# Threshold Estimation in the Preschool Period: The PEST Procedure

# *Estimation du seuil d'audition lors de la période préscolaire : la méthode PEST*

*Sandra E. Trehub, PhD and Bruce A. Schneider, PhD*
Centre for Research in Human Development
University of Toronto, Erindale Campus, Mississauga, Ontario

*Laurel J. Trainor, PhD*
Department of Psychology
McMaster University, Hamilton, Ontario

*Barbara A. Morrongiello, PhD*
Department of Applied Developmental Psychology
University of Guelph, Guelph, Ontario

**Key words:** auditory thresholds, preschool children, hearing

## Abstract

Preschool children's thresholds for narrow-band noises centred at 500 and 4000 Hz were estimated on two occasions with an adaptive procedure commonly used with adults (PEST). Averaged individual thresholds for 3-, 4-, and 5-year-old children were comparable to group thresholds obtained previously with the method of constant stimuli. As expected, thresholds were higher for 500-Hz than for 4000-Hz signals. Moreover, the thresholds of 3-year-olds were significantly higher than those of 4- and 5-year-olds, which did not differ from one another. Only the thresholds of 3-year-olds improved significantly from the first to the second day of testing. The present version of the PEST procedure, which necessitates a single test session of 30 trials or less, generates reasonably accurate threshold estimates for 4- and 5-year-olds, but underestimates the abilities of 3-year-olds.

## *Abrégé*

*Les seuils auditifs des enfants d'âge préscolaire pour les bruits à bande étroite de 500 Hz et de 4 000 Hz ont été estimés à deux reprises grâce à une méthode d'adaptation fréquemment utilisée pour l'adulte (PEST). Les seuils auditifs moyens des enfants de trois, de quatre et de cinq ans sont comparables aux seuils moyens obtenus précédemment selon la méthode du stimulus constant. Tel que prévu, les seuils est plus élevés pour les signaux de 500 Hz que pour les signaux de 4 000 Hz. D'autre part, Les seuils auditifs sont sensiblement plus élevés chez les enfants de trois ans que chez ceux de quatre et de cinq ans, qui ont des seuils identiques.Toutefois les seuils des enfants de trois ans s'améliorent de façon appréciable entre la première et la deuxième journée de l'épreuve. La version actuelle de la méthode PEST, qui exige une seule séance d'un maximum de 30 essais, débouche sur des estimations raisonnablement précises du seuil auditif des enfants de quatre ou de cinq ans, mais sous-estime les aptitudes des enfants de trois ans.*

Several researchers have been documenting age-related changes in auditory sensitivity (Allen & Wightman, 1994, 1995; Allen, Wightman, Kistler, & Dolan, 1989; Jensen & Neff, 1993; Olsho, Koch, Carter, Halpin, & Spetner, 1988; Schneider, Trehub, Morrongiello, & Thorpe, 1986, 1989; Trehub, Schneider, Morrongiello, & Thorpe, 1988, 1989) and temporal resolution (Elfenbein, Small, & Davis, 1993; Grose, Hall, & Gibbs, 1993; Hall & Grose, 1994; Trehub, Schneider, & Henderson, 1995; Werner, Marean, Halpin, Spetner, & Gillenwater, 1992; Wightman, Allen, Dolan, Kistler, & Jamieson, 1989). Recent interest in infant sensitivity, in particular (Schneider & Trehub, 1992; Werner, 1992), has spurred the development of techniques suitable for prelinguistic listeners with limited attention and motoric skill (e.g., Bull, Schneider, & Trehub, 1981; Olsho, 1985; Olsho, Koch, & Carter, 1988; Olsho et al., 1988; Olsho, Koch, & Halpin, 1987; Olsho, Koch, Halpin, & Carter, 1987; Schneider, Bull, & Trehub, 1988; Schneider & Trehub, 1992; Schneider, Trehub, & Bull, 1980; Schneider, Trehub, & Thorpe, 1991; Trehub, Bull, Schneider, & Morrongiello, 1986; Trehub, Schneider, & Endman, 1980; Trehub, Schneider, Thorpe, & Judge, 1991; Werner & Bargones, 1991; Werner & Gillenwater, 1990; Werner & Marean, 1991). As a result, there have been enormous strides in the specification of auditory sensitivity in the early months of life. With some notable exceptions (e.g., Allen & Wightman, 1992, 1995), however, there has been little methodological research with preschool children, despite the significant changes that occur during this period (e.g., Allen & Wightman, 1994; Jensen & Neff, 1993; Schneider et al., 1986) and the distinct possibility of noise exposure effects in early childhood (Mills, 1975; Roche, Siervogel, Himes, & Johnson 1978; Siervogel, Roche, Johnson, & Fairman, 1982).

The overall goal, in the present investigation, was to evaluate the applicability of an adult threshold-estimation procedure to preschool children. Although the adaptive procedure in question, Parameter Estimation by Sequential Testing, or PEST (Taylor & Creelman, 1967), has been used primarily with adults, it has had modest success in the assessment of infant visual (Lewis & Maurer, 1986) and auditory (Trehub et al., 1986) sensitivity. Adaptive procedures, which are used widely in the measurement of adult thresholds (e.g., Hall, 1981; Levitt, 1971; Madigan & Williams, 1987; Pentland, 1980; Shelton, Picardi, & Green, 1982; Stillman, 1989; Taylor & Creelman, 1967; Watson & Pelli, 1983), vary the intensity of the stimulus from trial to trial, attempting to zero in on a predetermined performance level such as 75% correct. The stimulus level on a given trial depends on performance over the last few trials. By contrast, the method of constant stimuli involves a number of trials at each of several predetermined stimulus levels, with the intensity of the stimulus on any trial being independent of performance. The threshold corresponding to a specified performance level, say 75% correct, can be interpolated from the function relating performance to stimulus level.

Adaptive procedures have a number of potential advantages over constant stimulus procedures, especially for preschool children. Many preschoolers are unable to complete the large number of trials usually required for threshold determination with the method of constant stimuli. Clinical situations, in particular, would place a special premium on efficiency. In the case of preschoolers, then, one of the key requirements of a suitable methodology is the ability to generate an accurate threshold estimate in a limited number of trials. Although the method of constant stimuli may give more extensive information (e.g., the slope of the psychometric function as well as the threshold location), adaptive procedures often yield more accurate threshold estimates in fewer trials (e.g., Kollmeier, Gilkey, & Sieben, 1988; Levitt, 1971; Lewis & Maurer, 1986; Raz & Wightman, 1984; Shelton et al., 1982; Taylor & Creelman, 1967; Trehub et al., 1986), particularly when information about the threshold region is limited and when the population in question exhibits large individual differences, as is the case with preschool children (Wightman & Allen, 1992).

Although adaptive procedures are typically used with experienced adults, there are indications that PEST or some modification of it may be especially suitable for inexperienced or young participants. For example, Stillman (1989) found that although the method of constant stimuli underestimated the capabilities of inexperienced adults, adaptive staircase and PEST procedures did not do so. Moreover, PEST estimates were relatively stable on retesting, in contrast to staircase estimates, which improved somewhat. Lewis and Maurer (1986) found a modified PEST procedure to be more efficient than the method of constant stimuli in estimating visual acuity thresholds in infants. Similarly, Trehub et al. (1986) successfully used the PEST procedure to estimate auditory thresholds in 6-month-old listeners. The resulting threshold estimates, obtained in 20 to 25 trials, were similar to group threshold estimates obtained previously with the method of constant stimuli (Trehub et al., 1980).

The purpose of the present investigation was to evaluate the utility of the PEST procedure for auditory threshold measurement in 3-, 4-, and 5-year olds. First, we sought to ascertain whether threshold estimates could be obtained readily from preschoolers in a limited number of trials. Second, we attempted to establish the comparability of the resulting thresholds to group thresholds obtained previously with the method of constant stimuli (Schneider et al., 1986). To facilitate such comparisons, we used the response measure of Schneider et al. (1986), which involved children judging the signal location (left or right side) in a two-alternative, forced-choice task. Third, we sought to determine the test-retest reliability, or stability, of thresholds obtained from preschoolers with the PEST procedure.

## Method

### Participants

The group of participants was made up of 20 3-year-olds (up to 3 years, 3 months; mean age = 3 years), 20 4-year-olds (± 3 months; mean age = 4 years, 1 day), and 20 5-year-olds (± 3 months; mean age = 5 years, 13 days), who completed two sessions on each of two days of testing. Mean age of the three groups at the second day of testing was 3 years, 9 days; 4 years, 6 days; and 5 years, 29 days. A further 13 children were eliminated from the sample for the following reasons: failure to meet the performance requirement for probe trials (one 5-year-old); equipment failure (one 3-year-old); failure to return on the second test day (four 3-year-olds and one 4-year-old); and failure to complete all four sessions (six 3-year-olds).

### Apparatus and Stimuli

The two test signals were octave-band noises centred at 4000 Hz and 500 Hz, which were created by passing white noise from a noise generator (General Radio, Model 1381) through a filter (General Radio, Model 1952). The rate of energy falloff was 30 dB per octave on each side, and the output was split into two channels for the right and left speakers. Each channel consisted of an electronic switch (rise/decay time = 25 ms) followed by a programmable attenuator, a preamplifier, a stereo amplifier (SAE 2600), and a loudspeaker (ESS-Heil, Model AMT1AM) inside the test booth. The switch and attenuator were controlled by a microcomputer (Commodore 4032), which controlled the intensity of the sound on each trial as well as its location (i.e., left or right speaker). Sound pressure levels were calibrated with an impulse precision sound level meter (Bruel and Kjaer, Model 2204) at the approximate location of the listener's head (without the listener present). Readings on the linear scale were taken with a 0.5-in microphone directed at the loudspeaker producing the signal. Sound-pressure variation within a 6-in radius of the calibration locations never varied by more than ± 2dB at any frequency and was typically less than 1 dB. The background noise level (measured with a 2.54-cm microphone) was approximately 16 dBA. Training and initial test levels were determined on the basis of pilot testing.

The child sat on a chair in one corner of a double-walled sound-attenuating chamber (Industrial Acoustics, 3 x 2.8 x 2 metres); the tester (who wore headphones with masking noise) sat facing the child in the opposite corner. The loudspeakers were positioned approximately 1.8 metres from the child, 45 degrees to the child's left and right. Under each loudspeaker was a four-chamber box with a smoked Plexiglas front containing lights and mechanical toys used for reinforcing correct responses. When the lights were off, children could not see the toys in the box. A television located above the speakers could also reinforce correct responses with 4-second presentations of cartoon segments. Each arm of the child's chair contained a large button corresponding to each speaker (left or right). The experimenter (who could not hear the signals) called for trials and relayed the child's responses to the computer via a hand-held button box.

## Procedure

On each of the two visits to the laboratory, every child was tested in two sessions, one with the 4000-Hz signal and the other with the 500-Hz signal. Half of the children were tested on 4000 Hz first, the other half on 500 Hz first on both test days. The experimenter pressed a button to initiate a trial, which consisted of a signal presented at a particular intensity on one of the two loudspeakers. The signal remained on until the child attempted to identify its location by pressing a button on the appropriate arm (left or right side) of the chair (two-alternative, forced-choice), which the tester relayed to the computer (by means of her button box). If the child responded correctly, the computer automatically initiated reinforcement for 4 seconds. During the first session of any day, reinforcement consisted of the illumination and activation of an animated toy in the box under the appropriate speaker. During the second session, reinforcement consisted of the presentation of an animated cartoon on the TV monitor on top of the speaker. Each session consisted of three phases: a training phase, a phase designed to move the listener quickly to the general threshold region (QUIR), and a PEST phase designed to estimate the actual threshold.

During the training phase, the sound was presented well above threshold and participants were required to achieve four successive correct responses at one stimulus intensity (65 dB SPL for 4000 Hz; 68 dB SPL for 500 Hz) followed by two successive correct responses at a lower intensity (55 dB SPL for 4000 Hz; 58 dB SPL for 500 Hz). Children who failed to achieve this training criterion within 20 trials were excluded from the test phase (see Participants). The QUIR (quickly into range) phase began 8 dB lower than the second training level (47 dB SPL for 4000 Hz; 50 dB SPL for 500 Hz). When the child responded correctly on two of three successive trials at a particular intensity, the intensity was lowered by 8 dB. The QUIR phase ended when the child responded incorrectly on two of three successive trials or when the programmed intensity dropped below 15 dB SPL. The intensity was then raised by 8 dB and the child entered the PEST phase. Every fifth trial (starting with the first PEST trial) was a probe trial in which the stimulus intensity was equivalent to the first training trial (65 dB SPL for 4000 Hz; 68 dB SPL for 500 Hz). These trials were not part of the PEST rules, being used simply to monitor the child's attention during the task. Failure to respond on probe trials resulted in exclusion of a child's data from the subsequent analysis.

The PEST procedure is designed to estimate the stimulus level (intensity, in this case) that corresponds to a predetermined performance criterion, *P* (75% correct, in this case). Trials are conducted at a particular intensity level until evidence accumulates about whether the threshold corresponding to *P* is higher or lower than the current stimulus (intensity) level. A decision is then made to either raise or lower the stimulus level. The magnitude of the change is determined by the decision history, or preceding series of intensity changes, that is, whether previous intensity changes were in the same direction (i.e., increases or decreases) or the reverse direction (i.e., an increase followed by a decrease or vice versa). A decision to change stimulus level occurs in the following way. PEST presents successive trials at one stimulus level, keeping track of the number of correct responses out of N trials. If this stimulus level is at the listener's threshold, then the expected number of correct responses would be $P \times N$, where $N$ is the number of successive trials presented at this intensity. When the number of correct responses exceeds $P \times N$ by more than a criterion value, $W$, the intensity level is decreased. When $P \times N$ exceeds the number of correct responses by more than $W$, the intensity is increased. In the present study, $W$ equaled 1.0.

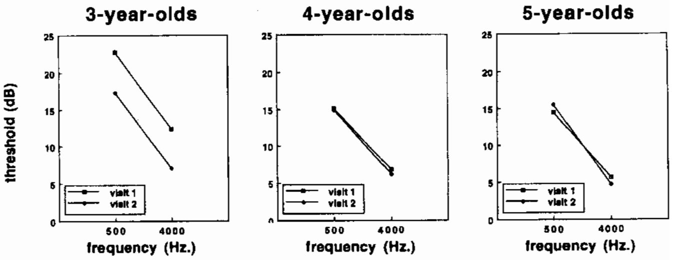
When a decision was made to change stimulus level, the magnitude of the change, or step size, was determined by means of Taylor and Creelman's (1967) decision rules. The basic idea is to decrease the step size when changing direction. A change in direction occurs when the current stimulus level is near threshold, at which time the smaller step size permits a more precise threshold estimate. At the same time, a series of stimulus changes in the same direction indicates that the stimulus level is not near threshold, at which time an increase in step size is warranted. The specific rules are as follows. *Rule 1*: On every directional change in intensity (i.e., reversal), halve the step size. If the previous change in stimulus level lowered the intensity by 8 dB and the current decision is to raise it, the current change raises it by 4 dB. *Rule 2*: The second successive step in a given direction is the same magnitude as the first. *Rule 3*: Whether a third successive step in a given direction is the same as or double the second depends on the sequence of steps leading to the most recent reversal. If the step immediately preceding the reversal results from a doubling of step size, then the third step in the same direction should not be doubled. If, however, the step leading to the most recent reversal did not result from doubling the step size, then this third step should be double the second. The function of this complex rule is to proceed quickly to the threshold region (i.e., double the step size) without getting too far away if a listener is temporarily inattentive or lucky (i.e., do not double the step size if previous performance provides evidence contrary to the current evidence). *Rule 4*: For the fourth step in a given direction, always double the third step size.

In addition to the aforementioned rules, stopping rules and boundary step sizes must be defined. In the present study, the initial step size and the maximum step size were both 4 dB. Thus, when the step size was 4 dB and the procedure called for a doubling of step size, it remained at 4 dB. The procedure was terminated with a threshold estimate when the step size called for on the next trial was less than the minimum step size (2 dB in the current study). Obviously, the smaller the minimum step size, the more precise the threshold estimate. This increase in accuracy must be balanced, however, against the additional test time required to meet the termination rule. In the present experiment, the maximum number of trials was set at 30. If the PEST stopping rule had not yet been invoked, threshold was defined as the intensity at which the next trial would have been presented. At all age levels, only 15% of children reached the PEST stopping rule, the default stopping rule (30 trials) applying to the remaining children and resulting in a test session of approximately 15 minutes.

## Results

There were five children (one 5-year-old, two 4-year-olds, and two 3-year-olds) who exhibited considerably greater performance differences (24 dB) than the other children when retested with same stimulus. On the assumption that such differences likely reflected motivational problems, the data from these children were excluded from all analyses. Thresholds as a function of signal frequency (500 Hz, 4000 Hz) and visit (1, 2) are plotted for each age group in Figure 1. An analysis of variance with age (3, 4, 5 years) and stimulus order (500-first, 4000-first) as between-subjects factors and frequency (500 Hz, 4000 Hz) and visit (test, retest) as within-subjects factors revealed main effects of age, $F(2, 49) = 5.63$, $p < .007$ (older children performed better than younger), frequency, $F(1, 49) = 162.68$, $p < .0001$ (thresholds higher at 500 than at 4000 Hz), and visit, $F(1, 49) = 8.58$, $p < .005$ (thresholds lower on retest), and an interaction between age and visit, $F(2, 49) = 6.9$, $p < .003$. No other main effects or interactions were significant. Newman-Keuls tests indicated that 3-year-olds had significantly higher thresholds than 4- and 5-year-olds, who did not differ from one another (see Figure 1).

Figure 1. Thresholds as a function of centre frequency and visit (1 or 2) for each age group.



Separate analyses for each age group examined the visit-by-age interaction. ANOVAs with stimulus order as a between-subjects factor and visit and frequency as within-subjects factors revealed frequency as the only significant effect for 4- and 5-year-olds. In contrast, there was a significant effect of visit (improvement at retest) for 3-year-olds, $F = 28.75$, $p < .0001$, in addition to the effect of stimulus frequency. Thus, 4- and 5-year-olds showed no systematic change in performance from their first to second day of testing, in contrast to the 3-year-olds, whose performance improved significantly at retesting. Increased familiarity with the task likely enhanced the performance of 3-year-olds on the second test day.

To examine whether the PEST procedure yielded a relatively stable ranking of individual thresholds across the two test days, correlations were calculated between test and retest thresholds at each frequency. For the combined group of 3-, 4-, and 5-year-olds, correlations across days were significant at 500 Hz, $r = .42$, $p < .001$, and at 4000 Hz, $r = .59$, $p < .001$, indicating moderate test-retest reliability. Correlations of performance across test days were considerably higher for the 3-year-olds, $r = .58$, $p < .01$ at 500 Hz, and $r = .76$, $p < .001$ at 4000 Hz. Thus, despite the practice effects that were evident for 3-year-olds, the rank-ordering of performance remained largely intact across the two test days. Presumably, the narrower range of performance among older children accounted for the more modest correlations of the overall group compared to those of the 3-year-olds.

The present PEST thresholds for 4000-Hz octave band noise can be compared with group thresholds for the same stimulus obtained with the method of constant stimuli (Schneider et al., 1986). The latter group thresholds, defined as 75% correct performance, for 3-, 4-, and 5-year- olds are shown in Table 1 together with averaged PEST thresholds from the present investigation. Included in the PEST averages were all children who completed both days of testing (including the five who had shown large variations from one test day to another). As can be seen in Table 1, average thresholds are fairly similar with the two procedures, generating confidence in the estimates obtained with either method.

**Table 1. Averaged thresholds in dB SPL (75% correct performance) for 4000-Hz octave-band noise obtained with the PEST procedure (second visit in parentheses) and with the method of constant stimuli (Schneider et al., 1986).**

| | PEST | Constant Stimuli |
|---|---|---|
| 3-year-olds | 12.4 (8.5) | 11.9 |
| 4-year-olds | 7.9 (7.6) | 9.8 |
| 5-year-olds | 5.4 (4.4) | 4.7 |

## Discussion

We demonstrated that thresholds for 500- and 4000-Hz octave-band noises could be obtained from 3-, 4-, and 5-year-old children by means of the PEST procedure in 30 trials or less. The resulting thresholds were orderly, as reflected in lower thresholds for older than for younger children and higher thresholds for low- compared to high-frequency signals, as would be expected on the basis of previous research (e.g., Allen & Wightman, 1994, 1995; Allen et al., 1989; Schneider et al., 1989). Performance was reasonably stable across the two days of testing, with some improvement evident, especially for the 3-year-olds. Despite this improvement, the 3-year-olds' performance was highly and significantly correlated across the two test days. In other words, the relative ranking of children was largely preserved. Finally, when the PEST thresholds were averaged and compared with group thresholds obtained with the method of constant stimuli (Schneider et al., 1986), threshold values were found to be very similar. These findings, taken together, provide assurance that the threshold estimates obtained with the PEST procedure are reasonably accurate and reliable.

Whether the PEST procedure yields individual threshold estimates that are more reliable (i.e., smaller test-retest differences), more efficient (i.e., fewer test trials required), or more conservative (i.e., higher threshold values) than those generated by alternative procedures remains to be determined. Specific comparisons with alternative procedures could be accomplished in a number of ways. One possibility would involve assessing children's thresholds for the same stimuli, first, with one procedure, and subsequently, with another. On the basis of the present findings, it would be reasonable to assume that 4- and 5-year-olds would exhibit reasonably stable performance over time and that any differences obtained would be primarily attributable to procedural factors as opposed to practice. Another possibility is to test one group of children on two occasions with the PEST procedure and another group with a comparison procedure. One could then determine which procedure minimized differences from test to retest and was most successful in retaining the relative rank ordering of children across test sessions.

With respect to the efficacy of the PEST procedure for estimating individual thresholds, either for clinical or research purposes, PEST would appear to yield reasonably accurate estimates, at least for 4- and 5-year-olds. In the case of 3-year-olds, the administration of a single test session will obviously result in an underestimate of the "true" threshold. If precision is essential, then additional test trials and testing at one rather than two frequencies might be preferable. Because motivational concerns are likely to be greater with 3-year-olds than with older children, adaptive procedures might exacerbate attentional problems because of progressively decreasing test levels. The method of constant stimuli or an ascending procedure may be suitable alternatives in such instances. For some purposes, particularly where rough estimates of hearing are adequate, it might be possible to reduce the number of trials, establishing simply that performance at a particular test level or levels is better than chance.

Some investigators (e.g., Allen & Wightman, 1994, 1995; Wightman & Allen, 1992) have suggested that the large inter- and intra-individual differences that are characteristic of the preschool period obscure "true" sensitivity levels. They believe that such variability accounts for the observed sensitivity differences between young children and adults. Indeed, it is not unusual to find individual preschoolers who perform at adult levels (e.g., Wightman & Allen, 1992).

Wightman and Allen (1992) go so far as to suggest that nonsensory factors such as memory and attention may be entirely responsible for the apparent age differences. Nevertheless, our findings of relatively stable and correlated performance across days of testing and of similar threshold estimates by means of two very different procedures would seem to indicate that sensory factors make an important contribution, but not necessarily the only contribution, to the reported measurements. One aspect of our methodology that may reduce the impact of such nonsensory factors is our practice of having the signal remain on until a child actually responds. In the typical test situation, with its fixed signal duration, lapses of attention would have more drastic consequences than they would in our test situation. In any case, we acknowledge the need to continue searching for means of minimizing the influence of extraneous (i.e., nonsensory) factors that interfere with the accurate measurement of auditory sensitivity in early childhood.

***Please address all correspondence to:*** Sandra E. Trehub, Centre for Research in Human Development, University of Toronto, Erindale Campus, Mississauga, Ontario, L5L 1C6. E-mail: trehub@credit.erin.utoronto.ca

## Acknowledgments

This research was supported by a grant to S. E. Trehub and B. A. Schneider from the Medical Research Council of Canada. The authors thank Marilyn Barras for her assistance in testing.

### *References*

Allen, P., & Wightman, F. L. (1992). Spectral pattern discrimination by children. *Journal of Speech and Hearing Research, 35,* 222-233.

Allen, P., & Wightman, F. L. (1994). Psychometric functions for children's detection of tones in noise. *Journal of Speech and Hearing Research, 37,* 205-215.

Allen, P., & Wightman, F. L. (1995). Effects of signal and masker uncertainty on children's detection. *Journal of Speech and Hearing Research, 38,* 503-511.

Allen, P., Wightman, F., Kistler, D., & Dolan, T. (1989). Frequency resolution in children. *Journal of Speech and Hearing Research, 32,* 317-322.

Bull, D., Schneider, B. A., & Trehub, S. E. (1981). The masking of octave-band noise by broad-spectrum noise: A comparison of infant and adult thresholds. *Perception & Psychophysics, 30,* 101-106.

Elfenbein, J. L., Small, A. M., & Davis, J. M. (1993). Developmental patterns of duration discrimination. *Journal of Speech and Hearing Research, 36,* 842-849.

Grose, J. H., Hall, J. W., & Gibbs, C. (1993). Temporal analysis in children. *Journal of Speech and Hearing Research, 36,* 351-356.

Hall, J. L. (1981). Hybrid adaptive procedure for estimation of psychometric functions. *Journal of the Acoustical Society of America, 69,* 1763-1769.

Hall, J. W., & Grose, J. H. (1994). Development of temporal resolution in children as measured by the temporal modulation transfer function. *Journal of the Acoustical Society of America, 96,* 150-154.

Jensen, J. K., & Neff, D. L. (1993). Development of basic auditory discrimination in preschool children. *Psychological Science, 4,* 104-107.

Kollmeier, B., Gilkey, R. H., & Sieben, U. K. (1988). Adaptive staircase techniques in psychoacoustics: A comparison of human data and a mathematical model. *Journal of the Acoustical Society of America, 83,* 1852-1862.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49,* 467-477.

Lewis, T. L., & Maurer, D. (1986). Preferential looking as a measure of visual resolution in infants and toddlers: A comparison of psychophysical methods. *Child Development, 57,* 1062- 1075.

Madigan, R., & Williams, D. (1987). Maximum-likelihood psychometric procedures in two-alternative forced-choice: Evaluation and recommendations. *Perception & Psychophysics, 42,* 240-249.

Mills, J. H. (1975). Noise and children: A review of the literature. *Journal of the Acoustical Society of America, 58,* 767-779.

Olsho, L. W. (1985). Infant auditory perception: Tonal masking. *Infant Behavior and Development, 8,* 371-384.

Olsho, L. W., Koch, E. G., & Carter, E. A. (1988). Nonsensory factors in infant frequency discrimination. *Infant Behavior and Development, 11,* 205-222.

Olsho, L. W., Koch, E. G., & Halpin, C. F. (1987). Level and age effects in infant frequency discrimination. *Journal of the Acoustical Society of America, 82,* 454-464.

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,* 627-640.

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

Pentland, A. (1980). Maximum likelihood estimation: The best PEST. *Perception & Psychophysics, 28,* 377-379.

Raz, I., & Wightman, F. L. (1984). Adaptive estimation of phoneme boundaries and selective adaptation for speech. *Perception & Psychophysics, 36,* 21-24

Roche, A. F., Siervogel, R. M., Himes, J. H., & Johnson, D. L. (1978). Longitudinal study of hearing in children: Baseline data concerning auditory thresholds, noise, exposure, and biological factors. *Journal of the Acoustical Society of America, 64,* 1593-1601.

Schneider, B. A., Bull, D., & Trehub, S. E. (1988). Binaural unmasking in infants. *Journal of the Acoustical Society of America, 83*, 1124-1132.

Schneider, B. A., & Trehub, S. E. (1992). Sources of developmental change in auditory sensitivity. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 3-46). Washington, D.C.: American Psychological Association.

Schneider, B. A., Trehub, S. E., & Bull, D. (1980). High-frequency sensitivity in infants. *Science, 207,* 1003-1004.

Schneider, B. A., Trehub, S. E., Morrongiello, B. A., & Thorpe, L. A. (1986). Auditory sensitivity in preschool children. *Journal of the Acoustical Society of America, 79,* 447-452.

Schneider, B. A., Trehub, S. E., Morrongiello, B. A., & Thorpe, L. A. (1989). Developmental changes in masked thresholds. *Journal of the Acoustical Society of America, 86,* 1733-1742.

Schneider, B. A., Trehub, S. E., & Thorpe, L. A. (1991). Developmental perspectives on the localization and detection of auditory signals. *Developmental Psychology, 27,* 40-47.

Shelton, B. R., Picardi, M. C., & Green, D. M. (1982). Comparison of three adaptive psychophysical procedures. *Journal of the Acoustical Society of America, 71,* 1527-1533.

Siervogel, R. M., Roche, A. F., Johnson, D. L., & Fairman, T. (1982). Longitudinal study of hearing in children II: Cross-sectional studies of noise exposure as measured by dosimetry. *Journal of the Acoustical Society of America, 71,* 372-377.

Stillman, J. A. (1989). A comparison of three adaptive psychophysical procedures using inexperienced listeners. *Perception & Psychophysics, 46,* 345-350.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates of probability functions. *Journal of the Acoustical Society of America, 41,* 782-787.

Trehub, S. E., Bull, D., Schneider, B. A., & Morrongiello, B. A. (1986). PESTI: A procedure for estimating individual thresholds in infant listeners. *Infant Behavior and Development, 9,* 107-118.

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., & Henderson, J. L. (1995). Gap detection in infants, children, and adults. *Journal of the Acoustical Society of America, 98,* 2532-2541.

Trehub, S. E., Schneider, B. A., Morrongiello, B. A., & Thorpe, L. A. (1988). Auditory sensitivity in school-age children. *Journal of Experimental Child Psychology, 46,* 273-285.

Trehub, S. E., Schneider, B. A., Morrongiello, B. A., & Thorpe, L. A. (1989). Developmental changes in high-frequency sensitivity. *Audiology, 28,* 241-249.

Trehub, S. E., Schneider, B. A., Thorpe, L. A., & Judge, P. (1991). Observational measures of auditory sensitivity in early infancy. *Developmental Psychology, 27,* 40-47.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33,* 113-120.

Werner, L. A. (1992). Interpreting developmental psychoacoustics. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 47-88). Washington, DC: American Psychological Association.

Werner, L. A., & Bargones, J. Y. (1991). Sources of auditory masking in infants: Distraction effects. *Perception & Psychophysics, 50,* 405-412

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 the Acoustical Society of America, 90,* 1867-1875.

Werner, L. A, Marean, G. C., Halpin, C. F., Spetner, N. B., & Gillenwater, J. M. (1992). Infant auditory temporal acuity: Gap detection. *Child Development, 63,* 260-272.

Wightman, F., & Allen, (1992). Individual differences in auditory capability among preschool children. In L. A. Werner & E. W. Rubel (Eds.), *Developmental Psychoacoustics* (pp. 113-133). Washington, DC: American Psychological Association.

Wightman, F., Allen, P., Dolan, T., Kistler, D., & Jamieson, D. (1989). Temporal resolution in children. *Child Development, 60,* 611-624.