Sensitivity to auditory-tactile colocation in early infancy

Rhiannon L. Thomas[1], Reeva Misra[2], Emine Akkunt[1], Cristy Ho[2],

Charles Spence[2], & Andrew J. Bremner[1]

[1]Sensorimotor Development Research Unit, Department of Psychology, Goldsmiths, University of London, UK

[2]Crossmodal Research Laboratory, Department of Experimental Psychology, University of Oxford, Oxford, UK

ACCEPTED FOR PUBLICATION (8th JUNE 2017) IN *DEVELOPMENTAL SCIENCE*

THIS IS A POST-PRINT (I.E. POST-REVIEW) VERSION OF THE ACCEPTED MANUSCRIPT

ABSTRACT WORD COUNT: 249 (250 max.)

WORD COUNT: 5586

CORRESPONDENCE TO: Professor Andrew J. Bremner, Department of Psychology, Goldsmiths University of London, London, SE14 6NW, United Kingdom; Tel. +44 (0) 207 078 5142; email: a.bremner@gold.ac.uk.

ACKNOWLEDGEMENTS: This research was supported by an award from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) (ERC Grant agreement no. 241242) to AJB. The authors would like to thank the parents and infants who took part in the study.

HIGHLIGHTS

1) At 4 months of age, infants are sensitive to the spatial colocation between auditory and tactile stimuli presented on their hands.

2) From 4 months of age at least, infants can represent auditory and tactile stimuli in a common spatial frame of reference.

3) Preference for auditory-tactile colocation at four months changes to a preference for non-colocation at six months of age.

4) The foundations of multisensory abilities underpinning the link between the body and external space exist within the first half-year of life.

ABSTRACT

An ability to detect the common location of multisensory stimulation is essential for us to perceive a coherent environment, to represent the interface between the body and the external world, and to act on sensory information. Regarding the tactile environment "at hand", we need to represent somatosensory stimuli impinging on the skin surface in the same spatial reference frame as distal stimuli, such as those transduced by vision and audition. Across two experiments we investigated whether 6- (n = 14; Experiment 1) and 4-month-old (n = 14; Experiment 2) infants were sensitive to the colocation of tactile and auditory signals delivered to the hands. We recorded infants' visual preferences for spatially congruent and incongruent auditory-tactile events delivered to their hands. At 6 months, infants looked longer toward incongruent stimuli, whilst at 4 months infants looked longer toward congruent stimuli. Thus, even from 4 months of age, infants are sensitive to the colocation of simultaneously-presented auditory and tactile stimuli. We conclude that 4- and 6-month-old infants can represent auditory and tactile stimuli in a common spatial frame of reference. We explain the age-wise shift in infants' preferences from congruent to incongruent in terms of an increased preference for novel crossmodal spatial relations based on the accumulation of experience. A comparison of looking preferences across the congruent and incongruent conditions with a unisensory control condition indicates that the ability to perceive auditory-tactile colocation is based on a crossmodal rather than a supramodal spatial code by 6 months of age at least.

Sensitivity to auditory-tactile spatial colocation in early infancy

The ability to detect the common source of multisensory external stimuli is essential for us to be able to perceive a coherent environment, to represent the interface between the body and the external world, and to act on sensory stimuli. Adults are able to solve this "crossmodal binding problem" (e.g., Parise, Spence, & Ernst, 2012), and also combine sensory stimuli in a manner that approaches statistical optimality (e.g., Ernst & Banks, 2002; Alais & Burr, 2004). However, the apparent ease with which this task is achieved in adulthood belies the true complexity of the feat of multisensory integration.

Due to the noise present in both the biological system and the environment, an internal signal originating from a stimulus can only ever provide an approximate estimate of the true properties of its source. Furthermore, each modality receives quantitatively and qualitatively different inputs (e.g., photons, molecules, pressure, etc.), which are represented using different neural codes, and transmitted at different speeds, therefore arriving in the central nervous system at different points in time (Harris, Harrar, Jaekl, & Kopinska, 2010). As such, the perceptual system must overcome a number of challenges in order to appropriately combine signals arriving via different senses. This is of particular concern in a developing system given the changing nature of its fundamental characteristics such as its physical size and in the speed and precision of neural processing. In this study we investigate whether young infants can overcome some of these challenges in order to determine whether auditory and tactile stimuli share the same location in external space.

Although the importance of crossmodal links between audition and touch may not be immediately apparent, there are numerous reasons why we might expect associations between these senses. Physiologically, both audition and touch rely on the mechanical

displacement of receptors to transduce external physical events into neural events. Anecdotally, there are a number of widely experienced sensations that appear to involve both audition and touch (e.g., the physical sensation on your skin from hearing nails scratching on a chalkboard or an insect buzzing; Kitagawa & Spence, 2006). Supporting this observation, there is also evidence of a patient with a form of synaesthesia that results in auditory stimuli producing tactile sensations (Ro, Ellmore, & Beauchamp, 2013). This patient has a right ventrolateral thalmic lesion, which deprives her somatosensory cortex of normal somatosensory input. Ro et al. (2013) suggest that this phenomenon may be explained by increased innervation of the somatosensory cortex by auditory projections due to a lack of normal thalamic input. There is also considerable empirical evidence for close links between touch and hearing in typically-developing individuals. Butler, Foxe, Fiebelkorn, Mercier, and Molholm (2012) provide evidence that representations of somatosensory and auditory frequency share similar preattentive neural circuits. In behaviour, Jousmäki and Hari (1998) reported that altering the sound that participants heard when rubbing their hands together influenced the tactile sensations that they experienced, leading them to change their judgements of roughness (see also Guest, Catmur, Lloyd, & Spence, 2002). Furthermore, there is evidence to suggest that auditory input can also influence tactile frequency discriminations (Yau, Olenczak, Weber, & Bensmaia, 2009).

Given the extensive links discussed above between auditory and tactile perception, including at early (preattentive) stages of information processing (e.g., Butler et al., 2012; Murray et al., 2005), it is pertinent to consider the question of when and how such links become available in early development. As infants begin to interact manually with objects

and surfaces at around 5 to 6 months of age, anecdotal observation suggests that they appear to make use of auditory-tactile links in their explorations, seemingly attending to the noises produced as they touch, hit or scratch surfaces. The ready availability of infant toys which provide simultaneous auditory and tactile stimulation as babies play with them, points similarly to the importance of such multisensory information in early life.

However, as already discussed, the task of colocating tactile and auditory stimuli in external space is by no means straightforward. Tactile stimuli (e.g., on the hands) are encoded in somatotopic (e.g., hand-centred) coordinates and auditory stimuli in head-centred coordinates. Due to this discrepancy, auditory-tactile spatial correspondences change every time the head and hands move with respect to one another. Adult humans and primates are equipped with neural circuits that continuously update the location of tactile stimuli in external space across changes in limb and body posture (e.g., Lloyd, Shore, Spence, & Calvert, 2002; Rigato, Bremner, Mason, Pickering, Davis, & Van Velzen, 2013). However, research with infants indicates that an ability to incorporate information about posture into sensory processing develops gradually across the first year of life (e.g., Begum Ali, Spence, & Bremner, 2015; Bremner, Mareschal, Lloyd-Fox, & Spence, 2008; Rigato, Begum Ali, Van Velzen, & Bremner, 2014), and continues into early childhood (e.g., Begum Ali, Cowie, & Bremner, 2014; Pagel, Heed, & Röder, 2009). This gradual development may be partly explained by the dramatic changes in early life of the relative sizes and shapes of the limbs, body, and head (Lampl, Veldhuis, & Johnson, 1992). Such physical changes along with the increasing number and variety of postural changes which an infant can readily and spontaneously execute (e.g., Van Hof, Van der Kamp, &

Savelsbergh, 2002), may lead to even greater difficulty in aligning sensory reference frames in early life.

Nonetheless, there is now a large body of empirical research suggesting that space (amongst other "amodal" properties) may be represented in a redundant manner across the senses and processed preferentially by infants (e.g., Bahrick & Lickliter, 2014). However, most of what we know about perception of these putative amodal properties in infancy has focused on representations of synchrony rather than colocation. For example, Lewkowicz, Leo, and Simion (2010) and Filippetti, Johnson, Lloyd-Fox, Dragovic, and Farroni (2013) have both demonstrated that synchrony between, respectively, sound and vision and between vision and touch may be perceived from birth.

Where attention has been paid to the development of crossmodal spatial perception, this has tended to focus on infants' sensitivity to auditory-visual colocation (although see recent studies on visual-tactile colocation by Freier, Mason, & Bremner, 2016, Filippetti et al., 2013; and Filippetti, Orioli, Johnson, & Farroni, 2015). Crossmodal orienting of vision to sounds is not a basic innate function. There is some ability in visual orienting to sounds at birth (Muir, Clifton, & Clarkson, 1989; Wertheimer, 1961), but this is not robust across different testing conditions (e.g., Butterworth & Castillo, 1976; Clifton, Morrongiello, Kulig, & Dowd, 1981) and is followed by an extended trajectory of development (Muir & Field, 1979). The most recent studies of auditory-visual colocation (Morrongiello, Fenwick, & Chance, 1998a; Morrongiello, Fenwick, & Nutley, 1998b) have shown that colocation between sights and sounds can be used as a cue for learning a crossmodal association from birth onwards. Nonetheless, ability to detect auditory visual colocation in

the first months of life is not an entirely robust finding across studies (e.g., Aronson & Rosenbloom, 1971; McGurk & Lewis, 1974; Morrongiello et al., 1998a).

Here we report findings of two experiments in which we examined whether 6-month-old (Experiment 1) and 4-month-old (Experiment 2) infants were able to distinguish between occasions where auditory and tactile sensations presented synchronously on the hands occur in the same or different location(s) (i.e., whether they were able to represent auditory-tactile colocation on the hands). To this purpose we adapted a method recently used by Freier et al. (2016) to investigate infants' perception of visual-tactile colocation. This method examines whether infants show a spontaneous visual preference for spatially congruent or incongruent crossmodal stimulus pairs presented on the hands. In the present study, infants' looking behaviour was compared across three conditions (see Fig. 1). In all three conditions various uni- and bimodal stimulus pairs were presented alternating between the two hands. In the congruent condition, tactile and auditory stimuli were presented together on the same hand. In the incongruent condition, tactile and auditory stimuli were presented moving between the hands in a random order. In the control condition, auditory stimuli on both hands and were alternated with tactile stimuli presented together on both hands. Because the only variation between congruent and incongruent conditions concerns the common or separate locations of auditory and tactile stimuli within external space, a spontaneous visual preference for either display implies an ability to detect colocation across audition and touch. Freier et al. (2016), in their study of visual-tactile processing which used a similar protocol to that presented here, report that both 6- and 10-month-old infants preferred to look at their hands during the incongruent trials.

On the basis of a visual preference for spatially incongruent over congruent bimodal stimuli, it is possible to conclude that there is an appreciation of the spatial relations between tactile and visual stimuli at some level of multisensory information processing. However, it remains to be seen whether this achievement is based on a perception of a crossmodal relationship between two unisensory stimuli, or on more coarse aspects of the perceptual array. For instance, it may be that infants preferred the incongruent trials because the stimuli, regardless of modality, are spread across a larger portion of space (across two hands). The control trials, in which either auditory or tactile stimuli were presented on both hands synchronously, were included in this experiment in order to address this possibility. The prediction was that the infants would look longer at spatially incongruent than at congruent auditory-tactile trials, as in Freier et al. (2016) for visual-tactile stimuli. If the infants also exhibited significantly greater duration of looking in the Incongruent compared to the Control trials this would suggest that their visual preference for the Incongruent condition was due to the lack of colocation of the multisensory stimuli rather than because these signals are spread across a larger space irrespective of modality.

Experiment 1

Method

Design

Participants completed six trials, presented across two blocks. Each block consisted of a Congruent trial, an Incongruent trial, and a Control trial. Ten stimulus events were presented in each trial involving various combinations of auditory and tactile stimuli on the participant's two hands. Each stimulus event had a duration of 700 ms followed by a

1,500 ms inter-stimulus interval (total trial duration = 20,500 ms; see Fig. 1). During the Congruent and Incongruent trials, the participants received auditory and tactile stimuli simultaneously which moved between the two hands from stimulus to stimulus in a random order. The auditory and tactile stimuli were presented concurrently on the same hand in Congruent trials and concurrently on different hands in the Incongruent trials. In the Control condition, tactile only and auditory only stimulus events, in which both hands were stimulated synchronously for 700 ms, with 1,500 ms inter-stimulus interval, were presented in a random order. The order of Congruent, Incongruent, and Control trials repeated across the two blocks within participants, but was counterbalanced between participants using the six possible different order sequences. The infants' looking behaviour was recorded by video camera, and the total duration of their fixations of their left and right hands were calculated for each trial from offline video records.

Participants

Fourteen 6-month-old infants participated in the study (6 male, mean age 197 days; SD = 9.0 days). An additional 4 infants were excluded from the sample due to fussiness such as continuous crying and excessive movement (3 infants) or parental influence of behaviour (1 infant). The participants were recruited from the local community and infants born before 37 weeks' gestation were excluded from taking part. The experiment was run in accordance with APA ethical principles for conducting research with children, and ethical approval was granted from the institutional Research Ethics Committee. Informed consent was obtained before testing was initiated and parents and legal guardians received a debrief afterwards.

Apparatus and materials

The auditory stimuli were presented via two miniature loudspeakers (voice coil transducers, mounted on custom-made enclosures) sewn into infant "scratch" mittens. These were controlled by a purpose-built unit that gated auditory signals from the computer to the transducers. The auditory stimuli consisted of a 1,000 Hz sine wave pulsing on for 50 ms and off for 50 ms for the duration of 700 ms. This sound was presented through each loudspeaker at 73 dBA (ambient noise in the room was 34 dBA). The vibrotactile stimuli were delivered to the infants' palms using custom-built voice-coil transducers and were driven at 220 Hz by a pure tone generator and amplifier. They were secured to the infants' palms using self-adhesive bandage. The volume of one tactile vibration when enclosed within an infant's palm was 35 dBA. A speaker emitting white noise (50 dBA) was placed under the table in order to mask the sounds made by the vibrotactile stimulators. The volume of the tactile vibration could not be detected above the sound of the white noise using a sound level meter positioned at the typical location of an infants' head. In order to avoid capture (i.e., spatial ventriloquism) of one stimulus by another, the auditory stimulus was pulsed whilst the tactile stimulus was continuous (Caclin et al., 2002). An infrared camera was positioned behind the experimenter, approximately 1 metre away from the infant and was connected to a monitor that was used to record the experiment. During testing, the recording was displayed on a video-monitor to a second experimenter outside the room. A PC running the EPRIME script was used to trigger stimulus events. All videos were coded by a second observer to obtain inter-rater reliability ($r = .80$, $p = .001$).

--Insert Figure 1 about here--

Procedure

The infant was seated on his/her carer's lap in a dark, sound-isolated room with black curtaining covering the walls. The infant's wrists were held by Experimenter A at arm's length, around 25 cm from the infant's body and 10 cm apart (these positionings were maintained as precisely as possible. The carer sat on a chair (60 cm high) in front of a table (50 cm wide, 30 cm deep and 76 cm high). Experimenter A engaged with the infant whilst Experimenter B sat in an adjoining room observing the infant's behaviour via a video monitor, triggering the trials. The carer was asked to refrain from attracting or directing the infant's attention during the trials and simply stabilised the infant's body whilst allowing the infant to freely move his/her head, eyes, and arms. To ensure that the carer did not bias the infant's response, he/she was unaware of the condition of the trial and could not feel or hear the location of the vibrotactile stimuli.

Experimenter A held the infant's hands above the table and played three rounds of "peek-a-boo" to attract the infant's gaze towards his/her hands, before saying "woo" and ducking out of sight. This "woo" signalled that Experimenter B should trigger a trial. After each trial, Experimenter A ensured that the infant's attention was attracted centrally between his/her hands and once this had been established, three rounds of peek-a-boo were played to signal the beginning of the next trial. The study ended once two blocks had been completed.

Results and discussion

--Insert Figure 2 about here--

Figure 2a shows the average duration of looking for 6-month-olds by Condition and experimental Block. The infants exhibited a consistent preference for the Incongruent condition (over both the Congruent and Control conditions), across both Blocks. Looking times in ms across Condition and Block were compared using a 2 (Block: 1 / 2)_x 3 (Condition: Congruent / Incongruent / Control) repeated measures analysis of variance (ANOVA). Preliminary analyses revealed no significant effects involving gender or trial order and so these variables were excluded. A main effect of Condition, $F(2,26) = 4.37$, $p = .023$, $\eta_p^2 = .25$, was explained by reliably longer looking in the Incongruent ($M = 7558$ ms, $SE = 995$ ms) than in the Congruent condition ($M = 5567$ ms, $SE = 657$ ms), $t(13) = 2.625$, p = .021, $d_z = .63$., as well as longer looking in the Incongruent as compared to the Control condition ($M = 5684$ ms, $SE = 978$ ms), $t(13) = 2.638$, $p = .020$, $d_z = .51$. There was also a significant main effect of Block, $F(1,13) = 5.26$, $p = .039$, $\eta_p^2 = .29$, describing a reduction in looking between the first ($M = 7111$ ms, $SE = 955$ ms) and second block ($M = 5428$ ms, $SE = 750$ ms). There was no interaction ($F < 1$).

At six months of age, infants are able to detect whether auditory and tactile signals presented on their hands are collocated or not. When given trials in which auditory and tactile stimuli were either presented on the same hand (congruent trials) or on different hands (incongruent trials), they exhibited a significant visual preference for the incongruent trials. The ability draw this distinction requires an ability, at some level of multisensory processing, to represent tactile stimuli impinging on the surface of the skin in the same spatial reference frame as auditory stimuli.

Given that the 6-month-olds preferred the incongruent condition to the control condition, it can also be concluded that their ability to colocate auditory and tactile cues

was based in a crossmodal rather than a supramodal sensory code. In the control condition, the stimuli on their hands alternated between two tactile events (one on each hand), and two auditory events (one emanating from each hand). On the basis of a supramodal sensory code, this provides the same spatial stimulus arrangement as in the incongruent condition. However, despite this, the 6-month-olds showed a preference for the incongruent over the control condition, indicating that the ability to differentiate colocated and non-colocated auditory-tactile events is based on the crossmodal spatial relationship between the auditory and tactile stimuli.

It seems reasonable to propose that by 6 months of age infants will have gained a significant degree of experience of concurrent auditory and tactile events on their hands. The onset of successful reaching for objects typically occurs at around 4 months (Clifton, Muir, Ashmead, & Clarkson, 1993), with subsequent improvements in object manipulation and grasping (see Adolph & Berger, 2005) which would presumably further enrich the kinds of auditory-tactile experiences which the infant receives leading to expectations regarding the common location of auditory and tactile events on the hands. As such we might predict that younger babies would demonstrate less ability to perceive auditory-tactile colocation. To examine this possibility, in Experiment 2, the same auditory-tactile events were presented to 4-month-old infants.

Experiment 2

All aspects of the design, apparatus, and materials were the same as Experiment 1. Sixty percent of the videos used to determine looking duration data were coded by a second rater obtaining an inter-rater reliability of (r = .76, p = .029).

Method

Participants

Fourteen 4-month-old infants (4 male, mean age 130 days; SD = 8.5 days) participated. An additional 6 infants were excluded as they failed to complete two experimental blocks (6 infants) due to fussiness. As in Experiment 1, the participants were recruited from the local community and infants born before 37 weeks' gestation were not included. The experiment was run in accordance with the APA ethical principles for conducting research with children, and ethical approval was also granted from the institutional Research Ethics Committee. Informed consent was obtained before testing was initiated and parents and legal guardians received a debrief afterwards.

Results and discussion

Figure 2b shows the average duration of looking for 4-month-olds by Condition and experimental Block. The 4-month-olds, in contrast to the 6-month-olds, exhibited a consistent preference for the Congruent condition (over both Incongruent and Control conditions), across both Blocks. Looking times in ms across Condition and Block were compared using a 2 (Block: 1 / 2) x 3 (Condition: Congruent / Incongruent / Control) repeated measures ANOVA. Preliminary analyses revealed no significant effects involving gender or trial order and so these variables were excluded. A main effect of Condition, $F(2,26) = 10.40$, $p < 0.001$, $\eta_p^2 = .45$, was explained by reliably longer looking in the Congruent trials (M = 7135 ms, SE = 861 ms) than the Incongruent trials (M = 5822 ms, SE = 628 ms), $t(13) = 2.475$, p= .028, $d_z = 0.47$, longer looking during Congruent trials than during Control trials (M = 4675 ms, SE = 699 ms), $t(13) = 3.711$, p = .003, $d_z = 0.84$ and longer looking during Incongruent trials than Control trials, $t(13) = 2.933$, p = .012, d_z

= 0.46. There was also a significant main effect of Block, $F(1,13) = 12.99$, $p = .003$, $\eta_p^2 = .50$, describing reduction in looking from the first (M = 7,045 ms, SE = 696 ms) to the second block (M = 4,769 ms, SE = 783 ms). There was no interaction of Condition and Block ($F < 1$).

In contrast to 6-month-olds, the 4-month-olds showed a visual preference for colocated over non-colocated auditory-tactile events. In this study, a statistically reliable visual preference, irrespective of direction, necessitates an ability to differentiate between conditions. Thus 4-month-old infants, like 6-month-olds, are able at some level of multisensory information processing to represent tactile stimuli impinging on the surface of the skin in the same spatial reference frame as auditory stimuli.

General discussion

The ability to detect the colocation of auditory and tactile stimuli is an important component of our ability to perceive the multisensory interface between the personal (tactile) world of the body and the external (audiovisual) environment. Auditory-tactile interactions are known to play a role in representing the layout of the body in external space in adulthood (Tajadura-Jiménez, Väljamäe, Toshima, Kimura, Tsakiris, & Kitagawa, 2012). Here we show for the first time that the foundations of such auditory-tactile multisensory interactions are in place by 4 months of age at the latest.

In the two experiments reported here, 4- and 6-month-old infants demonstrated visual preferences which differentiated between auditory-tactile events which were colocated on the hands and those which were non-colocated (i.e., on separate hands). A preference for either co-located or non-colocated auditory-tactile events indicates an ability

to differentiate such events on that basis. However, it is of note that an ability to differentiate colocation from non-colocation was observed through a preference for colocation at 4 months of age and a preference for non-colocation at 6 months of age. This changing preference for spatially congruent as compared to incongruent stimuli between 4 and 6 months of age is broadly consistent with a number of accounts of multisensory development. Bahrick and Lickliter's (2014) intersensory redundancy hypothesis, for instance, predicts greater preference for spatially congruent displays in young infants, but allows for attention to move more towards incongruent (or non-redundant) multisensory stimuli after the first half year of age. Meanwhile, Gergely and Watson's (1999) account of early social-emotional development rests on the idea of a "contingency detection module" which switches from a preference for perfect contingency (which includes spatial congruency; see also Rochat, 1998) up until 3 months of age, to a preference for imperfect contingency (including spatial incongruency) beyond that point.

Our preferred explanation of the pattern of visual preferences observed in the present study is developed from a proposal made by Freier et al. (2016) in relation to their findings with 6-month-old infants' ability to detect visual-tactile colocation. Freier et al. argued that 6-month-olds' preference for non-colocated visual-tactile trials represented a novelty preference for a perceptual state of affairs which is not typically experienced in ecological circumstances. We propose the same explanation of the 6-month-olds' preference for non-colocated auditory-tactile events. As discussed earlier, it is likely that by six months of age, infants will have gained significant experience of concurrent auditory and tactile events on their hands. The developmentally recent onset of successful reaching with subsequent improvements in object manipulation and grasping (see Adolph & Berger,

2005) are likely to lead them to associate spatially specific tactile sensations with concomitant auditory location cues (e.g., in the context of banging a toy against a table surface). The idea that infants might learn to expect auditory and tactile events to be associated according to common spatial coordinates in a single frame of reference is consistent with studies demonstrating that adults expect bimodal events to originate from a single place in external space (e.g., Pavani, Spence, & Driver, 2000; Shore, Barnes, & Spence, 2006).

In Experiment 2, we found that 4-month-old infants are also able to differentiate between colocated and non-colocated auditory-tactile stimuli via a preference for colocated over non-colocated auditory-tactile events. Our interpretation of this preference for colocation is that it represents a "familiarity preference" for a perceptual state of affairs which is more typically experienced ecologically than non-colocation. But why exactly should infants' preferences switch from familiarity (colocation) to novelty (non-colocation) between four and six months of age? The factors determining familiarity vs. novelty preferences are well documented in the infant learning literature. Supported by a range of findings (e.g., Caron & Caron, 1968; Cohen, Gelber, & Lazar, 1971; Hunter, Ross, & Ames, 1982; Wetherford and Cohen, 1973), Hunter and Ames (1988) propose that three factors account for relative visual preferences towards novel and familiar stimuli (age, duration of familiarization, complexity of discrimination). All three of these factors can help to explain a greater preference for familiarity in the younger age group. First, consistent with our familiarity-novelty account, younger infants are more likely to demonstrate a familiarity preference. Second, a novelty preference is predicted by longer exposure to the familiarised stimulus (e.g., Hunter et al., 1982; Roder, Bushnell, &

Sasseville, 2000). If we treat prior experience of auditory-tactile colocation as the familiarised stimulus, and assume that exposure to auditory-tactile colocation increases with age, this predicts increasing preference for non-colocation with age. Lastly, Hunter and Ames's (1988) model also includes the complexity of the perceptual discriminative task as a factor predicting greater familiarity preference (e.g., Caron & Caron, 1968; Cohen et al., 1971). There are a number of reasons to believe that differentiation between colocation and non-colocation would be more complex (and thus more likely to yield a familiarity preference) for the younger age group. For one, auditory spatial localization is poorer at 4 than at 6 months of age (Morrongiello, Fenwick, & Chance, 1990). Furthermore, to detect auditory-tactile colocation, it may be that infants need to take the relative postures of the head and arms into account in order to align auditory and tactile frames of reference. The sensory abilities required to differentiate the proprioceptive/visual cues necessary for this are likely to be more limited in 4- than 6-month-old infants. Overall then, an account of our findings in terms of a switch between preferences for familiar auditory-tactile spatial pairings in younger infants and novel pairings in older infants is consistent with what the literature tells us about the development of familiarity and novelty preferences.

How do our findings relate to previous work on the development of multisensory spatial abilities in infancy? Given the protracted development of infants' orienting responses to auditory (Muir et al., 1989) and, in particular, tactile (Bremner et al., 2008) stimuli, the current results might seem surprising. There are, however, a number of reasons as to why infants may show ability to detect auditory-tactile colocation at an age where studies have shown that they would normally be unable to make reliable spatial orienting

responses to the component sensory cues. First, infants in the current study were presented with ten stimuli of a given type across the duration of a trial. This extended and repeated exposure to a particular stimulus type is likely to support the successful processing of correlated spatial features across the senses. It is also possible that the multisensory presentations which we made (bimodal auditory-tactile events) supported spatial perceptual processing in a way that unimodal presentations typically presented in crossmodal orienting studies do not.

Another group of studies present a seeming conflict with the reported findings. Fenwick and Morrongiello (1998), studying the ability to form associations between non-spatial auditory and visual stimuli, found that 6-month-old infants were insensitive to a 10 cm spatial discrepancy between auditory and visual events. In contrast to adults, Fenwick and Morrongiello's 6-month-old participants still formed crossmodal associations at this degree of spatial separation. In the present study, the separation between auditory and tactile stimuli in the incongruent condition was also 10 cm, and here infants were able to differentiate between the colocated and non-colocated conditions. One possible explanation for this discrepancy is that different senses are involved. It is possible that auditory-tactile spatial processing develops prior to audio-visual spatial processing given the relative paucity of visual information *in utero,* and the close links between auditory and tactile processing in the brain (Butler et al., 2012; Yau et al., 2009). A further possibility is that an ability to differentiate between colocated and non-colocated bimodal events does not necessarily go hand-in-hand with a tendency to form crossmodal associations on that basis. Certainly, as we explain next, it is not possible to determine on the basis of the current

findings whether or not infants perceived colocated events as originating from coherent multisensory objects.

So what kind of representations underlie 4- and 6-month-old infants' ability to identify colocated auditory-tactile events? As just stated, it is unclear whether or not the infants bound the separate auditory and tactile stimuli into a single perceptual event (see Körding, Beierholm, Ma, Quartz, Tenenbaum, & Shams, 2007; Spence & Bayne, 2015). We can conclude though that in at least the 6-month-old group of infants who we tested, the ability to colocate auditory and tactile cues was based on a crossmodal rather than a supramodal sensory code. In the control condition that we presented to both age groups, the stimuli on the hands alternated between two tactile events (one on each hand), and two auditory events (one on each hand). On the basis of a supramodal sensory code, this provides the same spatial stimulus spread as in the incongruent condition. However, despite this, the 6-month-olds showed a preference for the incongruent over the control condition, indicating that the ability to differentiate colocated and non-colocated auditory-tactile events is based on the crossmodal spatial relationship between the auditory and tactile stimuli. The 4-month-olds however showed a visual preference for the colocated stimuli. This visual preference was also evident in comparison to the control condition which unfortunately only controls for preferences due to greater spatial spread. It is therefore possible that the 4-month-olds showed a preference for the colocated events on the basis of greater supramodal clustering of stimuli. Thus, further investigation will be needed in order to determine whether the ability to differentiate colocation and non-colocation in auditory-visual events at 4 months of age is based on a crossmodal or supramodal sensory code.

But what kind of crossmodal relationship did the 6-month-old infants perceive? We have focused our discussions on perception of crossmodal colocation. However, as pointed out by an anonymous reviewer of an earlier version of this manuscript, it is also possible that the 6-month-olds showed a preference for the incongruent condition because it is more spatiotemporally complex (i.e., the tactile stimulus moves left to right whilst the auditory stimulus moves right to left). This is certainly a possibility which warrants further investigation (although note that this explanation would not account for the pattern of preferences shown by the 4-month-old infants in Experiment 2). However, both colocation or codirection explanations presuppose that the 6-month-old infants are representing the tactile and auditory information individually (e.g. that their preferences are determined by a crossmodal relationship), and that they are able to register spatial relations between the senses.

Further questions remain about the nature of the crossmodal perceptual phenomena which we have measured in infants. For instance, it is unclear whether the 6-month-olds integrated the separate auditory and tactile stimuli into a unified perceptual event. In adults, temporal synchrony and (in some circumstances; see Spence, 2013) spatially colocated stimuli result in the perception of a multisensory event with a single origin (e.g., Körding et al., 2007). On the basis of the current data we are unable to determine whether either 4- or 6-month-old infants integrated the auditory-tactile stimulation in this task as a single event. In other words, it is possible that the infants' ability (both 4- and 6-month-olds) to differentiate the colocated event was based on either a perception of two colocated auditory and tactile stimuli or a single integrated auditory-tactile stimulus. Interestingly, research across a range of multisensory situations has suggested that such integration may not

develop until around 8 months of age (Neil, Chee-Ruiter, Scheier, Lewkowicz, & Shimojo, 2006), or even later (Barutchu, Crewther, & Crewther, 2009; Nardini, Jones, Bedford, & Braddick, 2008; Nardini, Begus, & Mareschal, 2013; Gori, Del Viva, Sandini, & Burr, 2008). Important questions for future research therefore include addressing the extent to which perceived synchrony and spatial colocation between auditory and tactile stimuli affect infants' perception of unified multisensory events, and whether these constraints on multisensory integration change across early life.

The current study provides the first window into the previously unexplored domain of auditory-tactile spatial representations in early life. From 4 months of age, infants are already sensitive to the colocation of auditory and tactile stimuli. This indicates that from this age infants can represent sensory stimuli impinging on the surface of their skin in the same spatial reference frame as auditory stimuli. Given the availability of auditory and tactile stimulation *in utero*, it is interesting to speculate that auditory-tactile spatial representation may be available prior even to birth. Such an early developing crossmodal ability may form the foundations of the multisensory interactions which underlie our perceptions of our own bodies in the external environment (e.g., Gallace & Spence, 2014; Tajadura-Jimenez et al., 2012).

REFERENCES

Adolph, K. E., & Berger, S. E. (2005). Physical and motor development. In M. Lamb & M. Bornstein (Eds.), *Developmental science: An advanced textbook, 5th Edition* (pp. 223-281). Mahwah, NJ: Lawrence Erlbaum Associates.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*, 257-262.

Bahrick, L. E., & Lickliter, R. (2014). Learning to attend selectively: The dual role of intersensory redundancy. *Current Directions in Psychological Science*, *23*, 414-420.

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.

Begum Ali, J., Cowie, D., & Bremner, A. J. (2014). Effects of posture on tactile localization by 4 years of age are modulated by sight of the hands: evidence for an early acquired external spatial frame of reference for touch. *Developmental Science*, *17*, 935-943.

Begum Ali, J., Spence, C., & Bremner, A. J. (2015). Human infants' ability to perceive touch in external space develops postnatally. *Current Biology*, *25*, R978-R979.

Bremner, A. J., Mareschal, D., Lloyd-Fox, S., & Spence, C. (2008). Spatial localization of touch in the first year of life: Early influence of a visual spatial code and the development of remapping across changes in limb position. *Journal of Experimental Psychology: General, 137,* 149-162.

Butler, J. S., Foxe, J. J., Fiebelkorn, I. C., Mercier, M. R., & Molholm, S. (2012). Multisensory representation of frequency across audition and touch: high density electrical mapping reveals early sensory-perceptual coupling. *Journal of Neuroscience*, *32*, 15338-15344.

Butterworth, G., & Castillo, M. (1976). Coordination of auditory and visual space in newborn human infants. *Perception, 5,* 155-160.

Caclin, A., Soto-Franco, S., Kingstone, A., & Spence, C. (2002). Tactile "capture" of audition. *Perception & Psychophysics, 64,* 616-630.

Caron, R. F., & Caron, A. J. (1968). The effects of repeated exposure and stimulus complexity on visual fixation in infants. *Psychonomic Science, 10,* 207-208.

Clifton, R. K., Morrongiello, B. A., Kulig, J., & Dowd, J. (1981). Newborns' orientation toward sound: Possible implications for cortical development. *Child Development, 52,* 833-838.

Clifton, R. K., Muir, D. W., Ashmead, D. H., & Clarkson, M. G. (1993). Is visually guided reaching in early infancy a myth? *Child Development*, *64*, 1099-1110.

Cohen, L. B., Gelber, E. R., & Lazar, M. A. (1971). Infant habituation and generalization to differing degrees of stimulus novelty. *Journal of Experimental Child Psychology*, *11*, 379-389

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

Fenwick, K. D., & Morrongiello, B. A. (1998). Spatial co-location and infants' learning of auditory-visual associations. *Infants Behaviour & Development, 21,* 745-760.

Filippetti, M. L., Johnson, M. H., Lloyd-Fox, S., Dragovic, D., & Farroni, T. (2013). Body

perception in newborns. *Current Biology, 23,* 2413-2416.

Filippetti, M. L., Orioli, G., Johnson, M. H., & Farroni, T. (2015). Newborn body perception: Sensitivity to spatial congruency. *Infancy*, *20*, 455-465.

Freier, L., Mason, L. & Bremner, A. J. (2016). Perception of visual-tactile colocation in the first year of life. *Developmental Psychology, 52,* 2184-2190.

Gallace, A., & Spence, C. (2014). *In touch with the future: The sense of touch from cognitive neuroscience to virtual reality*. Oxford, UK: Oxford University Press.

Gergely, G., & Watson, J. S. (1999). Early socio-emotional development: Contingency perception and the social-biofeedback model. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life* (pp. 101-136). Mahwah, NJ: Lawrence Erlbaum Associates.

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.

Guest, S., Catmur, C., Lloyd, D., & Spence, C. (2002). Audiotactile interactions in roughness perception. *Experimental Brain Research*, *146*, 161-171.

Harris, L., Harrar, V., Jaekl, P., & Kopinska, A. (2010). Mechanisms of simultaneity perception. In R. Nijhawan (Ed.), *Space and time in perception and action* (pp. 232-253). Cambridge, UK: Cambridge University Press.

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In L. P. Lipsitt (Ed.), *Advances in child development and behavior* (pp. 69-95). New York, NY: Academic.

Hunter, M. A., Ross, H. S., & Ames, E. W. (1982). Preferences for familiar or novel toys: Effects of familiarization time in 1-year-olds. *Developmental Psychology, 18,* 519-

529.

Jousmäki, V., & Hari, R. (1998). Parchment-skin illusion: Sound-biased touch. *Current Biology, 8,* 190.

Kitagawa, N., & Spence, C. (2006). Audiotactile multisensory interactions in information processing. *Japanese Psychological Research*, *48*, 158-173.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, *2*, e943.

Lampl, M., Veldhuis, J. D., & Johnson, M. L. (1992). Saltation and stasis: A model of human growth. *Science*, *258*, 801-804.

Lawson, K. R. (1980). Spatial and temporal congruity and auditory-visual integration in infants. *Developmental Psychology, 16*, 185-192.

Lewkowicz, D. J., Leo, I., & Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. *Infancy*, *15*, 46-60.

Lloyd, D. M., Shore, D. I., Spence, C., & Calvert, G. A. (2003). Multisensory representation of limb position in human premotor cortex. *Nature Neuroscience*, *6*, 17-18.

Morrongiello, B. A., Fenwick, K. D., & Chance, G. (1990). Sound localization acuity in very young infants: An observer-based testing procedure. *Developmental Psychology*, *26*, 75-84.

Morrongiello, B. A., Fenwick, K. D., & Chance, G. (1998). Crossmodal learning in newborn infants: Inferences about properties of auditory-visual events. *Infant Behavior & Development, 21*, 543-554.

Morrongiello, B. A., Fenwick, K. D., & Nutley, T. (1998). Developmental changes in

associations between auditory-visual events. *Infant Behavior & Development, 21*, 613-626.

Muir, D. W., Clifton, R. K. & Clarkson, M. G. (1989). The development of human auditory localization responses: A u-shaped function. *Canadian Journal of Psychology, 43,* 199-216.

Muir, D. W., & Field, J. (1979). Newborn infants orient to sounds. *Child Development, 50,* 431-436.

Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2005). Grabbing your ear: Rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cerebral Cortex, 15,* 963-974.

Nardini, M., Begus, K., & Mareschal, D. (2013). Multisensory uncertainty reduction for hand localization in children and adults. *Journal of Experimental Psychology: Human Perception & Performance*, *39*, 773-787.

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

Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science, 9,* 454-464.

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

Parise, C. V., Spence, C., & Ernst, M. O. (2012). When correlation implies causation in multisensory integration. *Current Biology*, *22*, 46-49.

Pavani, F., Spence, C., & Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological Science, 11,* 353-359.

Rigato, S., Begum Ali, J., van Velzen, J., & Bremner, A. J. (2014). The neural basis of somatosensory remapping develops in human infancy. *Current Biology*, *24*, 1222-1226.

Rigato, S., Bremner, A. J., Mason, L., Pickering, A., Davis, R., & van Velzen, J. (2013). The electrophysiological time course of somatosensory spatial remapping: vision of the hands modulates effects of posture on somatosensory evoked potentials. *European Journal of Neuroscience*, *38*, 2884-2892.

Ro, T., Ellmore, T. M. & Beauchamp, M. S. (2013). A neural link between feeling and hearing. *Cerebral Cortex, 23*, 1724-1730.

Rochat, P. (1998). Self-perception and action in infancy. *Experimental Brain Research*, *123,* 102-109.

Shore, D. I., Barnes, M. E., & Spence, C. (2006). Temporal aspects of the visuotactile congruency effect. *Neuroscience Letters, 392,* 96-100.

Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences, 1296,* 31-49.

Spence, C., & Bayne, T. (2015). Is consciousness multisensory? In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities* (pp. 95-132). Oxford, UK: Oxford University Press.

Tajadura-Jiménez, A., Väljamäe, A., Toshima, I., Kimura, T., Tsakiris, M., & Kitagawa, N. (2012). Action sounds recalibrate perceived tactile distance. *Current Biology*, *22*, R516-R517.

Van Hof, P., van der Kamp, J., & Savelsbergh, G. J. P. (2002). The relation of unimanual and bimanual reaching to crossing the midline. *Child Development, 73*, 1353-1362.

Wertheimer, M. (1961). Psychomotor coordination of auditory and visual space at birth. *Science, 134,* 1692.

Wetherford, M. J., & Cohen, L. B. (1973). Developmental changes in infant visual preferences for novelty and familiarity. *Child Development, 44,* 416-424.

Yau, J. M., Olenczak, J. B, Dammann, J. F., & Bensmaia, S. J. (2009). Temporal frequency channels are linked across audition and touch. *Current Biology*, *19*, 561-566.

FIGURE CAPTIONS

Figure 1: Stimulus protocols. (A) The three trial Conditions: i) Congruent trials in which the auditory and tactile stimuli were colocated on the same hand, ii) Incongruent trials in which auditory and tactile stimuli were presented on different hands, and iii) Control trials in which unisensory auditory or tactile stimuli were presented to both hands. The arrows indicate that the stimuli move between the hands (order of hands is randomised). (B) The timeline of stimulus presentation within a single trial. Each trial (20500 ms duration) contains 10 stimulus events (each is 700 ms duration) separated by 1500 ms inter-stimulus interval.

Figure 2: Mean looking duration across Conditions (Incongruent / Congruent / Control). Looking duration was compared across Conditions in two sequential experimental Blocks of trials (Block 1 / Block 2). The total duration of each of the six trials was 20,500 ms. The error bars indicate the standard error of the mean. (A) Experiment 1: 6-month-olds (n = 14). (B) Experiment 2: 4-month-olds (n = 14).

Figure 1

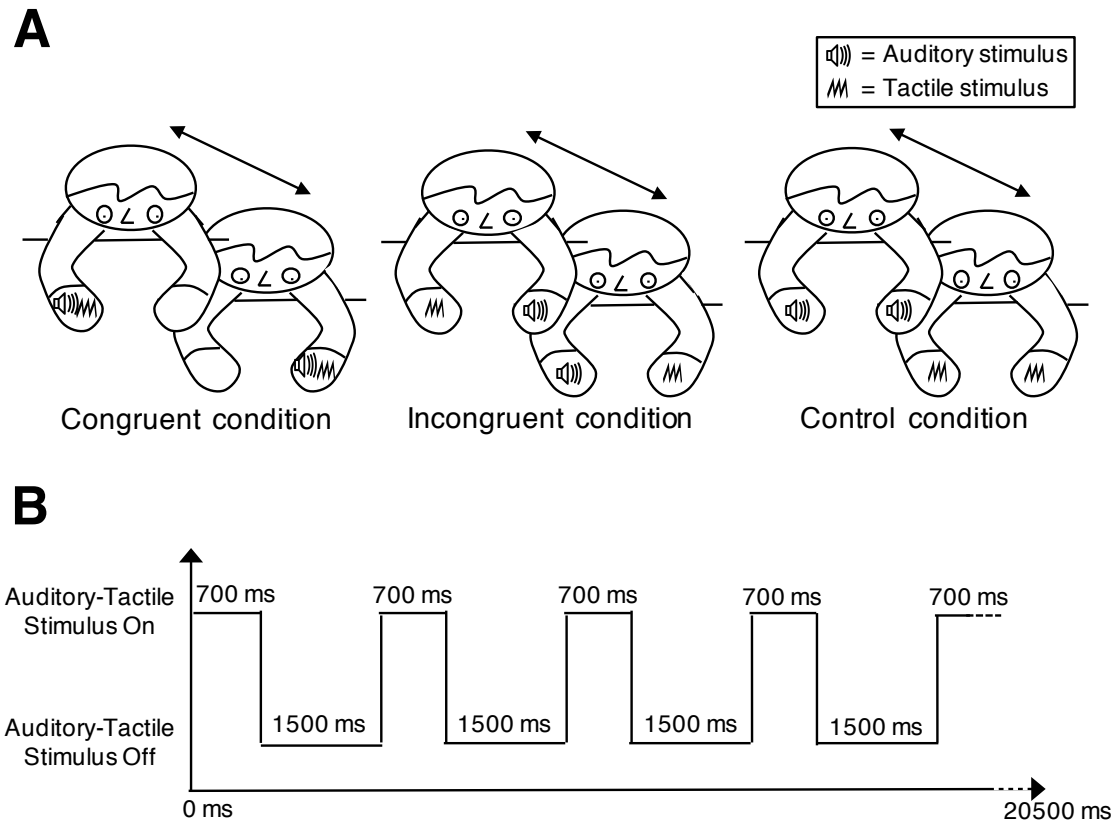
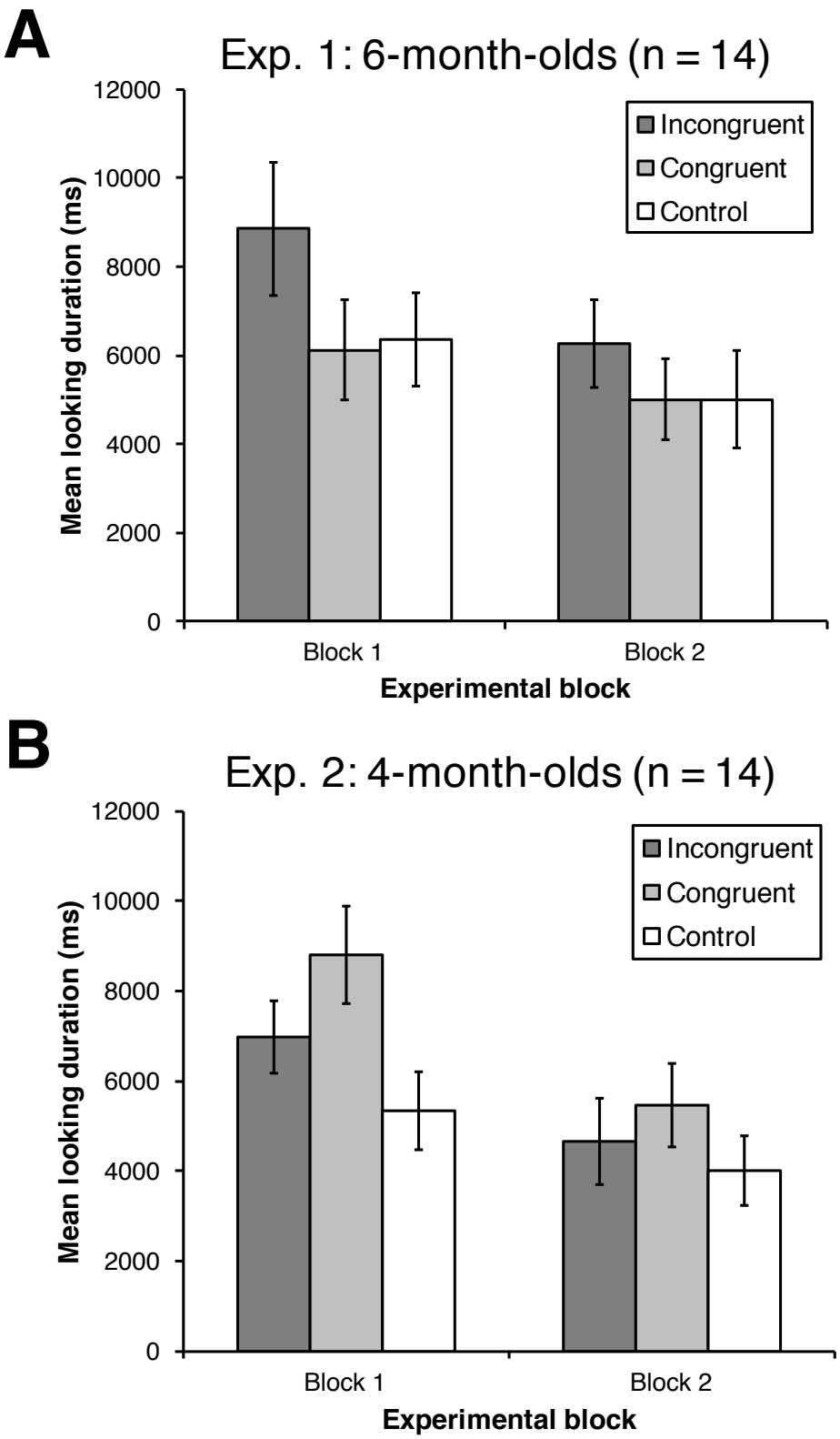

A
= Auditory stimulus
= Tactile stimulus
Congruent condition
Incongruent condition
Control condition
B
Auditory-Tactile Stimulus On
Auditory-Tactile Stimulus Off
700 ms
1500 ms
0 ms
20500 ms

Figure 2


A
Exp. 1: 6-month-olds (n = 14)
Mean looking duration (ms)
12000
10000
8000
6000
4000
2000
0
Incongruent
Congruent
Control
Block 1
Block 2
Experimental block
B
Exp. 2: 4-month-olds (n = 14)
Mean looking duration (ms)
12000
10000
8000
6000
4000
2000
0
Incongruent
Congruent
Control
Block 1
Block 2
Experimental block