous reports of the CCE in the spatial domain in which a larger CCE is often observed in response to viewing a body or body part versus an object (Aspell et al., 2009; Igarashi et al., 2004, 2007; Igarashi et al., 2008; Salomon et al., 2012; Thomas et al., 2006). However, it is consistent with evidence that temporal visuotactile perception can be unaffected by the body context (Keys et al., 2018; Smit et al., 2019). This finding may reflect a fundamental difference in how visual and tactile inputs interact across the spatial and temporal domains of perception in response to stimulus corporeality (e.g. Smit et al., 2019). For example, viewing a body context may have a higher relevance for spatial as opposed to temporal perception as vision can provide greater spatial precision for localising touch (Eimer, 2004; Margolis & Longo, 2015). Moreover, viewing the body may activate the body schema (Reed & Farah, 1995) which is known to facilitate tactile localisation specifically (Head & Holmes, 1911). Importantly, a link between body schema and patterns of spatial visuotactile interactions in PPS has been demonstrated (Holmes & Spence, 2004). It seems unlikely that the body does not provide an important context for perceiving sequences of events, given the sensitivity of the tactile system in perceiving temporal information (Green, Reese, Pegues, & Elliott, 1961). However, the nature of the body representation required to mediate crossmodal information of sequences of events may be less precise than that required for localisation.

We also found no evidence that viewing an image of one's own hand further enhanced the CCE over viewing the hand of an unfamiliar other, even when the participants correctly identified an image of their own hand, from both an allocentric and egocentric view. This failure to find an effect of self-identification on visuotactile interactions is inconsistent with the reported findings from previous work using the spatial CCT with hand stimuli (Maravita et al., 2002; Zopf et al., 2010) or full-body stimuli (Aspell et al., 2009; Salomon et al., 2012), as well as evidence for changes to temporal visuotactile interactions in response to self-identified hands (Maselli et al., 2016) and the prioritization of self-relevant information (Frassinetti et al., 2011; Humphreys & Sui, 2016; Schäfer et al., 2016; Sui & Rotshtein, 2019). In our CCT, all images of hands (and the object) were oriented so as to be compatible with the position of each participant's hand (i.e. all were shown from either a 'mirrored' third or typical first person view); therefore our results suggest that a visual representation of self within PPS did not influence visuotactile interactions, despite clear evidence for the explicit recognition of one's hand in our task (Frassinetti et al., 2011). Evidence for a specific effect of body identity on the spatial CCE or on temporal visuotactile interactions is not consistent (Maselli et al., 2016; Salomon et al., 2012; Keys et al., 2018) and methodological differences may be an important factor to consider. For example, Salomon et al. (2012) presented a full body, including the face, to participants. Other studies have explicitly manipulated body or rubber hand ownership (Aspell et al., 2009; Maselli et al., 2016; Zopf et al., 2010) and/or posture (Ide & Hidaka, 2013; Keys et al., 2018; Pavani et al., 2000; Smit et al., 2019). Thus, it remains possible that visual context may affect temporal visuotactile interactions with other stimulus manipulations such as changing the posture of viewed hands or increasing sense of hand ownership (although see Keys et al., 2018; Smit et al., 2019). For example, there is evidence for self-advantage effects when processing one's own hand which may stem from sensorimotor influences, thereby also relying on the body schema rather than visual recognition per se (Frassinetti et al., 2011). These factors could also, at least partially, explain why explicit visual hand recognition alone was insufficient to produce a change to implicit visuotactile interactions in the

CCT. Nonetheless, the absence of a difference in non-spatial CCEs between self and other, even in the presence of an image of one's own face (Experiment 3) or with clear identification with a self-hand only (Experiments 2 and 3; see Supplementary materials), suggests no robust effect of visual self-identification or self-prioritization for temporal visuotactile discrimination. This finding is consistent with previous evidence for an overlap in the representations of PPS and the body representation across self and other (Brozzoli, Gentile, Bergouignan, & Ehrsson, 2013; Haggard, 2006; Reed & Farah, 1995; Van Den Bos & Jeannerod, 2002; Whiteley et al., 2008).

It should be noted that the CCE in the spatial domain may be unrelated to multisensory integration per se, and more to do with late, decision making processes and cognitive factors. For example, there is some evidence to suggest that the spatial CCE may arise due to a spatial conflict of responses across modalities, with a smaller role for multisensory integration (Forsberg et al., 2019; Forster & Pavone, 2008; Marini, Romano, & Maravita, 2017). A role of spatial response conflict in the CCE may account for the lack of consistency in the effects of body contexts across spatial and temporal perception.

We noted that the overall number of errors made during the tactile discrimination task when viewing the object specifically, increased in Experiments 2 and 3 relative to Experiment 1. The overall increase in errors partly reflected a general cost to tactile perception on congruent trials. This is consistent with evidence for better tactile processing when viewing a body than an object, even when vision is non-informative to the tactile task. This effect is often referred to as the visual enhancement of touch (VET) (Kennett et al., 2001) and has been found for task of tactile spatial acuity (Kennett et al., 2001; Haggard, 2006) and tactile detection (Harris, Arabzadeh, Moore, & Clifford, 2007; Tipper et al., 1998). It is argued that the VET occurs when viewing a body prepares the brain for upcoming tactile stimulation (Fiorio & Haggard, 2005). Indeed, in the current experiments, the participants previewed the hand before the occurrence of a crossmodal event and a VET may have occurred during this period. Indeed, the VET has been found to arise in response to a relatively brief glimpse of a hand (Cardini, Longo, Driver, & Haggard, 2012). The VET may be linked to the multisensory representation of PPS in which tactile detection as well as visuotactile interactions are enhanced (Cléry, Guipponi, Odouard, Wardak, & Hamed, 2015; Fiorio & Haggard, 2005; Harris et al., 2007; Konen & Haggard, 2012; Làdavas et al., 1998; Rizzolatti et al., 1997; Spence et al., 2008). The absence of a difference between the hand conditions of Experiments 2 and 3 was ambiguous based on the Bayes Factor. However, as with the CCE, the VET can occur when viewing either one's own or another's body (Haggard, 2006), without reinforcement of body ownership, as evidenced in the present findings.

The results of Experiment 1 were somewhat surprising and it remains unclear to us why a cost in tactile perception was not found in response to viewing the object relative to viewing the hands, as found in the other experiments. Others have reported equivalent performance to viewing a hand or object when detecting tactile taps and a disadvantage in tactile gap detection when viewing a hand compared to an object (Press, Taylor-Clarke, Kennett, & Haggard, 2004). We suggest two possible reasons for our result. First, the results may be related to attentional processes in that the images presented in Experiment 1, particularly the object, may have drawn less attention to the screen relative to the images of Experiment 3, which involved an identical set-up. As a participant's physical hand was always visible in Experiment 1 during the task, a VET may have occurred across all stimulus contexts (Cardini et al., 2012). The absence of a similar effect for Experiment 3 may reflect the higher visual saliency of this display and a stronger capture of attention towards the screen and away from the physical hand, thus facilitating the dominance of the visual distractors over touch. An alternative explanation of our findings is that image manipulations such as the egocentric view of the object as an extension of the body (Experiment 2) or the additional face-like configurations presented with the object (Experiment 3) may have been more likely to violate a participant's body representation or perceptual expectations than the object shown in Experiment 1, thus appearing less convincingly 'hand-like'. Indeed, an egocentric view of the body seems to be important for bodily self-consciousness and self-other discrimination (Blanke, 2012; Brady et al., 2011; Conson et al., 2010; Saxe et al., 2005) while information from faces is perceptually prioritised (Langton, Watt, & Bruce, 2000; Scheller, Büchel, & Gamer, 2012). Stimuli perceived as broadly 'hand-like' in shape or form can also visually 'capture' touch (Rosén et al., 2009; Tsakiris, Carpenter, James, & Fotopoulou, 2010), albeit less convincingly than a realistic hand (Tsakiris et al., 2010). Although both attentional and perceptual processes may be implicated in the difference in results across experiments, future research is required to elucidate these influences on visuotactile interactions when viewing objects.

Finally, our results suggest a consistent effect of the number of visual distractors on performance on the tactile task. The presentation of 2 distractors with 1 tactile event was associated with significantly more errors than when 1 distractor was presented with 2 tactile events, across all experiments. This implies a stronger 'fission-like' than 'fusion-like' effect which is compatible with previous findings in crossmodal integration (Andersen et al., 2004) including for vision and touch (Holmes et al., 2006), demonstrating a robust effect of modality discontinuity on task performance (Shams et al., 2002). However, it is important to note that, due to the timing of the visual and tactile events during the CCT, there was a relatively small temporal overlap (20 ms) between a single visual distractor and a second vibration on incongruent trials with 2 vibrations, thus possibly explaining the lack of strong 'fusion-like' effects in the data.

The results of Experiment 2 and 3 suggest that there was a larger difference between hand and object contexts in response to 1 distractor than 2 distractors, largely due to relatively poor task performance when viewing the object on all visuotactile combinations. The presence of larger interference from the distractors on incongruent trials in response to an object contrasts with the findings of spatial CCT studies which reported larger crossmodal interference in a body context (Igarashi et al., 2007, 2008; Salomon et al., 2012; Thomas et al., 2006) and the expected influence of shared context on multisensory interactions (Chen & Spence, 2017; Johnson et al., 2006; Mirams et al., 2010; Welch & Warren, 1980). While there is evidence for an influence of visual distractors on the discrimination of tactile sequences of events, this is typically smaller than the influence of tactile distractors on visual discrimination due to the higher reliability of touch for temporal judgements (Bresciani, Dammeier, & Ernst, 2006; Ernst & Banks, 2002; Philippi, van Erp, & Werkhoven, 2008). As such, the difference in rates of visuotactile interference across contexts may be attributable to differences in modality weighting based on the relative reliability of the tactile information in response to a hand or an object. More specifically, in Experiments 2 and 3, a VET occurs when viewing the hands relative to the object, in that tactile discrimination is benefited and more reliable, accounting for the improved discrimination performance on all congruent trials and incongruent trials with 2 sequential vibrations, and the interference from the task-irrelevant distractors is reduced, accounting for the smaller 'fission-like' effects in particular.

Our results show no evidence for a role of visual body recognition on the CCE, even though participants were clearly able to identify their own hand and were offered cues to self-identity. Indeed, analyses included data only from participants who recognised their own hand in an array of "other" hands and the distinction between identifying self and other in Experiments 2 and 3 was supported by the results of a ratings task. As such, we made every effort to create a convincing illusion of looking at oneself in a mirror (Experiments 1 and 3) or directly at one's own hand (Experiment 2); therefore concerns relating to stimulus authenticity or plausibility were unlikely to result in a lack of distinction between self- and other-hands and/or hands and objects in the CCE. Moreover, the use of images of hands rather than real hands is consistent with previous CCT work contrasting visuotactile interactions in response to bodies and objects (Igarashi et al., 2004, 2007; Igarashi et al., 2008; Salomon et al., 2012) and self and other bodies (Salomon et al., 2012). As is typical in such CCT research we did not include a unimodal condition, in which only touch or vision was presented to the participants, or a condition in which the stimulus context is absent, but we acknowledge that these may be limitations of this work. Our study highlights the need for future research to further investigate the role of stimulus context and visual body identity on visuotactile interactions in the perception of sequences of events.

6. Conclusions

The findings of three experiments demonstrate a robust crossmodal congruency effect in a task involving the perception of sequences of events. Specifically, the presentation of visual distractors resulted in a larger number of errors in tactile perception when the number of events was incongruent compared to congruent across modalities. Our second main finding was that there was no evidence that the magnitude of the non-spatial CCE was affected by viewing images of a hand or object. However, tactile discrimination was significantly impaired when distractors were presented with an object shown from the first person view (Experiment 2) or when additional markers of identity were added to a display relative to viewing distractors on an image of a hand shown from the third person view (Experiment 3). Finally, we found no evidence that visual hand identity affected the CCE: this effect was similar across identified images of a participant's own hand and images of an unfamiliar other's hand.

Our results provide novel evidence that, unlike the spatial CCE, the CCE for sequences of events may not be sensitive to stimulus context and that performance on the tactile discrimination of sequences of events is impaired when viewing an object relative to a hand under certain visual contexts. That is, we found reduced tactile discrimination when viewing an object presented from the first person view and an object accompanied by additional visual information from the third person view, but not in response to the same object presented in isolation from the third person view. These results shed light on the interaction between tactile and visual information about external stimuli in the temporal domain and the influence of the body representation on the nature of visuotactile interaction in the perception of events.

CRediT authorship contribution statement

Alan O' Dowd: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Francesca Sorgini:** Methodology, Software, Resources, Writing - original draft, Writing - review & editing. **Fiona N. Newell:** Conceptualization, Methodology, Resources, Data curation, Writing - original draft, Writing - review & editing, Supervision.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.concog.2020.102900.

References

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research, 21*(3), 301–308. https://doi.org/10.1016/j.cogbrainres.2004.06.004.

Aspell, J. E., Lenggenhager, B., & Blanke, O. (2009). Keeping in touch with one's self: Multisensory mechanisms of self-consciousness. *PloS One, 4*(8), e6488. https://doi.org/10.1371/journal.pone.0006488.

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience, 13*(8), 556. https://doi.org/10.1038/nrn3292.

Brady, N., Maguinness, C., & Choisdealbha, Á. N. (2011). My hand or yours? Markedly different sensitivity to egocentric and allocentric views in the hand laterality task. *PloS One, 6*(8), e23316. https://doi.org/10.1371/journal.pone.0023316.

Bresciani, J. P., Dammeier, F., & Ernst, M. O. (2006). Vision and touch are automatically integrated for the perception of sequences of events. 2 2 *Journal of Vision, 6*(5), https://doi.org/10.1167/6.5.2 doi: 10.1007/s00221-004-2128-2.

Brozzoli, C., Gentile, G., Bergouignan, L., & Ehrsson, H. H. (2013). A shared representation of the space near oneself and others in the human premotor cortex. *Current Biology, 23*(18), 1764–1768. https://doi.org/10.1016/j.cub.2013.07.004.

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica, 51*(1), 5–13. https://biblio.ugent.be/publication/2001824.

Chen, Y. C., & Spence, C. (2017). Assessing the role of the 'unity assumption'on multisensory integration: A review. *Frontiers in Psychology, 8*, 445. https://doi.org/10.3389/fpsyg.2017.00445.

Cardini, F., Longo, M. R., Driver, J., & Haggard, P. (2012). Rapid enhancement of touch from non-informative vision of the hand. *Neuropsychologia, 50*(8), 1954–1960.

https://doi.org/10.1016/j.neuropsychologia.2012.04.020.
Cléry, J., Guipponi, O., Odouard, S., Wardak, C., & Hamed, S. B. (2015). Impact prediction by looming visual stimuli enhances tactile detection. *Journal of Neuroscience, 35*(10), 4179–4189. https://doi.org/10.1523/JNEUROSCI.3031-14.2015.
Conson, M., Aromino, A. R., & Trojano, L. (2010). Whose hand is this? Handedness and visual perspective modulate self/other discrimination. *Experimental Brain Research, 206*(4), 449–453. https://doi.org/10.1007/s00221-010-2418-9.
Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*, 781. https://doi.org/10.3389/fpsyg.2014.00781.
Eimer, M. (2004). Multisensory integration: How visual experience shapes spatial perception. *Current Biology, 14*(3), R115–R117. https://doi.org/10.1016/j.cub.2004.01.018.
Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429. https://doi.org/10.1038/415429a.
Fiorio, M., & Haggard, P. (2005). Viewing the body prepares the brain for touch: Effects of TMS over somatosensory cortex. *European Journal of Neuroscience, 22*(3), 773–777. https://doi.org/10.1111/j.1460-9568.2005.04267.x.
Forsberg, A., O'Dowd, A., & Gherri, E. (2019). Tool use modulates early stages of visuo-tactile integration in far space: Evidence from event-related potentials. *Biological Psychology, 145*, 42–54. https://doi.org/10.1016/j.biopsycho.2019.03.020.
Forster, B., & Pavone, E. F. (2008). Electrophysiological correlates of crossmodal visual distractor congruency effects: Evidence for response conflict. *Cognitive, Affective, & Behavioral Neuroscience, 8*(1), 65–73. https://doi.org/10.3758/CABN.8.1.65.
Frassinetti, F., Ferri, F., Maini, M., Benassi, M. G., & Gallese, V. (2011). Bodily self: An implicit knowledge of what is explicitly unknown. *Experimental Brain Research, 212*(1), 153–160. https://doi.org/10.1007/s00221-011-2708-x.
Green, J. B., Reese, C. L., Pegues, J. J., & Elliott, F. A. (1961). Ability to distinguish two cutaneous stimuli separated by a brief time interval. 1006 1006 *Neurology, 11*(11), https://doi.org/10.1212/WNL.11.11.1006.
Haggard, P. (2006). Just seeing you makes me feel better: Interpersonal enhancement of touch. *Social Neuroscience, 1*(2), 104–110. https://doi.org/10.1080/17470910600976596.
Harris, J. A., Arabzadeh, E., Moore, C. A., & Clifford, C. W. (2007). Noninformative vision causes adaptive changes in tactile sensitivity. *Journal of Neuroscience, 27*(27), 7136–7140. https://doi.org/10.1523/JNEUROSCI.2102-07.2007.
Head, H., & Holmes, G. (1911). Sensory disturbances from cerebral lesions. *Brain, 34*(2–3), 102–254. https://doi.org/10.1093/brain/34.2-3.102.
Holmes, N. P., Sanabria, D., Calvert, G. A., & Spence, C. (2006). Multisensory interactions follow the hands across the midline: Evidence from a non-spatial visual–tactile congruency task. *Brain Research, 1077*(1), 108–115. https://doi.org/10.1016/j.cub.2005.08.058.
Holmes, N. P., Sanabria, D., Calvert, G. A., & Spence, C. (2007). Tool-use: Capturing multisensory spatial attention or extending multisensory peripersonal space? *Cortex, 43*(3), 469–489. https://doi.org/10.1016/S0010-9452(08)70471-4.
Holmes, N. P., & Spence, C. (2004). The body schema and multisensory representation (s) of peripersonal space. *Cognitive Processing, 5*(2), 94–105. https://doi.org/10.1016/j.brainres.2005.11.010.
Holmes, N. P., & Spence, C. (2005). Multisensory integration: Space, time and superadditivity. *Current Biology, 15*(18), R762–R764. https://doi.org/10.1007/s10339-004-0013-3.
Humphreys, G. W., & Sui, J. (2016). Attentional control and the self: The Self-Attention Network (SAN). *Cognitive Neuroscience, 7*(1–4), 5–17. https://doi.org/10.1080/17588928.2015.1044427.
Ide, M., & Hidaka, S. (2013). Visual presentation of hand image modulates visuo–tactile temporal order judgment. *Experimental Brain Research, 228*(1), 43–50. https://doi.org/10.1007/s00221-013-3535-z.
Igarashi, Y., Kitagawa, N., & Ichihara, S. (2004). Vision of a pictorial hand modulates visual-tactile interactions. *Cognitive, Affective, & Behavioral Neuroscience, 4*(2), 182–192. https://doi.org/10.3758/CABN.4.2.182.
Igarashi, Y., Kitagawa, N., Spence, C., & Ichihara, S. (2007). Assessing the influence of schematic drawings of body parts on tactile discrimination performance using the crossmodal congruency task. *Acta Psychologica, 124*(2), 190–208. https://doi.org/10.1016/j.actpsy.2006.03.004.
Igarashi, Y., Kimura, Y., Spence, C., & Ichihara, S. (2008). The selective effect of the image of a hand on visuotactile interactions as assessed by performance on the crossmodal congruency task. *Experimental Brain Research, 184*(1), 31–38. https://doi.org/10.1007/s00221-007-1076-z.
Jeffreys, H. (1961). *Theory of Probability* (3rd ed.). Oxford: Oxford University Press.
Johnson, R. M., Burton, P. C., & Ro, T. (2006). Visually induced feelings of touch. *Brain Research, 1073*, 398–406. https://doi.org/10.1016/j.brainres.2005.12.025.
Kennett, S., Taylor-Clarke, M., & Haggard, P. (2001). Noninformative vision improves the spatial resolution of touch in humans. *Current Biology, 11*(15), 1188–1191. https://doi.org/10.1016/S0960-9822(01)00327-X.
Keys, R. T., Rich, A. N., & Zopf, R. (2018). Multisensory temporal processing in own-body contexts: Plausibility of hand ownership does not improve visuo-tactile asynchrony detection. *Experimental Brain Research, 236*(5), 1431–1443. https://doi.org/10.1007/s00221-018-5232-4.
Konen, C. S., & Haggard, P. (2012). Multisensory parietal cortex contributes to visual enhancement of touch in humans: A single-pulse TMS study. *Cerebral Cortex, 24*(2), 501–507. https://doi.org/10.1093/cercor/bhs331.
Làdavas, E., Pellegrino, G. D., Farnè, A., & Zeloni, G. (1998). Neuropsychological evidence of an integrated visuotactile representation of peripersonal space in humans. *Journal of Cognitive Neuroscience, 10*(5), 581–589. https://doi.org/10.1162/089892998562988.
Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences, 4*(2), 50–59. https://doi.org/10.1016/S1364-6613(99)01436-9.
Lawrence, M. A. Easy analysis and visualization of factorial experiments. 2016. https://CRAN.R-project.org/package=ez.
Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5.
Makin, T. R., Holmes, N. P., & Zohary, E. (2007). Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *Journal of Neuroscience, 27*(4), 731–740. https://doi.org/10.1523/JNEUROSCI.3653-06.2007.
Maravita, A., Spence, C., Sergent, C., & Driver, J. (2002). Seeing your own touched hands in a mirror modulates cross-modal interactions. *Psychological Science, 13*(4), 350–355. https://doi.org/10.1111/j.0956-7976.2002.00463.x.
Margolis, A. N., & Longo, M. R. (2015). Visual detail about the body modulates tactile localisation biases. *Experimental Brain Research, 233*(2), 351–358. https://doi.org/10.1007/s00221-014-4118-3.
Marini, F., Romano, D., & Maravita, A. (2017). The contribution of response conflict, multisensory integration, and body-mediated attention to the crossmodal congruency effect. *Experimental Brain Research, 235*(3), 873–887. https://doi.org/10.1007/s00221-016-4849-4.
Maselli, A., Kilteni, K., López-Moliner, J., & Slater, M. (2016). The sense of body ownership relaxes temporal constraints for multisensory integration. *Scientific Reports, 6*, 30628. https://doi.org/10.1038/srep30628.
McGovern, D. P., Astle, A. T., Clavin, S. L., & Newell, F. N. (2016). Task-specific transfer of perceptual learning across sensory modalities. *Current Biology, 26*(1), R20–R21. https://doi.org/10.1016/j.cub.2015.11.048.
McGovern, D. P., Roudaia, E., Stapleton, J., McGinnity, T. M., & Newell, F. N. (2014). The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration. *Frontiers in Aging Neuroscience, 6*, 250. https://doi.org/10.3389/fnagi.2014.00250.
Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons I. Temporal factors. *Journal of Neuroscience, 7*(10), 3215–3229. https://doi.org/10.1523/JNEUROSCI.07-10-03215.1987.
Mirams, L., Poliakoff, E., Brown, R. J., & Lloyd, D. M. (2010). Vision of the body increases interference on the somatic signal detection task. *Experimental Brain Research, 202*(4), 787–794. https://doi.org/10.1007/s00221-010-2185-7.
Morey, R. D., Rouder, J. N., & Jamil, T. (2018). BayesFactor: Computation of Bayes Factors for common designs. R package version 0.9. 12-4.2.
Pavani, F., Spence, C., & Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological Science, 11*(5), 353–359. https://doi.org/10.1111/1467-9280.00270.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y.

Philippi, T. G., van Erp, J. B., & Werkhoven, P. J. (2008). Multisensory temporal numerosity judgment. *Brain Research, 1242*, 116–125. https://doi.org/10.1016/j.brainres.2008.05.056.

Poole, D., Couth, S., Gowen, E., Warren, P. A., & Poliakoff, E. (2015). Adapting the crossmodal congruency task for measuring the limits of visual–tactile interactions within and between groups. *Multisensory Research, 28*(3–4), 227–244. https://doi.org/10.1163/22134808-00002475.

Press, C., Taylor-Clarke, M., Kennett, S., & Haggard, P. (2004). Visual enhancement of touch in spatial body representation. *Experimental Brain Research, 154*(2), 238–245. https://doi.org/10.1007/s00221-003-1651-x.

Reed, C. L., & Farah, M. J. (1995). The psychological reality of the body schema: A test with normal participants. *Journal of Experimental Psychology: Human Perception and Performance, 21*(2), 334. https://psycnet.apa.org/doiLanding?doi=10.1037%2F0096-1523.21.2.334.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1997). The space around us. *Science, 277*(5323), 190–191. https://doi.org/10.1126/science.277.5323.190.

Rosén, B., Ehrsson, H. H., Antfolk, C., Cipriani, C., Sebelius, F., & Lundborg, G. (2009). Referral of sensation to an advanced humanoid robotic hand prosthesis. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery, 43*(5), 260–266. https://doi.org/10.3109/02844310903113107.

Rosenthal, O., Shimojo, S., & Shams, L. (2009). Sound-induced flash illusion is resistant to feedback training. *Brain Topography, 21*(3–4), 185–192. https://doi.org/10.1007/s10548-009-0090-9.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. Journal of Mathematical Psychology, 56(5), 356-374. Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. Journal of Mathematical Psychology, 56(5), 356–374.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225.

Salomon, R., Van Elk, M., Aspell, J. E., & Blanke, O. (2012). I feel who I see: Visual body identity affects visual–tactile integration in peripersonal space. *Consciousness and Cognition, 21*(3), 1355–1364. https://doi.org/10.1016/j.concog.2012.06.012.

Saxe, R., Jamal, N., & Powell, L. (2005). My body or yours? The effect of visual perspective on cortical body representations. *Cerebral Cortex, 16*(2), 178–182. https://doi.org/10.1093/cercor/bhi095.

Schäfer, S., Wesslein, A. K., Spence, C., Wentura, D., & Frings, C. (2016). Self-prioritization in vision, audition, and touch. *Experimental Brain Research, 234*(8), 2141–2150. https://doi.org/10.1007/s00221-016-4616-6.

Scheller, E., Büchel, C., & Gamer, M. (2012). Diagnostic features of emotional expressions are processed preferentially. *PloS One, 7*(7), e41792. https://doi.org/10.1371/journal.pone.0041792.

Serino, A. (2019). Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. *Neuroscience & Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2019.01.016.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature, 408*(6814), 788. https://doi.org/10.1038/35048669.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*(1), 147–152. https://doi.org/10.1016/S0926-6410(02)00069-1.

Shore, D. I., Barnes, M. E., & Spence, C. (2006). Temporal aspects of the visuotactile congruency effect. *Neuroscience Letters, 392*(1–2), 96–100. https://doi.org/10.1016/j.neulet.2005.09.001.

Smit, S., Rich, A. N., & Zopf, R. (2019). Visual body form and orientation cues do not modulate visuo-tactile temporal integration. *PLoS ONE, 14*(12), e0224174. https://doi.org/10.1371/journal.pone.0224174.

Spence, C. (2019). Evaluating the spatial rule of multisensory integration: When exactly does spatial coincidence matter? In T. Cheng, O. Deroy, & C. Spence (Eds.). *Spatial senses: Philosophy of perception in an age of science* (pp. 284–306). New York, NY: Routledge.

Spence, C., Pavani, F., & Driver, J. (2004). Spatial constraints on visual-tactile cross-modal distractor congruency effects. *Cognitive, Affective, & Behavioral Neuroscience, 4*(2), 148–169. https://doi.org/10.3758/CABN.4.2.148.

Spence, C., Pavani, F., Maravita, A., & Holmes, N. P. (2008). Multi-sensory interactions. In M. C. Lin, & M. A. Otaduy (Eds.). *Haptic rendering: Foundations, algorithms, and applications* (pp. 21–52). Wellesley, MA: AK Peters.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General, 130*(4), 799. https://doi.org/10.1037/0096-3445.130.4.799.

Spence, C., & Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Current Biology, 13*(13), R519–R521. https://doi.org/10.1016/S0960-9822(03)00445-7.

Stein, B. E., & Meredith, M. A. (1990). Multisensory integration: Neural and behavioral solutions for dealing with stimuli from different sensory modalities. *Annals of the New York Academy of Sciences, 608*(1), 51–70. https://doi.org/10.1111/j.1749-6632.1990.tb48891.x.

Sui, J., & Rotshtein, P. (2019). Self-prioritization and the attentional systems. *Current Opinion in Psychology*. https://doi.org/10.1016/j.copsyc.2019.02.010.

Taylor-Clarke, M., Kennett, S., & Haggard, P. (2002). Vision modulates somatosensory cortical processing. *Current Biology, 12*(3), 233–236. https://doi.org/10.1016/S0960-9822(01)00681-9.

Team, R. (2015). RStudio: integrated development for R. RStudio, Inc., Boston, MA URL http://www. rstudio.com, 42, 14.

Team, R. C. (2017). R: A language and environment for statistical computing. R Found. Stat. Comput. Vienna, Austria. URL http://www. R-project. org/., page R Foundation for Statistical Computing.

Thomas, R., Press, C., & Haggard, P. (2006). Shared representations in body perception. *Acta Psychologica, 121*(3), 317–330. https://doi.org/10.1016/j.actpsy.2005.08.002.

Tipper, S. P., Lloyd, D., Shorland, B., Dancer, C., Howard, L. A., & McGlone, F. (1998). Vision influences tactile perception without proprioceptive orienting. *Neuroreport, 9*(8), 1741–1744. https://journals.lww.com/neuroreport/fulltext/1998/06010/vision_influences_tactile_perception_without.13.aspx.

Townsend, J. T., & Ashby, F. G. (2014). Methods of modeling capacity in simple processing systems. *Cognitive Theory* (pp. 211–252). Psychology Press. https://doi.org/10.1016/j.neuropsychologia.2009.09.034.

Tsakiris, M., Carpenter, L., James, D., & Fotopoulou, A. (2010). Hands only illusion: Multisensory integration elicits sense of ownership for body parts but not for non-corporeal objects. *Experimental Brain Research, 204*(3), 343–352. https://doi.org/10.1007/s00221-009-2039-3.

Tsakiris, M., Haggard, P., Franck, N., Mainy, N., & Sirigu, A. (2005). A specific role for efferent information in self-recognition. *Cognition, 96*(3), 215–231. https://doi.org/10.1016/j.cognition.2004.08.002.

Tsakiris, M., Hesse, M. D., Boy, C., Haggard, P., & Fink, G. R. (2007). Neural signatures of body ownership: A sensory network for bodily self-consciousness. *Cerebral Cortex, 17*(10), 2235–2244. https://doi.org/10.1093/cercor/bhl131.

Van Den Bos, E., & Jeannerod, M. (2002). Sense of body and sense of action both contribute to self-recognition. *Cognition, 85*(2), 177–187. https://doi.org/10.1016/S0010-0277(02)00100-2.

Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W., & Schirillo, J. A. (2004). Unifying multisensory signals across time and space. *Experimental Brain Research, 158*(2), 252–258. https://doi.org/10.1007/s00221-004-1899-9.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*(3), 638. https://doi.org/10.1037/0033-2909.88.3.638.

Whiteley, L., Spence, C., & Haggard, P. (2008). Visual processing and the bodily self. *Acta Psychologica, 127*(1), 129–136. https://doi.org/10.1016/j.actpsy.2007.03.005.

Zopf, R., Savage, G., & Williams, M. A. (2010). Crossmodal congruency measures of lateral distance effects on the rubber hand illusion. *Neuropsychologia, 48*(3), 713–725. https://doi.org/10.1016/j.neuropsychologia.2009.10.028.