Article

A Longitudinal Investigation of Infant Auditory Sensitivity

Anne Marie Tharpe
Daniel H. Ashmead
Vanderbilt Bill Wilkerson Center for Otolaryngology and Communication Sciences, Vanderbilt University Medical Center, Nashville, TN

The behavioral evaluation of hearing in very young infants has been fraught with procedural and interpretive problems. Despite the introduction of current physiological techniques of estimating hearing sensitivity, such as otoacoustic emissions and auditory brainstem-evoked responses, behavioral hearing assessment of young infants remains of interest to researchers of infant behavior and to clinicians who need to use a battery of tests in their assessment of infant hearing. The objective of this study was to provide the first longitudinal investigation of infant auditory sensitivity, using a new procedure for behavioral testing of neonates and infants. Behavioral responses to speech noise stimuli were obtained monthly from birth to 12 months of age. During each trial, the signal increased from an inaudible level in 2-dB steps until the infant responded. Therefore, a threshold estimate was obtained on each trial, and the average threshold could be computed across trials within a test session. Threshold estimates were in good agreement with previously reported infant behavioral thresholds based on cross-sectional designs. The age-related changes in threshold were fit with exponential functions for individual infants and for the group data. There was good agreement in the shape of these functions across infants, with asymptotic threshold level approached around 6 months of age. Therefore, this longitudinal study confirms that the age trend previously reported from cross-sectional findings is also observed in the development of individual infants.

Key Words: infant hearing, auditory development, longitudinal

---

Studies of developmental trends in infant hearing sensitivity have tended to be based on group levels of analysis, using cross-sectional designs or single age groups (Nozza & Wilson, 1984; Olsho, Koch, Carter, Halpin, & Spetner, 1988; Sinnott, Pisoni, & Aslin, 1983; Trehub, Schneider, & Endman, 1980; Trehub, Schneider, Thorpe, & Judge, 1991; Werner & Marean, 1991). Some investigators have used the method of constant stimuli with thresholds estimated from the group psychometric function. Others have obtained individual thresholds using adaptive psychophysical procedures and have averaged thresholds across infants. These studies have made significant contributions to our understanding of the development of hearing. For example, studies of absolute sensitivity (detection of sounds in quiet) based on group averages have shown rapid improvement in auditory thresholds, especially for high frequencies, between the neonatal period and approximately 6 months of age (Olsho, Koch, Carter, Halpin, & Spetner, 1988; Weir, 1979; Werner & Gillenwater, 1990). In fact, these studies demonstrate an infant-adult auditory threshold difference on the order of 20 to 70 dB in very young infants.

It is also of interest, however, to examine individual patterns of development. That is, because age trends have been obtained using group levels of analysis, it remains unclear whether these trends accurately reflect those for individual infants. As Werner and Marean (1991) note, there is a lack of longitudinal studies in auditory development and a reliance on cross-sectional observations. From a research perspective, methodological differences across studies and age groups could certainly contribute to observed differences in auditory behavior. From a clinical perspective, one wants to know how much variability in age trends there is within the population of normal-hearing infants. Despite the accuracy and efficiency of current physiological techniques available for assessment of hearing in young infants, clinicians still have cause to use behavioral assessment techniques. This is true especially when clinicians have reason to believe that physiological techniques might be limited as a result of the presence of certain pathologies (e.g., otitis media or auditory neuropathy). In such cases, the addition of behavioral results can improve the likelihood of accurate auditory threshold estimation.

In summary, the purpose of this study was to obtain individual, longitudinal measures of absolute hearing sensitivity for the first year of human infancy using a method of ascending limits for threshold estimation. Although the procedure developed for this study is not immediately practical for clinical purposes, it is appropriate for purposes of obtaining longitudinal data.

Methods

Participants

Seven infants (4 girls, 3 boys) participated in testing sessions at approximately monthly intervals. Testing was attempted during the newborn period, but the ages at first successful testing (i.e., the infant stayed awake through a series of trials) ranged from 29 to 84 days ($M = 53$, $SD = 22$). Before those ages, infants tended to be too sleepy to obtain useful data. Testing continued through about 12 months of age. All infants had full-term births (gestational ages $\geq$ 38 weeks), normal medical status by parental report, and no family history of hearing loss. They were screened for middle ear disorder on the days of threshold testing. For infants $\leq$7 months of age, tympanometry was obtained and attempts were made to elicit an acoustic reflex using a 660-Hz probe for a 90-dB 1-kHz pure tone. For infants greater than 7 months of age, tympanometry was conducted. An abnormal tympanogram (those with no measurable peak or peak pressure < -150 mmH_2O) or an absent reflex in either ear resulted in cancellation of the test session for infants $\leq$6 months of age. An abnormal tympanogram in one or both ears resulted in cancellation of the test session for the older infants. In addition to the seven infants who completed testing, four infants were enrolled but eventually eliminated from the study as a result of repeated tympanometric screening failures. It was not possible to obtain data at each month for each infant because of illness, fussiness during testing, or family scheduling complications. The mean number of test sessions was 8.0 ($SD = 2.0$) with a range of 5 to 10. In addition to the infant findings, data were collected from two young adults, one man and one woman, to provide a direct adult comparison.

Apparatus and Procedure

Testing occurred in a specially built 2 × 4-meter sound booth connected to an adjoining control room by two sound-insulated doors. In order to reduce visual distraction, the sound booth was completely dark during testing, except for a small flashing light and occasional video images as described below. The infant was held on the parent's lap, and loudspeakers (Panasonic 4-inch speaker cones) were located 90° to each side at a distance of 40 cm. The loudspeaker response was flat (±8 dB) between 0.25 and 14 kHz. Two video monitors (Panasonic CT13824) were located 45° to each side at a distance of 60 cm. Video images of the Disney film *Fantasia* could be shown to the infant on either monitor. A flashing light-emitting diode (LED) was positioned 75 cm directly in front of the infant, to keep the infant facing forward. This light continued flashing throughout the test period. An infrared-sensitive Panasonic WW-CD20 video camera was positioned just above the LED, with infrared light provided by two Kodak Model A darkroom lamps with 15-watt bulbs and Kodak No. 11 Safelight filters mounted above and behind the video monitors. The camera provided a close-up view of the infant's face, seen on a monitor in the control room.

The test protocol was implemented by a Zenith 386 computer. One experimenter ran all test sessions and voted on the presence of infant responses. On each trial the stimulus began at a low (inaudible) sound level and gradually increased during the trial. An experimenter in the control room watched the infant's image on the video monitor, especially noting orienting movements of the eyes and head toward one or the other loudspeaker, eye-widening, and changes in activity level. The experimenter was actually free to focus on any aspect of the infant's behavior and pressed a computer keyboard to indicate when and to which side the infant appeared to detect the stimulus. If the experimenter voted correctly (i.e., a response was observed during a stimulus interval), then the infant was presented with the Fantasia display on the video monitor corresponding to the side of the active loudspeaker. This display continued until the experimenter pressed a key to move on, typically for 3 to 5 s. Except when the *Fantasia* display was on, the monitor screens in the sound booth were completely darkened by a special blanking circuit. This procedure could result in a less than 100% reinforcement schedule for a given infant if the experimenter did not observe a response and vote during the stimulus trial. Although we did not evaluate the effects of reinforcement within this study, we do know that threshold estimates were equivalent whether the experimenter was correct or not (see Results).

The rationale underlying this procedure was that the time at which the experimenter voted corresponded to when the infant responded to sound from the active loudspeaker. The first train of sounds was at the lowest level; then the level was automatically increased by 2 dB for each succeeding train (e.g., 32, 34, 36, . . .). Thus, each trial provided an estimate of the infant's threshold. If the experimenter had not voted by the time the maximum sound level was reached, then the trial ended with no threshold estimation. On trials when the infant clearly became inattentive or off task (e.g., crying, fussing), the experimenter could press a computer key to abort the trial.

The stimulus used for behavioral threshold estimates was a series of trains of three digitally generated 282-ms speech-filtered (i.e., flat response between 0.25 and approximately 1 kHz with 15 dB/octave roll-off through 4 kHz) noise signals, each with 5-ms linear rise/fall times and with 51 ms between signals. Speech noise was used because it encompasses the frequency range of greatest practical interest in auditory development.[1] Two overall sound level ranges were used. For ages younger than ~5 months, the sound level (measured from the infant's position) ranged from 32 to 72 dB SPL (A scale), whereas for older ages it ranged from 18 to 48 dBA SPL. These ranges were based on pilot work and on previously reported estimates of infant hearing sensitivity (Olsho, Koch, Halpin, & Carter, 1987; Werner & Gillenwater, 1990). The stimulus was acoustically calibrated on each day of testing using a sound-level meter (Bruel & Kjaer 2209) with a 1-inch microphone (model 4145) to determine sound pressure levels for both speakers.

The experimenter could not hear the stimuli. To prevent the experimenter from estimating the current sound level based on elapsed time since the start of a trial, different time intervals between successive trains were used. On each trial, the interval was randomly chosen and unknown to the experimenter so as to give a fast, medium, or slow rate of ascent in sound level. For the fast rate of ascent, the time between trains varied randomly (for each intertrain interval within a trial) in the range of 25 to 275 ms. For the medium and slow rates of ascent these times were 275 to 775 and 775 to 1275 ms, respectively. With these temporal parameters in effect, the rates of ascent varied in the range of 0.97 to 2.22 s per increment in sound level. This was sufficient to render elapsed time a crude indicator of the current sound level during a trial.

When feasible, the parent holding the infant was masked from the stimuli via earphones emitting a broad band static noise or earplugs. Pilot work, however, indicated that with the room totally dark, parents needed to be able to hear their infants for reasons of safety and reassurance. Therefore, some parents were not masked for all sessions. All parents were instructed not to move or speak during the test session unless there was a need to stop the testing to comfort the infant. All of the parents understood that they were not to convey to their infants which side the sound was on. The stimuli were audible to the unmasked parents even at the lowest sound level presented. Therefore, if a parent were cueing the infant on the basis of the parent's ability to detect the sound, threshold estimates would have been at the lowest sound level. The results from infants whose parents were unmasked meshed well with the overall age functions for these infants.

Results

Number of Trials

Trials were presented during a test session as long as the infant was cooperative. The mean number of trials per session was 21.0 (SD = 6.4), with a range of 11 to 41. For purposes of these and several other analyses, data were grouped according to 3-month age intervals (1–3, 4–6, 7–9, and 10–12, using 30-day months). The mean number of trials per session and the standard deviations were consistent across the age range (18.9 ± 6.1, 22.3 ± 7.0, 21.3 ± 6.2, 18.1 ± 4.5, respectively).

Age-Related Changes in Thresholds

In principal, each trial produced a threshold estimate because the procedure was a method of limits. However, to obtain a more reliable threshold estimate for each test session, the median threshold across all trials in a session was used. For other approaches considered for obtaining an averaged threshold, see the Appendix. Age-related changes in thresholds are summarized in Figure 1.

The age function is obviously nonlinear in form, with an asymptotic level approached sometime beyond 200 days (~6 months of age) or so. To fit these data, we used an exponential function with a nonzero asymptotic floor level. The best-fitting curve, as shown in Figure 1, was chosen by a least squares criterion: $\text{threshold} = 17.1 + 48e^{(-0.0067 \cdot \text{age})}$. This curve accounts for 78.2% of the variance in threshold estimates, which is significant [$F(1,54) = 194.15$]. The type I error rate was set at $p < .05$ for all analyses reported herein.

Another way of characterizing the changes in thresholds was to group test sessions into 3-month increments. The mean thresholds and standard errors for each age range are illustrated in Figure 2.

As can be seen, the age-related changes level off at ~6 months of age.

Curves were also fit to the individual data, as illustrated in Figure 3. For all of the infants except infant 4, the exponential functions accounted for a significant proportion of the variance in thresholds, according to F tests [for the seven individuals, respectively, $F(1,4) = 34.86$, $F(1,8) = 23.78$, $F(1,3) = 171.82$, $F(1,6) = 2.95$, $F(1,7) = 45.84$, $F(1,7) = 69.55$, $F(1,7) = 14.18$]. It should be noted that these individual functions map well onto the group function in Figure 1.

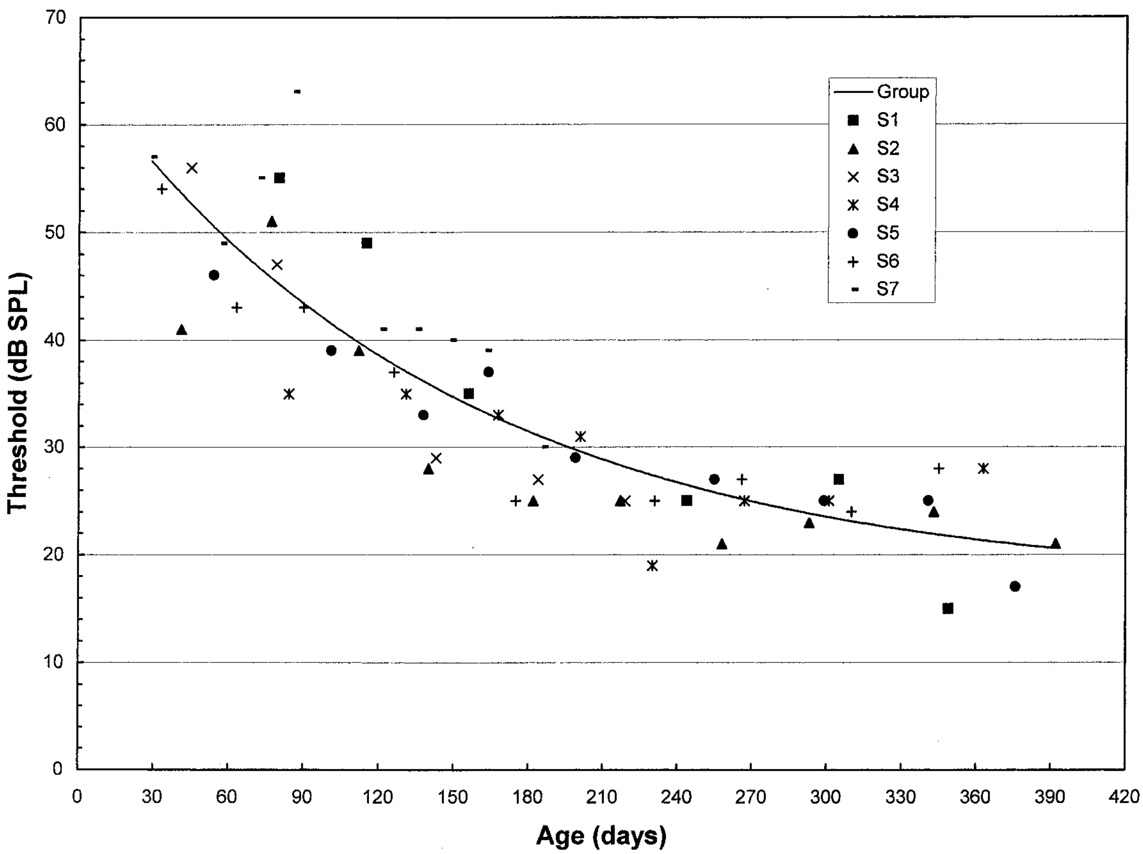
To determine the appropriateness of using an average of just a few test trials to provide individual threshold estimations, we compared the median of the first five trials of each test session with the median of all trials for the same session. The median of the first five trials was within 3 dB of the median of all trials for 88% of trials. Therefore, if we had used only the first five trials in a session to estimate thresholds, our results would have been similar to our reported findings. This suggests that reasonable threshold estimations can be made with only a small number of trials, thus, implying that this procedure could be clinically feasible.

Variability

The variability of the threshold measurements was assessed in two ways. The first was by analyzing the residuals around the best-fitting group age function, that is, the differences between the observed threshold estimates and

[1] The purpose of this experiment was to emphasize individual subject performance; therefore, the stimulus that would most likely result in robust responses was selected. Speech noise has been shown to be more response-eliciting than a narrow band or discrete frequency (Hoversten & Moncur, 1969; Thompson & Thompson, 1972). Given that this procedure holds promise for obtaining threshold estimations with a small number of responses, frequency-specific thresholds should be easily elicited. That was not, however, the purpose of the current study.

FIGURE 1. Threshold estimates by age for all individual test sessions. Thresholds for each subject are designated by unique symbols as shown in the key. The solid line shows the best fitting exponential function for the entire data set, as described in the text.



the estimates that would be predicted from the best-fitting exponential function for the group. This approach allows the "random" part of the variability to be assessed independently of systematic, age-related threshold changes. The difference score for each test session was defined as the observed threshold minus the predicted threshold. For example, if an infant's median threshold was 41.0 on day 112, then

$$\text{Difference score} = 41.0 - [17.1 + 48e^{(-0.0067 \cdot 112)}] = 1.2$$

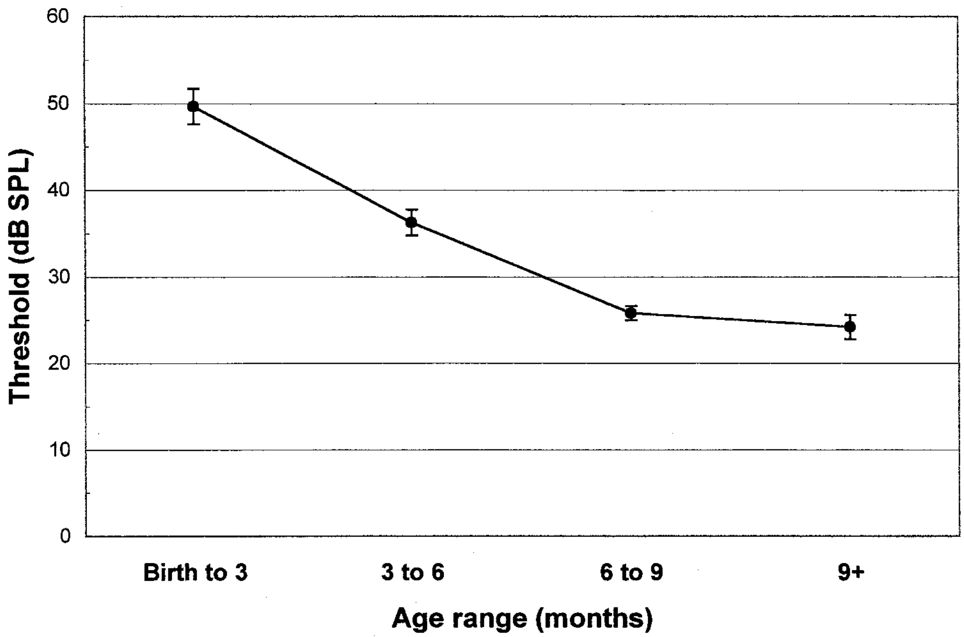
The mean difference score was zero (−0.027 to be exact), as would be expected if the residuals were distributed evenly around the best-fitting line. The standard deviation was 5.4 dB ($SE = 0.7$ dB), and the 10th and 90th percentiles were at −6.8 and +6.3 dB, respectively. The 90% confidence interval was ±6.5 dB, so we can summarize by saying that this testing method appears to yield thresholds for which we can have 90% confidence that they fall within 6.5 dB of the expected value for a particular age.

Second, we looked at within-session variability of threshold estimates across age. This was done by computing the standard error of threshold estimates across all trials in a test session. The magnitude of the standard errors did vary across age ($r = -0.30$, $p < .05$). However, this was largely attributable to three test sessions at younger ages that had standard errors of 3.1, 3.7, and 4.4 dB. The remaining standard errors were all ≤2.6 dB, and the correlation with age with those three test sessions removed was not significant ($r = -0.13$).

Discussion

To examine the developmental course of infant auditory sensitivity, we tested infants individually from approximately birth to 12 months of age and estimated behavioral thresholds at about 1-month intervals. For most infants, a strong relationship between age and threshold was demonstrated with an asymptotic level approached at ~6 months of age. The thresholds obtained with this procedure were comparable to threshold estimates from other studies of infant sensitivity. Despite the fact that differences exist across these studies in terms of procedures, stimuli used, and methods of calculating threshold, reasonably good agreement remains, as can be seen in Figure 4. Values shown in this figure were extrapolated from original published manuscripts. In studies where frequency-specific stimuli were used rather than speech stimuli, individual frequencies or averages of frequencies in the speech frequency range (0.5, 1, and 2 kHz) were used to

FIGURE 2. Mean of threshold estimates for each age range. Error bars designate standard errors.



plot data points. One specific contribution of the present findings is to demonstrate that the age function based on group data applies as well to individual infants. Thus, there seems to be a rather consistent developmental pattern of improvement in auditory sensitivity before 6 months of age.

In addition, we looked at the adult/infant threshold differences as opposed to only looking at absolute thresholds. This has the effect of adjusting for laboratory-to-laboratory variations in thresholds as a result of differences in equipment, procedures, and so forth. Again, as seen in Figure 5, the pattern of our results is in good agreement with other laboratories. That is, auditory responses from infants at ~3 months of age are greater than 20 dB above adult thresholds, whereas responses of 6 months old infants are ~10 to 15 dB higher than that of adults.

The following conclusions can be drawn from these results. First, the procedure described herein is capable of obtaining behavioral threshold data from infants between 1 or 2 and 12 months of age (so far, however, it cannot be used with neonates). The use of one procedure that can be used throughout most of the first year of life is desirable for behavioral longitudinal studies in that it reduces interpretive problems associated with changing test procedures across age. An additional advantage of this procedure is that it appears to require a small number of trials in order to estimate threshold, which is ideal for longitudinal assessment of individual infants. Further, prior knowledge of threshold is not as important with this procedure as with a method of constant stimuli. The present procedure based on the method of ascending limits proved to be workable with infants across a wide age range, but adaptive psychophysical methods such as "staircases" remain an important part of our repertoire for behavioral testing of infants (e.g., Gravel, 1989). One possible advantage of the ascending limits approach is that it tends to concentrate stimulus presentations in the subthreshold range, whereas adaptive methods and the method of constant stimuli are designed to bracket below and above threshold. Gray (1987) suggested that habituation to acoustic events is greatest at suprathreshold levels. Future research might examine whether habituation is less of a problem with the method of ascending limits than with adaptive procedures or the method of constant stimuli. Finally, unlike behavioral observation audiometry (Thompson & Weber, 1974) that tends to produce highly variable threshold estimates, this procedure allows threshold estimation within a confidence limit of ±6.5 dB.

It is interesting to consider observed developmental changes in behavioral responses to sound in the context of research findings in auditory anatomy and physiology. Auditory brainstem response (ABR) studies have shown a trend toward rapid auditory development from the newborn period to ~3 months of age, continuing more gradually to about 12 to 18 months of age (Gorga, Kaminski, Beauchaine, Jesteadt, & Neely, 1989; Hecox & Galambos, 1974; Jiang & Tierney, 1995; Salamy & McKean, 1976). These changes are reflected in reduced latency and threshold values with increasing age. Through ~6 months of age, ABR thresholds are lower than behavioral thresholds (Jacobson & Morehouse, 1984; Ruth, Horner, McCoy, & Chandler, 1983; Werner, Folsom, & Mancl, 1993); however, after 6 months of age the reverse is true (Gorga, Kaminski, Beauchaine, & Jesteadt, 1988). Absolute wave V latency values and interpeak latency differences also decrease as age increases from birth to at least 18 or 24 months (Gorga, Kaminski, Beauchaine, Jesteadt,

FIGURE 3. Individual subject threshold estimates by age with best fitting exponential function for each subject. The proportions of variance accounted for (R^2_{adj}) were, respectively, 0.871, 0.717, 0.977, 0.218, 0.849, 0.895, and 0.622.

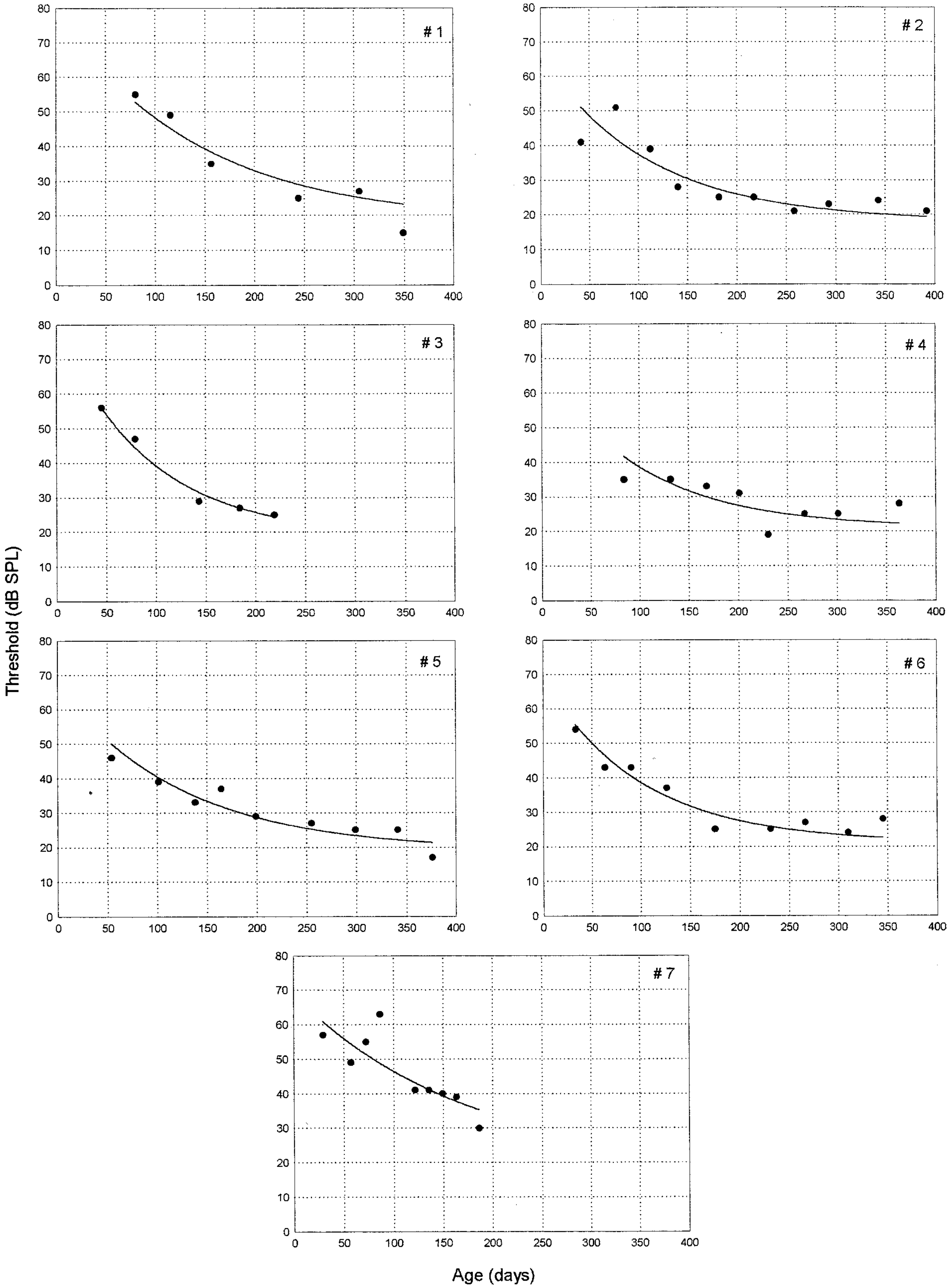
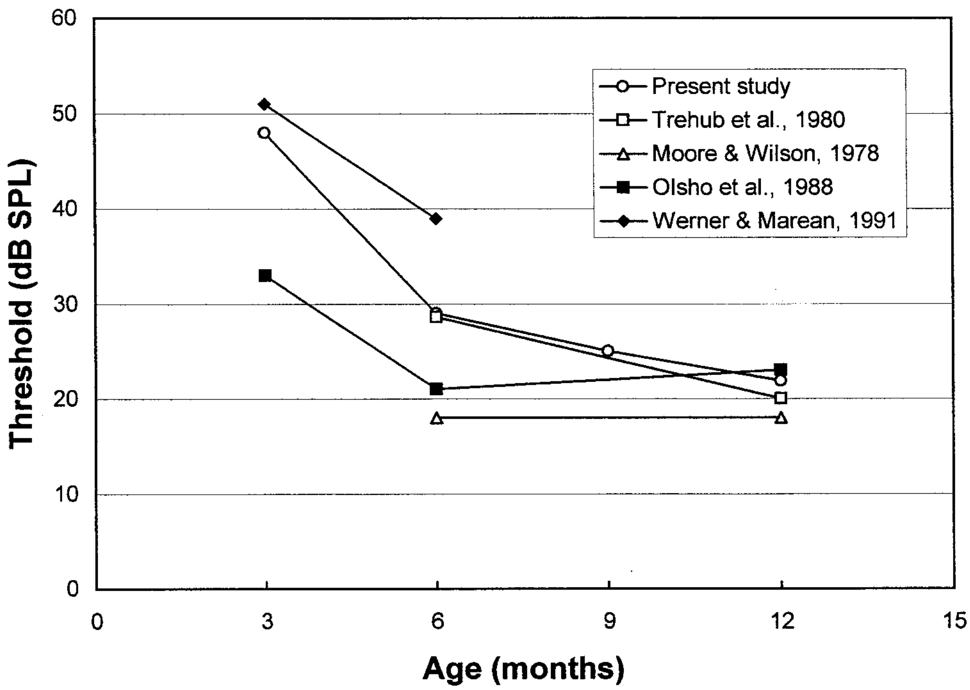
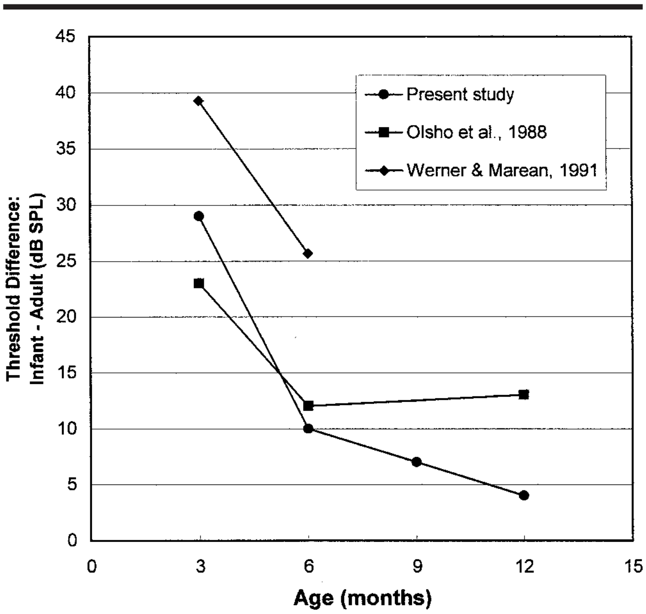


FIGURE 4. Comparison of absolute thresholds from present study to other studies of infant sensitivity.



& Neely, 1989; Jiang & Tierney, 1995). Several accounts have been proposed to explain the pattern of improvement in ABR and behavioral auditory thresholds during infancy (Nozza, 1995; Werner, 1996), emphasizing both sensory and nonsensory factors. Sensory factors include maturation of external and middle ear structures (Keefe, Bulen, Hoberg, & Burns, 1993; Keefe & Levi, 1996) and maturation of primary auditory pathways (Gorga, Kaminski, Beauchaine, Jesteadt, & Neely, 1989; Jiang & Tierney, 1995; Werner, Folsom, & Mancl, 1993), and nonsensory factors include methodological factors and immaturity of listening strategies (Werner, 1996). The test procedure described herein attempts to reduce the impact of nonsensory factors on infant threshold estimation by reducing infant distractibility and estimating threshold quickly before infant habituation. One must leave open the possibility, however, that improvement in nonsensory factors (e.g., attention, motivation, response time) as well as maturation of sensory factors (e.g., increased myelinization, changes in basilar membrane physiology) could have influenced the observed improvement in behavioral thresholds.

FIGURE 5. Comparison of infant/adult threshold differences from present study to other studies of infant sensitivity.



The nature of the procedure requires that the issue of response bias be considered carefully. The possibility of response bias with this procedure was addressed by several safeguards. First, the signal was inaudible to the experimenter. Second, the experimenter was unable to predict the sound level by considering elapsed time from the start of the trial because different time intervals between successive trains were used and were unknown to the experimenter. The low standard errors around the threshold medians suggest that these safeguards were effective in reducing response bias. That is, if the experimenter was using temporal cues to indicate a current sound level during a trial, one would expect to see considerably larger errors around the median because temporal parameters were changing on each trial. An acknowledged limitation of the method was that some parents were allowed to hear the stimuli during many test sessions. Although our original intention was to mask the stimuli from all parents, this was often not done, for two reasons. First, several parents were not comfortable holding their infants in a completely dark room without being able to hear their infants. Second, the stimuli were in free field and were often at a sound level high enough to make masking impractical. That is, the masker itself would have reached uncomfortable levels and could potentially have been audible to the infants. Although the parent holding the infant could usually hear the sounds, several aspects of the findings suggest that parents did not bias their infants or the experimenter. First, it seems extremely unlikely that parents would have introduced response biases at stimulus levels that produced the very orderly age-related change that was found. After all, the parent could hear the stimulus even at its lowest level. Parents received no feedback about their infants' thresholds until the entire study was completed. Therefore, parents would have had to know their infants' thresholds at each age, in order for the bias to produce the particular pattern of results. A second aspect of the findings also suggests that parental bias was minimal. Recall that the experimenter voted not only when the infant appeared to have responded to the stimulus, but also made a forced choice judgment about whether the stimulus came from the left or right loudspeaker. On average, the experimenter was correct about the side on 67% of the trials (see Appendix). If parents had introduced a bias, it is unlikely that they would have "chosen the wrong side" so often. The third feature of the findings that argues against a parental bias effect is that on those sessions when parents were masked, the threshold values agreed well with the overall age trend for those individual infants. Although we believe parental bias was unlikely in the present study, further development of the procedure should include provisions to keep parents from hearing the stimuli.

In summary, the procedure described herein has, to our knowledge, provided the first longitudinal investigation of infant auditory thresholds. It is noteworthy that the individual infant developmental trends found corroborate results from group analyses by other researchers. These results indicate that auditory development occurs most rapidly from the newborn period to ~6 months of age. This age-related trend was observed in all but one infant in this study.

Acknowledgments

The authors acknowledge the technical assistance provided by Drs. David Chandler, Wes Grantham, and Walt Murphy and gratefully acknowledge the parents and infants who generously gave their time to participate in this study. A.M.T. was funded in part by Maternal and Child Health Resources and Services Administration during the course of this study and D.H.A. was supported by grants from the National Institute of Child Health and Human Development.

References

Gorga, M. P., Kaminski, J. R., Beauchaine, K. A., & Jesteadt, W. (1988). Auditory brainstem responses to tone bursts in normally hearing subjects. *Journal of Speech and Hearing Research, 31,* 87–97.

Gorga, M. P., Kaminski, J. R., Beauchaine, K. L., Jesteadt, W., & Neely, S. T. (1989). Auditory brainstem responses from children three months to three years of age: Normal patterns of response II. *Journal of Speech and Hearing Research, 32,* 281–288.

Gravel, J.S. (1989). Behavioral assessment of auditory function. *Seminars in Hearing, 10,* 216–228.

Gray, L. (1987). Signal detection analyses of delays in neonates' vocalizations. *Journal of Acoustical Society of America, 82,* 1608–1614.

Hecox, K., & Galambos, R. (1974). Brain stem auditory evoked responses in human infants and adults. *Archives of Otolaryngology, 99,* 30–33.

Hoversten, G. H., & Moncur, J. P. (1969). Stimuli and intensity factors in testing infants. *Journal of Speech and Hearing Research, 12,* 687–702.

Jacobson, J. T, & Morehouse, C. R. (1984). A comparison of auditory brain stem response and behavioral screening in high risk and normal newborn infants. *Ear and Hearing, 5,* 247–253.

Jiang, Z. D., & Tierney, T. S. (1995). Development of human peripheral hearing revealed by brainstem auditory evoked potentials. *Acta Paediatrica, 84,* 1216–1220.

Keefe, D. H., Bulen, J. C., Hoberg, K., & Burns, E. M. (1993). Ear-canal impedance and reflection coefficient in human infants and adults. *Journal of the Acoustical Society of America, 94,* 2617–2638.

Keefe, D. H., & Levi, E. (1996). Maturation of the middle and external ears: Acoustic power-based responses and reflectance tympanometry. *Ear and Hearing, 17*(5), 361–373.

Nozza, R. J. (1995). Estimating the contribution of non-sensory factors to infant-adult differences in behavioral thresholds. *Hearing Research, 91,* 72–78.

Nozza, R., & Wilson, W. R. (1984). Masked and unmasked pure tone thresholds of infants and adults: Development of auditory frequency selectivity and sensitivity. *Journal of Speech and Hearing Research, 27,* 613–622.

Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., & Spetner, N. B. (1988). Pure-tone sensitivity of human infants. *Journal of Acoustical Society of America, 84*(4), 1316–1324.

Olsho, L. W., Koch, E. G., Halpin, C. F., & Carter, E. A. (1987). An observer-based psychoacoustic procedure for use with young infants. *Developmental Psychology, 23*(5), 627–640.

Ruth, R. A., Horner, J. S., McCoy, G. S., & Chandler, C. R. (1983). Comparison of auditory brainstem response and behavioral audiometry in infants. *Scandinavian Audiology Suppl, 17,* 94–98.

Salamy, A. & McKean, C. M. (1976). Postnatal development of human brainstem potentials during the first year of life. *Electroencephalography and Clinical Neurophysiology, 40,* 418–426.

Sinnott, J. M., Pisoni, D. B., & Aslin, R. M. (1983). A comparison of pure tone auditory thresholds in human infants and adults. *Infant Behavior and Development, 6,* 3–17.

Thompson, M., & Thompson, G. (1972). Response of infants and young children as a function of auditory stimuli and test methods. *Journal of Speech and Hearing Research, 15,* 699–707.

Thompson, G., & Weber, B. A. (1974). Responses of infants and young children to behavior observation audiometry (BOA). *Journal of Speech and Hearing Disorders, 39,* 140–147.

Trehub, S.E., Schneider, B.A., & Endman, M. (1980). Developmental changes in infants'sensitivity to octave-band noises. *Journal of Experimental Child Psychology, 29,* 282–293.

Trehub, S. E., Schneider, B. A., Thorpe, L. A., & Judge, P. (1991). Observational measures of auditory sensitivity in early infancy. *Developmental Psychology, 27*(1), 40–49.

Weir, C. (1979). Auditory frequency sensitivity of human newborns: some data with improved acoustic and behavioral controls. *Perception and Psychophysics, 26,* 287–294.

Werner, L.A. (1996). The development of auditory behavior (or what the anatomists and physiologists have to explain). *Ear & Hearing, 17*(5), 438–446.

Werner, L. A., Folsom, R. C., & Mancl, L. R. (1993). The relationship between auditory brainstem response and behavioral thresholds in normal hearing infants and adults. *Hearing Research, 77,* 88–98.

Werner, L. A., & Gillenwater, J. M. (1990). Pure-tone sensitivity of 2- to 5-week-old infants. *Infant Behavior and Development,* 13, 355–375.

Werner, L. A., & Marean, G. C. (1991). Methods for estimating infant thresholds. *Journal of Acoustical Society of America, 90*(4), 1867–1875.

Received April 10, 2001
Accepted September 4, 2001
First published (online) November 5, 2001
http://journals.asha.org
D.O.I: 10.1044/1059-0889 (2001/011)

Contact author: Anne Marie Tharpe, PhD, The Vanderbilt Bill Wilkerson Center for Otolaryngology and Communication Sciences, 1114 19th Avenue South, Nashville, TN 37212. Email: tharpeam@ctrvax.vanderbilt.edu

Appendix

Several approaches were considered for obtaining an averaged threshold estimate from a single test session. One issue was whether to include all trials or only those trials on which the observer correctly judged from which side the sound was presented. Another issue was which averaging method to use, the mean or the median. To investigate these issues, average thresholds were calculated across all 56 usable test sessions. For thresholds based on all trials in a test session, the mean (and standard error) of thresholds based on means was 34.8 ± 1.5, and the mean of thresholds based on medians was 34.2 ± 1.5. For thresholds based on only the correct trials within a test session, the mean of thresholds based on means was 35.7 ± 1.6, and the mean of thresholds based on medians was 35.3 ± 1.6. Thus, there was little difference between these ways of calculating averaged thresholds. We elected to report the averaged threshold for a single test session as the median across all trials. We preferred the median because it is less influenced by extreme values. In addition, it made sense to use all trials because an infant may occasionally respond to the stimulus, albeit in a spatially incorrect way.

Although the logic of our threshold estimates did not require that the observer correctly judge from which side the sound was presented, it would have been disturbing if the observer was not usually correct. Across the 56 test sessions, 67% of the observer's judgments were correct, which was significantly higher than the chance level of 50%, $t(55) = 8.65$. This indicates that on most trials infants made directionally appropriate responses that the observer noticed. We were interested in whether the percentage correct increased with age, because sound localization is known to improve during the first year after birth. Indeed, there was a significant correlation between age (in days) and percentage correct, $r(55) = 0.46$. Despite this positive correlation, it was age rather than percentage correct that showed a strong association with hearing thresholds. This was demonstrated with a multiple regression analysis in which threshold was the dependent variable. Percentage correct was included by forced entry on the first step of the regression analysis, with the result that $R^2_{adj} = 0.229$. Age was entered on the second step, resulting in a large increase to $R^2_{adj} = 0.705$. This was a significant increase in the variance accounted for, $F(2,53) = 66.62$. Therefore, even with the association between percentage correct and threshold taken into account, age accounted for an additional 47% of the variance in thresholds.