# Informational masking and musical training

Andrew J. Oxenham[a]
*Research Laboratory of Electronics, Massachusetts Institute of Technology,*
*Cambridge, Massachusetts 02139, and Harvard-MIT Division of Health Sciences and Technology,*
*Speech and Hearing Bioscience and Technology Program, Cambridge, Massachusetts 02139*

Brian J. Fligor, Christine R. Mason, and Gerald Kidd, Jr.
*Hearing Research Center and Programs in Communication Disorders, Boston University,*
*635 Commonwealth Avenue, Boston, Massachusetts 02215*

(Received 27 November 2002; revised 23 May 2003; accepted 16 June 2003)

The relationship between musical training and informational masking was studied for 24 young adult listeners with normal hearing. The listeners were divided into two groups based on musical training. In one group, the listeners had little or no musical training; the other group was comprised of highly trained, currently active musicians. The hypothesis was that musicians may be less susceptible to informational masking, which is thought to reflect central, rather than peripheral, limitations on the processing of sound. Masked thresholds were measured in two conditions, similar to those used by Kidd *et al.* [J. Acoust. Soc. Am. **95**, 3475–3480 (1994)]. In both conditions the signal was comprised of a series of repeated tone bursts at 1 kHz. The masker was comprised of a series of multitone bursts, gated with the signal. In one condition the frequencies of the masker were selected randomly for each burst; in the other condition the masker frequencies were selected randomly for the first burst of each interval and then remained constant throughout the interval. The difference in thresholds between the two conditions was taken as a measure of informational masking. Frequency selectivity, using the notched-noise method, was also estimated in the two groups. The results showed no difference in frequency selectivity between the two groups, but showed a large and significant difference in the amount of informational masking between musically trained and untrained listeners. This informational masking task, which requires no knowledge specific to musical training (such as note or interval names) and is generally not susceptible to systematic short- or medium-term training effects, may provide a basis for further studies of analytic listening abilities in different populations. © *2003 Acoustical Society of America.*

[DOI: 10.1121/1.1598197]

PACS numbers: 43.66.Dc, 43.75.Cd [MRL]

## I. INTRODUCTION

The ability to distinguish, or ‘‘hear out,’’ a sequence of tones, such as a melodic line, in the presence of other tones is fundamental to the appreciation of most forms of music. When individual tones are not perceived due to the presence of other sounds, they are said to be masked. The most commonly studied form of masking, sometimes referred to as ‘‘energetic masking’’ (Pollack, 1975; Leek *et al.*, 1991; Kidd *et al.*, 1994), is thought to be determined primarily by the frequency selectivity of the peripheral auditory system: when the masker and signal are sufficiently close in frequency, and when the masker is sufficiently intense, the peripheral neural representation of the masker dominates that of the signal to such an extent that no subsequent cognitive strategies are sufficient to extract the signal.

Masked thresholds can also be influenced by attention, listener expectations, and uncertainty about the signal’s characteristics. For instance, if the listener is presented with a background of white noise and, through experimental manipulations or instructions, is expecting a tone of a particular frequency or duration to be presented, tones differing in some way from expectations are often more poorly detected (Greenberg and Larkin, 1968; Scharf *et al.*, 1987; Hafter *et al.*, 1993; Wright and Dai, 1994).

Finally, masked thresholds in situations involving *masker* uncertainty are often much higher than predicted by energetic masking alone. For instance, if the masker consists of a number of tones having frequencies that are changed randomly from presentation to presentation, listeners often have great difficulty in detecting a signal of a fixed and known frequency, even if the masker frequencies are always far removed from that of the signal. Masking that cannot be explained in terms of peripheral frequency selectivity is often referred to as ‘‘informational masking’’ (Pollack, 1975; Watson, 1987; Neff *et al.*, 1993), and is almost certainly mediated at higher stages of perceptual processing.

Informational masking is often accompanied by large differences in performance between listeners (e.g., Lutfi *et al.*, 2003). In a comprehensive study using up to 49 listeners, Neff and Dethlefs (1995) found individual differences in masked thresholds as large as 50 dB. For detecting a 1-kHz signal in a masker comprised of ten random-frequency tones, the standard deviation around the mean was 11 dB. This is considerably larger than the standard deviations of 2 dB or less often found for the detection of the same signal in broad-

---

[a] Electronic mail: oxenham@mit.edu

band noise. The large effects of, and differences in, informational masking persist even after extended practice (Neff and Callaghan, 1988; Neff and Dethlefs, 1995).

One obvious question is whether listeners at the extreme ends of performance in informational masking employ different listening strategies. A cursory analysis performed by Neff and Dethlefs (1995) showed that this might indeed be the case. They found that their ‘‘best’’ listeners had thresholds that seemed to be determined by energetic masking. That is, their thresholds were similar to what would have been expected from an ‘‘optimal’’ single-channel model, where only the output of the auditory filter centered on the signal frequency was analyzed. At the other extreme, their ‘‘worst’’ listeners had thresholds that could be predicted based on the level of the overall stimulus: these listeners seemed unable to ‘‘hear out’’ the signal and instead may have based their judgments on an impression of the overall loudness of the total tone complex. In summary, a distinction could be made between ‘‘analytic’’ and ‘‘holistic’’ listeners, although the population formed a continuum rather than two distinct groups.

So far, it has proved challenging to identify predictors of how listeners will perform in an informational masking task. Neff *et al.* (1996) found a significant effect of sex in their group of 49 listeners, with females tending to have somewhat higher thresholds than males. Other measures of auditory performance, such as absolute thresholds or thresholds for a tone in broadband noise, have yielded no significant correlations with thresholds in informational masking situations (Neff *et al.*, 1993; Neff and Dethlefs, 1995).

Musicianship, both active and passive, relies in part on an ability to listen analytically. Thus, a reasonable hypothesis is that listeners with a high level of musical training and ability (musicians) should be able to listen more analytically, and hence perform better in informational masking tasks, than listeners with no formal musical training (nonmusicians). No study has yet directly tested this hypothesis. However, other studies have indicated that nonmusicians may have more difficulty in tasks requiring analytic listening. Soderquist (1970) found that nonmusicians were less able to identify individual partials within a harmonic tone complex than were musicians. He described this difference in terms of poorer frequency selectivity on the part of nonmusicians. Fine and Moore (1993) confirmed Soderquist’s finding that nonmusicians had more difficulty than musicians in identifying individual tones within a complex. However, they also measured auditory filter shapes, using the notched-noise method, whereby masked thresholds for a tone are measured in the presence of noise with a spectral notch (Patterson, 1976; Glasberg and Moore, 2000). They found no relationship between the bandwidth of the auditory filters and level of musical training. In short, while the ability to identify tones in a background of other tones might be poorer in nonmusicians, there is no evidence for poorer peripheral frequency selectivity, as originally implied by Soderquist (1970).

Another study comparing the performance of musicians and nonmusicians (Spiegel and Watson, 1984) investigated frequency discrimination. The tasks included a simple comparison of two successive tones and a more complex task, where the target tones formed part of a longer sequence of tones. The general conclusion was that the musicians’ performance was initially better but that, with extended training, the performance of nonmusicians could be brought up to that of musicians. The effects of training are particularly striking in the frequency discrimination and pattern recognition tasks employed by Spiegel and Watson (1984). In contrast, practice effects in informational masking tasks appear to be much less robust. The two studies that have examined learning effects in informational masking (Neff and Callaghan, 1988; Neff and Dethlefs, 1995) both found little evidence for systematic learning effects: although certain individuals did show improvements with time, there was generally no systematic trend when examining group data.

The assumption that good performance requires analytic listening, together with the finding that practice effects tend not to be particularly large or systematic, makes informational masking an attractive task for testing analytic listening abilities. The aim of this study was to establish whether musicians do in fact exhibit less informational masking than do nonmusicians, as might be expected if musicians are able to listen more analytically. Detection thresholds were measured in an informational masking task, similar to that used by Kidd *et al.* (1994), in two groups comprising musicians and nonmusicians, with equal numbers of males and females in each group. In the same group of listeners, frequency selectivity was measured using a version of the notched-noise method (Stone *et al.*, 1992; Glasberg and Moore, 2000). This enabled us to test for any differences in peripheral frequency selectivity in our two groups. Based on the data of Fine and Moore (1993), none was expected.

## II. METHODS

### A. Stimuli

All sounds were computer generated at a rate of 20 kHz, were played through 16-bit digital-to-analog converters (Tucker-Davis Technology), and were then low-pass filtered at 7500 Hz. The signal was a single 200-ms (total duration) burst of a 1000-Hz tone for the noise-masker conditions, or a sequence of 60-ms (total duration) 1000-Hz tone bursts for the multitone masking conditions. In both cases, the signals were gated on and off with 10-ms raised-cosine ramps. In all masking conditions, the signal was presented at a fixed level of 20 dB above absolute threshold, adjusted for each subject individually.

The tonal maskers were comprised of a series of random-frequency multitone complexes. The multitone complexes were played in a sequence of eight contiguous bursts having rise/steady-state/decay times of 10/40/10 ms, for a total duration of 480 ms. The signal, when present, was gated on and off synchronously with the masker bursts. There were two versions of the multitone masker that differed in the way the frequencies of the tones were randomized. For one masker, referred to as ‘‘multiple-bursts same’’ (MBS), the frequencies of the eight masker tones in the first burst were randomly drawn from a uniform distribution of frequencies, on a logarithmic scale, ranging from 200–5000 Hz excluding

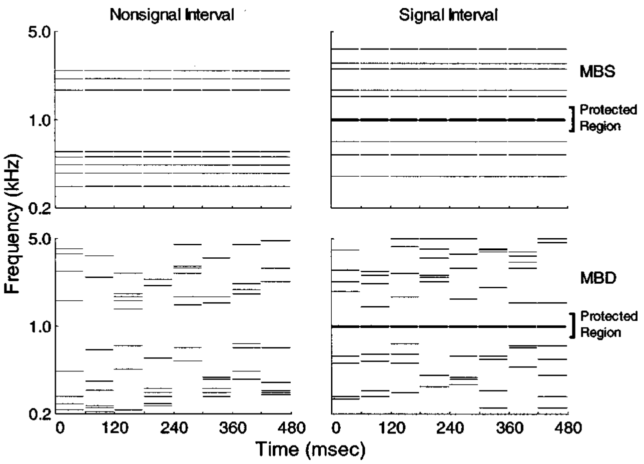

FIG. 1. Schematic diagram of the stimuli used in the random-frequency multitone masker conditions. The left panels show typical masker-alone samples that might occur in the nonsignal interval of a trial, while the right panels show typical signal-plus-masker samples that might occur in the signal interval of a trial. The signal is shown with heavy lines. The upper row illustrates the "multiple-bursts same" (MBS) masker while the lower row illustrates the "multiple-bursts different" (MBD) masker (see the text for details).

a "protected region" with a bandwidth of 400 Hz centered geometrically around the 1000-Hz signal frequency. This protected region is designed to reduce the amount of energetic masking produced when the masker and signal frequencies are close together. For the MBS masker, the frequencies chosen in the first burst in a sequence were repeated in the seven subsequent masker bursts for that sequence. A different set of random frequencies was chosen for each sequence. The other multitone masker is referred to as "multiple-bursts different" (MBD). For the MBD masker, the frequencies for every burst in a sequence were drawn randomly in exactly the same way as for the first burst of MBS. Thus, each sequence of MBD contained eight bursts that were different random frequency draws. The stimulus configuration was similar to that used in earlier studies (Kidd *et al.*, 1994; 2002b) and has been shown to provide a robust informational masking effect in the MBS condition, as well as a large masking difference between the MBS and MBD conditions. Figure 1 illustrates these two multitone maskers schematically in sound spectrogram form.

The figure shows typical draws of both MBS and MBD maskers with the right column indicating a masker plus signal and the left column illustrating the masker alone. Also shown is the protected region around the signal frequency where masker tones were not permitted to fall. The levels of the tones within a masker draw were equal.

A Gaussian noise was used to estimate peripheral frequency selectivity. The noise had a flat spectrum within the bandpass region and was filtered in many conditions to introduce a spectral notch. These conditions were used to estimate the characteristics of the auditory filter containing the signal and the processing efficiency of the listener for the tone-in-noise detection task (Patterson, 1976; Glasberg and Moore, 2000). The noise was presented as a single continuous burst having a duration of 300 ms. The 200-ms, 1000-Hz signal was temporally centered in the noise. The five notch widths were given by: 0.0 and 0.0 (no notch); 0.2 and 0.2; 0.4 and 0.4; 0.2 and 0.4; and 0.4 and 0.2, expressed as the difference between signal frequency and notch edge frequency, divided by signal frequency. This abbreviated version of the notched-noise test has been found in the past to provide reliable results when deriving auditory filter shapes (Stone *et al.*, 1992). The noise had a bandwidth of 400 Hz on either side of the notch.

### B. Subjects

A total of 24 adult subjects with normal hearing (thresholds of 15 dB HL or less at octave frequencies between 250 and 8000 Hz) served as listeners in these experiments. Their ages ranged from 19 to 47 years (mean age 24.7; median 22.5). Twelve subjects were trained musicians and 12 subjects identified themselves as nonmusicians. In both groups males and females were equally represented. The mean age of the musicians was 28.3 years (s.d. 8.3) and mean age of the nonmusicians was 21.1 years (s.d. 2.2). The mean absolute threshold for the multiple-burst signal was 3.2 dB SPL (s.d. 5.2 dB) and 3.3 dB SPL (s.d. 5.7 dB) for the 200-ms signal used in the noise-masking conditions. There were no significant differences in mean absolute threshold between the musician and nonmusician groups for either signal $[|t(22)|<1.6;\ p>0.1]$.

The selection criteria for inclusion in the musician group were as follows: first, all subjects had musical training at the college level in addition to 2 or more years of private lessons (virtually all reported beginning their musical training formally or informally as children). Nine were currently students in college-level music programs and two were graduates of music programs. Two reported holding graduate degrees in music. One other subject studied at the college level for 3 years and then became (and remains) a professional musician engaged in recording and performing. All currently play musical instruments regularly and had at least 2 semesters of formal ear training or, in one case, the subject tested out of the college ear-training requirement. In addition, all subjects were able to achieve 90% or higher accuracy on a relative pitch test. In the relative pitch test, the listener was given a pure tone of a frequency corresponding to the musical note "A" (440 Hz). Following that tone, another tone was presented having a frequency equaling that of one of the 12 notes on the semitone scale beginning with and extending above middle "C" (261.63 to 493.88 Hz) and the listener was asked to name the musical interval. The listener made 20 such judgments. For the nonmusicians, nine of the subjects reported that they had never had any musical training whatsoever. The other three subjects reported minimal experience attempting to learn to play musical instruments as children—including some private lessons or lessons in elementary school—but did not continue to play the instrument or take further lessons past the age of 10 years. None of the nonmusicians currently played any musical instrument and only one (not one who had taken lessons as a child) indicated some knowledge of how to read music (the subject gained some experience following written music as a dancer).

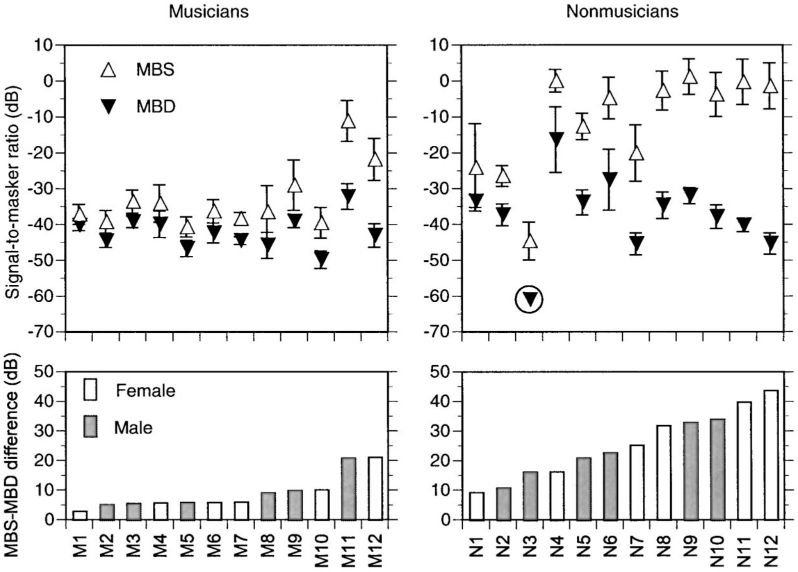

FIG. 2. Individual results from the random-frequency multitone masker conditions. The left panels show data from the musician group and the right panels show data from the nonmusician group. The upper panels show the signal-to-masker ratios at threshold in the multiple-burst same (MBS; upward-pointing open triangles) and the multiple-burst different (MBD; downward-pointing filled triangles) conditions. The symbol in a circle represents a point that could not be measured. Error bars denote ±1 s.d. of the mean. The lower panels show the difference in masking between the two conditions. The subjects are arbitrarily ordered according to the magnitude of the difference in masked threshold between MBS and MBD masking conditions. Shaded and unshaded bars in the lower panels denote male and female listeners, respectively.

### C. Procedures

Thresholds were measured using an adaptive 2-interval, 2-alternative forced-choice procedure that estimates the 70.7%-correct point on the psychometric function (Levitt, 1971). The level of the adaptively changing stimulus was varied initially in 6-dB steps, which was reduced to 3-dB steps after the first four reversals. There were 50 trials in each block of trials, and a threshold estimate was counted as valid only if at least nine reversals were obtained. The average of the reversals after discarding the first three or four reversals (whichever resulted in an even number) was then computed. Response feedback was provided after every trial. Initially, in each session, threshold was measured twice for the signal presented in quiet (no masker). If the two estimates differed by more than 3 dB a third estimate was obtained. After that the signal level was fixed at a level 20 dB above the average of the quiet threshold estimates (20 dB SL) for subsequent masking conditions.

In the masked conditions, the masker was varied using the same procedures as for the signal in quiet, except that the signal level was fixed and masker level was varied adaptively. The listeners were tested first on the Gaussian notched-noise masking task (except for one listener who was tested on the informational masking task first due to experimenter error) with a minimum of four estimates obtained for each notch width. After collection of those data was complete, the listeners were tested on the multitone masking conditions with the two types of maskers alternated in sets of two with at least eight estimates obtained for each type of multitone masker.

The stimuli were presented to one ear using a calibrated TDH-50 earphone in a double-walled IAC booth. The listeners were tested individually for two sessions lasting approximately 2 h each, except for one listener who required a brief third session to finish data collection. During the first session, the listeners' hearing was tested using standard pure-tone air-conduction audiometry and a questionnaire was administered on musical training and background. For the listeners in the musician group, the relative pitch test was also conducted at some point during the testing sessions.

## III. RESULTS

### A. Random-frequency multitone masking

For each listener in each condition, the signal-to-masker ratio at threshold was calculated by subtracting the overall masker level at threshold from the signal level (both in dB SPL), which was fixed at 20 dB SL individually for each listener, with a mean presentation level of about 23 dB SPL.[1] Because signal thresholds in quiet varied somewhat across listeners, expressing thresholds in terms of signal-to-masker ratio provided a way of comparing performance across listeners more directly. The data were first analyzed to test for learning effects across the eight runs completed by each listener in every condition. This was done by performing a within-subjects linear regression of threshold as a function of repetition number separately for the musicians and nonmusicians in the MBD and MBS conditions. For the musicians in the MBD condition, there was a significant trend for thresholds to improve with repetition number $[F(1,11)=63.6$, $p<0.001]$. However, the overall improvement was rather small, with a mean improvement of 0.5 dB per repetition. None of the other conditions showed a significant trend for threshold changes across the eight repetitions ($p>0.05$). Given that only eight runs were made, our data do not address the question of longer-term learning. However, they are generally consistent with earlier findings of rather weak learning effects in random-frequency multitone masking experiments (Neff and Callaghan, 1988; Neff and Dethlefs, 1995).

The individual mean signal-to-masker ratios at threshold are plotted in the top two panels of Fig. 2. The left panel shows data from the musicians, while the right panel shows data from the nonmusicians. The sex of individual subjects is indicated by the shading of the bars in the lower panels. Open upward-pointing triangles denote thresholds from the MBS condition; filled downward-pointing triangles denote thresholds from the MBD condition. The error bars represent

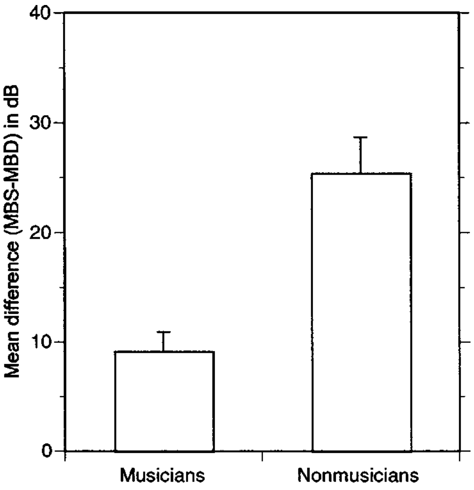

FIG. 3. Mean masking differences between the two types of random-frequency multitone maskers, MBS and MBD, for the two groups of listeners. Error bars denote +1 standard error of the mean.

±1 s.d. of the mean. Consider first the MBD condition (filled symbols). The results are reasonably similar both within and across the two groups. This impression was confirmed by a *t*-test, showing no significant difference in the mean signal-to-masker ratio for the MBD condition between the two groups [$t(22)=1.54$; $p>0.1$]. The generally good and relatively uniform performance in the MBD condition is consistent with the view that thresholds in this condition depend largely on energetic masking (Kidd *et al.*, 1994): the constant frequency of the signal repetitions assists the listener in perceptually segregating the signal from the maskers, thereby avoiding (or at least reducing) informational masking.

Consider next the MBS condition (open symbols). Thresholds are considerably higher in this condition (indicating poorer performance), as expected based on previous results (e.g., Kidd *et al.*, 1994). While there are some subjects in the nonmusician group whose performance is comparable to those in the musician group, the overall performance of the nonmusician group is considerably poorer. This impression was confirmed by a *t*-test showing a highly significant difference between the two group means for the MBS signal-to-masker ratios [$t(22)=4.44$; $p<0.001$].

If the interpretation that the MBD condition reflects primarily energetic masking is correct, then the *difference* between thresholds in the two conditions should provide a good measure of the amount of additional informational masking produced by the MBS condition. These results are shown in the lower panels of Fig. 2. Overall, the smaller differences observed in the musician group are consistent with the hypothesis, outlined in the Introduction, that musicians should exhibit less informational masking than nonmusicians. The mean differences for the two groups are shown in Fig. 3.

As mentioned in the Introduction, Neff *et al.* (1996) reported a sex effect in informational masking, with female listeners as a group exhibiting significantly higher thresholds (poorer performance) than male listeners. However, Neff *et al.* (1996) did not report on the musical abilities of their listeners. Our two groups (musicians and nonmusicians) were selected so that both sexes were equally and evenly represented. This enabled us to test both for effects of musical training and for sex. An examination of the distribution of males and females in the lower half of Fig. 2 suggests no clear effect of sex for either musicians or nonmusicians. A two-factors (musical training and sex) between-subject analysis of variance (ANOVA) was performed, with the MBS–MBD difference as the dependent variable. The results were clear: there was a significant effect of musical training [$F(1,20)=18.56$; $p<0.001$], confirming our earlier analysis, while neither the effect of sex [$F(1,20)<1$] nor the interaction between musical training and sex [$F(1,20)<1$] was significant. Sex remained a nonsignificant effect when the analysis was carried out using only the MBS or MBD signal-to-masker ratios.

Our results are not consistent with the findings of Neff *et al.* (1996), in that we find no effect of sex on performance. It is not clear what accounts for this difference. While we used a smaller number of subjects than did Neff *et al.* (24 vs 49), it is unlikely that simply adding more subjects would produce a significant result in our case, given that the effect of sex did not even approach significance. In light of the present study, one might speculate that there were more musicians among Neff *et al.*'s male listeners than among their female listeners. In any case, no hint of an effect of sex was found here, when listeners were balanced for musical training, suggesting that men are not necessarily better listeners than women.

As mentioned in the Methods section, the mean age of the musician group was somewhat higher than that of the nonmusician group. To test for any effect of age, a between-subjects analysis of covariance (ANCOVA) was performed with the MBS–MBD difference as the dependent variable, musicianship as a factor, and age as a covariate. As expected from the previous analyses, the effect of musicianship was highly significant [$F(1,21)=12.99$, $p<0.002$], but the effect of age was not ($F<1$). Thus, the difference in mean age between the two groups cannot account for the differences in informational masking.

Finally, some recent work has examined the effects of hearing impairment on informational masking (Micheyl *et al.*, 2000; Kidd *et al.*, 2002a). While all our listeners qualified as normal-hearing according to ANSI (1969) standards, there were substantial variations in absolute threshold for the signal, ranging from −9 dB to 15 dB SPL. However, absolute thresholds seemed to play no role in the amount of informational masking observed; another ANCOVA with absolute threshold as the covariate found no significant effect of absolute threshold ($F<1$). Separate analyses of musician and nonmusician data confirmed the lack of correlation between the MBS–MBD difference and absolute threshold (Pearson product-moment correlations of 0.03 and −0.23, respectively; $p>0.2$ for both).

### B. Notched-noise masking: Auditory filter shapes

The individual masker thresholds from the notched-noise conditions, using the fixed 1000-Hz 200-ms signal with the level set 20 dB above individual thresholds in quiet (mean presentation level of about 23 dB SPL), were ana-

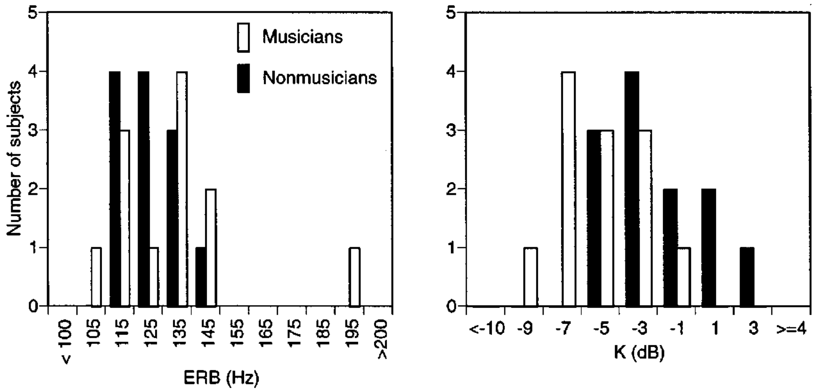

FIG. 4. Results from fitting auditory filters to the individual data from the notched-noise conditions. The data are plotted in the form of histograms, with the values on the abscissa representing the midpoint of each bin. The left panel shows estimates of the equivalent rectangular bandwidth (ERB) of the individual auditory filters, and the right panel shows estimates of detection efficiency, *K*.

lyzed and used to derive auditory filter shapes and detection efficiency values using the roex(pr) model of auditory filter shape (Patterson and Moore, 1986). The filter weighting function ($W$) is assumed to be of the form

$$W(g)=(1+p|g|)e^{-p|g|}+r, \tag{1}$$

where $g$ is the normalized frequency distance from the filter's center frequency (distance divided by center frequency), $p$ is a parameter defining the sharpness of filter tuning, and $r$ is a constant, designed to limit the dynamic range of the filter. The value of $p$ is allowed to be different below ($p_l$) and above ($p_u$) the filter's center frequency, thereby allowing asymmetry in the filter shape. The best-fitting values of $p$ were determined using a least-squares minimization routine. The detection efficiency, $K$, in dB is the mean signal-to-noise ratio (in dB) across all conditions for the best-fitting filter shape. For further details see, e.g., Glasberg and Moore (1990).

The resulting filter shapes were reasonably uniform across listeners, both in terms of their asymmetry and their bandwidths. Because of this, only the equivalent rectangular bandwidths (ERBs) of the filters are discussed. The values of the ERBs and $K$s derived from the individual data from notched-noise experiment are shown in Fig. 4 in the form of histograms. Bin widths of 10 Hz and 2 dB were used for the ERB and $K$ values, respectively. Higher ERB values imply poorer frequency selectivity; higher $K$ values imply poorer detection efficiency. Data from musicians are shown as open bars; data from the nonmusicians are shown as filled bars. Consider first the ERB values. The distributions of the two groups are rather similar, as are the mean ERB values of 134 Hz for the musicians and 126 Hz for the nonmusicians. These values are also in very good agreement with the value of 132 Hz at a center frequency of 1 kHz, given by the formula of Glasberg and Moore (1990), which was based on data from a number of earlier studies. A two-way ANOVA showed that neither the effect of sex nor musical training (or their interaction) was significant ($p>0.1$ in all cases). This confirms Fine and Moore's (1993) finding that musical training has no effect on peripheral frequency selectivity.

Consider next the values of detection efficiency, $K$ (Fig. 4, right panel). There seems to be a slight trend for $K$ values to be lower in musicians, indicating somewhat greater detection efficiency. This is reflected in the mean $K$ values of $-3.43$ and $-1.97$ dB for the musicians and nonmusicians, respectively, and is also consistent with the findings of Fine and Moore (1993). However, in contrast to that earlier study, the trend observed in our data was not statistically significant. In a two-way between-subjects ANOVA, neither the effect of sex nor musical training (or their interaction) reached significance ($p>0.1$ in all cases). When sex was ignored (as in the Fine and Moore study), a simple pooled $t$-test comparing the two groups of musicians and nonmusicians still revealed no significant difference [$t(22)=1.49$; $p>0.1$].

The distribution of $K$ values appears somewhat skewed towards positive values. For this reason, the analyses were repeated using the nonparametric Mann–Whitney U-test and the Kolmogorov–Smirnov Z-test (which do not assume normal distribution). Both tests also failed to show significant differences in either ERB or $K$ between musicians and nonmusicians. A comparison of our listener groups with those of Fine and Moore (1993) shows that the two musician groups performed similarly, whereas our nonmusician group showed somewhat lower (more efficient) $K$ values than did those of Fine and Moore. In summary, filter bandwidth and detection efficiency were not significantly related to musical training, at least in our group of 24 listeners.

## IV. DISCUSSION

This study investigated the effects of musical training on informational and energetic masking using random-frequency multitone and noise maskers. In the random-frequency multitone masking condition thought to involve informational masking (MBS condition), a large and statistically significant difference in performance was found between a group of musically trained listeners (musicians) and a group of listeners with no musical training (nonmusicians). In the random-frequency multitone condition, thought to rely more on energetic masking (MBD condition), no significant difference between musicians and nonmusicians was found. The reduced susceptibility to informational masking found in musicians may reflect superior analytic listening abilities in trained musicians. Here, "analytic listening" refers specifically to the ability to discern or "hear out" a predefined partial of a complex sound. In some ways this can be considered analogous to situations in which musicians must follow individual musical "voices" in unfamiliar pieces, although of course the stimuli used here do not resemble any traditional musical forms, nor do the tones (pure sinusoids with no amplitude fluctuations) resemble any physical instruments. In contrast to a previous study (Neff *et al.*, 1996), no effect of sex was found in any condition.

Peripheral frequency selectivity and detection efficiency were estimated using a version of the notched-noise technique. The data, as summarized using the ERB and detection ratio $K$, revealed no significant differences in performance between musicians and nonmusicians. This finding is consistent in part with the findings of an earlier study (Fine and Moore, 1993), although in that study detection efficiency was found to be significantly better in musicians than in nonmusicians.

The informational masking paradigm used in this study seems to tap into a genuine difference in listening abilities between people with and without a high degree of musical training. Considering the difference in thresholds between the MBS and MBD conditions (lower panels of Fig. 2 and Fig. 3), it can be seen that there is surprisingly little overlap in performance between the two groups. For instance, an attempt to discriminate musicians from nonmusicians solely by their MBS–MBD threshold difference by setting an arbitrary criterion of 10.5 dB leads to only two misidentified musicians and one misidentified nonmusician out of a total of 24. The finding of a robust difference between the groups is interesting because the task is not one in which musicians are specifically trained: the stimuli, comprising repeated bursts of random-frequency sinusoids gated in precise synchrony, are extremely unnatural. Furthermore, the task is not one that requires any explicit musical training, such as the ability to name notes or musical intervals.

In summary, we have shown a link between musical training and performance in an informational-masking task. Our study is by its nature correlational. Because of this, there is no way of telling whether it is musical training, a natural inclination towards pursuing music, or some other common factor that results in better performance in the musician group. However, the informational-masking task seems to provide a quantitative measure of analytic listening ability, without requiring knowledge specific to Western music, such as note names or music notation, or any high cognitive load, such as remembering a melodic line. The task may therefore prove to be a useful tool in future studies of analytic listening and musical ability in general.

## ACKNOWLEDGMENTS

This work was supported by NIH Grants R01 DC 04545 (G.K.) and R01 DC 03909 (A.J.O.). The first version of this paper was written while the first author was a fellow at the Hanse Institute for Advanced Study in Delmenhorst, Germany. We thank Tanya Arbogast, Christine Carter, and Sally Tressler for assistance. Christophe Micheyl, Bertram Scharf, Charles Watson, and Marjorie Leek provided helpful comments on earlier versions of the paper.

---

[1] For one nonmusician it was not possible to measure a threshold in the MBD condition, as he could still detect the signal when the masker was at the highest permitted level of 90 dB SPL. This is shown in the upper-right panel of Fig. 2 by the symbol in a circle. For the purposes of analysis, the listener's threshold was set to the maximum allowable level, as shown by the position of the symbol.

ANSI (**1969**). ''Specifications for audiometers,'' American National Standards Institute, ANSI S3.6-1969.

Fine, P. A., and Moore, B. C. J. (**1993**). ''Frequency analysis and musical ability,'' Music Percept. **11**, 39–53.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hear. Res. **47**, 103–138.

Glasberg, B. R., and Moore, B. C. J. (**2000**). ''Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise,'' J. Acoust. Soc. Am. **108**, 2318–2328.

Greenberg, G. S., and Larkin, W. D. (**1968**). ''Frequency-response characteristics of auditory observers detecting signals of a single frequency in noise: The probe-signal method,'' J. Acoust. Soc. Am. **44**, 1513–1523.

Hafter, E. R., Schlauch, R. S., and Tang, J. (**1993**). ''Attending to auditory filters that were not stimulated directly,'' J. Acoust. Soc. Am. **94**, 743–747.

Kidd, Jr., G., Arbogast, T. L., Mason, C. R., and Walsh, M. (**2002a**). ''Informational masking in listeners with sensorineural hearing loss,'' J. Assoc. Res. Otolaryngol. **3**, 107–119.

Kidd, Jr., G., Mason, C. R., and Arbogast, T. L. (**2002b**). ''Similarity, uncertainty, and masking in the identification of nonspeech auditory pattern,'' J. Acoust. Soc. Am. **111**, 1367–1376.

Kidd, G., Mason, C. R., Deliwala, P. S., Woods, W. S., and Colburn, H. S. (**1994**). ''Reducing informational masking by sound segregation,'' J. Acoust. Soc. Am. **95**, 3475–3480.

Leek, M. R., Brown, M. E., and Dorman, M. F. (**1991**). ''Informational masking and auditory attention,'' Percept. Psychophys. **50**, 205–214.

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Lutfi, R. A., Kistler, D. J., Oh, E. L., Wightman, F. L., and Callahan, M. R. (**2003**). ''One factor underlies individual differences in auditory informational masking within and across age groups,'' Percept. Psychophys. **65**, 396–406.

Micheyl, C., Arthaud, P., Reinhart, C., and Collet, L. (**2000**). ''Informational masking in normal-hearing and hearing-impaired listeners,'' Acta Oto-Laryngol. **120**, 242–246.

Neff, D. L., and Callaghan, B. P. (**1988**). ''Effective properties of multicomponent simultaneous maskers under conditions of uncertainty,'' J. Acoust. Soc. Am. **83**, 1833–1838.

Neff, D. L., and Dethlefs, T. M. (**1995**). ''Individual differences in simultaneous masking with random-frequency, multicomponent maskers,'' J. Acoust. Soc. Am. **98**, 125–134.

Neff, D. L., Dethlefs, T. M., and Jesteadt, W. (**1993**). ''Informational masking for multicomponent maskers with spectral gaps,'' J. Acoust. Soc. Am. **94**, 3112–3126.

Neff, D. L., Kessler, C. J., and Dethlefs, T. M. (**1996**). ''Sex differences in simultaneous masking with random-frequency maskers,'' J. Acoust. Soc. Am. **100**, 2547–2550.

Patterson, R. D. (**1976**). ''Auditory filter shapes derived with noise stimuli,'' J. Acoust. Soc. Am. **59**, 640–654.

Patterson, R. D., and Moore, B. C. J., (**1986**). ''Auditory filters and excitation patterns as representations of frequency resolution,'' in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London).

Pollack, I. (**1975**). ''Auditory informational masking,'' J. Acoust. Soc. Am. Suppl. 1 **57**, S5.

Scharf, B., Quigley, S., Aoki, C., Peachy, N., and Reeves, A. (**1987**). ''Focused auditory attention and frequency selectivity,'' Percept. Psychophys. **42**, 215–223.

Soderquist, D. R. (**1970**). ''Frequency analysis and the critical band,'' Psychonomic Sci. **21**, 117–119.

Spiegel, M. F., and Watson, C. S. (**1984**). ''Performance on frequency-discrimination tasks by musicians and nonmusicians,'' J. Acoust. Soc. Am. **76**, 1690–1695.

Stone, M. A., Glasberg, B. R., and Moore, B. C. J. (**1992**). ''Simplified measurement of impaired auditory filter shapes using the notched-noise method,'' Br. J. Audiol. **26**, 329–334.

Watson, C. S., (**1987**). ''Uncertainty, informational masking, and the capacity of immediate auditory memory,'' in *Auditory processing of complex sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, N.J.).

Wright, B. A., and Dai, H. (**1994**). ''Detection of unexpected tones with short and long durations,'' J. Acoust. Soc. Am. **95**, 931–938.