Discrimination of frequency variance for tonal sequences^{a)}

Andrew J. Byrne,^{b)} Neal F. Viemeister, and Mark A. Stellmack

Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455

(Received 10 July 2014; revised 4 September 2014; accepted 17 October 2014)

Real-world auditory stimuli are highly variable across occurrences and sources. The present study examined the sensitivity of human listeners to differences in global stimulus variability. In a two-interval, forced-choice task, variance discrimination was measured using sequences of five 100-ms tone pulses. The frequency of each pulse was sampled randomly from a distribution that was Gaussian in logarithmic frequency. In the non-signal interval, the sampled distribution had a variance of σ_{STAN}^2 , while in the signal interval, the variance of the sequence was σ_{SIG}^2 (with $\sigma_{\text{SIG}}^2 > \sigma_{\text{STAN}}^2$). The listener's task was to choose the interval with the larger variance. To constrain possible decision strategies, the mean frequency of the sampling distribution of each interval was randomly chosen for each presentation. Psychometric functions were measured for various values of σ_{STAN}^2 . Although the performance was remarkably similar across listeners, overall performance was poorer than that of an ideal observer (IO) which perfectly compares interval variances. However, like the IO, Weber's Law behavior was observed, with a constant ratio of $(\sigma_{\text{SIG}}^2 - \sigma_{\text{STAN}}^2)$ to σ_{STAN}^2 yielding similar performance. A model which degraded the IO with a frequency-resolution noise and a computational noise provided a reasonable fit to the real data.

© 2014 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4900825>]

PACS number(s): 43.66.Fe, 43.66.Mk [FJG]

Pages: 3172–3177

I. INTRODUCTION

How we extract useful information about stimuli in the face of considerably variable sensory inputs is a fundamental problem in perception. This issue has many manifestations but generally falls under the broad category of perceptual constancy. In audition, this has received the most attention in speech perception. How are the “objects” of speech, such as phonemes and syllables, identified despite considerable temporal and spectral variability in their acoustic signatures across utterances and speakers? The focus of the study described here was on a basic aspect of this broad and difficult problem; namely, on the perception of variability itself rather than perception in the face of variability.

There is a sparse literature on variability perception and little is known as to how well observers can actually calculate a statistic such as variance for comparison across stimuli. Morgan *et al.* (2008) investigated the discrimination of line orientation variance within visual textures. They found that, as the variance of the standard texture elements increased, the just-noticeable-difference for variance first decreased but then increased, thus showing a so-called “dipper” function. Their results also suggested that a trend toward Weber's Law might be occurring for discrimination of their largest standard variances. Wasserman *et al.* (2004) broadly considered issues related to variability discrimination and reviewed data from both humans and animals; however, in general, those experiments were more related to variability detection rather than on discrimination of variability.

In audition, past research on variability has primarily involved examination of its effects on detection, discrimination, and recognition (e.g., Watson *et al.*, 1976; Kidd and Watson, 1992; Lutfi, 1992; Durlach *et al.*, 2005) rather than the discrimination of variability *per se*. One study that is more directly relevant to the question of variability discrimination is a study of spectral variance discrimination by Lutfi *et al.* (1996). In that study, the level of each component in a multi-tone complex was randomly sampled from Gaussian distributions with different variances, and the listener chose which of the two complexes had the greater variance in component levels. Lutfi *et al.* proposed that the component weights derived in their experiments may vary over trials or that decisions were based on the maximum level of a single tone in each complex. There is also evidence that, in temporal pattern discrimination, listeners tend to give greater weight to the intervals that have the greatest variability (Sadralodabai and Sorkin, 1999). As in Lutfi *et al.* (1996), this may not reflect sensitivity to variability itself, but rather a strategy based on simple increases in the maxima or minima that are a by-product of increased variance.

Finally, there is fairly extensive published research on what might be called “microvariance” discrimination (i.e., discrimination in which short-term variability may determine performance), for example, detection of a noise signal within a noise masker such as studied in the classic work of Green (1960). For Gaussian noise, power (intensity) is proportional to variance and thus discrimination between independent noise bursts differing in mean intensity could be viewed as a type of variance discrimination. To what extent the intrinsic variability in the noise limits human performance is uncertain except for brief bursts, on the order of milliseconds.

The present work is a psychophysical examination of variance perception for moderately long sequences of

^{a)}A portion of these data was presented at the 161st Meeting of the Acoustical Society of America (Viemeister *et al.*, 2011).

^{b)}Author to whom correspondence should be addressed. Electronic-mail: byrn0050@umn.edu

sounds. The general question that is addressed concerns the ability of listeners to discriminate differences in the frequency variance of a sequence of temporally discrete tones. The relationship between performance and various factors such as the baseline (standard) variance was investigated, as well as the possible use of other related decision metrics. Finally, performance was compared with that of a simulated observer, which can perfectly calculate variance, and a model was proposed that incorporates a relatively simple decision statistic and a multiplicative “computational” noise.

II. METHODS

A. Stimuli

Discrimination of frequency variance was examined using sequences of tones drawn randomly from different frequency distributions. The listeners were presented with a sequence of five 100-ms, 70-dB-sound-pressure-level (SPL) pure tones in each of the two observation intervals. Each tone pulse was gated using 20-ms raised-cosine, on-and-off ramps. There was 30 ms of silence between each pulse within an interval and 500 ms of silence between each interval.

The frequency of each pulse within an interval was selected randomly and independently from a Gaussian (in logarithmic frequency) distribution with a given variance, the “underlying variance” for that interval. In different conditions, the “standard” interval variance (σ_{STAN}^2) was either 0, 2e-5, 7e-5, 45e-5, or 171e-5 in units of log(Hz), with σ corresponding to 0, 14e-3, 29e-3, 70e-3, and 138e-3 octaves, respectively. For the “signal” interval, the variance (σ_{SIG}^2) was greater than σ_{STAN}^2 and varied across experimental conditions.

In order to reduce the effectiveness of using decision statistics other than variance (such as listening for the interval with the highest frequency), a frequency rove was imposed on the mean frequency (M_f) of the underlying distribution of each interval prior to selecting the individual pulse frequencies. The M_f of the distribution for each interval was chosen randomly from a Gaussian (logarithmic frequency) distribution with a mean of 2 kHz and standard deviation (σ) of 0.176 in log(Hz) units. The selection of M_f was limited to within 2.5σ of 2 kHz (725–5510 Hz). To ensure high audibility of all frequencies sampled from the resulting distribution, frequency limits of 100 Hz and 12 kHz were set for the experiment. After choosing M_f randomly, if the $\pm 2.5\sigma$ cutoffs of the resulting frequency distribution exceeded those limits, the M_f value was shifted to either 725 or 5510 Hz such that the entire distribution was within a more acceptable range.

To summarize, for the first observation interval, an M_f was randomly selected from a distribution of possible mean frequency values. That selected value became the mean frequency of another distribution (with a variance determined by the particular condition) from which five frequencies were randomly and independently selected to make up the tonal sequence for that interval. That process was then repeated for the second interval. Figure 1 illustrates a sample trial with the signal interval in Interval 2.

FIG. 1. An illustration of the temporal sequence of a sample trial with the second (signal) interval having greater frequency variance than the first (standard) interval.

B. Procedure and apparatus

Two-interval, forced-choice (2-IFC) psychometric functions were measured on each individual listener using fixed σ_{STAN}^2 and σ_{SIG}^2 values, and the listeners were instructed to select the interval with the greater frequency variance. Feedback was given based on the variances of the underlying distributions from which samples were drawn, not the variances of the two sequences of sampled frequencies (referred to as the “sample variances” of each interval). Therefore, occasionally, listeners could have correctly selected the interval with the greater sample variance, but then received feedback that their response was incorrect.

A variety of σ_{SIG}^2 values was tested on each listener for each σ_{STAN}^2 value to determine the appropriate range that would produce a 4–6 point function with d' s ranging from 0.3 to 3. The σ_{SIG}^2 values were selected such that all listeners were tested on identical values, although some listeners ran additional conditions with values that were outside the range of the other listeners. Blocks of 50 trials were run until 200 trials were obtained for each condition, but with blocks of different conditions interleaved. Conditions with different σ_{STAN}^2 values, however, were not interleaved to help the listeners become familiar with the perceptual quality of the standard variance.

Listeners were run individually in a sound-attenuating chamber (Industrial Acoustics Company). The stimuli were generated digitally in MATLAB (The MathWorks) on a personal computer equipped with a 24-bit sound card (GINA 3G, Echo Digital Audio) and were presented over stereo headphones (MDR-V6, Sony Corporation) to only the listener’s left ear. Listeners were run in 2-h sessions over the course of approximately two months.

C. Listeners

Four listeners participated in the experiment (one male and three female). One was the first author who had previous experience with other psychoacoustics tasks, while the others were students from the University of Minnesota who were paid to participate. All listeners had normal hearing based on pure-tone thresholds of 15 dB hearing level (HL) or better at octave frequencies from 250 to 8000 Hz and were trained to stable performance in the present task before the final data were collected.

FIG. 2. Average psychometric functions across four listeners showing performance (d') as a function of the difference in interval variance [underlying variance of the signal interval (σ_{SIG}^2) minus the variance of the standard (σ_{STAN}^2)] for each standard variance function (different symbols). Error bars represent the standard errors of the mean d' 's.

III. RESULTS

Psychometric functions for the individual listeners were constructed from the different σ_{SIG}^2 conditions for each σ_{STAN}^2 value. The measure of performance chosen was d' , despite sampled variance not having equal-variance Gaussian distributions (see the Appendix). In the event that a listener did not obtain at least one of each response type (hit, miss, false alarm, and correct rejection) over the course of the 200 trials for each condition, 0.5 was added to all four values before calculating d' . Due to the two-interval task, d' was divided by the square-root of 2 for the final d' performance value (Macmillan and Creelman, 2005). Various metrics for the abscissa were tested, with the most intuitive way to illustrate the results initially being in terms of the difference between the underlying signal and standard variances. Due to the strong similarity between individual listeners, the psychometric functions shown in Fig. 2 are averaged across listeners, but only for the points for which all four listeners provided data. The standard errors of the mean were generally small and are shown by the error bars.

The different symbol types in Fig. 2 represent the different σ_{STAN}^2 values and show psychometric functions with roughly similar slopes. An orderly trend was observed between the different standard variances; as the standard variance increased, the difference between the variances of the signal and standard needed to be greater to produce the same level of performance. It was therefore reasonable to attempt to develop a metric comparable to Weber's Law to interpret the results. In Fig. 3, the individual listeners' functions are plotted in terms of d' as a function of the ratio between the interval variance difference and the standard variance [$(\sigma_{\text{SIG}}^2 - \sigma_{\text{STAN}}^2)/\sigma_{\text{STAN}}^2$]. There was considerable overlap of all four functions, suggesting this lawful relationship does account for the variance discrimination results. [The variance detection ($\sigma_{\text{STAN}}^2 = 0$) function cannot be plotted using this metric.]

The listener performance shown in Figs. 2 and 3 was based on the feedback of the 2-IFC task, i.e., which interval used the underlying σ_{SIG}^2 distribution, not necessarily the interval with larger sample variance. Additional analyses were performed on each function of each listener to determine what potential cues the listeners were using as their decision variable. During the running of the experiment, the frequencies of all tones presented in each trial were recorded for later analysis. This trial-by-trial information was used to

analyze responses for the following different definitions of a "correct" response: (1) the greater actual sample variance of the two intervals (s^2), (2) the greater frequency range of the intervals (max-min), (3) the interval with the greatest maximum frequency that was presented (max), or (4) the interval with the lowest minimum frequency (min).

The general pattern was similar across listeners and σ_{STAN}^2 functions, so only the average results across listeners for the largest standard variance is shown in Fig. 4. The filled-diamond function represents the performance the "average" listener would have obtained based on a comparison of the sample variances of the two intervals. The frequency range function (shaded squares), is just below the s^2 function due to the high correlation ($r = 0.94$) between these two measures. Finally, both the maximum- and minimum-frequency listening strategies produced substantially poorer performance relative to the previous metrics, strongly suggesting that the listeners were not using those strategies when making their responses. This analysis also served to

FIG. 3. Performance (d') as a function of the difference in variance relative to the standard variance for each of the four listeners (separate panels). (There were no negative d' values for any listener.) Error bars represent the standard errors of the mean d' 's.

FIG. 4. For the largest standard variance ($\sigma_{\text{STAN}}^2 = 171\text{e-}5$), d' based on listener responses compared to different decision variables, averaged across four listeners. The d' values are based on the level of performance the listeners would have achieved given the different (*post hoc*) definitions of a correct response.

confirm that the mean frequency rove that had been implemented did in fact reduce the reliance on such cues.

IV. MODELING THE DATA BY DEGRADING AN IDEAL OBSERVER

The ideal observer (IO) for the present task would be similar to that for detection of a noise signal within a noise masker (Green, 1960) where the underlying distributions of an instantaneous sample are Gaussian with identical means but with different variances (intensities) for signal-plus-noise and noise-alone trials. As was shown in that analysis, the sample variance computed over independent samples is monotonic with likelihood ratio and, in that sense, is the optimal decision statistic. That is also the case for the present task, and thus the IO can use sample variance to achieve optimal performance. For both tasks, the distributions of likelihood ratio are χ^2 with different means but are well-approximated with Gaussian distributions (see the Appendix).

In the present sample variance discrimination task, even an IO that could accurately calculate the variance of each interval would not achieve perfect performance because the correct answer feedback was based on the underlying variances of each interval, not the actual sample variances obtained on a given trial. In order to determine how well the IO would perform, computer simulations were run in which stimuli were generated as in the real experiment and the simulated observer responded based on the sample variances of each interval. The results of these simulations are plotted in Fig. 5 along with the average results across listeners (the average across the four panels of Fig. 3). The four functions of the IO overlap considerably using the Weber's Law metric and thus are plotted simply as solid lines. The smoothness of the IO curves resulted from running a total of 1000 trials per condition (while the data points from the real listeners represent only 200 trials).

The performance of the IO was roughly a factor of two (or more) better than the real listeners; therefore, a model

FIG. 5. Performance (d') as a function of the difference in variance relative to the standard variance. The average results across the four listeners from Fig. 3 (symbols, standard error bars) are plotted along with the simulated results of an IO whose decisions are based on the sample variance of each interval (overlapping solid lines).

was developed to better simulate the performance of the real listeners. This model consisted of degrading the performance of an IO by adding internal noise at two hypothesized stages of the processing of the stimuli.

A. Frequency-resolution noise

The IO originally was modeled such that it was able to use a perfect estimation of the frequencies of the sequence when calculating the variance of each interval. One reasonable way in which the performance of the IO could be degraded is by adding noise to the estimation of the frequency of each tone pulse, thus simulating an observer that does not have a perfect frequency representation of the sequence. [This is comparable to the internal noise implemented in the model of Morgan *et al.* (2008) to account for visual-orientation, variance-discrimination thresholds. See Berg and Robinson (1987) regarding internal noise for a sample frequency discrimination task.]

To facilitate modeling, the four real listeners in the present study performed an additional 2-IFC frequency discrimination experiment, the results of which were used to estimate the noise that should be added to the frequency estimation of each pulse by the IO. These measurements used an adaptive procedure to determine frequency discrimination thresholds from a 2-kHz standard frequency, using a single 100-ms pure tone for each interval. [For all other methodological details, refer to the frequency discrimination condition of Byrne *et al.* (2013).] These thresholds (ranging from 3.6 to 5.1 Hz) were used to determine a Gaussian “jitter” that was added to the frequency of each tone in the simulations using the IO. Specifically, the magnitude of the noise added to the frequency of each tone (in terms of the σ of the Gaussian distribution, σ_{FRN}) was manipulated for a simulated (but degraded) IO so that the IO would obtain the same level of performance in the frequency discrimination condition as each real listener. (The noise was added to the logarithmic frequency values; therefore, in units of Hz, the magnitude of the noise was also relative and proportional to the original frequency.)

Psychometric functions were generated for the degraded ideal observer (DIO) with this frequency-resolution noise present on the frequency value of each pulse before the DIO calculated the variance of each interval. The resulting DIO functions were different from the original IO functions, but performance was not reduced to that of the real listeners, primarily due to the frequency-resolution noise only substantially affecting the smaller standard variance functions.

B. Computational noise

Because additional noise would be necessary to degrade the IO to the level of performance of the real observers, it was assumed that noise was present in the real listeners' *calculations* of the variance of each interval. We also assumed that the error in the variance calculation should be relative, and on average proportional, to the actual variance because sample variance is distributed as χ^2 . As the mean of the χ^2 distribution increases (as σ_{STAN}^2 increases), the standard deviation of the sample distribution increases proportionally to the mean of the distribution. (This multiplicative-noise property accounts for the Weber's Law behavior shown by the IO.)

Calculating a realistic level of this proportional "computational" noise for the DIO involved manipulating the magnitude of this noise after the addition of the frequency-resolution noise. The DIO computer simulations systematically adjusted the amount of computational noise (but always proportional to the standard deviation of the actual simulated interval) and compared the resulting DIO functions to that of a real listener. A least sum-of-squares fit was used to determine the single best fitting value of the computational noise for all data from an individual listener. Table I lists the magnitudes of both the frequency-resolution noise (σ_{FRN}) and the computational noise (σ_{CN} , with the standard variance factored out), as well as the percent of variance accounted for (PAF) by the simulated DIO functions in modeling the results of each real listener.

The noise values listed in Table I are expressed in terms of the standard deviation of the Gaussian distributions (the amount of frequency or computational "jitter") applied, but the actual values and [log(Hz)] units may not be intuitive.¹ However, the amount of noise is quite similar for the first three listeners, with the values for the fourth listener being somewhat different due to a higher frequency-resolution

TABLE I. The estimated amounts (σ) of Gaussian noise for each listener resulting from the final modeling simulations along with the percentage of variance accounted for (PAF) by the fitted functions. For each standard variance, the actual computational noise value is $\sigma_{\text{CN}} * \sigma_{\text{STAN}}$, such that the noise is at a constant proportion to the standard. The frequency-resolution noise is added to log(Hz) values, so its magnitude is relative to the original frequency of each pulse.

Listener	Frequency-resolution noise [σ_{FRN} , log(Hz)]	Computational noise [σ_{CN} , log(Hz)]	PAF
S1	7.3e-4	0.61	86.6%
S2	7.6e-4	0.60	89.3%
S3	7.3e-4	0.62	82.0%
S4	10.2e-4	0.41	75.9%
Average data	8.1e-4	0.57	91.5%

noise (due to that listener having higher frequency discrimination thresholds measured in the aforementioned adaptive task.) The PAF values range from 76% to 89% for the individual listeners, indicating that the DIO simulations provided reasonably good fits to the real data. Finally, the same DIO noise simulations were performed on the average data across listeners, and the resulting functions are plotted along with the average data in Fig. 6. The model for the average-across-listeners results fit the functions with a PAF of 92%.

V. DISCUSSION

Discrimination of frequency variance over time was measured using sequences of tone pulses, and an orderly pattern of results was observed. It appears that listeners did in fact respond to the variability of the two sequences rather than less reliable cues, such as the maximum frequency, that potentially would be useful under less variable conditions (i.e., in the absence of an overall frequency rove). Additional research would be necessary to distinguish between the variance and frequency-range listening strategies. Since the decision statistics are essentially interchangeable for the present experiment (because of their high correlation), one or the other can be substituted into the proposed model to the same end.

There was also surprisingly low variability between the performance of the individual listeners, as shown by the small error bars in Figs. 2 and 4, as well as by the general similarity between panels of Fig. 3. Although two of the four listeners initially required extra training at the task to produce orderly psychometric functions, the final results of all four listeners show roughly the same sensitivity to variance using the present stimuli.

Despite frequency being the parameter varied, the present work may have implications for the discrimination of stimulus variance in general. In fact, another experiment was piloted using intensity variance, instead of frequency, and produced a similar pattern of results, albeit with the much smaller amount of data that was collected. It would have been interesting to compare the estimated computational noise values across the different domains, but the intensity variance task was abandoned due to practical limitations on the range of the sample intensity distributions. However, we

FIG. 6. Modeling results (solid lines) of the mean data (symbols, from Fig. 2, with standard error bars). Performance (d') as a function of interval variance difference (left panel) and variance difference relative to the standard variance (right panel). (See text for modeling details.)

would predict that the general (Weber's Law) trend observed with frequency variance discrimination should hold for other stimuli and variations of the paradigm, as was seen somewhat in vision by [Morgan et al. \(2008\)](#).

In summary, listeners can reliably detect differences in the frequency variability of sequences of tone pulses, and performance in terms of d' is only a factor of two poorer than that of an ideal observer which bases its decision on sample variance. Like the IO, the discrimination of variance by the real listeners followed Weber's Law; a constant ratio of the difference in interval variances to the standard variance yielded similar performance. A straightforward model was constructed (using a degraded IO) that incorporated a frequency-resolution noise (based on measured frequency-discrimination thresholds) and a multiplicative computational noise that was estimated in order to fit the data. The computational noise reflected the idea that listeners were imperfect in their computation of stimulus variance for each interval, and the resulting model accounted for the averaged listeners' data extremely well.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute on Deafness and Communication Disorders of the National Institutes of Health under award number R01DC00683. The authors would like to thank Dr. Frederick Gallun and Dr. Robert Lutfi who provided helpful comments and suggestions for improving this manuscript.

APPENDIX: RECEIVER OPERATING CHARACTERISTIC (ROC) CURVES

Because sampled variances are distributed as χ^2 (rather than Gaussian) distributions, receiver operating characteristic (ROC) curves were created to justify the use of d' as the measure of performance. Using the stimuli of the present experiment, including the M_f rove and frequency distribution ($\pm 2.5\sigma$) truncations, simulations were run and ROC curves were constructed based on the responses of a simulated observer that would obtain the specified d' values, as approximated from the IO functions from Fig. 5. The sample variance (s^2) decision statistic produced nearly identical ROC curves as those for equal-variance Gaussian distributions with specified means ([Macmillan and Creelman, 2005](#)), as shown in the left panel of Fig. 7.

The similarity between the ROC curves for variance discrimination (black solid lines) and those for sample discrimination using Gaussian distributions (gray dashed lines) is the result of the proportion of overlap between the distributions, not their shapes. Although χ^2 distributions (Fig. 7, right panel, upper curves) are obviously skewed, they are easily converted to more "normal" distributions with a logarithmic transform of variance (right panel, lower curves). [Due to a logarithmic transform being monotonic for all values greater than zero, the $\log(s^2)$ decision statistic would produce identical responses to those based on sample variance (s^2), though it is presented here for illustrative purposes only.] Thus, the

FIG. 7. Left panel: ROC curves for the present variance discrimination task (black solid lines) plotted against those that would be obtained from underlying equal-variance Gaussian distributions (gray dashed lines). Right panel: χ^2 distributions (upper distributions, s^2) which produced the ROC curves, with the solid line being the "standard," along with $\log(s^2)$ values (lower distributions) for comparison.

χ^2 distributions (for the present experiment) overlap in a proportionally similar manner as equal-variance Gaussian distributions. Any comparison of sample values or distributions of values (e.g., likelihood ratios) would also be similar to those for Gaussian distributions.

¹The average σ_{FRN} is 0.00081 in $\log(\text{Hz})$ units (i.e., linear standard deviation); therefore, in Hz, that standard deviation is actually a multiplier value of 1.0019. For the present frequency discrimination task, the $+1\sigma$ value of the noise at the 2-kHz standard frequency would be $2000 \times 1.0019 = 2000$, or 3.8 Hz. (Note that this noise value roughly corresponds to the actual thresholds.)

- Berg, B. G., and Robinson, D. E. (1987). "Multiple observations and internal noise," *J. Acoust. Soc. Am.* **81**, S33.
- Byrne, A. J., Viemeister, N. F., and Stellmack, M. A. (2013). "The effects of unmodulated carrier fringes on the detection of frequency modulation," *J. Acoust. Soc. Am.* **133**, 998–1003.
- Durlach, N. I., Mason, C. R., Gallun, F. J., Shinn-Cunningham, B., Colburn, S., and Kidd, G., Jr. (2005). "Informational masking for simultaneous non-speech stimuli: Psychometric functions for fixed and randomly mixed maskers," *J. Acoust. Soc. Am.* **118**, 2482–2497.
- Green, D. M. (1960). "Auditory detection of a noise signal," *J. Acoust. Soc. Am.* **32**, 121–131.
- Kidd, G. R., and Watson, C. S. (1992). "The 'proportion-of-the-total-duration rule' for the discrimination of auditory patterns," *J. Acoust. Soc. Am.* **92**, 3109–3118.
- Lutfi, R. A. (1992). "Informational processing of complex sound. III: Interference," *J. Acoust. Soc. Am.* **91**, 3391–3401.
- Lutfi, R. A., Doherty, K. A., and Oh, E. (1996). "Psychometric functions for the discrimination of spectral variance," *J. Acoust. Soc. Am.* **100**, 2258–2265.
- Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*, 2nd ed. (Erlbaum, Mahwah, NJ), pp. 7–184.
- Morgan, M., Chubb, C., and Solomon, J. A. (2008). "A 'dipper' function for texture discrimination based on orientation variance," *J. Vis.* **8**(11), 9, 1–8.
- Sadrulodabai, T., and Sorkin, R. D. (1999). "Effect of temporal position, proportional variance, and proportional duration on decision weights in temporal pattern discrimination," *J. Acoust. Soc. Am.* **105**, 358–365.
- Viemeister, N. F., Stellmack, M. A., and Byrne, A. J. (2011). "Discrimination of stimulus variance," *J. Acoust. Soc. Am.* **129**, 2588.
- Wasserman, E. A., Young, M. E., and Cook, R. G. (2004). "Variability discrimination in humans and animals," *Am. Psych.* **59**, 879–890.
- Watson, C. S., Kelly, W. J., and Wroton, H. W. (1976). "Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty," *J. Acoust. Soc. Am.* **60**, 1176–1186.