

Published in final edited form as:

*Dev Cogn Neurosci.* 2011 July 1; 1(3): 351–360. doi:10.1016/j.dcn.2011.01.003.

## Attention modifies sound level detection in young children

Elyse S. Sussman<sup>a</sup> and Mitchell Steinschneider<sup>b</sup>

<sup>a</sup>Department of Neuroscience Albert Einstein College of Medicine 1300 Morris Park Avenue Bronx, NY 10461

<sup>b</sup>Department of Neurology Albert Einstein College of Medicine 1300 Morris Park Avenue Bronx, NY 10461

### Abstract

Have you ever shouted your child's name from the kitchen while they were watching television in the living room to no avail, so you shout their name again, only louder? Yet, still no response. The current study provides evidence that young children process loudness changes differently than pitch changes when they are engaged in another task such as watching a video. Intensity level changes were physiologically detected only when they were behaviorally relevant, but frequency level changes were physiologically detected without task relevance in younger children. This suggests that changes in pitch rather than changes in volume may be more effective in evoking a response when sounds are unexpected. Further, even though behavioral ability may appear to be similar in younger and older children, attention-based physiologic responses differ from automatic physiologic processes in children. Results indicate that 1) the automatic auditory processes leading to more efficient higher-level skills continue to become refined through childhood; and 2) there are different time courses for the maturation of physiological processes encoding the distinct acoustic attributes of sound pitch and sound intensity. The relevance of these findings to sound perception in real-world environments is discussed.

### Keywords

Development; Sound Intensity; Sound Frequency; Event-related potentials (ERPs); Mismatch Negativity (MMN); Attention

## 1. INTRODUCTION

Intensity level perception, such as whether or not a sound is heard as louder than or softer than another sound, is evident from early childhood (Berg & Boswell, 2000; Sinnott & Aslin, 1985). However, little else is known about the development of sound level perception. In natural environments, we perceive stable auditory events despite great variation in the intensity level of the sounds around us. However, there has been very little focus on how intensity is represented and used within complex auditory scenes, and even less on the development of such processes during maturation. Previous studies have focused on complex scene processing in adults by investigating the influence of frequency level or

© 2011 Elsevier Ltd. All rights reserved

**Corresponding Author:** Elyse S. Sussman, PhD Department of Neuroscience Albert Einstein College of Medicine 1300 Morris Park Avenue Bronx, NY 10461 USA elyse.sussman@einstein.yu.edu Phone: 1-718-430-3313 Fax: 1-718-430-8821.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

spatial location on stream perception (Bregman, Ahad, Crum, & O'Reilly, 2000; Bregman & Campbell, 1971; Carlyon, Cusack, Foxton, & Robertson, 2001; Eramudugolla, McAnally, Martin, Irvine, & Mattingley, 2008; Ihlefeld & Shinn-Cunningham, 2008; Muller, Widman, Schröger, 2005; Shinn-Cunningham, Lee, & Oxenham, 2007; Snyder, Alain, & Picton, 2006; Sussman & Steinschneider, 2009; Sussman, 2005, Sussman, Ritter, & Vaughan, 1999), with little attention to how other sound features contribute to auditory object recognition. Much is still unknown about how the brain analyzes and perceives the dynamically changing complex auditory scene, when both task-relevant and task-irrelevant sounds occur at the same time.

Intensity variation is not always directly implicated in sound event processing, as, for example, understanding the word “apple” does not depend on whether it's spoken in a soft or a loud voice. The neural basis for intensity coding and the relationship of intensity coding mechanisms to processing other sound features is still debated.

Neurons in auditory cortex that are tuned to sound frequency and spatial location can be coded across different levels of stimulus intensity (Sadagapan & Wang, 2008; Recanzone, Schreiner, & Merzenich, 1993; Miller & Recanzone, 2009). Recently, it has been proposed that level-invariant coding at the cortical level is an important mechanism for detecting sound objects in complex or noisy auditory scenes (Sadagapan & Wang, 2008). Sadagapan & Wang (2008) found three types of cells in primary auditory cortex of the marmoset: 1) those in which frequency tuning was dependent upon the sound level; 2) those in which frequency was unrelated to sound level (cells have the same bandwidth at all sound levels); and 3) those that were relatively independent of one another, being narrowly but separately tuned to frequency and intensity. Neurons that are coded with some independence between frequency and intensity allow frequency to be coded without the influence of level, whereas neurons involved in level-invariant processing, with broad bandwidths, contribute to analyzing the more global scene, such as is needed to follow a sound object against a wide background of frequency and intensity levels. Thus, the range of tuning properties of these different types of neurons would allow a greater computational capacity contributing to the perception of a dynamically changing auditory scene with concurrently occurring sound streams.

Maturation of sound intensity processing for simple level changes within a single sound stream was tested by comparing automatic (irrelevant) and attention-based (relevant) processes in two groups of children and one group of adults. The goal was to determine whether attention to sound level (in the form of task relevance) would alter the neurophysiologic response in a way that would be concordant with the behavioral ability to detect sound level changes. We hypothesized that 1) automatic sound processes are not fully developed in childhood, but rather are shaped progressively by experience with sounds during development. This hypothesis is consistent with our previous data showing discordance between passive and active listening in complex scenes (Sussman & Steinschneider, 2009); and 2) with the great variety of tuning properties for intensity coding that occurs in mature animals, there will be a longer developmental time course for intensity coding than frequency coding. Thus, we expected to find a difference in the passive (automatic) processing of simple detection of intensity level changes compared to simple frequency level changes, and a difference between intensity level detection in younger and older children.

In the current study, we compared suprathreshold intensity level detection in younger (6–9 years) and older (10–12 years) children and young adults (22–46 years) using an auditory oddball paradigm. Behavioral measures and event-related brain potentials (ERPs) were used to compare active and passive processing of the oddball. In the Passive listening condition

(Passive), sound level detection was irrelevant and participants watched a silent video. In the active listening condition (Active), sound level was relevant to the task of pressing a response key to louder or softer sounds. The main dependent physiologic measures were three ERP components: MMN, N2, and P3. Each of these components reflects a neurophysiologic process of deviance detection (Sussman, 2007). MMN is elicited when the deviant is physiologically detected in both passive and active listening conditions, whereas N2 and P3 are elicited only in active listening conditions (Novak, Ritter, Vaughan, & Wiznitzer, 1990), such as when loudness changes are targets. The MMN component peaks earlier than the attention-based components (approx. 150 ms from deviance detection) and reflects early sensory stages of sound processing. MMN is a modality-specific component, with neural generators within auditory cortices (Alho, 1995). The N2/P3 components peak later and reflect integration of information from earlier processing stages (Picton, 1992). The components associated with active target-detection (N2/P3) are largely non-modality specific, with widespread neural generators, giving rise to a centro-parietal distribution in adults that can be seen at the scalp (Perrault & Picton, 1984). Thus, together these ERP components reflect automatic and controlled processing of sound change detection.

Concordance between behavioral and physiological indices of sound level detection would be shown by good performance in detecting sound changes (e.g., seen in hit rate) along with elicitation of the MMN, N2, and P3 components in the Active listening (target detection) condition. Ability of the brain to detect frequency or sound level deviations in the passive state would be indicated by elicitation of MMN when attention was focused on watching a video. Thus, a difference between passive (automatic) and active processing of sound level changes could be physiologically observed if, for example, MMN were elicited in the Active but not in the Passive conditions.

## 2. METHODS

### 2.1 Participants

Ten right-handed children ranging in age between 6–9 years ( $M=8$ ,  $SD=0.9$ ), ten children ages 10–12 years ( $M=11$ ,  $SD=0.9$ ), and ten adults ages 22–46 years ( $M=31$ ,  $SD=7$ ) participated in the study (Table 1). Participants were recruited by flyers posted in the immediate medical/research community and in local area schools in the Bronx. Among children and adults, participants were 30% Caucasian, 15% Asian, 50% Hispanic, and 5% African American. Children gave written assent and their accompanying parent gave written consent after the protocol was explained to them. The protocol was approved by the Committee for Clinical Investigations at the Albert Einstein College of Medicine and The New York City Department of Education. All procedures were carried out according to the Declaration of Helsinki. Recruits were pre-screened by phone interviews, or interviews with parents if under 18 years of age, to exclude past or present diagnoses of learning, speech/language, hearing, behavioral, or neurological disorders, or report of special educational services or school grade retention. Recruits who met the pre-screen criteria underwent a two-hour psychometric testing session with a licensed psychologist<sup>1</sup>. All participants passed a hearing screening at 20 dB HL for 500, 1000, 2000, and 4000 Hz.

<sup>1</sup>Standard scores on all screening tests for cognitive and language function, reading/decoding skills, and phonological processing abilities had to be at least within the average range (standard scores of at least 85) on the following instruments: Wechsler Abbreviated Scale of Intelligence (WASI) for cognitive function; Woodcock-Johnson-III Tests of Achievement (WJ-III) Letter/Word Identification and Word Attack for reading/decoding; WJ-III Understanding Directions and Children's Essentials of Language Fundamentals-3 (CELF-3); The Phonological Awareness Test (PAT) Rhyming; and The Comprehensive Test of Phonological Processing (CTOPP) core screening tests for phonological awareness, phonological memory, and rapid naming. In addition, participants or their parents had to report fewer than six symptoms of inattention and hyperactivity/impulsivity on a DSM-IV checklist.

## 2.2 Stimuli

Stimuli were complex harmonic tones (containing 5 partials), 50ms in duration (5ms rise/fall times), created using Adobe Audition® software and presented binaurally via E-a-rtunes® 3A insert earphones with NeuroStim (Compumedics Inc., Texas USA) software and hardware. Two different tones were presented in each sound sequence at a 300 ms stimulus onset asynchrony (SOA), one frequent (.88, called the 'standard') and one infrequent (.12, called the 'deviant'). The fundamental frequency for the standard tone was 440 Hz, 50.5 dB(A). In one condition (Intensity), the deviant was 15 dB louder than the standard, and all other dimensions of the two complex tones were the same. In the other condition (Frequency), the fundamental frequency of the deviant was 2637 Hz and all other dimensions of the two complex tones were identical. Stimuli were pseudo-randomized so that deviants did not occur successively, and at least two standard tones intervened between deviants. Tones were calibrated using a sound-level meter (Brüel & Kjær 2209) with an artificial ear (Brüel & Kjær 4152). The intensity oddball data for the adult participants were originally obtained from a different study. These data were reanalyzed for the purposes of comparison with the child intensity oddball conditions in the current study. Therefore, there are no frequency oddball data for the adults.

## 2.3 Control for delineating MMN

MMN is delineated by subtracting the ERP evoked by the standard from the ERP evoked by the deviant. A control condition was conducted for the Intensity oddball conditions so that a standard control stimulus could be obtained with the same physical characteristics as the deviant stimulus in the main blocks. Thus, the standard ERP to be subtracted from the deviant ERP would be evoked by a stimulus with the exact same physical characteristics as the deviant but they would differ in their role in the stimulus block. To do this, the intensity values of the two tones were reversed from the main condition. These conditions were conducted so that the deviant intensity in the control condition was 15 dB softer than the standard. In this control condition, when attended, participants pressed a response key for the softer tone.

## 2.4 Procedures

Participants sat in a reclining chair in an electrically-shielded and sound-attenuated booth. Passive conditions were recorded first, in which participants watched a captioned silent video of their choosing, with the order of presentation between intensity and frequency randomized across participants. 2010 tones were presented each for the Intensity and Frequency Oddball Passive condition and 1005 in the control condition. 1600 tones were presented in the Active condition, and 600 stimuli in the control condition. The experimenter monitored eye saccades in the EEG to ensure that participants were watching the movies and reading the captions. In the active condition, participants were instructed to listen to the sounds and press the response key every time they detect a louder (or softer in the control condition) sound. The experimental session lasted approximately one hour (including electrode placement and breaks).

## 2.5 Electroencephalogram (EEG) recording and Data analysis

EEG recordings were obtained using a 32-channel electrode cap that incorporated a subset of the International 10–20 system. Additional electrodes were placed over the left and right mastoids (LM and RM, respectively). The tip of the nose was used as the reference electrode during recordings. F7 and F8 electrode sites were used in a bipolar configuration to monitor the horizontal electro-oculogram (EOG). FP1 and an electrode placed below the left eye were used in a bipolar configuration to monitor the vertical EOG. All impedances were maintained below 5 k $\Omega$ . The EEG and EOG were digitized (Neuroscan Synamps amplifier,

Compumedics Corp., Texas, USA) at a sampling rate of 1000 Hz (0.05–200 Hz bandpass). EEG was filtered off-line with a lowpass of 30 Hz (zero phase shift, 24 dB rolloff). Artifact rejection was set to exclude activity exceeding 100  $\mu$ V after EEG epochs were baseline corrected. Approximately 16% of all trials were rejected, prior to further averaging, in each condition. Epochs were 600 ms in duration, starting 100 ms pre-stimulus onset and ending 500 ms post-stimulus onset.

Peak latency of the ERP components was determined in the grand mean difference waveforms at the electrode site of greatest signal-to-noise ratio for each respective age group, and used to measure the amplitude in a 40 ms (MMN and N2) or 50 ms (P3) interval centered on the peak (Fig. 3 displays the peak latency of these components). The P1 peak latency was determined from the standard ERPs, for both the intensity (louder/softer), and condition (Active/Passive) factors separately, and used to measure the 40 ms interval centered on its grand mean peak.

To determine the presence of response components, one-sample *t* tests were used to verify whether the mean amplitude was significantly different from zero. Two-way repeated measures ANOVA were calculated to further determine the presence of the MMN, N2 and P3 (stimulus type: deviant, control standard) and scalp distribution (electrode site: Fz, Cz, F3, F4, FC1, FC2, C3, C4 for MMN; and Fz, Cz, Pz, F3, F4, C3, C4, P3, P4 for N2 and P3 components). These electrodes were chosen to cover frontal, central, and parietal scalp sites, which also include those with the greatest signal-to-noise ratio for the components of interest according to previous studies in children and adults (Ponton et al., 2000; Sussman et al., 2008; Sussman & Steinschneider, 2009). Amplitude and latency of the components were compared with repeated measures ANOVA on the difference waveforms. Mixed model ANOVA was used for all group comparisons. To calculate group comparison of component latencies, the electrode site of greatest signal-to-noise ratio for each age group. For the P1 component, Fz was used for child groups, and Cz for adults. For the MMN, Fz was used for all groups, and for P3, Pz for all groups. Greenhouse-Geisser corrections were used and reported. Post hoc tests were calculated using Tukey HSD.

Behavioral responses were calculated for reaction time (RT), hit rate (HR), and false alarm rate (FAR) for each participant in each condition separately. Mixed model repeated measures ANOVA were used to compare RT, HR, and FAR using a within factor of intensity (louder vs. softer deviants) and a between factor of age (6–9 years/10–12 years/adult). Tukey HSD was used to calculate post hoc tests.

### 3. RESULTS

#### 3.1 Behavioral results

Table 2 summarizes the behavioral data for the louder and softer intensity deviants in all ages. Both adults and children were able to reliably detect the softer and louder intensity deviants as reflected in RT, HR, and FAR. For HR, there was a main effect of age group ( $F_{2,27} = 30.19, p < 0.001$ ), with post hoc calculations revealing that younger children had a significantly lower HR (.72) than older children (.87), and older children a significantly lower HR than adults (.95). There was a main effect of intensity on HR ( $F_{1,27} = 42.99, p < 0.001$ ), which was due to an overall higher HR to louder (.91) than softer deviants (.79). There was also an interaction between factors ( $F_{2,27} = 13.03, p < 0.001$ ). Post hoc analysis showed 1) a significantly higher HR to louder than softer deviants in the children, but no significant difference in HR for loud vs. soft in the adults; 2) HR increased with increasing age for softer deviants: younger children (.60) < older children (.82) < adults (.94), however, there was no significant difference in HR among the age groups for the louder deviants.



FAR was low for all age groups. There was a significant main effect of age group on FAR ( $F_{2,27} = 4.36, p < 0.023$ ). Younger children had a significantly higher false alarm rate (.026) than adults (.003). FAR of older children (.011) did not differ significantly from either younger children or adults. There was both a main effect of intensity ( $F_{1,27} = 14.27, p < 0.001$ ), with significantly more false alarms to softer (.016) than louder (.009) deviants, and an interaction between age group and intensity ( $F_{2,27} = 3.78, p < 0.036$ ). Post hoc analysis revealed 1) a significantly higher FAR to softer than louder deviants for each group of children, but no significant difference in FAR to softer vs. louder deviants in the adults; 2) false alarms did not differ for the softer sounds between the two child groups, but younger children had significantly more false alarms (.03) than adults (.002), whereas there was no significant difference in false alarms among the age groups for the louder deviants.

There was also a main effect of age group on RT ( $F_{2,27} = 18.56, p < 0.001$ ), with RT being slower in younger children (.445 ms) than both older children (.357 ms) and adults (.308 ms). There was no significant difference between RT in older children and adults. There was also an overall main effect of intensity on RT ( $F_{1,27} = 112.53, p < 0.001$ ), with a shorter RT to louder (.333 ms) than softer (.407 ms) deviants. The interaction between age group and intensity did not quite reach significance ( $F_{2,27} = 3.07, p = 0.063$ ).

### 3.2 ERP results

**3.2.1 P1 obligatory response**—P1 is a modality-specific component, generated within auditory cortices (Liégeois-Chauvel et al., 1994), and reflects early perceptual processes associated with obligatory onset detection of an acoustic event (Näätänen & Picton, 1987). Auditory P1 peaks approximately 50 ms after stimulus onset in adults and 100 ms in children.

Figure 1 displays the event-related potentials elicited by the standard sounds ( $p = .88$ ) for all age groups and conditions. This figure shows the obligatory responses across conditions and age groups, displayed at Fz for children and at Cz for adults. This is due to a difference in P1 scalp distribution observed during maturation. P1 amplitude is observed with a maximum frontally in children, but at the vertex in adults (Sussman et al., 2008). Figure 2 displays the midline (Fz, Cz, Pz) ERPs elicited by the deviant overlain with the ERPs elicited by the standard for the child groups so that the differences between the standard (obligatory) ERP responses and deviant ERP responses can be seen before delineating the MMN component. The midline electrodes are displayed to show the scalp distribution of the ERP responses.

In children, a prominent P1 component was elicited by both standard and deviant tones peaking approximately 100 ms from stimulus onset, with maximal amplitude at the Fz electrode site (Fig 1, left column; Fig 2, Fz electrode). With the 600 ms epoch used for display, the obligatory responses to two tones can be seen, based on the 300 ms onset-to-onset rate of presentation. Thus, the second positive peak peaking at approximately 400 ms is a P1 evoked by the next tone. The obligatory N2 component that is a hallmark of the child obligatory response, is greatly reduced at this fast stimulus rate (Sussman, Steinschneider, Gumenyuk, Grushko, & Lawson, 2008)<sup>2</sup>.

P1 amplitude and latency factors were compared separately for child groups and adults. P1 latency decreased with age (87 vs. 107 ms) (main effect of group,  $F_{1,18} = 16.17, p < 0.001$ ), consistent with other developmental studies (Gilley, Sharma, Dorman, & Martin, 2005; Ponton, Eggermont, Kwong, & Don, 2000; Shafer, Morr, Kreuzer, & Kurtzberg, 2000; Sharma et al., 2005; Sussman et al., 2008; Wunderluch & Cone-Wesson, 2006), and was

<sup>2</sup>The target-based N2 component, evoked only by deviants in active listening conditions will hence be referred to as the target N2 component to distinguish it from the obligatory N2 that is evoked by both standard and deviant sounds.

shorter when the sounds were attended (92 vs. 102 ms, main effect of attention,  $F_{1,18}=9.67$ ,  $p<0.01$ ), with no significant interactions among factors. P1 amplitude had a fronto-central distribution (main effect of electrode,  $F_{2,36}=126.97$ ,  $\epsilon=0.72$ ,  $p<0.001$ ), with post hoc calculations showing that Pz (2.2  $\mu\text{V}$ ) was significantly smaller in amplitude than Fz (6.2  $\mu\text{V}$ ) and Cz (5.9  $\mu\text{V}$ ). The amplitude at Fz and Cz did not significantly differ. However, there was an interaction between electrode and attention ( $F_{2,36}=13.51$ ,  $\epsilon=0.67$ ,  $p<0.001$ ). Post hoc analysis showed that P1 amplitude was larger to the attended tones but only at Fz (6.6 vs. 5.8  $\mu\text{V}$ ). When attended, there was a stronger frontal distribution ( $Fz>Cz>Pz$ ) compared to when the tones were ignored ( $Fz=Cz>Pz$ ). There was no main effect of group ( $p=0.33$ ) or of attention ( $p=0.32$ ) and no interactions among factors.

In adults, there were no main effects and no interactions on P1 latency as a function of attention or intensity (mean peak latency varied between 50–56 ms across conditions). The P1 amplitude was larger to louder than softer tones (0.91 vs. 0.54  $\mu\text{V}$ , respectively), with a fronto-central scalp distribution (smallest at Pz), but these differences did not reach statistical significance. There were no significant main effects of attention ( $p>0.16$ ), intensity ( $p>0.06$ ), or electrode ( $p>0.09$ ), and no significant interactions on P1 amplitude.

The obligatory response patterns for latency and scalp topography observed across the age groups in the current study is consistent with finding from previous studies testing the same age groups (Gilley et al., 2005; Ponton et al., 2000; Shafer et al., 2000; Sharma et al., 2005; Sussman et al., 2008; Wunderlich & Cone-Wesson, 2006). This indicates that any higher level differences observed in the MMN/N2/P3 components cannot be attributed to differences occurring at the basic level of response indexed by P1.

**3.2.2 Passive and active target detection components (MMN/N2/P3)**—Table 3 shows the mean amplitude of the difference waveforms in the time interval that the ERP components were measured and examined statistically. Figure 3 displays the difference waveforms (deviant ERP minus the control standard ERP), focusing on the MMN, N2, and P3 components. A clear MMN component can be seen in the Passive Frequency (Fig 3A, top row) for the two child groups (Table 3 for significance testing). In contrast, MMN was absent in the younger child group, but was observed in the older children and adults in the Passive Intensity conditions (Fig. 3B, second and third rows). In further contrast, MMN was present in the Active Intensity conditions in all age groups (Fig 3C, bottom two rows). The MMN and N2, both evoked by the deviant tone in the active listening conditions, overlap at the frontal electrode sites when the louder sound was the target. The target N2 evoked by the deviant is the largest negative peak deflection in the epoch seen at Fz, with a peak latency varying from 165–205 ms (Fig 3C). The inversion at the mastoid (indicated with an arrow) helps delineate the MMN, and distinguishes the MMN from the target N2 in the active conditions, which does not invert at the mastoid (Novak et al., 1990). The target N2 component is followed by a partly overlapping target P3 component. The target P3, evoked by the deviant in the active conditions, is best observed at the Pz electrode (Fig. 3C, red traces). The partial overlap of N2 and P3 components in children is consistent with the results of Sussman & Steinschneider (2009), who also found this overlap for the two target-detection components with stimuli presented at a rapid rate.

In summary, the main finding of the study was that when the sounds were unattended, MMN was elicited by frequency deviants in 6–9-year-olds and 10–12-year-olds and by intensity deviants in 10–12-year-olds, whereas MMN to intensity deviants was absent in 6–9-year-olds. In contrast, when the sounds were attended and the louder or softer intensity oddball sound was the primary target, MMN was elicited by the intensity deviants in both 6–9-year-olds and 10–12-year-olds.

### 3.2.3 Group comparison of MMN amplitude and latency

**3.2.31 MMN amplitude:** To compare the MMN amplitude evoked by softer and louder deviants across the three age groups, mixed model ANOVA was used with factors of group (younger child, older child, adult), intensity (louder, softer), and electrode (left and right mastoid) in the Active conditions only (i.e., where MMN was elicited by both intensity deviants in all three age groups). The mastoid amplitudes (LM and RM) were used to compare MMN because a “true” measure of the MMN amplitude cannot be obtained at the Fz electrode due to overlap with the N2 component in the Active conditions (Fig. 3). The mastoid provides a measure of the MMN amplitude (i.e., the polarity inversion generated by the dipoles within auditory cortices), not overlapped by the N2 component. There was no group main effect on MMN amplitude ( $F_{2,27}=1.39$ ,  $p=.27$ ) (mean amplitude at the mastoids in the young:  $0.78\mu\text{V}$ ; old:  $1.37\mu\text{V}$ , and adults:  $1.62\mu\text{V}$ ) in this analysis. There was a main effect of intensity ( $F_{1,27}=4.98$ ,  $p=.034$ ), with the mastoid amplitude significantly larger to the softer ( $1.61\mu\text{V}$ ) than the louder ( $0.91\mu\text{V}$ ) deviants. There was also a main effect of electrode ( $F_{1,27}=22.37$ ,  $p<0.001$ ), due to the amplitude at RM ( $1.60\mu\text{V}$ ) being significantly larger than at LM ( $0.92\mu\text{V}$ ). A significant interaction of electrode with group ( $F_{2,27}=11.75$ ,  $p<0.001$ ) showed that the larger RM than LM was only found for the child groups; there was no significant amplitude difference at the mastoids for adults. There were no other significant interactions among factors.

**3.2.32 MMN latency:** The MMN elicited by the louder deviant peaked earlier than that to the softer deviant (138 vs. 166 ms; main effect of intensity,  $F_{1,18} = 47.29$ ,  $p < 0.001$ ).

### 3.2.4 Group comparison of the active target detection components (N2/P3) amplitude and latency

**3.3.41 N2/P3 amplitude:** In addition to automatic change detection (indexed by MMN), active detection of intensity increments elicited target N2 and target P3 components in all groups (Table 3). Comparing the amplitude of the target N2 component with adults, there were no significant main effects of group ( $p>0.8$ ) or intensity ( $p>0.50$ ), and no interactions ( $p>0.20$ ). In contrast, the P3 component was significantly larger in the older children ( $10.51\mu\text{V}$ ) than the younger children ( $5.43\mu\text{V}$ ) and the adults ( $5.27\mu\text{V}$ ) (main effect of group: ( $F_{2,27}=5.04$ ,  $p=0.014$ ). There was no main effect of intensity on the P3 amplitude ( $p>0.24$ ), and no interaction between factors ( $p>0.14$ ).

Comparing the two child groups, the N2 component had a frontal distribution in 6–9-year and 10–12-year-olds (main effect of electrode:  $F_{2,36} = 20.02$ ,  $\epsilon=.73$ ,  $p < 0.001$ ), with the amplitude largest at the Fz electrode ( $Fz>Cz=Pz$ ). There was an interaction between electrode and intensity ( $F_{2,36} = 8.93$ ,  $\epsilon=.58$ ,  $p < 0.001$ ), with post hoc analysis showing no difference between loud and soft N2 amplitude at Fz, but more negative amplitude for soft at Cz and Pz. The overall amplitude was more negative for the softer deviants (main effect of intensity,  $F_{1,18} = 7.48$ ,  $p = 0.01$ ). There was no child group effect ( $p=0.56$ ) and no other interactions.

For target P3, amplitude was largest at the Pz electrode in both 6–9 and 10–12-year-old children (main effect of electrode:  $F_{1,18} = 21.96$ ,  $p < 0.001$ ). P3 amplitude was largest at Pz, but with a stronger parietal distribution in the older group, as Pz amplitude was larger than Cz in the older but not in the younger group (intensity  $\times$  electrode  $\times$  group interaction,  $F_{5,90} = 5.95$ ,  $\epsilon=.68$ ,  $p < 0.001$ ).

**3.2.42 N2/P3 Latency:** N2 latency was not compared due to overlap at frontal electrodes with the MMN component. The P3 component peaked earlier to the louder (252 ms) than



softer (355 ms) deviants (main effect of intensity,  $F_{1,9} = 361.73$ ,  $p < 0.001$ ), with no group difference in latency ( $p=0.37$ ) and no interaction of intensity and group ( $p=0.45$ ).

#### 4. DISCUSSION

Neural indicators of active change detection paralleled the behavioral ability to detect and respond to the louder and softer sounds. When loudness was a relevant behavioral cue, MMN was elicited in both younger and older age groups of children, as were the higher-level, non-modality-specific, attention-based indicators of target detection (N2/P3) (Picton, 1992). In contrast, when loudness wasn't a relevant behavioral cue (in the passive condition), the same intensity decrements and increments evoked no automatic neural indicator of change detection in younger (6–9-year-olds) children (i.e., no MMN). This finding is in contrast to the presence of an MMN index for automatic change detection along the frequency dimension in this age group. Frequency deviants elicited a large amplitude MMN ( $> 4\mu\text{V}$  observed at the Fz electrode) in the passive oddball condition. These results were also in contrast to the results of the older children, in that, in 10–12-years-old, significant MMNs were elicited by both frequency and intensity deviants in the passive conditions. Thus, we found a difference in the maturational trajectory for automatic sound change detection for frequency and intensity in children.

These results are consistent with our previous study showing that 1) attention modulates task-specific physiologic responses that are concordant with behavioral detection of deviants in complex environments; and 2) that attention-based physiologic responses modulate automatic physiologic processes in children (Sussman & Steinschneider, 2009). Moreover, the results indicate that the developmental time course of automatic intensity level detection is delayed relative to frequency detection. This may be due to a difference in the way frequency information is used in forming auditory objects, as well as in the coding mechanisms for the different tone attributes. For instance, frequency tuning in core areas of auditory cortex occurs along spatially organized maps that is not present for intensity tuning (Recanzone, Guard, & Phan, 2000; Takahashia, Nakaoa, & Kaga, 2005).

A difference between active and passive physiologic processing was previously found between older children (10–12-year-olds) and adults in a study of auditory stream segregation ability (Sussman & Steinschneider, 2009). However, no difference was found for simple feature detection (i.e., the intensity oddball), which is consistent with the results of the current study in these age groups. Thus, the current results, along with our previous results, suggest that the developmental time course of automatic sound processing differs among various simple sound attributes (e.g., frequency versus intensity) and when more complex acoustic features are examined. Moreover, the results suggest an important role for attention in developing automatic sound processing capabilities during development. Even though behavior may appear to be similar in children as adults, the automatic processes supporting the higher-level skills continue to develop through adolescence.

For frequency MMN, there have been many studies conducted in children of all ages demonstrating passive elicitation of MMN to considerably smaller frequency differences than used in the current study (e.g., 20%  $\Delta f$ , Shafer et al., 2000; Morr, Shafer, Kreuzer, & Kurtzberg, 2002; Uwer, Albrecht, & von Suchodoletz, 2002). Thus, the size of the difference for frequency or intensity was not likely an explanation as to why MMN was elicited in the younger group for frequency but not for intensity. Moreover, the crucial findings of the study pertained to the intensity dimension. In one case, the between-subjects factor showed differential processing to the same 15 dB intensity difference as a function of age; and in the other case, the within-subjects factor showed in the younger group that only with attention to detect the deviant sounds was MMN evoked. Thus, the crucial comparisons

were not based on the size of the stimulus difference, as they were identical in these comparisons

Despite the fact that both deviations were suprathreshold with respect to their dimension: 15 dB difference in intensity value and 31 ST difference in frequency separation, the size of the MMN evoked by the loudness change was significantly smaller than that evoked by unattended frequency deviants. The amplitude of the MMN response in children does not appear, on the face of it, to reflect the degree of saliency indicated by the physical difference between the standard and deviant sounds.

Detecting the softer sounds was significantly more difficult for both age groups of children, which was indicated by a lower hit rate and longer reaction time. One possibility is that the softer sounds were confused with silence (e.g., as longer ISI), which could occur with a fast pace where there is not enough time to fully contemplate each sound individually.

Alternatively, it is possible that in general loudness decrements are harder to detect than increments, as this difference is also observed in adulthood (Cusack & Carlyon, 2003).

## 5. CONCLUSIONS

Younger children process loudness changes differently than pitch changes when they are engaged in another task such as watching a video. In 6–9-year-old children, intensity deviants were physiologically detected only when they were behaviorally relevant, despite the fact that frequency deviants were physiologically detected without task relevance.

However, by age 10 years, physiologic deviance detection was automatic for both frequency and intensity changes. These results suggest two important characteristics of development.

1) Attention and performance of auditory-based tasks modulates neural activity associated with automatic processes; and 2) Neural processing of different sound attributes undergoes distinct developmental trajectories. These results thus suggest that even though we may be born with complex acoustic processing capabilities, experience and cortical maturation continue to shape the neural processes underlying these perceptual abilities throughout development. The current results also suggest that neural frequency tuning is more highly specified at earlier developmental stages, whereas intensity coding becomes more highly specified after a greater degree of experience and maturity. Neural processes become optimized for sounds we frequently encounter, and influence the acuity with which we automatically engage in the sound environment. Thus, even though basic sound processing mechanisms for scene analysis may be initiated at birth, the fine tuning seen in adulthood is a sequence of developmental processes that progressively alters our perceptual apparatus. It is likely that our specific experience with sounds and maturation of the auditory system helps drive these maturational changes. When considering the young child who doesn't respond to the louder calling of their name when engaged in another activity, current data suggest that the automatic system is not detecting the difference in loudness.

### Research Highlights

>>loudness changes physiologically detected only when behaviorally relevant in younger children >> pitch changes physiologically detected without task relevance in younger and older children >> automatic auditory processes continue to become refined through childhood >>neural processing of sound frequency and sound intensity undergo distinct developmental trajectories

## Acknowledgments

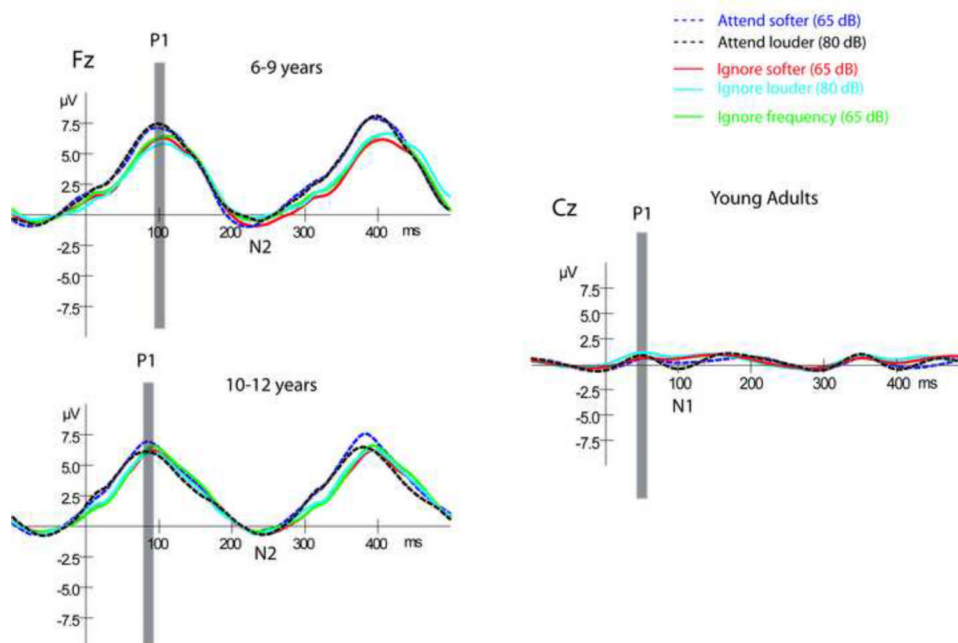
This research was funded by the National Institutes of Health (grant #DC006003). We thank Dr. Katherine Lawson for performing and analyzing the psychometric tests. We thank Jean DeMarco for recruiting and scheduling the participants and Wei Wei Lee for collecting the data.

## References

- Alho K. Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and Hearing*. 1995; 16(1):38–51. [PubMed: 7774768]
- Berg KM, Boswell AE. Noise increment detection in children 1 to 3 years of age. *Perception & Psychophysics*. 2000; 62(4):868–873. [PubMed: 10883590]
- Bregman AS, Ahad PA, Crum PA, O'Reilly J. Effects of time intervals and tone durations on auditory stream segregation. *Perception & Psychophysics*. 2000; 62(3):626–636. [PubMed: 10909253]
- Bregman AS, Campbell J. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*. 1971; 89(2):244–249. [PubMed: 5567132]
- Carlyon RP, Cusack R, Foxton JM, Robertson IH. Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology. Human Perception and Performance*. 2001; 27(1):115–127. [PubMed: 11248927]
- Cusack R, Carlyon RP. Perceptual asymmetries in audition. *Journal of Experimental Psychology. Human Perception and Performance*. 2003; 29(3):713–725. [PubMed: 12848335]
- Eramudugolla R, McAnally K, Martin R, Irvine D, Mattingley J. The role of spatial location in auditory search. *Hearing Research*. 2008; 238:139–146. [PubMed: 18082346]
- Gilley PM, Sharma A, Dorman M, Martin K. Developmental changes in refractoriness of the cortical auditory evoked potential. *Clinical Neurophysiology*. 2005; 116:648–657. [PubMed: 15721079]
- Ihlefeld A, Shinn-Cunningham B. Disentangling the effects of spatial cues on selection and formation of auditory objects. *The Journal of the Acoustical Society of America*. 2008; 124(4):2224–2235. [PubMed: 19062861]
- Miller LM, Recanzone GH. Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106(14):5931–5935. [PubMed: 19321750]
- Morr ML, Shafer VL, Kreuzer JA, Kurtzberg D. Maturation of mismatch negativity in typically developing infants and preschool children. *Ear and Hearing*. 2002; 23(2):118–136. [PubMed: 11951848]
- Muller D, Widmann A, Schröger E. Auditory streaming affects the processing of successive deviant and standard sounds. *Psychophysiology*. 2005; 42(6):668–676. [PubMed: 16364062]
- Novak GP, Ritter W, Vaughan HG Jr, Wiznitzer ML. Differentiation of negative event-related potentials in auditory discrimination task. *Electroencephalography and Clinical Neurophysiology*. 1990; 75:255–275. [PubMed: 1691075]
- Perrault N, Picton TW. Event-related potentials recorded from the scalp and nasopharynx, II, N2, P3 and slow wave. *Electroencephalography and Clinical Neurophysiology*. 1984; 59(4):261–278. [PubMed: 6203716]
- Picton TW. The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society*. 1992; 9(4):456–479. [PubMed: 1464675]
- Ponton CW, Eggermont JJ, Kwong B, Don M. Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical Neurophysiology*. 2000; 111(2):220–236. [PubMed: 10680557]
- Recanzone GH, Guard DC, Phan ML. Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *Journal of Neurophysiology*. 2000; 83(4):2315–31. [PubMed: 10758136]
- Recanzone GH, Schreiner CE, Merzenich MM. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *The Journal of Neuroscience*. 1993; 13(1):87–103. [PubMed: 8423485]

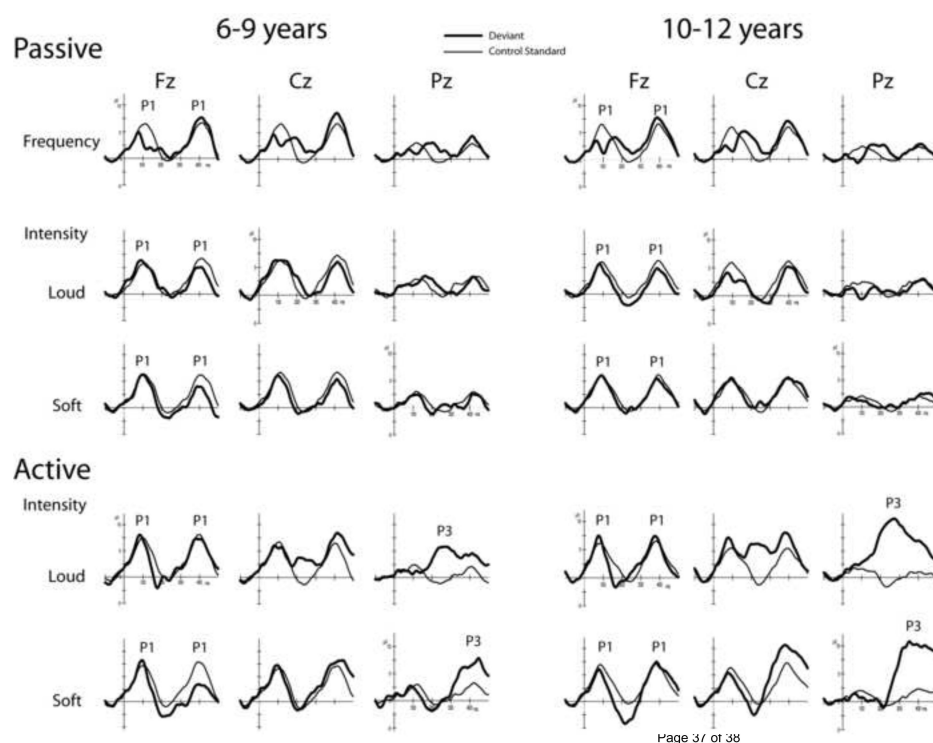
- Sadagopan S, Wang X. Level invariant representation of sounds by populations of neurons in primary auditory cortex. *The Journal of Neuroscience*. 2008; 28(13):3415–3426. [PubMed: 18367608]
- Shafer VL, Morr ML, Kreuzer JA, Kurtzberg D. Maturation of mismatch negativity in school-age children. *Ear & Hearing*. 2000; 21:242–251. [PubMed: 10890733]
- Sharma A, Martin K, Roland P, Bauer P, Sweeney MH, Gilley P, Dorman M. P1 latency as a biomarker for central auditory development in children with hearing impairment. *Journal of the American Academy of Audiology*. 2005; 16(6):564–573. [PubMed: 16295243]
- Shinn-Cunningham BG, Lee AK, Oxenham AJ. A sound element gets lost in perceptual competition. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104(29):12223–12227. [PubMed: 17615235]
- Sinnott JM, Aslin RN. Frequency and intensity discrimination in human infants and adults. *Journal of the Acoustical Society of America*. 1985; 78(6):1986–1992. [PubMed: 4078175]
- Snyder JS, Alain C, Picton TW. Effects of attention on neuroelectric correlates of auditory stream segregation. *Journal of Cognitive Neuroscience*. 2006; 18(1):1–13. [PubMed: 16417678]
- Sussman ES. Integration and segregation in auditory scene analysis. *The Journal of the Acoustical Society of America*. 2005; 117(3 Pt 1):1285–1298. [PubMed: 15807017]
- Sussman E. A new view on the MMN and attention debate: Auditory context effects. *Journal of Psychophysiology*. 2007; 21(3–4):164–175.
- Sussman E, Ritter W, Vaughan HG Jr. An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiology*. 1999; 36(1):22–34. [PubMed: 10098377]
- Sussman E, Steinschneider M. Attention effects on auditory scene analysis in children. *Neuropsychologia*. 2009; 47(3):771–785. [PubMed: 19124031]
- Sussman E, Steinschneider M, Gumenyuk V, Grushko J, Lawson K. The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hearing Research*. 2008; 236(1–2):61–79. [PubMed: 18207681]
- Takahashia H, Nakaoa M, Kaga K. Interfield differences in intensity and frequency representation of evoked potentials in rat auditory cortex. *Hearing Research*. 2005; 210:9–23. [PubMed: 16213681]
- Uwer R, Albrecht R, von Suchodoletz W. Automatic processing of tones and speech stimuli in children with specific language impairment. *Developmental Medicine and Child Neurology*. 2002; 44(8):527–532. [PubMed: 12206618]
- Wunderlich JL, Cone-Wesson BK. Maturation of CAEP in infants and children: a review. *Hearing & Research*. 2006; 212(1–2):212–223.

## STANDARD WAVEFORMS

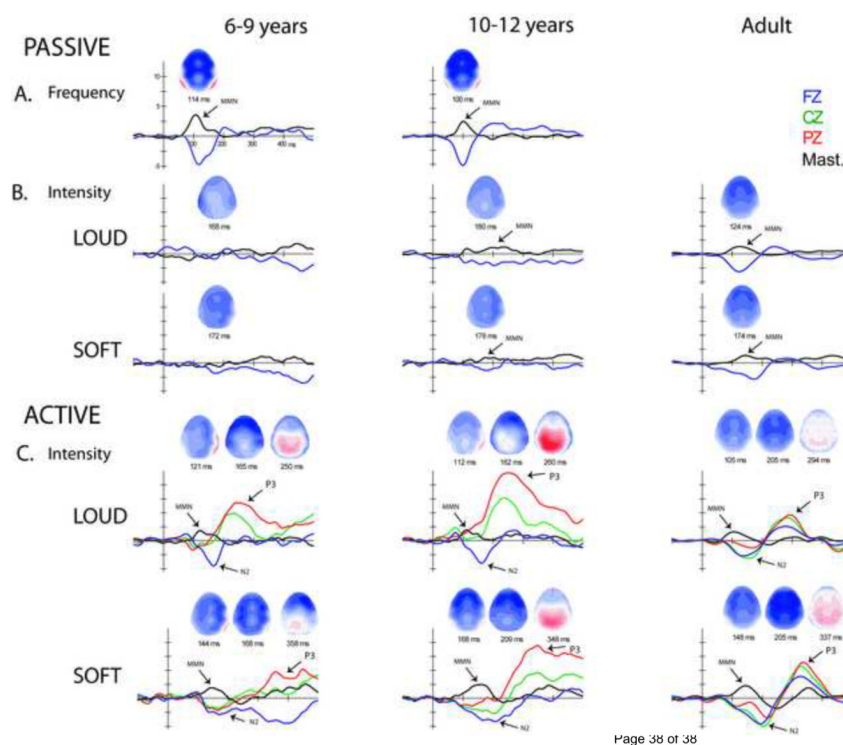
**Fig. 1.**

Standard waveforms for all three age groups. Event-related brain potentials (ERPs) evoked by the standard ( $p=.88$ ) tones in the oddball sequence are displayed for each condition and age group. Waveforms for children are displayed at the Fz electrode (left column), where the obligatory responses have the largest signal to noise ratio, and are displayed where the signal is largest in adults, at Cz (right column). The dashed lines depict the waveforms in the Active conditions, softer sound (dashed purple line) when the deviant was the louder sound, and louder sound (dashed black line) when the deviant was the softer sound. The solid lines depict the waveforms in the Passive conditions: The softer standard (red line) when the deviant was louder; the louder standard (light blue) when the deviant was softer; and softer sound (green line) when frequency was deviant. The latency of the P1 component is marked with a gray bar, illustrating the significantly shorter latency with maturation. Note the difference in amplitude of the overall waveforms in adults compared to children, especially at this rapid (300 ms onset-to-onset) stimulus presentation rate. There is no frequency oddball condition for adults.



**Fig. 2.**

ERPs evoked by the deviant (thick solid line) are overlain with their respective control standard (thin solid line) in the younger (left three columns) and older (right three columns) children in all conditions: Passive Frequency (top row), Passive Intensity (middle two rows), and Active Intensity (bottom two rows) conditions. The P1 component evoked by both deviant and standard tones is best seen at Fz and is labeled. The epoch displays the tone onset response to two tones, concordant with the 300 ms stimulus presentation rate. The P3 component, evoked during active detection, is clearly seen at the parietal electrode (Pz) and is also labeled.

**Fig. 3.**

Difference waveforms. MMN, N2b and P3b components evoked by the deviants are delineated in the difference waveforms (deviant-minus-control standard) and are labeled with arrows where they were statistically significantly present for the younger (left column) and older (middle column) children, and for the Adults (right column). A. Passive Frequency condition (top row) shows a robust MMN in both age groups, with a clear inversion at the mastoid. B. Passive Intensity conditions (middle two rows) show MMNs only in the older group (middle column) and adults (right column). C. Active Intensity conditions (bottom two rows) show MMNs, and active target detection components N2 and P3 in all age groups. Fz (blue line), Cz (green line), Pz (red line), and the mastoids (black line) show the topography of the three components. Due to overlap with N2 in active conditions, MMN is delineated by the inversion at the mastoid (arrows), whereas the peak of the N2 is seen at Fz (blue trace). Voltage maps computed at the latency used for statistical measurements, shows the scalp voltage distribution for each component. MMN has a more frontal distribution, N2 a more central distribution, and P3 a more parietal distribution in all age groups. There is no Frequency oddball condition for adults.

Table 1

Participant Information

Age	6	7	8	9	10	11	12	22–46	Total
<i>n</i>	1	2	5	2	4	1	5	10	30
Gender	1f	1f	3f	2f	2f	1f	3f	4f	17f

Table 2

Behavioral Results

Age Group	Reaction Time (ms)		Hit Rate		False Alarm Rate	
	Louder	Softer	Louder	Softer	Louder	Softer
6–9 years	398 (41)	492 (63)	.85 (.09)	.60 (.17)	.02 (.032)	.03 (.026)
10–12 years	320 (48)	394 (37)	.92 (.03)	.82 (.05)	.006 (.003)	.016 (.009)
22–46 years	281 (56)	334 (61)	.96 (.02)	.94 (.05)	.003 (.002)	.002 (.002)

**Table 3**

Mean amplitudes of MMN, N2b, and P3b components (SD in parentheses) for Passive and Active conditions in all age groups.

AGE	Component	Frequency	Passive		Active	
			Intensity		Intensity	
			Louder	Softer	Louder	Softer
6-9 years	MMN	Fz	-3.88 (1.2)**	-0.56 (1.0)	-1.59 (3.2)	-2.38 (2.7)**
		RM	2.72 (1.69)**	0.03 (1.5)	1.32 (1.8)*	1.61 (2.2)*
	N2	Cz	•	•	-3.69 (3.1)**	-2.70 (2.8)**
	P3	Pz	•	•	6.64 (3.4)**	4.21 (4.8)**
	MMN	Fz	-4.18 (3.2)**	-1.75 (2.4)*	-1.17 (2.6) <sup>+</sup>	-3.27 (3.2)**
		RM	1.98 (2.3)**	0.85 (1.4)*	1.04 (1.2)**	2.28 (1.9)**
10-12 years	N2	Cz	•	•	-3.34 (2.8)**	-3.89 (3.5)**
	P3	Pz	•	•	11.96 (6.3)**	9.05 (7.3)**
	MMN	Fz	—	-3.00 (1.1)**	-2.71 (2.1)**	-2.85 (1.99)**
		RM	—	1.28 (0.51)**	1.06 (0.90)**	1.62 (0.86)**
	N2	Cz	—	•	•	-4.72 (3.3)**
	P3	Pz	—	•	4.43 (2.7)**	6.12 (4.2)**

\*\*  
 $p \leq 0.01$

\*  
 $p \leq 0.05$

<sup>+</sup>  
 $p \leq 0.10$  (one-sample  $t$ -test against zero)

• these components not elicited in Passive conditions

— no data