Journal of Experimental Child Psychology 207 (2021) 105094

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

The development of visuotactile congruency effects for sequences of events

Alan O'Dowd [a], Sarah M. Cooney [a,b], Francesca Sorgini [a,c], Eugene O' Rourke [d], Richard B. Reilly [d], Fiona N. Newell [a], Rebecca J. Hirst [a,*]

[a] *School of Psychology and Institute of Neuroscience, Trinity College Dublin, Dublin 2, Ireland*
[b] *School of Psychology, University College Dublin, Dublin 4, Ireland*
[c] *BioRobotics Institute, Scuola Superiore Sant'Anna, 56127 Pisa, PI, Italy*
[d] *Centre for BioEngineering, Institute of Neuroscience, and School of Engineering, Trinity College Dublin, Dublin 2, Ireland*

ARTICLE INFO

Article history:
Received 27 May 2020
Revised 11 December 2020
Available online 11 March 2021

Keywords:
Crossmodal congruency task
Development
Multisensory
Sequences of events
Visuo-tactile interactions
Visual context

ABSTRACT

Sensitivity to the temporal coherence of visual and tactile signals increases perceptual reliability and is evident during infancy. However, it is not clear how, or whether, bidirectional visuotactile interactions change across childhood. Furthermore, no study has explored whether viewing a body modulates how children perceive visuotactile sequences of events. Here, children aged 5–7 years (n = 19), 8 and 9 years (n = 21), and 10–12 years (n = 24) and adults (n = 20) discriminated the number of target events (one or two) in a task-relevant modality (touch or vision) and ignored distractors (one or two) in the opposing modality. While participants performed the task, an image of either a hand or an object was presented. Children aged 5–7 years and 8 and 9 years showed larger crossmodal interference from visual distractors when discriminating tactile targets than the converse. Across age groups, this was strongest when two visual distractors were presented with one tactile target, implying a "fission-like" crossmodal effect (perceiving one event as two events). There was no influence of visual context (viewing a hand or non-hand image) on visuotactile interactions for any age group. Our results suggest robust interference from discontinuous visual information on tactile discrimination of sequences of events during early and middle childhood. These findings are discussed with respect to age-related

---

* Corresponding author.
E-mail address: hirstr@tcd.ie (R.J. Hirst).

https://doi.org/10.1016/j.jecp.2021.105094

changes in sensory dominance, selective attention, and multisensory processing.

Introduction

Multisensory interactions and sensory weighting are known to change across childhood (Hirst, Cragg, & Allen, 2018; Nava & Pavani, 2013), a period when the brain undergoes significant structural and functional maturation (Gogtay et al., 2004; Lebel, Walker, Leemans, Phillips, & Beaulieu, 2008). To date, evidence from behavioral work suggests that children younger than 8 years do not exhibit adult-like multisensory processing, for example, when making audiovisual judgments (Adams, 2016), visuohaptic judgments (Gori, Del Viva, Sandini, & Burr, 2008), and visuomotor judgments (Nardini, Jones, Bedford, & Braddick, 2008) (see Burr & Gori, 2012, and Ernst, 2008, for reviews). However, findings by Negen et al. (2019) suggested that providing feedback on the reliability and biases of visual and auditory estimates can facilitate audiovisual integration in 7- to 10-year-olds. When provided with multisensory information, younger children typically rely on one input or the other to guide behavior as opposed to optimally integrating the sensory inputs as seen in older children and adults (Burr & Gori, 2012; Ernst, 2008; Ernst & Banks, 2002). The importance of measuring multisensory processing during childhood was recently highlighted by Denervaud, Gentaz, Matusz, and Murray (2020), who found that multisensory gains on an audiovisual detection task significantly predicted higher-order cognitive abilities in schoolchildren (see also Dionne-Dostie, Paquette, Lassonde, & Gallagher, 2015).

Crossmodal interactions relating to the temporal synchronicity of visual and tactile signals on the body (Filippetti, Johnson, Lloyd-Fox, Dragovic, & Farroni, 2013) and visual and haptic object information (Sann & Streri, 2007, 2008) are believed to emerge early in infancy. Furthermore, additional sensory cues are used to aid tactile judgments during early childhood. For example, the use of external reference frames to localize tactile signals appears to be evident around 5 years of age (Pagel, Heed, & Röder, 2009). However, sensitivity to visuotactile synchronicity (Greenfield, Ropar, Themelis, Ratcliffe, & Newport, 2017) and simultaneity (Chen, Lewis, Shore, Spence, & Maurer, 2018) appear to reach adult-like maturity during later childhood. In the Rubber Hand Illusion (RHI; see Botvinick & Cohen, 1998), children under 10 years of age experience a more pronounced dominance of vision over proprioception compared with children older than 10 years and adults (Cowie, Makin, & Bremner, 2013; Cowie, Sterling, & Bremner, 2016). These findings may indicate that the weighting of visual information relative to tactile and proprioceptive information shifts across childhood and into adulthood, perhaps depending on modality combination and task demands (e.g. see Gori et al., 2008). Temporal multisensory interactions may arise early in childhood, with the ability to identify visuotactile correspondence for sequential events being evident by 5 months of age (Féron, Gentaz, & Streri, 2006). In adults, touch appears to dominate vision in the discrimination of sequences of events (Bresciani, Dammeier, & Ernst, 2006; Philippi, van Erp, & Werkhoven, 2008). To shed further light on this, we aimed to investigate how vision and touch interact in children and adults when judging the number of sequential events presented (one or two).

The way in which we use multisensory signals can be influenced by contextual factors (Chen & Spence, 2017; Welch & Warren, 1980). For example, the body may constitute a unique context that reinforces interactions between visual and tactile information. Presenting a flash of light on the body at the same time as a tactile sensation can impair tactile detection (Hartcher-O'Brien, Levitan, & Spence, 2010) and interfere with the spatial discrimination of tactile sensations (Igarashi, Kitagawa, Spence, & Ichihara, 2007; Salomon, Van Elk, Aspell, & Blanke, 2012). In the absence of tactile input, a flash of light presented on the body can create an illusion of a corresponding tactile sensation (Johnson, Burton, & Ro, 2006; Mirams, Poliakoff, Brown, & Lloyd, 2010). This may be underpinned by multisensory regions in the frontoparietal cortex that map the body surface and the space within reach of the body (Makin, Holmes, & Zohary, 2007; Rizzolatti, Fadiga, Fogassi, & Gallese, 1997).

Although it is known that viewing a body can augment visuotactile interactions for spatial judgments (Igarashi et al., 2007; Salomon et al., 2012), there is conflicting evidence as to whether this generalizes to visuotactile temporal processing (Ide & Hidaka, 2013; Smit, Rich, & Zopf, 2019). O'Dowd, Sorgini, and Newell (2020) asked young adults to judge the number of sequential tactile targets delivered to the hand (one or two) while ignoring either a congruent or incongruent number of visual distractors (nonspatial crossmodal congruency task [CCT]; Holmes, Sanabria, Calvert, & Spence, 2006; Poole, Couth, Gowen, Warren, & Poliakoff, 2015). Visual stimuli were superimposed onto an image of either a hand or an object. Tactile discrimination performance was improved when viewing an image of a hand compared with an object, consistent with previous evidence for a body-based "visual enhancement of touch" (VET; Kennett, Taylor-Clarke, & Haggard, 2001). For crossmodal temporal interactions, two kinds of effect can be considered: "fusion-like" effects, perceiving one event from a target modality when two events were presented along with one event from another different modality (e.g., Andersen, Tiippana, & Sams, 2004) and "fission-like" effects, perceiving two events when one event from a target modality is presented with two events from a different modality (e.g., Shams, Kamitani, & Shimojo, 2000, 2002). Both fission and fusion have been observed in the visuotactile domain (Bresciani et al., 2006; Chen, Maurer, Lewis, Spence, & Shore, 2017; Violentyev, Shimojo, & Shams, 2005). O'Dowd et al. (2020) observed lower rates of fusion interference compared with fission interference when the event stimuli were presented on a viewed hand than on an object.

An interesting question is at what age during childhood interactions between body representation and visuotactile judgments emerge. It has been argued that the effect of context (e.g. viewing a body part) on multisensory processing is experience dependent; therefore, it might be expected that such effects would strengthen with age (Chen & Spence, 2017; Murray, Lewkowicz, Amedi, & Wallace, 2016); however, to our knowledge, this has yet to be empirically tested. Therefore, we compared visuotactile interactions in children and adults when asked to judge the number of sequentially presented events and also investigated whether context influenced these interactions.

An additional question is how visuotactile fusion-like and fission-like[1] phenomena manifest across childhood. Fission, in particular, has been forwarded as a robust measure of crossmodal integration that can shed light on the broader characteristics of multisensory perception (Hirst, McGovern, Setti, Shams, & Newell, 2020; Keil, 2020). However, in the sound-induced flash illusion, fusion effects have been less frequently reported relative to fission effects and therefore have been relatively less studied, and it remains unclear to what extent fusion and fission reflect similar phenomena. For example, in the visuotactile domain, fission and fusion peak with peripheral and foveal presentation, respectively, possibly indicating differing neural mechanisms (Chen et al., 2017). Although rates of audiovisual fusion and fission have been shown to differentially change across the lifespan (during development: Innes-Brown et al., 2011; during aging: McGovern, Roudaia, Stapleton, McGinnity, & Newell, 2014), it is currently not known whether the asymmetrical influence of visual context on rates of tactile fission-like and fusion-like effects reported by O'Dowd et al. (2020) is a product of early developmental processes or stems from extended perceptual experience. By addressing this ambiguity, we hoped to expand on O'Dowd et al.'s (2020) work and explore the role of visual context in specific temporal patterns of crossmodal interactions for vision and touch.

In the current study, we used a nonspatial CCT to investigate how visual distractors, superimposed onto images of a hand or an object, influence the discrimination of tactile sequences of events on the fingertip (as in O'Dowd et al., 2020). Extending this paradigm, we also examined how tactile distractors influenced the discrimination of visual sequences of events across the same visual contexts. The aim of this study, therefore, was to assess bidirectional visuotactile interactions during the discrimination of sequences of events and how this is modulated by visual context and to compare these effects across childhood age groups.

---

[1] The terms "fusion-like" and "fission-like" refer to instances where visuotactile events feature longer temporal durations than those in studies of illusory "fusion" and "fission" effects (e.g., Shams, Kamitani, & Shimojo, 2000, 2002).

Our hypotheses were as follows:

1. Lower accuracy would occur when the number of distractors conflicts with the number of targets (incongruent trials) relative to when the number of distractors is compatible with the number of targets (congruent trials) (i.e., a significant crossmodal congruency effect).
   If viewing a hand enhances the tactile discrimination of events (as in O'Dowd et al., 2020), then we expected the following:
2. The effect of visual distractors would be reduced, and the effect of tactile distractors would be enhanced, when viewing an image of a hand compared with an image of an object (i.e., a significant target modality by stimulus context interaction).
3. Rates of vision-induced tactile fusion would be lower than fission effects when viewing an image of a hand compared with an image of an object (as in O'Dowd et al., 2020) (i.e., a significant interference type by stimulus context interaction).

In addition to these hypotheses, we aimed to characterize how, or whether, these effects change with development.

Method

Participants

In total, 82 children (mean age = 8.61 years, range = 4.26–12.10; 49 boys) and 20 adults (mean age = 21.6 years, range = 18–47; 8 men) took part in this study. Adults were recruited from the student population of Trinity College Dublin, Ireland, and received course credit for their time. Children were voluntary attendants of a public engagement event titled Summer Scientist Week held at the University of Nottingham, UK (www.summerscientist.org). The number of child participants in this study was restricted by the number of children attending the event. However, to gauge optimal sample size per age group, an a priori power analysis was conducted using PANGEA (Westerfall, 2016). This analysis indicated that $n = 31$ per age group would provide 95% power to detect a medium effect size for a four-way Age × Congruency × Object × Modality mixed analysis of variance (ANOVA) design. Previous related research by O'Dowd et al. (2020, Experiment 2) observed large effect sizes for the effects of both congruency ($\eta^2_G = .30$) and stimulus context ($\eta^2_G = .27$). The number of children attending the public engagement event who partook in our experiment did not meet the initial sampling requirements set out by the power analysis. To assess whether further data collection was required, we calculated Bayes factors to assess whether null findings evidenced genuine support for the null or whether they could be attributed to reduced statistical sensitivity (as recommended by Dienes, 2014). We also performed a sensitivity analysis to determine the minimally detectable effect sizes for each interaction term with the smallest group size recruited ($n = 19$). The results indicated that we should be able to reliably detect medium effect sizes for these terms and are illustrated in an interactive plot available at https://chart-studio.plotly.com/~PsychPlots/1/#/.

All participants were fluent English speakers (children's parents/guardians confirmed that English was spoken at home). All children under 5 years of age ($n = 6$; 4 boys) exhibited difficulty in understanding and completing the task; therefore, we focused our analysis on children aged 5 years and over ($n = 76$).

As part of the wider public engagement event, all parents/guardians of participating children were first asked to complete a number of questionnaires, including the Autism Spectrum Quotient (AQ; Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005), Strength and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN; Swanson et al., 2001), and questions on sensory processing (visual, tactile, vestibular, and proprioception). These questionnaires and assessments were administered independent of our study (i.e., these were not study-specific screening tools for our investigation). Nevertheless, these measures enabled exploratory analyses for the final sample of children (see online supplementary material) to examine whether these factors related to performance on our task given evidence of atypical tactile sensitivity in autism spectrum disorder (ASD) and attention-deficit/hyperactivity disorder (ADHD) (e.g., see Ide, Yaguchi, Sano, Fukatsu, & Wada, 2019, and Puts et al., 2017; see also Ropar, Greenfield, Smith, Carey, & Newport, 2018).

Data from 12 additional children were excluded as follows: *n* = 5 (2 boys) had a formal diagnosis of a developmental disorder (ASD) and/or ADHD based on parent/guardian report, *n* = 6 (4 boys) completed less than 20% of the CCT, and *n* = 1 boy performed at 50% accuracy for judging unimodal stimuli. This yielded a final sample of 64 child participants. Table 1 shows the demographic information of all child and adult participants (total *N* = 84). Ethical approval was obtained from the research ethics committees of Trinity College Dublin and the University of Nottingham. The study was carried out in compliance with the Declaration of Helsinki, and data were collected in line with European General Data Protection Regulations. Written consent was provided by all adult participants and by the parents or guardians of each child participant attending the public engagement event.

Materials and apparatus

Fig. 1 illustrates the task setup and stimuli used. All stimuli were presented on an Alienware laptop (Alienware 15 R4, 64-bit, Intel Core i5-8300 at 2.30 GHz, Nvidia GeForce GTX 1060 graphics) and were programmed in PsychoPy Version 3.6.6 (Peirce et al., 2019). The laptop was positioned perpendicular to participants, and the screen (34.5 × 19.4 cm, 1920 × 1080 resolution, 60-Hz refresh rate) was folded backward, thereby obscuring the view of the hand receiving tactile stimulation. Participants viewed the screen from above at approximately a 40-cm distance. Participants placed their right hand beneath the laptop screen and rested their index finger on the coin motor (10 × 3 mm, 12,000 rpm). All vibrations were presented suprathreshold. The coin motor was connected to a custom-built electrical circuit and was connected to the laptop via USB. The circuit was occluded from view and was controlled by a combination of Arduino Version 1.8.9 and Python (via PsychoPy). Participants placed the tip of their index finger on top of the coin motor to receive the tactile stimulus. All participants wore sound-attenuating headphones, and white noise was played for the duration of the experiment (at a level of 40 dB) to mask the sound of the vibrations. A mirror was used to allow the experimenter to ensure that participants kept their finger on the coin motor throughout the task.

On the screen, a fixation dot was presented midway along the horizontal axis and approximately 6 cm below the vertical center (from the observer's perspective). Throughout each block, an image of either a right hand or an object (arrow) remained on the screen to the right of the fixation point (as in O'Dowd et al., 2020). These images were superimposed on a plain black background, with the tip of the finger/arrow positioned approximately 2 cm to the right of the fixation point. The hand was presented from a first-person view, palm side down with the index finger pointing away and the remaining digits held beneath the palm. The object image of an arrow, approximately matched in color to the color of the hand, was pointed in the same direction as the index finger of the hand (Fig. 1B). Thus, both stimuli contained directional cues.

The visual target/distractor was a yellow disc (1 cm in diameter) that appeared on the tip of the index finger of the hand or on the tip of the arrow (Fig. 1B). The coin motor was positioned beneath the laptop and was spatially aligned with the position of the visual target. Participants' physical right hand was positioned directly underneath the part of the screen on which the stimuli were displayed, and the hand and wrist areas were occluded by an opaque black cloth throughout the experiment.

Fig. 2 illustrates an example time course of incongruent and congruent trials of the CCT. On a trial with two events (e.g., two flashes, two vibrations), each event lasted 100 ms. The duration of the gap between two events was adjusted based on each participant's individual gap detection threshold and was presented one frame (16.66 ms) above threshold. In instances where visual and tactile thresholds were not identical, the gap between visual and tactile events was presented 16.66 ms above the larger of the two thresholds. Visual and tactile gap detection thresholds were quantified with a method of adjustment (described in the "Procedure" section below). On trials with one event, the event lasted 200 ms. Double and single vibration trials followed the same time course as double and single visual trials, respectively. On congruent trials, visual and tactile events occurred synchronously. On incongruent trials with two events in one modality and one event in the other modality, the single event occurred directly in between the two events (e.g., one flash appeared in between two vibrations; see Fig. 2). The timing of the visual event and the timing of the tactile event were verified with a pho-

Table 1
Demographic information of final selected sample ($N = 84$) following exclusion criteria.

Age group	*N*	Age in years [*M* (*SD*)]	Age range	Number of male participants
5–7 years	19	6.7 (0.88)	5.1–7.9	10
8–9 years	21	8.9 (0.63)	8.1–9.9	9
10–12 years	24	11.1 (0.63)	10.1–12.1	16
Adult	20	21.6 (6.52)	18–47	8

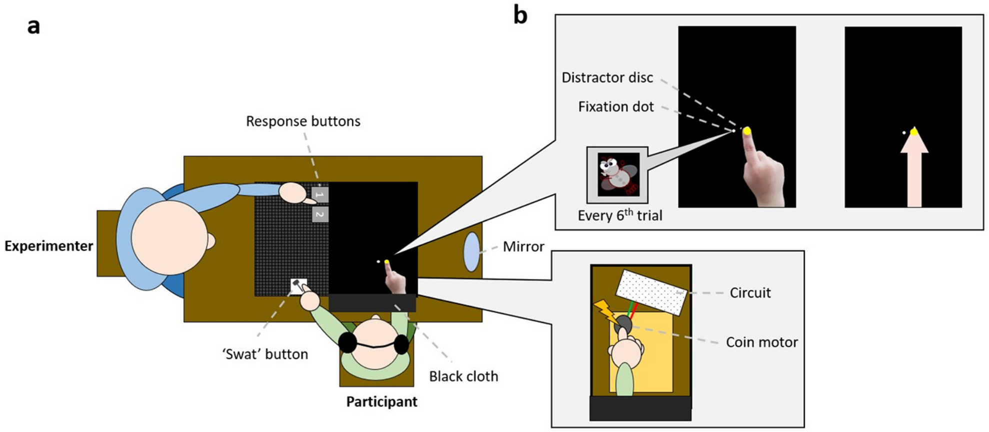

Fig. 1. Task setup. (A) Participants placed their right hand beneath the laptop screen and rested their index finger on the coin motor. (B) They fixated on a fixation dot throughout the experiment and were presented with either a hand or an arrow context on which a yellow distractor disc flashed. Vocal responses were recorded by the experimenter, who pressed the response buttons. Participants pressed a "swat" button with their left hand whenever a cartoon insect appeared on the fixation dot alongside a concurrent vibration to the right index finger.

todiode and oscilloscope, respectively. The event stimuli used in the threshold detection task were identical to those described in the main experiment (although no hand or arrow images were presented).

All child participants were tested in a brightly lit quiet room alongside other studies at the University of Nottingham (with these studies separated by partitions to prevent visual distraction). Adult participants were tested in a separate venue, specifically in a brightly lit dedicated laboratory for behavioral testing at the Trinity College Institute of Neuroscience.

Procedure

Participants completed two tasks: a method of adjustment gap detection task and the CCT. Testing time with each participant was limited to approximately 15 min.

Gap detection task

The gap detection task followed a simple method of adjustment procedure (similar to Hirst, Kicks, Allen, & Cragg, 2019). Participants were asked to determine the point at which they detected two stimuli (either the visual flash or the vibration stimuli described). Participants completed both an ascending and descending method of adjustment task, the order of which was counterbalanced across participants. In the ascending method, participants started with the smallest temporal gap between two tactile/visual events (an interstimulus interval [ISI] of 16.66 ms). The ISI was progressively increased until participants identified two events (ascending method). In the descending method, participants started with a large suprathreshold ISI between two events that was progressively decreased

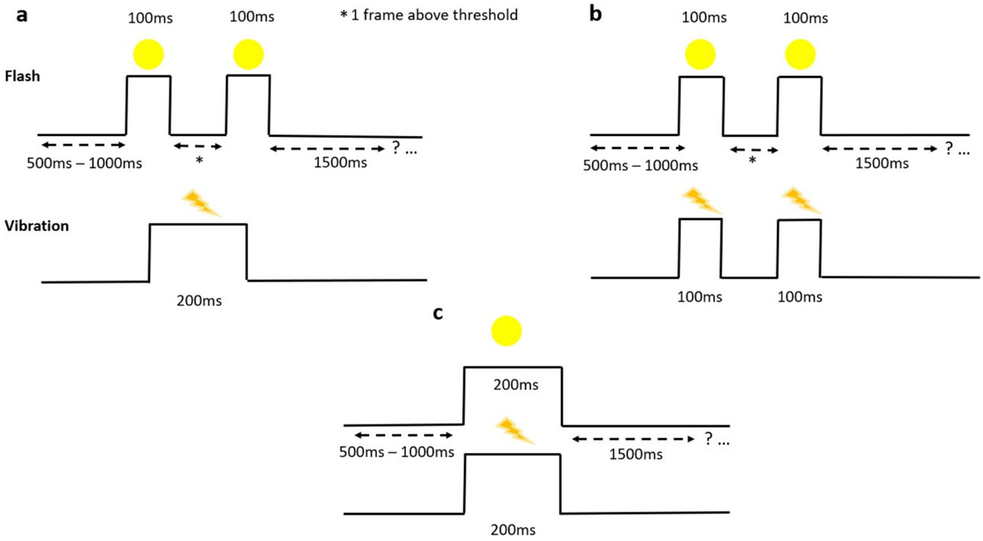

Fig. 2. Time course of stimulus presentation during a typical incongruent visuotactile trial with two visual events and one tactile vibration (A), congruent trial with two flashes and two tactile vibrations (B), and congruent trial with one flash and one vibration (C). See main text for further details. Note that the durations of the stimulus events are longer than those of Shams et al. (2000, 2002); therefore, these effects are referred to as "fusion-like" and "fission-like" effects.

until they could no longer detect the occurrence of two events. Here, the size of the gap was set at 200 ms for children aged 5–7 years, 167 ms for children aged 8–12 years, and 117 ms for adults (Adams, 2016). The largest gap was detectable for all participants. On each trial, participants provided a verbal response as to whether they detected one or two stimuli, enabling the experimenter to adjust the gap appropriately. The order of protocol (ascending or descending first and visual or tactile first) was counterbalanced across participants. Each final threshold (i.e., visual gap and tactile gap) was taken as the average of the ascending and descending trials.

Crossmodal congruency task

Participants were asked to determine either the number of visual flashes or the number of vibrations presented while ignoring the irrelevant modality (see Table 2). The task followed a 2 (Target Modality: vision or touch) × 3 (Congruency: congruent, incongruent, or unimodal) × 2 (Stimulus Context: hand or object) design, resulting in 24 unique trial combinations. Target modality and stimulus context were blocked, with stimulus context nested within the target modality. Congruency was randomized within each block. There were 96 trials in total for the entire congruency task.

An image of a hand or an arrow and the fixation dot were presented on the screen throughout the task (depending on stimulus context block). Participants fixated for 500–1000 ms before the visuotactile (or unimodal) stimuli were presented. Participants were instructed to focus on the fixation point and to verbally report the number of targets (one or two) felt or seen (depending on the target modality). Participants were given 1500 ms to respond before a question mark appeared, signaling the need for a response. A response was required before the next trial commenced. Participants were encouraged to respond as quickly and accurately as possible. Responses were recorded by the experimenter, who pressed the corresponding button on the keyboard. All participants were explicitly instructed to ignore the distracting information in the opposite modality and to keep their finger on the coin motor at all times. A "checkpoint" break was offered every 12 trials, where participants selected a number on the screen (1 or 2) to receive a clue as to where their "token" was hidden in the room (at the public engagement event, children gain tokens through research participation). For consistency, adults were

Table 2
Trial types presented as congruency and relevant modality.

Congruency	Target modality	Number of tactile events	Number of visual events
Unimodal	Touch	1	0
		2	0
	Vision	0	1
		0	2
Congruent	Touch	1	1
		2	2
	Vision	1	1
		2	2
Incongruent	Touch	1	2
		2	1
	Vision	1	2
		2	1

Note. These trials were presented four times each per stimulus context (hand and object), eight times total, resulting in 96 trials altogether (two 48-trial blocks). Each block was preceded by a practice block containing 6 trials per condition with no background stimulus.

given the same clues and narrative context. A reminder of the task instructions was provided prior to commencing each block.

To ensure that fixation was maintained on the fixation dot, a cartoon insect appeared over the fixation dot, accompanied by a persistent vibration on the index finger, at a time point that was halfway between each checkpoint. Participants were instructed to press a "swat" button on the keyboard as quickly as possible when this happened with their left hand. The trials would not continue until participants pressed this button. Participants were monitored throughout by the experimenter to ensure that fixation and contact with the coin motor were maintained.

Data analysis

The primary focus of our study was to examine the effect of age on visuotactile congruency. Therefore, the analysis focused on congruency within this section. To assess the extent of crossmodal congruency effects (CCEs) between groups, while controlling for baseline performance, our dependent variable was the difference in percentage accuracy between bimodal and unimodal trials; for example, when the task was to judge touch, CCEs would be calculated by comparing accuracy in the following conditions:

For congruent trials:

$$(1 \text{ vibration} + 1 \text{ flash}) - (1 \text{ vibration} + 0 \text{ flashes}) \tag{1}$$

or

$$(2 \text{ vibrations} + 2 \text{ flashes}) - (2 \text{ vibrations} + 0 \text{ flashes}) \tag{1}$$

For incongruent trials:

$$(1 \text{ vibration} + 2 \text{ flashes}) - (1 \text{ vibration} + 0 \text{ flashes}) \tag{2}$$

or

$$(2 \text{ vibrations} + 1 \text{ flash}) - (2 \text{ vibrations} + 0 \text{ flashes}) \tag{2}$$

Thus, difference scores were calculated using trials with a comparable number of events. Unimodal trials, therefore, served as the baseline to which participants' performance was corrected. This correction was performed separately for congruent and incongruent conditions, allowing us to examine facilitation effects (i.e., higher accuracy for congruent visuotactile vs. unimodal events) (Barutchu, Crewther, & Crewther, 2009) and interference effects (i.e., lower accuracy for incongruent vs. unimodal events). Difference scores were analyzed using a mixed-factors 2 (Target Modality) × 2 (Congruency) × 2 (Stimulus Context) × 4 (Age Group) ANOVA. We interpreted negative difference scores as indicating interference effects and positive difference scores as indicating facilitation effects.

Finally, we examined rates of fission-like and fusion-like effects. A fission effect occurred when participants responded "two" when a single stimulus was presented in the target modality along with two stimuli in the irrelevant modality. Conversely, a fusion effect occurred when participants responded "one" to the presence of two stimuli in the relevant modality with one stimulus in the irrelevant modality. These effects were compared across target modalities, stimulus contexts, and age groups in a separate secondary analysis to explore the specific influence of different incongruent combinations on accuracy performance (as in a visuotactile study by Holmes et al., 2006, and an audiovisual study by Parker & Robinson, 2018).

All analyses were conducted using R (R Core Team, 2017) via RStudio Version 3.5.0 (R Core Team, 2015). The *ez* package (Lawrence, 2016) was used for performing ANOVAs with Type 3 sum of squares to test for main effects and interactions. All assumptions of ANOVA were satisfied. Significance values are reported with Bonferroni correction in cases of multiple, post hoc comparisons. To ensure that null results were not due to statistical tests lacking sufficient sensitivity, we also performed Bayesian ANOVAs and *t* tests and report Bayes factors (BF_{01}) to provide an indication of the support for the null hypotheses (Dienes, 2014). In line with Jeffreys (1961), a cutoff of $BF_{01} > 3$ was taken as indication of robust support for the null hypothesis. Bayes factors were determined using the *BayesFactor* package with default prior settings (Morey, Rouder, & Jamil, 2018; Rouder, Morey, Speckman, & Province, 2012; Rouder, Speckman, Sun, Morey, & Iverson, 2009). The study was preregistered on the Open Science Framework (https://osf.io/e7h6s).

Results

Gap detection thresholds

The gap detection thresholds for vision and touch of the participants are displayed in Fig. 3. We hypothesized that visual and tactile gap detection thresholds would narrow with development. Threshold scores were first corrected for violations of normality using a Tukey power of ladders transformation. Corrected tactile and visual gap detection thresholds were entered into a mixed-factors 2 (Modality) × 4 (Age Group) ANOVA. The analysis yielded significant main effects of modality, $F(1, 80) = 62.41$, $p < .001$, $\eta^2_G = .21$, and age group, $F(3, 80) = 6.67$, $p < .001$, $\eta^2_G = .13$, and a significant interaction between modality and age group, $F(3, 80) = 15.11$, $p < .001$, $\eta^2_G = .16$ (Fig. 3). Post hoc pairwise *t* tests (Bonferroni corrected) confirmed that the adult group had significantly smaller tactile gap detection thresholds ($M = 21.71$ ms, $SD = 11.05$) than the 5- to 7-year-olds ($M = 57.26$ ms, $SD = 37.15$; $p < .001$), 8- and 9-year-olds ($M = 47.72$ ms, $SD = 19.14$; $p < .001$), and 10- to 12-year-olds ($M = 41.79$ ms, $SD = 23.20$; $p = .001$). There were no significant differences among the tactile gap detection thresholds of the 5- to 7-year-olds, 8- and 9-year-olds, and 10- to 12-year-olds (all *ps* > .05) or between the visual gap detection thresholds of any group (all *ps* > .05).

Crossmodal congruency task

Descriptive statistics for accuracy difference scores are illustrated in Table 3.

A mixed-factors 2 (Target Modality) × 2 (Congruency) × 2 (Stimulus Context) × 4 (Age Group) ANOVA on accuracy difference scores (congruent–unimodal and incongruent–unimodal; see Equations 1 and 2 in Method) revealed a main effect of congruency, $F(1, 80) = 301.17$, $p < .001$, $\eta^2_G = .33$, with improved accuracy for congruent trials ($M = 2.60\%$, $SD = 13.68$), which was not observed for incongruent trials ($M = -22.36\%$, $SD = 23.08$). A comparison of absolute accuracy difference scores for congruent and incongruent trials (+22.36% vs. +2.60%) via a paired *t* test also confirmed a significant congruency effect ($p < .001$). The effects of target modality ($p = .14$; $BF_{01} = 4.17$), stimulus context ($p = .53$; $BF_{01} = 11.1$), and age group ($p = .29$; $BF_{01} = 33.3$) failed to reach significance. The only two-way interactions to reach significance were between target modality and age group, $F(3, 80) = 3.30$, $p = .02$, $\eta^2_G = .02$,

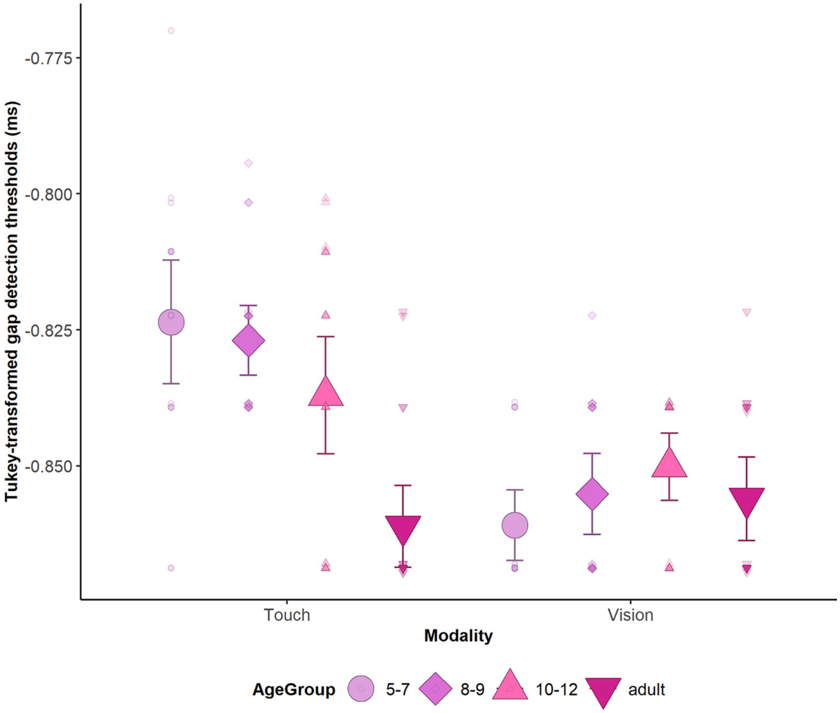

Fig. 3. Plot showing the Tukey-transformed gap detection threshold per age group and across modalities (touch and vision). Error bars show 95% confidence intervals. Smaller points show individual participant thresholds.

Table 3
Means (and standard deviations) for accuracy difference scores (%) per target modality (vision or touch), congruency (congruent or incongruent), stimulus context (hand or object), and age group (5–7 years, 8–9 years, 10–12 years, or adult) of the crossmodal congruency task.

	Vision				Touch			
	5–7 years	8–9 years	10–12 years	Adult	5–7 years	8–9 years	10–12 years	Adult
Overall								
Congruent	2.30 (18.58)	2.98 (10.98)	−1.30 (13.46)	0.00 (6.33)	7.24 (15.55)	3.87 (12.80)	5.73 (17.67)	0.31 (7.21)
Incongruent	−14.47 (20.24)	−14.88 (17.07)	−16.67 (17.36)	−26.88 (22.57)	−27.30 (29.62)	−37.50 (25.75)	−20.57 (23.42)	−21.56 (18.99)
Hand								
Congruent	−0.66 (21.44)	2.98 (12.44)	0.52 (13.53)	−1.25 (6.91)	3.95 (12.54)	4.76 (12.17)	6.25 (16.89)	−0.63 (7.56)
Incongruent	−13.16 (22.62)	−19.05 (17.06)	−18.23 (16.06)	−29.38 (22.68)	−26.97 (26.11)	−34.52 (29.02)	−21.88 (22.19)	−18.75 (18.36)
Object								
Congruent	5.26 (15.20)	2.98 (9.61)	−3.13 (13.42)	1.25 (5.59)	10.53 (17.81)	2.98 (13.64)	5.21 (18.77)	1.25 (6.91)
Incongruent	−15.79 (18.09)	−10.71 (16.43)	−15.10 (18.79)	−24.38 (22.75)	−13.16 (22.62)	−40.48 (22.33)	−19.27 (24.99)	−24.38 (19.65)

and between target modality and congruency, $F(1, 80) = 14.57$, $p < .001$, $\eta^2_G = .03$. The only three-way interaction to reach significance was among target modality, congruency, and age group, $F(3, 80) = 3.86$, $p = .01$, $\eta^2_G = .02$, as shown in Fig. 4.

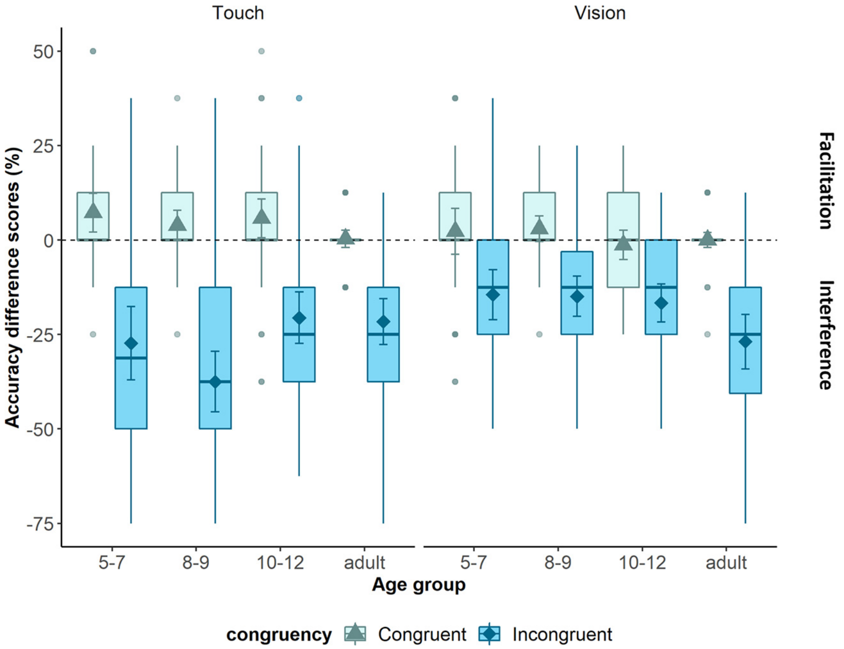

Fig. 4. Plot showing the interaction among target modality, congruency, and age group for accuracy difference scores (%). Positive difference scores show bimodal facilitation (higher accuracy on bimodal trials than on unimodal trials). Negative difference scores show bimodal interference (lower accuracy on bimodal trials than on unimodal trials). Mean scores are illustrated with large triangle (congruent) and diamond (incongruent) points, and error bars represent 95% confidence intervals for the means. Small circles show outliers. The dashed horizontal line indicates cases where accuracy on unimodal trials was equal to accuracy on bimodal trials.

Post hoc tests were conducted to explore the three-way interaction among target modality, congruency, and age group. The data were split by age group, and two within-participants 2 (Target Modality) × 2 (Congruency) ANOVAs were performed. This analysis showed a significant interaction between target modality and congruency for the 5- to 7-year-olds, $F(1, 18) = 5.87$, $p = .03$, $\eta^2_G = .06$, and 8- and 9-year-olds, $F(1, 20) = 20.09$, $p < .001$, $\eta^2_G = .15$, but not for the 10- to 12-year-olds ($p = .09$; $BF_{01} = 1.01$) or adults ($p = .40$; $BF_{01} = 3.23$). Post hoc pairwise t tests on the 5- to 7-year-olds' accuracy confirmed a significant CCE that was smaller for visual targets ($M_{diff} = 16.25\%$, $SD = 16.19$; $p < .001$) than for tactile targets ($M_{diff} = 34.54\%$, $SD = 27.45$; $p < .001$). Post hoc pairwise t tests on the 8- and 9-year-olds' accuracy confirmed a significant CCE that was also smaller for visual targets ($M_{diff} = 17.86\%$, $SD = 10.20$; $p < .001$) than for tactile targets ($M_{diff} = 41.37\%$, $SD = 22.81$; $p < .001$). The size of the CCE for tactile targets did not significantly differ between the 5- to 7-year-olds and 8- and 9-year-olds ($p = .82$; $BF_{01} = 3.13$).

Secondary analysis: Fusion and fission

A mixed-factors 2 (Target Modality) × 2 (Interference Type: fission or fusion) × 2 (Stimulus Context) × 4 (Age Group) ANOVA conducted on accuracy difference scores yielded a significant main effect of target modality, $F(1, 80) = 9.08$, $p = .003$, $\eta^2_G = .02$, and interference type, $F(1, 80) = 91.27$, $p < .001$, $\eta^2_G = .17$. More errors were made in response to the tactile targets ($M = -26.56\%$, $SD = 38.09$) than to the visual targets ($M = -18.15\%$, $SD = 30.51$) and on fission-like trials ($M = -36.09\%$, $SD = 34.47$) than on fusion-like trials ($M = -8.63\%$, $SD = 29.17$). The main effects of stimulus context ($p = .80$; $BF_{01} = 11.10$) and age group ($p = .25$; $BF_{01} = 25.00$) did not reach significance. The only interactions to reach significance were between target modality and age group, $F(3, 80) = 4.44$, $p = .006$, $\eta^2_G = .03$,

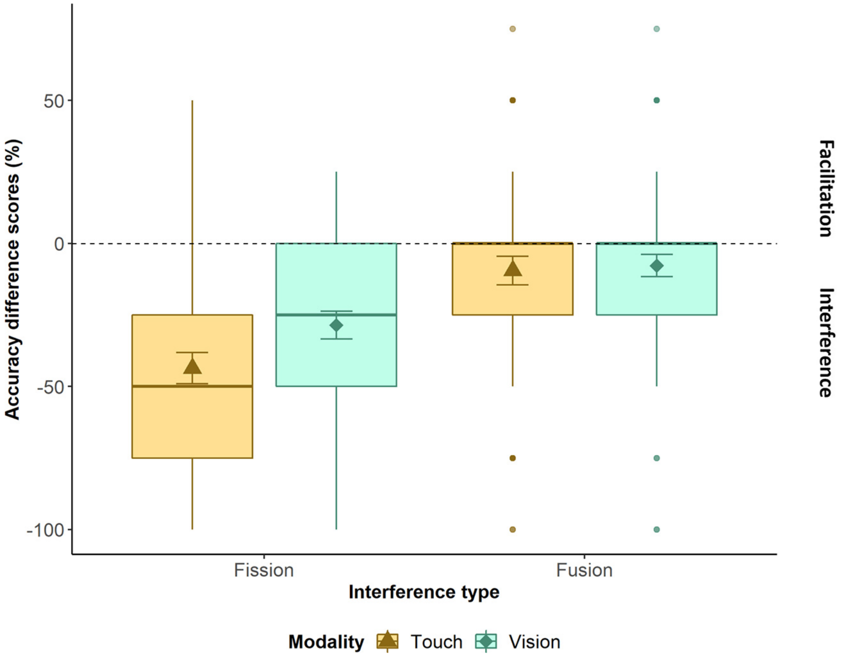

Fig. 5. Plot showing the interaction between interference type and target modality for accuracy difference scores. Mean scores are illustrated with large triangle (touch) and diamond (vision) points, and error bars represent 95% confidence intervals for the means. Small circles show outliers.

as shown in the incongruent values of Fig. 4 (see also supplementary material) and between interference type and target modality, $F(1, 80) = 7.47$, $p = .008$, $\eta^2_G = .01$, as shown in Fig. 5.

Post hoc paired *t* tests confirmed a significant difference in task performance based on target modality only for the 8- and 9-year-olds ($p < .001$), with more errors committed when attending to tactile targets than to visual targets ($M_{diff} = -22.62\%$, $SD = 22.04$).

Paired *t* tests confirmed significantly higher rates of fission-like effects than of fusion-like effects in response to the tactile targets ($M_{diff} = -34.08\%$, $SD = 36.57$; $p < .001$) and, to a lesser extent, to the visual targets ($M = -20.83\%$, $SD = 33.53$; $p < .001$). To investigate this further, a paired *t* test was performed to compare difference scores between fission-like and fusion-like interference for the tactile and visual targets (34.08% vs. 20.83%). This confirmed a significantly larger difference score for tactile targets compared with visual targets ($p < .009$).

Discussion

The aim of this study was to examine the developmental differences in bidirectional visuotactile interactions during the discrimination of sequences of events and how this is modulated by visual context. Children aged 5–7 years, 8 and 9 years, and 10–12 years and adults discriminated sequences of events in a target modality (touch or vision) and attempted to ignore distractors in the opposite modality while viewing a hand or an object presented from a first-person perspective. We found significant CCEs for both visual and tactile targets, consistent with findings in the adult literature (Holmes et al., 2006; O'Dowd et al., 2020; Poole et al., 2015; Spence & Walton, 2005; Walton & Spence, 2004). Whereas children aged 10–12 years and adults showed comparable CCEs across target modalities, children aged 5–7 years and 8 and 9 years showed larger CCEs when ignoring visual distractors as compared with tactile distractors. This was driven by trials in which a single target stimulus was presented with two sequential distractors in the irrelevant modality (i.e., fission-like effects).

The dominance of discontinuous sensory information on perceptual judgments in adults has been reported in studies of visuotactile interactions (Bresciani, Dammeier, & Ernst, 2008; Holmes et al., 2006; O'Dowd et al., 2020) and audiovisual interactions (Shams et al., 2000, 2002). Audiovisual fission is stronger in children under 9 years of age than in older children and adults (Adams, 2016; Nava & Pavani, 2013), consistent with reports for greater auditory dominance over vision during childhood than during adulthood (Hirst et al., 2018; Parker & Robinson, 2018). This finding indicates that modality weighting differs across development and there is a need to investigate multisensory interactions in a bidirectional fashion. The findings reported here suggest similar developmental changes in sensory weighting in the visuotactile domain, with younger children showing a stronger influence of vision on tactile judgments than vice versa, and this relative weighting was not observed in older children and adults. The weaker fusion-like effects across age groups is consistent with previous findings (Innes-Brown et al., 2011; for a review, see Hirst, McGovern, Setti, Shams, & Newell, 2020).

The pronounced interference from visual distractors but not tactile distractors during early and middle childhood suggests that vision was afforded a higher weighting than touch, perhaps because it was relatively more reliable (Beauchamp, Pasalar, & Ro, 2010; Ernst & Banks, 2002). This is consistent with the threshold data; tactile gap detection thresholds improved with development, whereas visual thresholds were mature by 5–7 years of age. These findings are also congruent with reports of early maturing temporal visual perception (Ellemberg, Lewis, Liu, & Maurer, 1999) and support a protracted development of tactile temporal acuity, in line with the findings of developmental work on tactile spatial acuity (e.g., Peters & Goldreich, 2013). Although the biasing effect of visual distractors on tactile judgments, and reduced susceptibility to bias from tactile distractors on visual judgments, is consistent with optimally weighted integration (Bresciani et al., 2006), our data might not reflect multisensory integration per se, in part due to the timing of the stimulus parameters.

Although some studies have reported that touch is more reliable than vision for numerosity judgments within the adult literature (Bresciani et al., 2006, 2008; Philippi et al., 2008), Werkhoven, van Erp, and Philippi (2009) found the opposite effect when considering purely bottom-up influences. That is, vision biased numerosity judgments more than touch only when participants were not directed to attend to one specific modality. When instructed to concentrate on a single modality (touch or vision), touch biased visual judgments more, suggesting that visual information was more easily suppressed. Given this evidence and the nature of the CCT, in which participants are explicitly instructed to ignore one modality and attend to another, a role of selective attention in our findings is important to consider.

There is evidence that the ability to selectively attend to a specific stream of information while ignoring distractors is well developed by 10 years of age (Klenberg, Korkman, & Lahti-Nuuttila, 2001; Klimkeit, Mattingley, Sheppard, Farrow, & Bradshaw, 2004; Passler, Isaac, & Hynd, 1985), with executive functioning reaching maturity during later childhood (Abundis-Gutiérrez, Checa, Castellanos, & Rueda, 2014; Romine & Reynolds, 2005; Stuss, 1992). Although evidence suggests that selectively attending to each sense does not eliminate temporal multisensory interactions (Odegaard, Wozny, & Shams, 2016), this could reduce the magnitude of crossmodal facilitation and interference. For example, during an audiovisual task in which adult participants were discriminating either visual or auditory cues while ignoring information in the other modality or dividing attention between both modalities simultaneously, selective attention to one modality reduced integration for matching signals across modalities (Mozolic, Hugenschmidt, Peiffer, & Laurienti, 2008; see also Badde, Navarro, & Landy, 2020, and Talsma, Doty, & Woldorff, 2007). Although Mozolic et al. (2008) noted no effect of attention on crossmodal conflict, Werkhoven et al. (2009) found that selectively attending to touch minimized the biasing effect of numerically incongruent visual events, demonstrating the capacity for top-down modulation of crossmodal interference.

Given the protracted development of neural regions associated with executive functions (Abundis-Gutiérrez et al., 2014; Romine & Reynolds, 2005; Stuss, 1992), it might be expected that the ability to suppress task-irrelevant visual information during a tactile numerosity judgment task should improve from early to late childhood, particularly as touch becomes more reliable, reducing both crossmodal interference and facilitation. Our findings show a target modality by age group interaction. It is possible that this could map onto a reduced ability to ignore task-irrelevant information during early to mid-childhood when this information is presented in the visual modality compared with the tactile modality. Conversely, children aged 10–12 years and adults do not show this distinction between modalities. However, whether these findings can be attributed to differences in selective attention, modality reliability, potential developmental differences in coupling prior width (i.e., prior knowledge of the typical correspondence of sensory signals; Bresciani et al., 2006), or a combination of these factors remains a question for future research.

Interference in general can arise at different levels of the processing stream: at early stimulus encoding (stimulus interference), later during response selection (response interference), or both (Hirst et al., 2019). It is possible that the origin of distraction, and in particular crossmodal distraction, fluctuates across development. For example, Cragg (2016) showed that enhanced interference during childhood is attributable to stimulus interference, and Hirst et al. (2019) also suggested that this is the case for crossmodal (audiovisual) interference. Crossmodal (visuotactile) interference in the spatial CCT is associated with response conflict (Forster & Pavone, 2008), which may also be the case for the temporal CCT (Forsberg, O'Dowd, & Gherri, 2019). Based on Werkoven et al.'s (2009) findings for stronger bottom-up visual interference on touch during the discrimination of sequences of events, it is possible that visuotactile interference occurs at early processing stages if both sensory sources are sufficiently attended. One approach to separating perceptual-based selection processes from response-based selection processes within numerosity judgments tasks is to implement signal detection theory (e.g., Rosenthal, Shimojo, & Shams, 2009). However, to implement signal detection analyses, a large number of trials are typically required, and owing to time restrictions we had a relatively small number of trials, thereby restricting the use of signal detection analysis. Such an approach remains a direction for future research.

We found no effect of stimulus context (hand vs. object) on task performance. This was surprising because our task maintained important similarities with O'Dowd et al.'s (2020) study, including presenting all visuotactile events within peripersonal space boundaries (Serino et al., 2015) and from a first-person perspective. However, several differences exist between the current paradigm and that implemented by O'Dowd et al. (2020) given that the task was adapted for use with children. First, to maintain consistency in the visual context condition, we used a standard image of a child-sized hand for all age groups. Importantly, O'Dowd et al. (2020) observed an effect of context when adult participants viewed their own hand or another's hand (see also Haggard, 2006). Thus, our hypothesized effect of stimulus context was not contingent on the embodiment of the image of the hand. Nonetheless, we recognize that hand size may influence embodiment in adults (e.g., Pavani & Zampini, 2007); thus, calibrating the hand image to that of a participant's hand could be adopted by future work. Further methodological differences, included fewer trials, vocal rather than manual responses, and a longer response period may have made the task easier. This is important to consider given that more robust visual enhancement of touch effects have been found when the task is difficult and performance is close to threshold (Press, Taylor-Clarke, Kennett, & Haggard, 2004; Serino, Farnè, Rinaldesi, Haggard, & Làdavas, 2007). Finally, unlike O'Dowd et al. (2020), here we also adjusted the gap between sequential stimuli based on each participant's threshold, which was necessary for implementing the task effectively with children and allowing for comparisons across age groups. It is possible that when the timing of sequential events is adjusted to be optimal for each participant, the tactile-relevant visual context provides no further advantage to task performance. These methodological factors should be considered in future work examining the influence of visual body- and object-based contexts on crossmodal interactions.

Conclusion

We examined visuotactile interactions during the discrimination of sequences of events in children aged 5–7 years, 8 and 9 years, and 10–12 years and adults across visual contexts. There was no effect of context on visuotactile interactions across age groups. However, children aged 5–7 years and 8 and 9 years showed large interference from visual distractors on tactile targets. Moreover, all children showed significantly larger tactile gap detection thresholds than adults, whereas visual gap detection thresholds were similar across age groups. Across age groups, fission-like effects were larger than fusion-like effects and fission-like effects were stronger when judging tactile targets than when judging visual targets. Our findings shed light on the development of numerosity judgments for visuotactile sequences of events.

Acknowledgement

This work was supported by the European Commission under grant no. 73239, Science Foundation Ireland Future Frontiers Award under grant no. 19/FFP/6812 and a TCD post-graduate scholarship (A. O'D.).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2021.105094.

References

Abundis-Gutiérrez, A., Checa, P., Castellanos, C., & Rueda, M. R. (2014). Electrophysiological correlates of attention networks in childhood and early adulthood. *Neuropsychologia, 57*, 78–92.

Adams, W. J. (2016). The development of audio-visual integration for temporal judgements. *PLoS Computational Biology, 12* e1004865.

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research, 21*, 301–308.

Badde, S., Navarro, K. T., & Landy, M. S. (2020). Modality-specific attention attenuates visual-tactile integration and recalibration effects by reducing prior expectations of a common source for vision and touch. *Cognition, 197* 104170.

Barutchu, A., Crewther, D. P., & Crewther, S. G. (2009). The race that precedes coactivation: Development of multisensory facilitation in children. *Developmental Science, 12*, 464–473.

Beauchamp, M. S., Pasalar, S., & Ro, T. (2010). Neural substrates of reliability-weighted visual-tactile multisensory integration. *Frontiers in Systems Neuroscience, 4*. https://doi.org/10.3389/fnsys.2010.00025.

Botvinick, M., & Cohen, J. (1998). Rubber hands "feel" touch that eyes see. *Nature, 391*, 756.

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

Bresciani, J. P., Dammeier, F., & Ernst, M. O. (2008). Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. *Brain Research Bulletin, 75*, 753–760.

Burr, D., & Gori, M. (2012). Multisensory integration develops late in humans. In M. M. Murray & M. T. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 345–362). Boca Raton, FL: CRC Press/Taylor & Francis.

Chen, Y. C., Lewis, T. L., Shore, D. I., Spence, C., & Maurer, D. (2018). Developmental changes in the perception of visuotactile simultaneity. *Journal of Experimental Child Psychology, 173*, 304–317.

Chen, Y. C., Maurer, D., Lewis, T. L., Spence, C., & Shore, D. I. (2017). Central-peripheral differences in audiovisual and visuotactile event perception. *Attention, Perception, & Psychophysics, 79*, 2552–2563.

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

Cowie, D., Makin, T. R., & Bremner, A. J. (2013). Children's responses to the rubber-hand illusion reveal dissociable pathways in body representation. *Psychological Science, 24*, 762–769.

Cowie, D., Sterling, S., & Bremner, A. J. (2016). The development of multisensory body representation and awareness continues to 10 years of age: Evidence from the rubber hand illusion. *Journal of Experimental Child Psychology, 142*, 230–238.

Cragg, L. (2016). The development of stimulus and response interference control in midchildhood. *Developmental Psychology, 52*, 242–252.

Denervaud, S., Gentaz, E., Matusz, P. J., & Murray, M. M. (2020). Multisensory gains in simple detection predict global cognition in schoolchildren. *Scientific Reports, 10*, 1394.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00781.

Dionne-Dostie, E., Paquette, N., Lassonde, M., & Gallagher, A. (2015). Multisensory integration and child neurodevelopment. *Brain Sciences, 5*, 32–57.

Ellemberg, D., Lewis, T. L., Liu, C. H., & Maurer, D. (1999). Development of spatial and temporal vision during childhood. *Vision Research, 39*, 2325–2333.

Ernst, M. O. (2008). Multisensory integration: A late bloomer. *Current Biology, 18*, R519–R521.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*, 429–433.

Féron, J., Gentaz, E., & Streri, A. (2006). Evidence of amodal representation of small numbers across visuo-tactile modalities in 5-month-old infants. *Cognitive Development, 21*, 81–92.

Filippetti, M. L., Johnson, M. H., Lloyd-Fox, S., Dragovic, D., & Farroni, T. (2013). Body perception in newborns. *Current Biology, 23*, 2413–2416.

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.

Forster, B., & Pavone, E. F. (2008). Electrophysiological correlates of crossmodal visual distractor congruency effects: Evidence for response conflict. *Cognitive, Affective, & Behavioral Neuroscience, 8*, 65–73.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., ... Rapoport, J. L. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 8174–8179.

Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Current Biology, 18*, 694–698.

Greenfield, K., Ropar, D., Themelis, K., Ratcliffe, N., & Newport, R. (2017). Developmental changes in sensitivity to spatial and temporal properties of sensory integration underlying body representation. *Multisensory Research, 30*, 467–484.

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.

Hartcher-O'Brien, J., Levitan, C., & Spence, C. (2010). Extending visual dominance over touch for input off the body. *Brain Research, 1362*, 48–55.

Hirst, R. J., Cragg, L., & Allen, H. A. (2018). Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. *Neuroscience & Biobehavioral Reviews, 94*, 286–301.

Hirst, R. J., Kicks, E. C., Allen, H. A., & Cragg, L. (2019). Cross-modal interference-control is reduced in childhood but maintained in aging: A cohort study of stimulus- and response-interference in cross-modal and unimodal Stroop tasks. *Journal of Experimental Psychology: Human Perception and Performance, 45*, 553–572.

Hirst, R. J., McGovern, D. P., Setti, A., Shams, L., & Newell, F. (2020). What you see is what you hear: 20 years of research using the Sound-Induced Flash Illusion. *Neuroscience & Biobehavioral Reviews, 118*, 759–774.

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*, 108–115.

Ide, M., & Hidaka, S. (2013). Visual presentation of hand image modulates visuo-tactile temporal order judgment. *Experimental Brain Research, 228*, 43–50.

Ide, M., Yaguchi, A., Sano, M., Fukatsu, R., & Wada, M. (2019). Higher tactile temporal resolution as a basis of hypersensitivity in individuals with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 49*, 44–53.

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*, 190–208.

Innes-Brown, H., Barutchu, A., Shivdasani, M. N., Crewther, D. P., Grayden, D. B., & Paolini, A. G. (2011). Susceptibility to the flash-beep illusion is increased in children compared to adults. *Developmental Science, 14*, 1089–1099.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Johnson, R. M., Burton, P. C., & Ro, T. (2006). Visually induced feelings of touch. *Brain Research, 1073*, 398–406.

Keil, J. (2020). Double flash illusions: Current findings and future directions. *Frontiers in Neuroscience, 14*. https://doi.org/10.3389/fnins.2020.00298.

Kennett, S., Taylor-Clarke, M., & Haggard, P. (2001). Noninformative vision improves the spatial resolution of touch in humans. *Current Biology, 11*, 1188–1191.

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology, 20*, 407–428.

Klimkeit, E. I., Mattingley, J. B., Sheppard, D. M., Farrow, M., & Bradshaw, J. L. (2004). Examining the development of attention and executive functions in children with a novel paradigm. *Child Neuropsychology, 10*, 201–211.

Lawrence, M. A. (2016). ez: Easy analysis and visualization of factorial experiments. https://CRAN.R-project.org/package=ez.

Lebel, C., Walker, L., Leemans, A., Phillips, L., & Beaulieu, C. (2008). Microstructural maturation of the human brain from childhood to adulthood. *NeuroImage, 40*, 1044–1055.

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*, 731–740.

McGovern, D. P., Roudaia, E., Stapleton, J., McGinnity, T. M., & Newell, F. M. (2014). The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration. *Frontiers in Aging Neuroscience, 6*. https://doi.org/10.3389/fnagi.2014.00250.

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*, 787–794.

Morey, R. D., Rouder, J. N., & Jamil, T. (2018). BayesFactor: Computation of Bayes factors for common designs. https://richarddmorey.github.io/BayesFactor/.

Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., & Laurienti, P. J. (2008). Modality-specific selective attention attenuates multisensory integration. *Experimental Brain Research, 184*, 39–52.

Murray, M. M., Lewkowicz, D. J., Amedi, A., & Wallace, M. T. (2016). Multisensory processes: A balancing act across the lifespan. *Trends in Neurosciences, 39*, 567–579.

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology, 18*, 689–693.

Nava, E., & Pavani, F. (2013). Changes in sensory dominance during childhood: Converging evidence from the Colavita effect and the sound-induced flash illusion. *Child Development, 84*, 604–616.

Negen, J., Chere, B., Bird, L. A., Taylor, E., Roome, H. E., Keenaghan, S., ... Nardini, M. (2019). Sensory cue combination in children under 10 years of age. *Cognition, 193* 104014.

Odegaard, B., Wozny, D. R., & Shams, L. (2016). The effects of selective and divided attention on sensory precision and integration. *Neuroscience Letters, 614*, 24–28.

O'Dowd, A., Sorgini, F., & Newell, F. N. (2020). Seeing an image of the hand affects performance on a crossmodal congruency task for sequences of events. *Consciousness and Cognition, 80* 102900.

Pagel, B., Heed, T., & Röder, B. (2009). Change of reference frame for tactile localization during child development. *Developmental Science, 12*, 929–937.

Parker, J. L., & Robinson, C. W. (2018). Changes in multisensory integration across the life span. *Psychology and Aging, 33*, 545–558.

Passler, M. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology, 1*, 349–370.

Pavani, F., & Zampini, M. (2007). The role of hand size in the fake-hand illusion paradigm. *Perception, 36*, 1547–1554.

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*, 195–203.

Peters, R. M., & Goldreich, D. (2013). Tactile spatial acuity in childhood: Effects of age and fingertip size. *PLoS One, 8*(12) e84650.

Philippi, T. G., van Erp, J. B., & Werkhoven, P. J. (2008). Multisensory temporal numerosity judgment. *Brain Research, 1242*, 116–125.

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*, 227–244.

Press, C., Taylor-Clarke, M., Kennett, S., & Haggard, P. (2004). Visual enhancement of touch in spatial body representation. *Experimental Brain Research, 154*, 238–245.

Puts, N. A., Harris, A. D., Mikkelsen, M., Tommerdahl, M., Edden, R. A., & Mostofsky, S. H. (2017). Altered tactile sensitivity in children with attention-deficit hyperactivity disorder. *Journal of Neurophysiology, 118*, 2568–2578.

R Core Team (2015). *RStudio: Integrated development for R*. Boston: RStudio, Inc. http://www.rstudio.com.

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing http://www.R-project.org/.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1997). The space around us. *Science, 277*, 190–191.

Romine, C. B., & Reynolds, C. R. (2005). A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Applied Neuropsychology, 12*, 190–201.

Ropar, D., Greenfield, K., Smith, A. D., Carey, M., & Newport, R. (2018). Body representation difficulties in children and adolescents with autism may be due to delayed development of visuo-tactile temporal binding. *Developmental Cognitive Neuroscience, 29*, 78–85.

Rosenthal, O., Shimojo, S., & Shams, L. (2009). Sound-induced flash illusion is resistant to feedback training. *Brain Topography, 21*, 185–192.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*, 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*, 225–237.

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*, 1355–1364.

Sann, C., & Streri, A. (2007). Perception of object shape and texture in human newborns: Evidence from cross-modal transfer tasks. *Developmental Science, 10*, 399–410.

Sann, C., & Streri, A. (2008). Inter-manual transfer of object texture and shape in human neonates. *Neuropsychologia, 46*, 698–703.

Serino, A., Farnè, A., Rinaldesi, M. L., Haggard, P., & Làdavas, E. (2007). Can vision of the body ameliorate impaired somatosensory function?. *Neuropsychologia, 45*, 1101–1107.

Serino, A., Noel, J. P., Galli, G., Canzoneri, E., Marmaroli, P., Lissek, H., & Blanke, O. (2015). Body part-centered and full body-centered peripersonal space representations. *Scientific Reports, 5*, 18603.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature, 408*, 788.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*, 147–152.

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) e224174.

Spence, C., & Walton, M. (2005). On the inability to ignore touch when responding to vision in the crossmodal congruency task. *Acta Psychologica, 118*, 47–70.

Stuss, D. T. (1992). Biological and psychological development of executive functions. *Brain and Cognition, 20*, 8–23.

Swanson, J., Deutsch, C., Cantwell, D., Posner, M., Kennedy, J. L., Barr, C. L., ... Wasdell, M. (2001). Genes and attention-deficit hyperactivity disorder. *Clinical Neuroscience Research, 1*, 207–216.

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: Is attending to both modalities a prerequisite for early integration?. *Cerebral Cortex, 17*, 679–690.

Violentyev, A., Shimojo, S., & Shams, L. (2005). Touch-induced visual illusion. *NeuroReport, 16*, 1107–1110.

Walton, M., & Spence, C. (2004). Cross-modal congruency and visual capture in a visual elevation-discrimination task. *Experimental Brain Research, 154*, 113–120.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*, 638–667.

Werkhoven, P. J., van Erp, J. B., & Philippi, T. G. (2009). Counting visual and tactile events: The effect of attention on multisensory integration. *Attention, Perception, & Psychophysics, 71*, 1854–1861.

Westerfall, J. (2016). PANGEA [computer program]. https://jakewestfall.shinyapps.io/pangea/.

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for Asperger syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders, 35*, 331–335.