

# Design of a Multisensory Stimulus Delivery System for Investigating Response Trajectories in Infancy

Dayi Bian<sup>1(✉)</sup>, Zhaobo Zheng<sup>2</sup>, Amy Swanson<sup>5</sup>, Amy Weitlauf<sup>3,5</sup>,  
Zachary Warren<sup>3,4,5</sup>, and Nilanjan Sarkar<sup>1,2</sup>

<sup>1</sup> Department of Electrical Engineering,  
Vanderbilt University, Nashville, TN, USA  
{dayi.bian, nilanjan.sarkar}@vanderbilt.edu

<sup>2</sup> Department of Mechanical Engineering,  
Vanderbilt University, Nashville, TN, USA

<sup>3</sup> Department of Pediatrics, Vanderbilt University, Nashville, TN, USA

<sup>4</sup> Department of Psychiatry, Vanderbilt University, Nashville, TN, USA

<sup>5</sup> Treatment and Research Institute of Autism Spectrum Disorders (TRIAD),  
Vanderbilt University, Nashville, TN, USA

**Abstract.** Sensory processing differences, including auditory, visual, and tactile, are ideal targets for early detection of neurodevelopmental risk. However, existing studies focus on the audiovisual paradigm but ignore the sense of touch. In this work, we present a multisensory delivery system that can deliver audiovisual stimuli and precisely controlled tactile stimuli to infants in a synchronized manner. The system also records multi-dimensional data including eye gaze and physiological data. A pilot study of six 3–8 month old infants was conducted to investigate the tolerability and feasibility of the system. Results have shown that the system is well tolerated by infants and all the data were collected robustly. This work paves the way for future studies charting the meaning of sensory response trajectories in infancy.

**Keywords:** Affective touch · Multisensory stimulation · Tactile stimulation · Physiological data · Eye tracking · Infancy

## 1 Introduction

Autism spectrum disorder (ASD) is a common neurodevelopmental disability characterized by social and communication impairments and is associated with costly human experience and financial impact [1, 2]. Despite the fact that a reliable diagnosis of ASD can be made by the age of 2 years, with many symptoms evident much earlier, most children are not accurately identified with ASD until after age four due to multiple factors, including difficulties accessing care and a lack of trained providers [2]. Consequently, these children do not receive early intervention in the first years of life, a time period recognized as optimal for enhancing developmental outcomes due to neural plasticity [3]. Although the neural basis of complex social and communicative behaviors develops over the course of childhood, brain response to more basic sensory

stimuli – e.g., touch, sight, smell - are present much earlier, even during the first few months of life – long before the observable behavioral and communication symptoms of ASD become apparent [4]. Given that hypo- and hyper-responsiveness to sensory input is a core diagnostic feature of ASD that can cause significant impairment over time [5, 6], it logically follows that children at risk of ASD or other neurodevelopmental disorders may show subtle sensory differences earlier than ASD can reliably be diagnosed at present, within the first year of life, identifying those children that may benefit from closer developmental monitoring.

Existing prospective studies of high-risk infant siblings of children with ASD (Sibs-ASD) suggest that sensory differences related to visual processing clearly emerge in the first two years of life [5]. A growing number of studies have investigated visual attention to faces and other social stimuli in Sibs-ASD [6–9]. For both high- and low-risk infants, most of these studies have described early point-in-time group level similarities on simple performance measurements of visual scanning and preferential looking to core facial features. However, these studies have suggested that high-risk infants may show subtle processing differences in the brain-based mechanisms for responding to these stimuli. These subtle processing differences may contribute to neurodevelopmental impairment over time. For example, for those infants who are eventually diagnosed with ASD, attention directed to eyes during infant-directed audiovisual speech initially appears intact but declines from 2 to 6 months of age, a pattern not observed in infants who do not develop ASD [7].

These important findings identify a potentially critical developmental trajectory of decreased visual attention in high-risk infants. However, existing social attention paradigms are limited by their focus on solely audiovisual modalities of early sensory learning. We know that neural mechanisms for processing various sensory inputs, such as tactile, vestibular, and auditory inputs, start to come online prenatally, playing an immediate role in the postnatal social-sensory experiences that lay a foundation for multisensory processing and social learning over time [10]. Paradigms using additional sensory processing channels to augment existing visual attention findings may provide more robust methods for detecting actionable neurodevelopmental risk at earlier time points.

One sensory processing channel that could augment existing visual attention work is related to tactile perception, or sense of touch. The sense of touch is widely known for the role it plays in discriminating and identifying external stimuli. However, there is growing evidence that the sense of touch has another dimension, also known as “affective touch,” which conveys social information just like what someone sees and hears [11]. Affective touch is often defined as a form of pleasant touch involving mutual skin-to-skin contact between individuals [11]. Previous research has demonstrated that infants are sensitive to affective touch [12] and that compared to other forms of touch, stroking an infant can not only induce positive emotions, but also modulate negative ones [13]. In work with adults with ASD, Croy et al. found that they show atypical perception and processing of affective touch. Additionally, the authors hypothesized that the affective touch functionality, which is based on C tactile fiber activation [14], is impaired to some extent in individuals with ASD [15]. Furthermore, Kaiser et al. demonstrated that in the presence of affective touch stimuli, individuals with ASD exhibit reduced brain activity in social-emotional-related brain regions

compared to typically developing (TD) individuals [16]. As such, affective touch represents an identified area of atypical sensory processing related to ASD that can also influence infant response, making it an optimal target for early detection of neurodevelopmental risk.

Although it is not practical to produce affective touch in laboratory settings, an analogous tactile stimulation which is produced by a mechanical source (e.g., soft brushing) is comparable to affective touch that is manually produced by hand [17]. Previous tactile stimulation work in infants has utilized trained human confederates to administer pleasant social touch via dorsal forearm stroking with Hake brushes at predetermined velocities and pressures [12]. Although adequate for documenting generalized physiological response to stimuli, this manual control has several limitations: speed and pressure are not precisely controlled or measured, stroking is hard to coordinate with other stimuli/measurements, and the human presence may confound certain experimental paradigms.

None of the aforementioned work investigated tactile perceptions when investigating how the infants perceive and process sensory stimuli. A few studies solely described how adults with and without ASD process affective touch differently. In this paper, we present a multisensory stimulation and data capture system (MADCAP) for infants that delivers multiple sensory stimulations and simultaneously captures multi-dimensional data. To the best of our knowledge, this is the first work to demonstrate a multimodal technological system incorporating affective touch that has the potential to meaningfully chart differences in coordinating visual, auditory, and tactile processing in infancy. To deliver simulated tactile stimuli, we developed a novel tactile stimulation device for infants that utilizes precisely controlled speed and pressure of brush strokes.

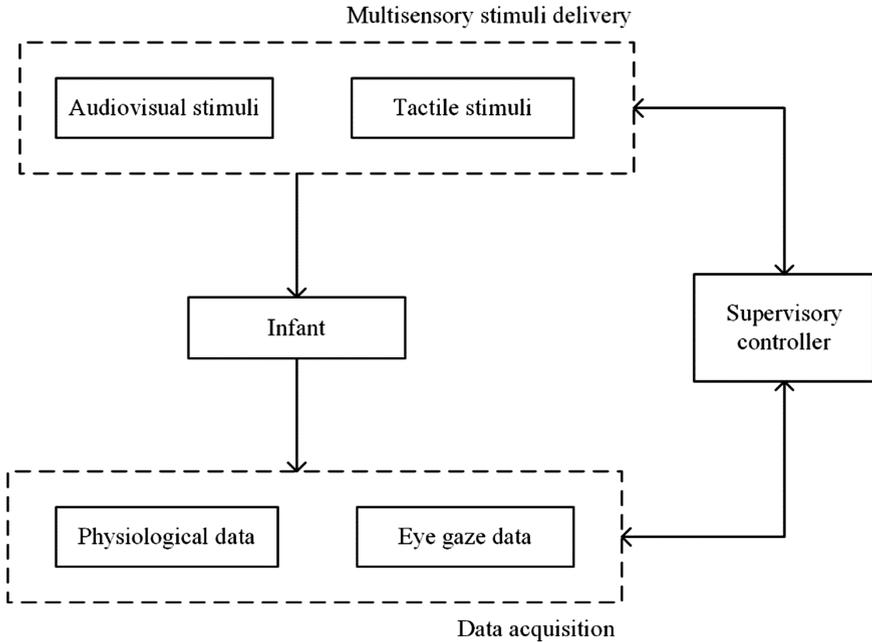
This paper is organized as follows. Section 2 demonstrates the system design of MADCAP. Section 3 describes the evaluation of the system with a pilot study. Section 4 presents the results of the system evaluation. In the final section, we conclude the work with a discussion of study results and future work.

## 2 System Design

MADCAP includes three main components: a multisensory stimulation delivery module, a multi-dimensional data capture module, and a supervisory controller module to synchronize the connections between modules. Figure 1 shows the overall system diagram.

### 2.1 Multisensory Stimulation Delivery Module

We developed an intelligent mechatronic device to simulate affective touch on the forearm of an infant. To our knowledge, this is the first computer-controlled tactile stimulation device used to study infants with ASD. This device was designed to provide precisely controlled brush stroking with variable speed and pressure. This device has three compartments from top to bottom. The top compartment contains two stepper

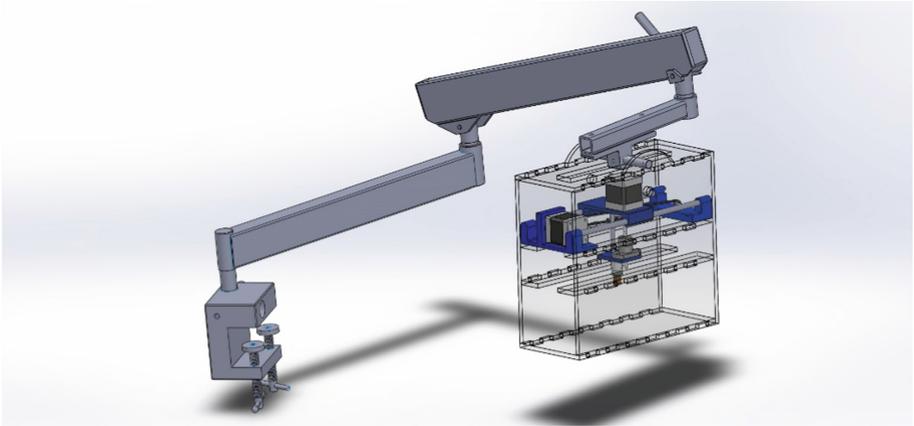


**Fig. 1.** System diagram

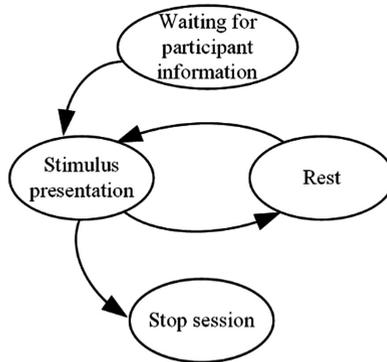
motors and belts which control the brush, allowing it to move in both horizontal and vertical directions. The middle compartment includes a replaceable soft brush. A pressure sensor is attached to the bottom of the brush so that the pressure applied on the arm from the brush can be measured and modulated. The infant’s arm rests in the bottom compartment. The infant’s forearm is placed into a soft strap to hold it in place and guarantee that the brush will contact the forearm.

The device is attached to an articulating arm which gives the infant a certain degree of freedom to move his/her arm while the relative position between the arm and the device remains the same (Fig. 2). The two stepper motors are controlled by an Arduino-based microcontroller. The speed, stroke length, stroking direction and number of stroking cycles, can be precisely controlled. The pressure on the arm can be maintained within a certain range (soft, medium, and high).

We designed a custom audiovisual stimulation delivery submodule which has the flexibility to communicate with other stimulation delivery submodules and the data capture module. Unity (<https://unity3d.com>) was used to implement the module. A finite state machine (Fig. 3) was used to control the logic of the audiovisual stimuli presentation. All parameters of the stimuli delivery such as the content of the audiovisual stimuli, the duration of rest between each stimulus, and the dose of the stimuli, could be adjusted effortlessly according to the experimental protocol. Also, the user-defined event markers, such as start/end experiment and start/end stimulus, were logged into files using the JSON ([www.json.org](http://www.json.org)) format.



**Fig. 2.** CAD model of the tactile stimulation device



**Fig. 3.** Audiovisual delivery module FSM. The program starts in the “waiting for participant information” state. After the experimenter inputs the participant information, the program starts the “Stimulus presentation” state. The participant has a certain amount of time to rest between each stimulus. After several rounds of stimulus presentation according to the protocol, the program will come to the “Stop session” state and record the data.

This audiovisual stimulation delivery module also communicated with the tactile stimulation device via supervisory controller to make sure multisensory stimuli were properly synchronized. Furthermore, this module sent the user-defined event markers to data capture module for later data analysis. We will discuss the details of the inter-module communication in Sect. 2.3.

## 2.2 Data Capture Module

We tracked the eye gaze positions of infants when they looked at audiovisual stimuli. The Tobii X120 eye tracker ([www.tobii.com](http://www.tobii.com)), which has a sampling rate of 120 Hz,

was used to measure gaze position across defined regions within the audiovisual presentation. The (X, Y) coordinates of the gaze position—(0, 0) for upper left corner—as well as time stamps and event markers were recorded.

In addition to gaze data, we also measured the infant's physiological data, including blood volume pulse (BVP) and electrodermal activity (EDA). The E4 wristband (<https://www.empatica.com/e4-wristband>), which is an unobtrusive device highly suitable for infant study, was used to record the physiological data while worn on the ankle. The sampling rate for BVP and EDA is 64 Hz and 4 Hz respectively. By using the hardware API provided by the E4 wristband, we developed a custom program for the E4 wristband to record time-stamped physiological data within the protocol. The physiological data measured by the E4 wristband were streamed to the custom program wirelessly via Bluetooth.

### 2.3 Supervisory Controller Module

As you can see from the system diagram (Fig. 1), the supervisory controller module served as a bridge between the multisensory stimulation delivery and data capture modules. It played a crucial role in MADCAP system, making sure that all the stimulations were presented in a time-synchronized manner and all the data captured were properly time stamped.

The supervisory controller program communicates with the tactile stimulation device through the serial port at a baud rate of 9600 bits/s. A command is sent to the tactile stimulation device to initiate the brush stroking when a tactile stimulus is needed. The brush stroking does not start immediately because it must first be moved vertically to come into contact with the infant's arm, as described in Sect. 2.1. As soon as the brush touches the infant's arm and the pressure sensor detects the touch, the tactile stimulation device sends a notification message to the supervisory controller. Then the supervisory controller initiates the audiovisual stimulus. In this way, we guaranteed the audiovisual and tactile stimuli fired at the same time.

Throughout the experiments, the supervisory controller monitored the stimulation delivery modules. When a user-defined event occurred (e.g., start brush stroking/audiovisual stimulus), the supervisory controller sent an event marker in JSON format to data capture modules over a TCP/IP based socket in a LAN.

## 3 System Validation

In order to demonstrate the tolerability and feasibility of the MADCAP system we tested our integrated audio, visual, and tactile protocol across a sample of 6 infants from 3–8 months of age (3 girls, 3 boys; mean age = 5.45 months, SD = 1.60). The protocol was reviewed and approved by the Institutional Review Board (IRB) at Vanderbilt University. After receiving a thorough explanation of the experiment, parents gave written informed consent for their children's participation.

Within a sound-attenuated room, infants were seated in an infant/toddler seat (appropriate to age) and positioned 50 cm from the LCD monitor, video recording

device, and eye tracker. If parents requested to have the infant on the lap or if the infant refused to sit in an infant seat, the parent would be permitted to hold their infant in their lap but was asked to minimally distract the infant. Then the infant's left forearm was positioned within the bottom compartment of the tactile stimulation device and velcro strapped. The E4 wristband was attached to the right ankle. Subsequently, the eye tracker calibration was accomplished by presenting audio-enabled, animated cartoon pictures. We appropriately used 2- or 5-point calibration procedures depending on the infant's age [18]. The infants then participated in a single session lasting approximately 10 min where they were exposed to three distinct presentations of audiovisual speech stimuli (each 50 s in length).

The audiovisual speech presentations were the same stimuli utilized by Lewkowicz [19] to demonstrate the multisensory coherence of fluent audiovisual speech. Specifically, the stimuli presented included clips of an adult female narrating a short story in English (native tongue), Spanish (non-native tongue), and audio-asynchronous native (English, with timing of audio presentation delayed 500 ms seconds relative to video). Each audiovisual speech recording was presented with and without a concurrent tactile stimulation. In the presence of tactile stimulation, the infant's dorsal forearm was continuously stroked (i.e., back and forth) at a speed of 3 cm/s, the pressure of 0.2 N, with a soft makeup brush. The stroking speed and pressure were chosen based on existing studies which have shown that gentle touch at medium speed produces the most pleasant effect [20]. The task stimuli were presented in pseudorandom order and the infant received 10 s of rest between each stimulus to reduce sensory habituation.

## 4 Results

### 4.1 Tolerability and Feasibility of the System

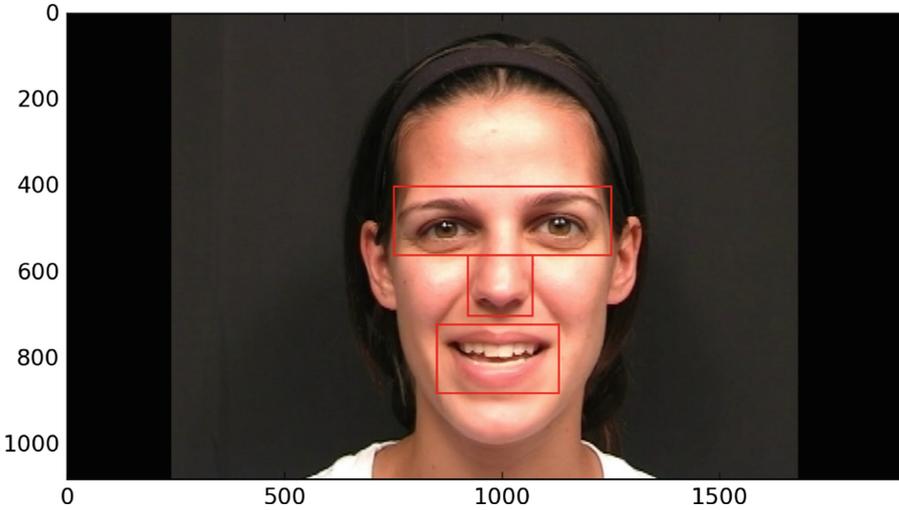
Results of this study demonstrated that all infants tolerated the approximate 10-min. protocol. All 6 infants completed the experiments. None demonstrated tolerability issues with either eye tracker calibration or stimuli presentation. Both the eye gaze data and physiological data were robustly collected.

### 4.2 Eye Tracker Data and Physiological Data Analysis

In this study, we defined three Regions of Interest (ROI) surrounding eyes, nose, and mouth [9] (Fig. 4). We focused on how much time infants spent fixating on these ROIs. Since infants often varied in the amount of time they spent looking at the stimuli, the fixation data were normalized (looking time to ROIs divided by total looking time to the stimuli) for each stimulus presentation.

Results indicated that, across all the sessions, participants looked at the stimulus screen 27% of the time with gaze toward demarcated ROI for 57% of this time.

Multiple features were extracted from physiological signals (PPG and EDA). Heart rate was calculated by detecting peaks in the PPG signal. Tonic and phasic components of EDA were decomposed separately from the original signal. The tonic component is the baseline level of EDA and is generally referred to as skin conductance level (SCL).



**Fig. 4.** Defined ROIs in speech stimuli

The phasic component is the part of the signal that changes when stimuli take place and is known as skin conductance response (SCR). The following table shows the features extracted from two conditions with and without tactile stimulation Table 1.

**Table 1.** Results for two conditions with and without tactile stimulation

	Heart rate (/min)	scr_rate (response peaks/min)	scr_mean ( $\mu$ S)	scl_mean ( $\mu$ S)
With touch	127.02	3.50	0.08	0.55
Without touch	133.74	3.74	0.09	0.58

scr\_rate: Rate of skin conductance response  
 scr\_mean: Mean of skin conductance response  
 scl\_mean: Mean of skin conductance level

## 5 Discussion

In the current study, we presented a multisensory stimuli delivery and data capture system (MADCAP) for infants. To the best of our knowledge, this is the first multisensory stimuli-delivery system capable of delivering auditory, visual, and precisely-controlled tactile stimuli in a synchronized manner. MADCAP was used to collect infants’ eye gaze and physiological data together with user-defined event markers. A pilot study validating the system by demonstrating that: (1) the tactile stimulation device delivered precisely controlled brush stroking; (2) infants under the age of 12 months tolerate the system and 10-min. protocol; and (3) the eye gaze and physiological data as well as the user-defined event markers could be collected robustly.

We defined three ROIs (areas around eyes, nose, and mouth) of the speech stimuli and calculated the infants' normalized attention time in these ROIs. This measurement is important for future study because it is an informative feature to investigate individual's social attention. The results from physiological data demonstrated that the heart rate and skin conductance response rates were both lower when the affective touch stimuli were presented, indicating a decrease in arousal. Although none of the observed differences reached statistical significance, the results show trends similar to those of earlier work, which revealed that affective touch resulted in infants' arousal decreasing [12].

The presented work paves the way for the future research into multisensory perception and processing in infancy. This system could be utilized to explore the multisensory processing difference between the infants who will and will not develop ASD later. Our future work includes conducting a longitudinal study for children from their early infancy until the age when they can be clinically diagnosed with ASD.

## References

1. Buescher, A.V., et al.: Costs of autism spectrum disorders in the United Kingdom and the United States. *JAMA Pediatr.* **168**(8), 721–728 (2014)
2. Christensen, D.L.: Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2012. *MMWR. Surveill. Summ.* **37**(1), 1–8 (2016)
3. Veenstra-VanderWeele, J., Warren, Z.: Intervention in the context of development: pathways toward new treatments. *Neuropsychopharmacology* **40**(1), 225–237 (2015)
4. Murray, M.M., et al.: Multisensory processes: a balancing act across the lifespan. *Trends Neurosci.* **39**(8), 567–579 (2016)
5. Germani, T., et al.: Brief report: assessment of early sensory processing in infants at high-risk of autism spectrum disorder. *J. Autism Dev. Disord.* **44**(12), 3264–3270 (2014)
6. Weitlauf, A.S., et al.: *Therapies for Children with Autism Spectrum Disorder*. Createspace Independent Publishing Platform, North Charleston (2014)
7. Jones, W., Klin, A.: Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature* **504**(7480), 427–431 (2013)
8. Sacrey, L.-A.R., Bryson, S.E., Zwaigenbaum, L.: Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder: a longitudinal study from 6 to 36 months of age. *Behav. Brain Res.* **256**, 441–450 (2013)
9. Shic, F., Macari, S., Chawarska, K.: Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. *Biol. Psychiatry* **75**(3), 231–237 (2014)
10. Lewkowicz, D.J., Leo, I., Simion, F.: Intersensory perception at birth: newborns match nonhuman primate faces and voices. *Infancy* **15**(1), 46–60 (2010)
11. Löken, L.S., Olausson, H.: The skin as a social organ. *Exp. Brain Res.* **204**(3), 305–314 (2010)
12. Fairhurst, M.T., Löken, L., Grossmann, T.: Physiological and behavioral responses reveal 9-month-old infants' sensitivity to pleasant touch. *Psychol. Sci.* **25**(5), 1124–1131 (2014)
13. Peláez-Nogueras, M., et al.: The effects of systematic stroking versus tickling and poking on infant behavior. *J. Appl. Dev. Psychol.* **18**(2), 169–178 (1997)

14. McGlone, F., Wessberg, J., Olausson, H.: Discriminative and affective touch: sensing and feeling. *Neuron* **82**(4), 737–755 (2014)
15. Croy, I., et al.: Affective touch awareness in mental health and disease relates to autistic traits – an explorative neurophysiological investigation. *Psychiatry Res.* **245**, 491–496 (2016)
16. Kaiser, M.D., et al.: Brain mechanisms for processing affective (and nonaffective) touch are atypical in autism. *Cereb. Cortex* **26**(6), 2705–2714 (2016)
17. Tricoli, C., et al.: CT-optimized skin stroking delivered by hand or robot is comparable. *Front. Behav. Neurosci.* **7**, Article 208 (2013)
18. Gredebäck, G., Johnson, S., von Hofsten, C.: Eye tracking in infancy research. *Dev. Neuropsychol.* **35**(1), 1–19 (2009)
19. Lewkowicz, D.J., et al.: Perception of the multisensory coherence of fluent audiovisual speech in infancy: Its emergence and the role of experience. *J. Exp. Child Psychol.* **130**, 147–162 (2015)
20. Löken, L.S., Evert, M., Wessberg, J.: Pleasantness of touch in human glabrous and hairy skin: order effects on affective ratings. *Brain Res.* **1417**, 9–15 (2011)