

Tests of human auditory temporal resolution: Simulations of Bayesian threshold estimation for auditory gap detection

Mori, Shuji

Department of Informatics, Faculty of Information Science and Electrical Engineering, Kyushu University

Murata, Yuto

Department of Informatics, Graduate School of Information Science and Electrical Engineering, Kyushu University

Morimoto, Takashi

Medical Equipment Division, RION CO., LTD

Okamoto, Yasuhide

Department of Otorhinolaryngology, Tokyo Saiseikai Central Hospital

他

<https://hdl.handle.net/2324/7329786>

出版情報 : JASA Express Letters. 4 (9), pp.094402-1-094402-8, 2024-09-03. Acoustical Society of America

バージョン :

権利関係 : © 2024 Author(s).



Tests of human auditory temporal resolution: Simulations of Bayesian threshold estimation for auditory gap detection

Shuji Mori,^{1,a)} Yuto Murata,² Takashi Morimoto,³ Yasuhide Okamoto,^{4,5} and Sho Kanzaki^{5,6}

¹Department of Informatics, Faculty of Information Science and Electrical Engineering, Kyushu University, Fukuoka 819-0395, Japan

²Department of Informatics, Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka 819-0395, Japan

³Medical Equipment Division, RION CO., LTD, Tokyo 185-8533, Japan

⁴Department of Otorhinolaryngology, Tokyo Saiseikai Central Hospital, Tokyo 108-0073, Japan

⁵Keio University School of Medicine, Tokyo 160-8582, Japan

⁶Division of Hearing and Balance Research, National Institute of Sensory Organs, National Hospital Organization Tokyo Medical Center, Tokyo 152-8902, Japan

mori@inf.kyushu-u.ac.jp, oba91321@gmail.com, t-morimoto@rion.co.jp, okamotoy.10@kdt.biglobe.ne.jp, shokanza@yahoo.co.jp

Abstract: In an attempt to develop tests of auditory temporal resolution using gap detection, we conducted computer simulations of Zippy Estimation by Sequential Testing (ZEST), an adaptive Bayesian threshold estimation procedure, for measuring gap detection thresholds. The results showed that the measures of efficiency and precision of ZEST changed with the mean and standard deviation (SD) of the initial probability density function implemented in ZEST. Appropriate combinations of mean and SD values led to efficient ZEST performance; i.e., the threshold estimates converged to their true values after 10 to 15 trials. © 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

[Editor: Qian-Jie Fu]

<https://doi.org/10.1121/10.0028501>

Received: 30 June 2024 Accepted: 15 August 2024 Published Online: 3 September 2024

1. Introduction

Auditory temporal resolution, the ability to detect temporal changes in sounds, is vital for listening to complex acoustic patterns in daily life. Gap detection offers a way to assess one aspect of an individual's auditory temporal resolution. The minimum detectable gap duration, or gap threshold, is seen as a direct measure of auditory temporal resolution. It usually ranges from 1 to 5 ms for people with normal hearing (Mori *et al.*, 2015; Phillips and Smith, 2004) and increases with age (Heinrich and Schneider, 2006). The gap thresholds of hearing-impaired individuals are much higher and more variable than those of their normal-hearing counterparts in the same experimental conditions (Glasberg *et al.*, 1987; Heinrich and Schneider, 2006).

Attempts have been made to use gap detection as a standardized clinical test of auditory temporal resolution. There are commercially available tests, e.g., the Gap-in-Noise (GIN) test, Auditory Fusion Test-Revised, and the Random Gap Detection Test (RGDT), all of which employ gap detection in a non-adaptive manner; i.e., stored sound files with silent gaps of varying lengths are presented in a fixed order. Lister *et al.* (2006) proposed the Adaptive Test of Temporal Resolution (ATTR), which was designed to run on any desktop computer a transformed up-down procedure of gap detection, involving stored broadband noise files. The ATTR enables efficient measurement of gap detection thresholds. While non-adaptive tests like the GIN and RGDT take ≥ 10 min to complete, the ATTR takes ≤ 3 min (Lister *et al.*, 2006). Such short measurement times are vital for clinical applications.

To further develop efficient methods of measuring gap detection thresholds, we used Zippy Estimation by Sequential Testing (ZEST) (King-Smith *et al.*, 1994), an adaptive Bayesian threshold estimation procedure (see below). ZEST has not previously been used for gap detection, although it is considered to be one of the most efficient threshold estimation methods, and is more efficient than the transformed up-down procedure used in the ATTR (Marvit *et al.*, 2003; Treutwein, 1995). When the relevant parameters are properly set, ZEST only requires around 10 trials to obtain a reliable threshold measurement (Marvit *et al.*, 2003; Turpin *et al.*, 2003).

^{a)} Author to whom correspondence should be addressed.

1.1 ZEST

In ZEST, the threshold measurement proceeds by updating the probability density function (*pdf*) of the target threshold, $q_i(T)$, based on the observer’s response in each trial,

$$q_i(T) = p(r_i, x_i|T)q_{i-1}(T), \tag{1}$$

where T and i denote the threshold and the trial number and $p(r_i, x_i|T)$ is the probability of the observer making response r_i (1 for “yes” or correct, 0 for “no” or incorrect) to the presented stimulus, x_i , given that the observer’s threshold is T . For the correct response ($r_i = 1$), $p(1, x_i|T)$ is given by the psychometric function, $\Psi(X)$, whereas $p(0, x_i|T)$ for the incorrect response is given by $1 - \Psi(X)$. In ZEST, the stimulus dimension is represented by a logarithmic scale, since ZEST assumes that $\Psi(X)$ has the same shape along a logarithmic scale of stimulus values (King-Smith *et al.*, 1994).

In each trial, the stimulus value, x_i , is set at the mean of $q_{i-1}(T)$, i.e., *pdf* obtained in trial $i - 1$. $q_0(T)$, which is called the initial *pdf*, is determined by the experimenters based on their prior knowledge about possible threshold values [for more details, see King-Smith *et al.* (1994) and Mori *et al.* (2023)]. The measurement is terminated when the variance of the *pdf* or the number of trials reaches a predetermined value. The mean of the *pdf* obtained in the final trial is taken as the threshold value for that measurement.

ZEST places x_i at the current best estimate of the true threshold value of the target population (i.e., normal-hearing listeners), when the psychometric function, $\Psi(X)$, and the initial *pdf*, $q_0(T)$, reflect the true nature of that population. King-Smith *et al.* (1994) demonstrated by computer stimulations that ZEST is most efficient; i.e., requires the fewest trials to obtain a threshold estimate that is reasonably close to the true value, when $q_0(T)$ approximates the threshold distribution of the target population, which can be inferred from a representative set of samples (Turpin *et al.*, 2003). On the other hand, the ZEST performance is relatively unaffected by deviations of $\Psi(X)$ or $q_0(T)$ from their true forms, unless the deviations are very large (King-Smith *et al.*, 1994; Mori *et al.*, 2023). We followed the approaches outlined in these previous studies to use existing data and computer simulations to identify appropriate forms of $\Psi(X)$ and $q_0(T)$ for gap detection thresholds.

2. Method

2.1 Psychometric function

Based on previous studies of ZEST (King-Smith *et al.*, 1994; Mori *et al.*, 2023), we used a Weibull distribution for the psychometric function of gap detection,

$$\Psi(x - T) = 1 - \delta - (1 - \gamma - \delta)\exp(-10^{\beta(x-T+\varepsilon)}), \tag{2}$$

where δ and γ are the false-negative and false-positive rate, respectively, β is a shape parameter, which determines the slope of the psychometric function, and ε is a parameter for the sweat factor, which determines the threshold-detection probability. Note that the stimulus value, x , which in the present study is the gap duration (ms), is scaled logarithmically. Previous studies have shown that the psychometric functions for gap detection have very similar, if not identical, slopes along a logarithmic scale of gap duration over a range between 1 to 50 ms (Florentine *et al.*, 1999; Mori *et al.*, 2018).

In the present study, we set the values of these parameters in the same way as we did in our previous study (Mori *et al.*, 2023) of amplitude modulation (AM) detection. δ was set to 0.02 based on the previous studies (King-Smith *et al.*, 1994; Marvit *et al.*, 2003; Mori *et al.*, 2023). γ was set to 0.33 because we planned to use a 3AFC task in experiments in which ZEST would be used to obtain gap detection thresholds from human observers. For β , we estimated the slope of psychometric function by fitting Eq. (2) to the data of Mori *et al.* (2018). The data were collected by a method of constant stimuli with a cued yes-no task. The stimuli consisted of two 300-ms sinusoidal tones at 70 dB SPL with a silent interval (gap) between them. There were seven gap durations, separated by an equidistant logarithmic steps, and each gap duration was presented in 50 trials each, and the presented durations were randomly mixed in a single session of 350 trials. The participants were three male students of 23 years of age with normal hearing [for details, see Mori *et al.* (2018)].

Mori *et al.* (2018) measured psychometric functions in several conditions, involving the frequency pairing of the two sounds, and the slopes of these functions did not vary much among the conditions. In the present study, we used psychometric functions obtained with 800-Hz tones. After they have been fitted to Eq. (2), the slopes of these functions for the three participants were 5.59, 5.25, and 4.68, respectively. Their mean, 5.17, was used for the value of β in the present study.

The value of ε was set at the ideal sweat factor; i.e., the minimum value produced by the following equation:

$$\frac{\Psi(x)[1 - \Psi(x)]}{[d\Psi(x)/dx]^2}. \tag{3}$$

We substituted the Weibull distribution for $\Psi(x)$ and plotted the values of Eq. (3) as a function of x (the logarithm of the gap duration) and found the minimum value of 0.036, which was used as the value of ε . The detection probability

corresponding to $\epsilon = 0.036$ was 84%. The ideal sweat factor is the minimum measurement error associated with the psychometric function, which enabled acceptable threshold estimates to be obtained with the fewest number of trials.

2.2 Initial pdf

Since the threshold distribution for gap detection is unknown and, as far as we know, has never been examined empirically, we adopted a Gaussian distribution for the initial pdf (Mori *et al.*, 2023; Watson and Pelli, 1983). As for the psychometric function, the Gaussian distribution was represented in a logarithmic scale, i.e., \log_{10} of the gap detection thresholds (ms). The mean and variance of the distribution were inferred from the datasets collected in previous studies (Morimoto, 2019; Okamoto *et al.*, 2014). The datasets consisted of the gap detection thresholds of 17 normal-hearing (NH) listeners and 30 hearing-impaired (HI) listeners, which were measured using a 1-up, 2-down adaptive procedure in a three-alternative forced-choice task involving 500-ms broadband noise as stimuli [for further details, see Okamoto *et al.* (2014)]. The mean and standard deviation (SD) values of the gap detection thresholds were 0.4 (2.5 ms) and 0.09 for the NH listeners and 1.0 (10 ms) and 0.22 for the HI listeners, respectively.

When these means and SD values were used in the simulations, ZEST did not exhibit good performance for measuring gap detection thresholds (see Fig. 1 below). Therefore, we also tested ZEST in the simulations involving combinations of the following mean and SD; mean: 0.4, 0.8 (6.3 ms), 1.0 or 1.1 (12.6 ms); SD: 0.2, 0.4, or 0.6. A mean of 0.8 corresponded to the midpoint between the minimum (0.2) and maximum (1.4) values of the gap detection thresholds in our dataset; i.e., 1.5 and 25 ms, while 1.1 corresponded to the \log_{10} transformed midpoint between 1.5 and 25 ms.

2.3 Simulations

The simulations were run on personal computers (DELL, OptiPlex 7070; ASUS, ZenBook UX305) using MATLAB (Mathworks, Inc.). Measurements for up to 20 trials were simulated with the assumed true threshold, T , ranging from 0.2

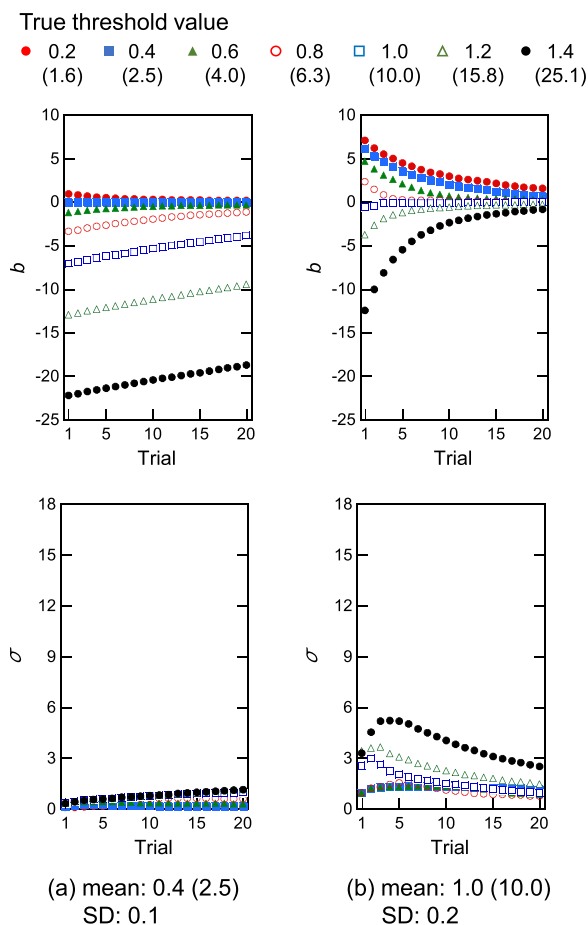


Fig. 1. b (top) and σ (bottom) values obtained in simulations performed with the mean and SD set at the values for (a) the 17 NH listeners or (b) 30 HI listeners. The values in parentheses are shown in ms.

to 1.4 in 0.2 steps in a log₁₀ scale. Those values were chosen because they encompassed the range of gap detection thresholds seen in our dataset.

The simulations were conducted using an exact enumeration technique (King-Smith *et al.*, 1994; Mori *et al.*, 2023), in which all of the possible sequences of responses (either 1 or 0) from trial 1 to N were simulated, and the posterior *pdf*, $q_{Nj}(T)$, for sequence j in trial N was computed for all sequences and the mean of that *pdf*, E_j , was taken as the threshold estimate for sequence j . When $N=1$, for example, there are only two sequences, either a correct (1) or an incorrect response (0), yielding two threshold estimates. When $N=2$, there are four possible sequences of responses from trial 1 to 2; i.e., 1 and 1, 1 and 0, 0 and 1, and 0 and 0. There are 2^N sequences for trial N ; $2^3=8$ for $N=3$, ..., $2^{20}=1,048,576$ for $N=20$, and these sequences yield different threshold estimates.

For all the simulations conducted in this study, the psychometric function [Eq. (2)] parameter values were fixed: $\delta=0.02$, $\gamma=0.33$, $\beta=5.17$, $\varepsilon=0.036$ (see Sec. 2.1). The parameter values of the initial *pdf*, i.e., the mean and SD of the Gaussian distribution, were varied as stated in the Sec. 2.2.

To evaluate the ZEST performance, we used two measures, the measurement bias and SD of the threshold estimates E_j . The measurement bias, b , is the difference between the expected value, \hat{E} , of E_j in the N th trial, and the true threshold T ,

$$b = \hat{E} - T, \tag{4}$$

where

$$\hat{E} = \sum_{j=1}^{2^N} E_j p(E_j|T),$$

$$p(E_j|T) = \frac{q_{Nj}(T)}{q_0(T)}.$$

The SD, σ , of E_j from \hat{E} is given by

$$\sigma = \sqrt{\sum_{j=1}^{2^N} (E_j - \hat{E})^2 p(E_j|T)}. \tag{5}$$

b and σ were computed for each combination of the mean and SD of the initial *pdf* as a function of the number of trials. b reflects the efficiency of ZEST; i.e., how quickly the threshold estimates approach the true threshold as the number of trials increases. σ reflects the precision of these estimates; i.e., how variable they are. Both measures were computed in ms.

3. Results

Figure 1 shows the performance of ZEST, as shown by b and σ , when the mean and SD of the initial *pdf* were set at those of the 17 NH or 30 HI listeners in our dataset. When the mean and SD of the NH listeners were used [Fig. 1(a)], b (top panel) was almost 0, indicating little measurement bias, when the true threshold T was set at 0.4, which was equal to the mean of the initial *pdf*. The value of b became smaller as T increased, and it increased somewhat with the number of trials, but it still deviated from 0 in trial 20. The σ value (bottom panel), the measure of precision, was less than 0.5 for all T values in trial 1 and increased slightly as the number of trials increased, although it was still around 1.0 in trial 20.

Very different patterns were observed when the mean and SD of the initial *pdf* were set at those of the HI listeners [Fig. 1(b)]. In trial 1, the value of b appeared to reflect the differences between the mean (10.0 ms) and T ; it deviated more from 0 as T became more distant from the mean. As the number of trials increased, b gradually approached 0 for all T values. The σ value was relatively small for $T \leq 0.8$ and changed little as the number of trials increased. For $T \geq 1.0$, the σ was around 3 in trial 1, then it increased for the next few trials, after which it eventually decreased as the number of trials increased further.

Figures 2 and 3 present some of the simulation results obtained when the mean or SD of the initial *pdf* fixed and the other parameter was varied (the results for all combinations of the mean and SD values examined in this study are reported in supplemental materials). In the simulations shown in Fig. 2, the SD was fixed at 0.2 (the value for the HI listeners) while the mean was set at 0.4 (the value for the NH listeners), 0.8, or 1.1. Combining these findings with the results shown in Fig. 1(b) for a mean of 1.1, demonstrated that the b value (top panels) showed clear mean-dependent differences as shown in Fig. 1. The b value was equal to or very close to 0 from trial 1, when T was close to the mean, whereas it deviated increasingly from 0 as the difference between T and the mean increased.

The value of σ (bottom panels) also differed depending on the mean, particularly for large T values. When the mean was 0.4 [Fig. 1(a)], the σ value for $T=1.4$ increased up to trial 18 and then levelled off. As the mean changed from 0.8 to 1.1, σ was large at first and started to decrease in the early trial. As a result, it became smaller for higher means by trial 20. Such up-down patterns were also seen for $T=0.8$ to 1.2 but these changes were small and mean-dependent. For $T \leq 0.6$, the value of σ was small and remained relatively unchanged across the trials.

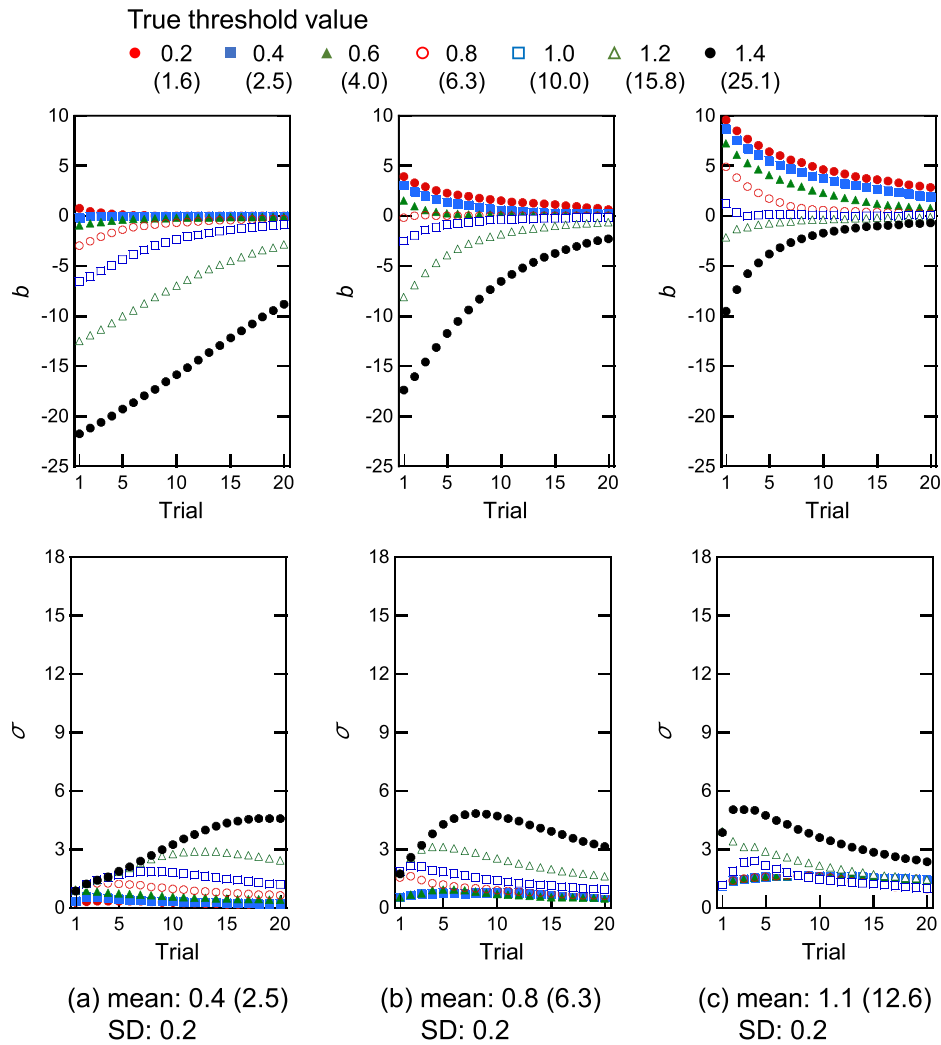


Fig. 2. b (top) and σ (bottom) values obtained in simulations performed with the SD fixed at 0.2 and the mean set at (a) 0.4, (b) 0.8, or (c) 1.1. The values in parentheses are shown in ms.

Figure 3 presents the results obtained when the mean was fixed at 1.1 and the SD varied from 0.4 to 0.6. They can be compared with the results shown in Fig. 2(c), which were obtained with a mean of 1.1 and an SD of 0.2. The effect of SD on the b value was apparent when T was distant from the mean (12.6 ms). As the SD increased from 0.2 to 0.6, the b values for T values other than 1.0 (10.0 ms) and 1.2 (15.8) became less distant from 0 in trial 1, and they approached 0 more rapidly as the number of trials increased. The σ value was also clearly affected by the SD. As the SD increased, the σ value in trial 1 increased, to a larger degree for relatively large T values. As the number of trials increased, σ decreased, more rapidly for large T values, but it was more distant from 0 in trial 20 for relatively large SD values.

4. Discussion

Our simulations demonstrated that the performance of ZEST as a method for measuring gap detection thresholds depends on the mean and SD of the initial *pdf*. The efficiency of ZEST, evaluated by b , was largely dependent on the difference between the mean and the true threshold, T . When the mean was equal to or close to T , b was almost 0 in trial 1 or converged to 0 within the first few trials, indicating that the threshold estimates reached the true threshold values quickly. As the mean became increasingly distant from T , the b value differed more from 0 in trial 1, and it required more trials to converge to 0 (top panels of Fig. 2). In some cases, it did not come close to 0, even by trial 20. The SD also contributed to the efficiency of ZEST; i.e., as the SD increased, the b value approached 0 more quickly (top panels of Fig. 3).

The precision of ZEST was evaluated using σ , which exhibited dependence on the SD, rather than the mean, of the initial *pdf*. When the SD was 0.1, the σ was less than 0.5 in trial 1 and only slightly increased with the number of trials [bottom panel of Fig. 1(a)]. As the SD increased, the σ value became larger, particularly for T values distant from the

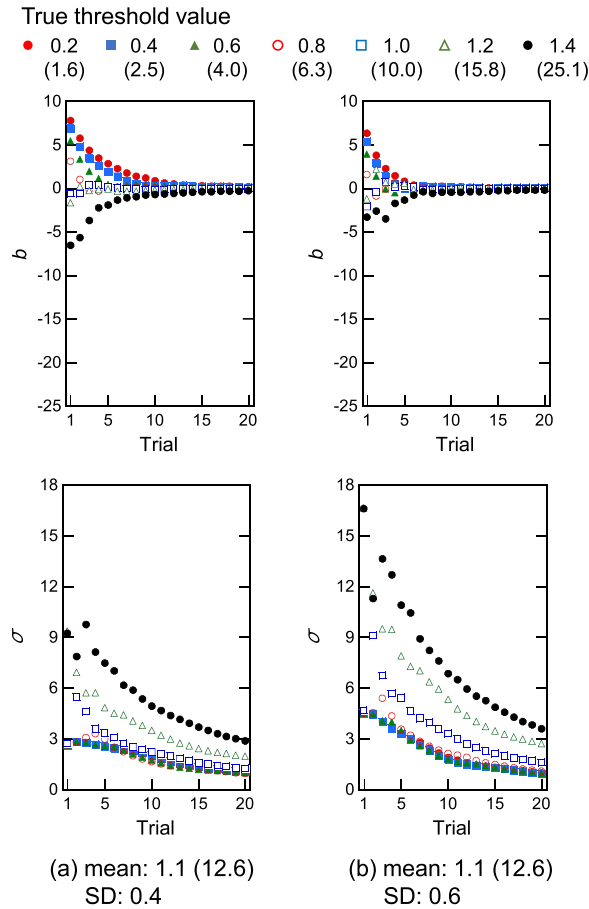


Fig. 3. b (top) and σ (bottom) values obtained in simulations performed with the mean fixed at 1.1 and the SD set at (a) 0.4 or (b) 0.6. The values in parentheses are shown in ms.

mean, and it decreased as the number of trials increased but was still much larger than those obtained when the SD was 0.1 by trial 20 (bottom panel of Fig. 3).

These results are consistent with the findings of Mori *et al.* (2023) where they conducted simulations of ZEST as a measurement method of AM detection thresholds. They found that the SD of the initial *pdf* affected both the efficiency and precision of ZEST, while the mean of *pdf* mainly affected the efficiency of ZEST. Regarding the effects of SD, they reasoned that the SD of the initial *pdf* would determine the size of the variance and shift in posterior *pdf* calculated from Eq. (1) in the first and subsequent trials. As the SD increases, the variance of the posterior *pdf* increases, resulting in greater variance in the threshold estimates (σ) in subsequent trials. The posterior *pdf* moves over a long distance along the stimulus axis when its variance is relatively large, so the *pdf* moves further toward the true threshold value, T , resulting in faster convergence of b to 0. The same reasoning is likely to apply to the findings of the present study.

Such up-down patterns of σ values with an increasing number of trials [see the bottom panel of Fig. 1(b)] were also observed in Mori *et al.* (2023). We reasoned that they were due to the nature of the exact enumeration technique used for the simulations. The exact enumeration technique was used to calculate the threshold estimates for all sequences of responses at trial N , and these sequences included a sequence of wrong responses for all trials, which deviated from the expected estimate value. The deviation was particularly large when the standard deviation of the initial *pdf* was large. In such cases, the value of σ increased for the first few trials. On the other hand, the likelihood of such all-wrong sequences became increasingly lower with increasing number of trials. The calculation of σ involved a multiplication of the deviation and the likelihood of the estimate [Eq. (3)] and σ started to decrease as the number of trials increased further.

Our simulation results suggest that the mean and SD obtained in our previous study (Morimoto, 2019; Okamoto *et al.*, 2014) would not be appropriate parameter values for the initial *pdf*. When the mean and SD were set at those of the 17 NH listeners, the b value for $T = 0.8$ to 1.4 remained far away from 0 in trial 20 [Fig. 1(a)]. When the mean and SD of the 30 HI listeners were used, the b value approached 0 by trial 20 for all T values, although the σ value for $T = 1.4$ was somewhat large in trial 20 and it did not appear that it would soon approach 0 in further trials [Fig. 1(b)]. In theory, ZEST is most efficient when the initial *pdf* approximates the threshold distribution of the target population, which can be

Table 1. Maximum and minimum values of b and maximum values of σ obtained at trial 20.

Mean	SD	b		Maximum σ
		Maximum	Minimum	
0.4	0.1	0.17	-18.70	1.15
1.0	0.2	1.55	-0.85	2.53
0.4	0.2	-0.03	-8.80	4.56
0.4	0.4	-0.03	-1.63	3.80
0.4	0.6	-0.03	-0.79	3.72
0.8	0.2	0.65	-2.28	3.12
0.8	0.4	0.01	-0.52	3.02
0.8	0.6	-0.01	-0.30	3.63
1.0	0.2	2.04	-1.02	2.50
1.0	0.4	0.05	-0.33	2.92
1.0	0.6	0.01	-0.23	3.53
1.1	0.2	2.84	-0.69	2.36
1.1	0.4	0.16	-0.27	2.86
1.1	0.6	0.02	-0.17	3.60

inferred from existing data. The reason why ZEST exhibited low efficiency and/or precision of ZEST when the abovementioned means and SD values were used is that they do not represent the mean and SD of the threshold distributions of the NH and HI listeners. The numbers of listeners in these groups were rather small and neither group constituted an appropriate sample of the target distribution encompassing a wide range of threshold values.

We examined which combinations of the mean and SD of the initial *pdf* would lead to high efficiency and precision. Table 1 summarizes the maximum and minimum b values and the maximum σ values obtained in trial 20 for each combination of mean and SD values used in our simulations. Ideally the b value for the final threshold estimate of gap detection should be < 1 ms, as such thresholds are often measured in 1-ms steps, particularly for clinical applications (Keith, 2000; Lister et al., 2006), so a b value of < 1 ms implies that the threshold estimate is indistinguishable from its true value. The following combinations of mean and SD values exhibited differences of < 1 ms between their maximum and minimum b values, which indicated that the b values for all T values were < 1 ms in trial 20; (0.4, 0.6), (0.8, 0.4), (0.8, 0.6), (1.0, 0.4), (1.0, 0.6), (1.1, 0.4), and (1.1, 0.6). Among these combinations, (1.1, 0.4) yielded the smallest σ value, 2.86. Therefore, a mean of 1.1 and an SD of 0.4 are most appropriate among the examined parameter values for the initial *pdf*.

As a final remark, recent theoretical developments in Bayesian adaptive methods have shown ways to estimate the asymptotic parameter values of psychometric functions and posterior *pdf* (Kuss et al., 2005; Li and Zhu, 2022). By combining ZEST with these approaches, we may be able to predict, without simulations, the number of trials required to reach a true threshold value. This will be the focus of a future study.

Supplementary Material

See supplementary material for the graphs showing the b and σ values (as a function of the number of trials) obtained in simulations with various combinations of mean and SD values for the initial *pdf*.

Acknowledgments

The simulations reported in this paper were conducted by the second author as part of a graduation thesis submitted to Kyushu University. This research was supported by Grant-in-Aid from the Japan Society for the Promotion of Science for Scientific Research (A) 24H00174 and Challenging Research (Exploratory) 20K2066, the Soda Toyoji Memorial Foundation, the TERUMO Life Science Foundation, AMED under Grant No. 24ym0126816j0003, and Center for Clinical and Translational Research of Kyushu University Hospital to S.M.

Author Declarations

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Florentine, M., Buus, S., and Geng, W. (1999). "Psychometric functions for gap detection in a yes-no procedure," *J. Acoust. Soc. Am.* **106**, 3512–3520.
- Glasberg, B. R., Moore, B. C. J., and Bacon, S. P. (1987). "Gap detection and masking in hearing-impaired and normal-hearing subjects," *J. Acoust. Soc. Am.* **81**, 1546–1556.
- Heinrich, A., and Schneider, B. (2006). "Age-related changes in within- and between-channel gap detection using sinusoidal stimuli," *J. Acoust. Soc. Am.* **119**, 2316–2326.
- Keith, R. W. (2000). *Random Gap Detection Test*, Auditec, Inc., St Louis, www.auditec.com (Last viewed 2023/3/3).
- King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., and Supowit, A. (1994). "Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation," *Vision Res.* **34**, 885–912.
- Kuss, M., Jäkel, F., and Wichmann, F. A. (2005). "Bayesian inference for psychometric functions," *J. Vision* **5**, 8–492.
- Li, M., and Zhu, J. (2022). "Bayesian adaptive estimation with theoretical bound: An exploration-exploitation approach," *Comput. Intell. Neurosci.* **2022**, 1143056.
- Lister, J. J., Roberts, R. A., Shackelford, J., and Rogers, C. L. (2006). "An adaptive clinical test of temporal resolution," *Am. J. Audiol.* **15**, 133–140.
- Marvit, P., Florentine, M., and Buus, S. (2003). "A comparison of psychophysical procedures for level-discrimination thresholds," *J. Acoust. Soc. Am.* **113**, 3348–3361.
- Mori, S., Kikuchi, Y., Hirose, N., Lepage, H., and Wong, W. (2018). "Auditory gap detection: Psychometric functions and insights into the underlying neural activity," *Biol. Cybern.* **112**, 575–584.
- Mori, S., Morimoto, T., Murata, Y., Okamoto, Y., and Kanzaki, S. (2023). "Tests of human auditory temporal resolution: Preliminary investigation of ZEST parameters for amplitude modulation detection," *Front. Neurosci.* **17**, 1148476.
- Mori, S., Oyama, K., Kikuchi, Y., Mitsudo, T., and Hirose, N. (2015). "Between-frequency and between-ear gap detections and their relation to perception of stop consonants," *Ear Hear.* **36**, 464–470.
- Morimoto, T. (2019). "Estimation methods of auditory temporal resolution for clinical application," Wakayama University, <https://repositor-y.center.wakayama-u.ac.jp/ja/3575> (Last viewed 2023/3/3).
- Okamoto, Y., Kanzaki, S., Kanno, A., Nakaichi, T., Morimoto, T., Harada, K., Kubota, E., and Ogawa, K. (2014). "Temporal resolution measurement in presbycusis," *Audiol. Jpn.* **57**, 694–702.
- Phillips, D. P., and Smith, J. C. (2004). "Correlations among within-channel and between-channel auditory gap-detection thresholds in normal listeners," *Perception* **33**, 371–378.
- Treutwein, B. (1995). "Adaptive psychophysical procedures," *Vision Res.* **35**, 2503–2522.
- Turpin, A., McKendrick, A. M., Johnson, C. A., and Vingrys, A. J. (2003). "Properties of perimetric threshold estimates from full threshold, ZEST, and SITA-like strategies, as determined by computer simulation," *Invest. Ophthalmol. Vis. Sci.* **44**, 4787–4795.
- Watson, A. B., and Pelli, D. (1983). "QUEST: A Bayesian adaptive psychophysical method," *Percept. Psychophys.* **33**, 113–120.