Comparing behavioral and physiological measures of combination tones: Sex and race differences

Dennis McFadden^{a)} and Edward G. Pasanen

Department of Psychology and Center for Perceptual Systems, University of Texas, Austin, 108 East Dean Keeton, A8000, Austin, Texas 78712-1043

Erin M. Leshikar and Michelle D. Hsieh

Department of Communication Sciences and Disorders, University of Texas, Austin, 2504-A Whitis Avenue, A1100, Austin, Texas 78712-0144

Mindy M. Maloney

Department of Psychology and Center for Perceptual Systems, University of Texas, Austin, 108 East Dean Keeton, A8000, Austin, Texas 78712-1043

(Received 8 July 2011; revised 6 June 2012; accepted 11 June 2012)

Both distortion-product otoacoustic emissions (DPOAEs) and performance in an auditory-masking task involving combination tones were measured in the same frequency region in the same ears. In the behavioral task, a signal of 3.6 kHz (duration 300 ms, rise/fall time 20 ms) was masked by a 3.0-kHz tone (62 dB SPL, continuously presented). These two frequencies can produce a combination tone at 2.4 kHz. When a narrowband noise (2.0–2.8 kHz, 17 dB spectrum level) was added as a second masker, detection of the 3.6-kHz signal worsened by 6–9 dB (the Greenwood effect), revealing that listeners *had* been using the combination tone at 2.4 kHz as a cue for detection at 3.6 kHz. Several outcomes differed markedly by sex and racial background. The Greenwood effect was substantially larger in females than in males, but only for the White group. When the magnitude of the Greenwood effect was compared with the magnitude of the DPOAE measured in the 2.4 kHz region, the correlations typically were modest, but were high for Non-White males. For many subjects, then, most of the DPOAE measured in the ear canal apparently is not related to the combination-tone cue that is masked by the narrowband noise.

© 2012 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4731224]

PACS number(s): 43.66.Dc, 43.66.Ki, 43.64.Jb, 43.64.Kc [EB]

Pages: 968–983

I. INTRODUCTION

Otoacoustic emissions (OAEs) now have been studied extensively for several decades, yet little is known about how these fascinating cochlear phenomena are related to behavioral measures of hearing. Here we report that the strength of distortion-product otoacoustic emissions (DPOAEs) is only moderately correlated with psychophysical performance on a behavioral task that is known to provide a distortion tone as a possible cue for detection. That is, there appears to be only a modest relationship between cochlear physiology and performance for the behavioral task chosen. This modest relationship was somewhat counterintuitive, as discussed in Sec. IV B.

DPOAEs are produced in normal mammalian cochleas whenever two tones (f_{lower} and f_{higher}) of the appropriate frequencies and sound-pressure levels (SPLs) are presented simultaneously (Kemp, 1979; Kim, 1980; Wit *et al.*, 1981). The underlying nonlinear mechanism is thought to be different from the mechanisms underlying other forms of OAE (Shera and Guinan, 1999). The DPOAE that has received the greatest attention to date is the so-called cubic difference tone (CDT) which lies at the frequency $2f_{\text{lower}} - f_{\text{higher}}$

(commonly called $2f_1 - f_2$). Because this combination tone is present even when the primary tones are not intense, the underlying mechanism is not simply an overdriving nonlinearity. In mammals, DPOAEs typically are strongest when the frequency ratio between the two primary tones ($f_{\text{higher}}/f_{\text{lower}}$) is about 1.15–1.25 (e.g., Gaskill and Brown, 1990). In some species, the level of the DPOAE is only about 30 dB weaker than the primary tones used to produce it, but in humans, DPOAEs typically are much weaker, being as much as 60 dB down from the primaries (for a discussion, see McFadden *et al.*, 2009).

On the perceptual side, musicians have known for centuries that when two tones of appropriate frequencies are simultaneously present, listeners often can hear frequencies other than those two—most notably a simple difference tone and the CDT. These combination tones (CTs) have been studied experimentally from a number of perspectives (e.g., Greenwood, 1971; Smoorenburg, 1972a,b, 1974; Hall, 1972, 1975; Wilson, 1980; Shannon and Houtgast, 1980; Furst *et al.*, 1988; Zwicker and Harris, 1990; Purcell *et al.*, 2007). Of greatest interest here is $2f_{\text{lower}} - f_{\text{higher}}$ and its role in particular auditory-masking tasks. For the remainder of this paper, the generic terms CT and DPOAE will be used to denote $2f_{\text{lower}} - f_{\text{higher}}$ specifically.

When a single tonal signal is masked by a single tonal masker (e.g., Wegel and Lane, 1924), under certain

^{a)}Author to whom correspondence should be addressed. Electronic mail: mcfadden@psy.utexas.edu

circumstances the two tones will interact and produce CTs that can aid a listener attempting to detect the weak signal (e.g., [Licklider, 1951](#)). For example, when the signal lies above the masker by an appropriate frequency difference, the upward asymmetry of masking can render the signal less audible than the CDT, making the CDT a potentially valuable cue for detection. [Greenwood \(1971\)](#) elegantly demonstrated this fact by measuring detectability in two ways. The signal frequency was higher than the masker frequency by a ratio ($f_{\text{higher}}/f_{\text{lower}}$) of about 1.2. In one condition, a band of noise low-pass filtered just below the masker frequency was present as a second masker, and in another condition that noise was absent. For Greenwood's listeners, the signal was 15–20 dB less detectable when the noise was present than when it was not, and this was true even though control measurements revealed that the noise band itself exerted little or no direct masking on the signal. Here we will call this difference in detectability the Greenwood effect. Greenwood's explanation was that the listeners were (perhaps unintentionally) listening at the frequency of the CDT, not the signal; the low-pass noise prevented that off-frequency listening and forced them to listen at the signal frequency, where detection was considerably more difficult.

Here we use a variant of the [Greenwood \(1971\)](#) task, and we also measure DPOAEs in the same ears and in the same frequency region. The motivation was to explore the relationship between physiology and behavior; specifically, do the individual differences in DPOAE strength underlie the individual differences in signal detection in the Greenwood task? Simple intuition suggests that there ought to be some relationship, perhaps quite a strong one, because similar stimuli are used for both the physiological and behavioral tasks, meaning that similar patterns of stimulation must exist on the basilar membrane for both types of measurement. Also, the explanation for the existence of the DPOAE is generally similar to the explanation for the CT—a nonlinearity in cochlear mechanics. However, there is precedent for the physiological and behavioral measures of the distortion tone *not* being strongly related. [Furst et al. \(1988\)](#) and [Zwicker and Harris \(1990\)](#), among others, noted that a DPOAE still could be measured in the ear canal after the listener had eliminated the perceived CT by canceling it with a properly adjusted tone at the frequency of the CT. That is, not all of what can be recorded in the ear canal can be perceived. This dissociation presumably is attributable in part to the fact that, while the DPOAE recorded in the ear canal is a sum of sounds of the same frequency originating from multiple sites in the cochlea (e.g., [Kim, 1980](#); [Shera and Guinan, 1999](#); [Konrad-Martin et al., 2001](#); [Martin et al., 2009](#)), the listener has access to (“hears”) only the component originating from that frequency's characteristic place on the basilar membrane. Also, both auditory sensitivity and the strength of DPOAEs can show microstructure (marked and reliable variation across small changes in frequency) that differs across individual ears ([Gaskill and Brown, 1990](#); [Heitmann et al., 1998](#); [Talmadge et al., 1999](#)). That microstructure has the potential to affect any estimates of the relationship between behavior and physiology in the present context. Thus, before the measurements were made, we knew that the relationship

between the Greenwood effect and the DPOAEs measured here might not be simple.

The question about relationships between OAEs and perception can be studied using different research strategies. Investigators can examine group differences in OAEs (means) and compare them with group differences on some behavioral measure (other means). [McFadden and Mishra \(1993\)](#) used this strategy when they reported that, as a group, people having several spontaneous otoacoustic emissions (SOAEs) had better hearing sensitivity than people with no SOAEs, and that this was true for both males and females, as groups, and in both ears, as groups. Alternatively, one can compare OAE and behavioral measures for *individual* subjects. When small numbers of subjects are tested, such comparisons are necessarily only qualitative (do the data look similar?—e.g., see [Long and Tubis, 1988](#)), but when larger samples are tested, correlations can be used to evaluate how the individual differences in OAEs compare with the individual differences in behavioral performance (e.g., [Micheyl and Collett, 1996](#); [Dorn et al., 1999](#)). Here we used a combination of these two strategies: Group means are compared using effect sizes (see Sec. [IIC](#)), and covariation in individual differences is evaluated using correlations between DPOAE strength and behavioral performance.

II. METHODS

The data reported here were collected as part of a larger study on the relationships between OAEs and psychoacoustics. Subjects were hired in crews of four to eight listeners who worked about 2 h per day, 5 days a week, for several weeks while performance was measured for a number of common auditory tasks, including the Greenwood task. During that time, various OAE and other physiological measures also were collected, sometimes more than once. Informed consent was obtained from all subjects prior to the study, and all were paid for their participation. The research protocol was approved by the Institutional Review Board at The University of Texas.

Past work suggested that, for various tasks, the data collected from females during menses are *more* similar to the data for males, while data collected during the midluteal phase are *less* similar to those of males (reviewed in [McFadden, 1998, 2011](#)). Accordingly, the menstrual cycle was monitored in the female subjects, and the data obtained were analyzed separately for their menses and midluteal phases. Because of our interest in the effect of the menstrual cycle on auditory performance, females using oral contraception were excluded from the larger study, and thus from this one.

Following the initial analyses of the data, some other outcomes from the larger study ([McFadden et al., 2012](#)) led us to suspect that our Greenwood results also might be dissimilar across subjects from different racial backgrounds, and that did prove to be true. Accordingly, the data here are reported both pooled across race and within two racial groupings (described in the following).

A. Subjects

Reported here are data from 43 female and 27 male listeners. All were screened to have audiometrically normal

hearing sensitivity (≤ 15 dB Hearing Level) for the standard audiometric frequencies between 0.25 and 8.0 kHz, and normal middle-ear function as measured by a clinical audiometric screening device (Auto Tym 38, GSI, Inc.). Subjects were paid an hourly wage plus a bonus at the end of testing, which lasted about 6–8 weeks for males and 8–10 weeks for females. The average age for both males and females was approximately 21 years. All were students at The University of Texas.

After being hired, each subject completed an extensive questionnaire that included questions about past exposures to intense sound, drug use, sexual experience and fantasies, and physical characteristics. Because OAEs have shown differences for females of differing sexual orientations (McFadden and Pasanen, 1998, 1999), here we excluded the data of the few subjects of both sexes whose questionnaire responses revealed them to be nonheterosexual. The questionnaire included the two items on race/ethnicity required by the National Institutes of Health (NIH): 1. (Yes or No) Are you of Hispanic, Latino, or Spanish Origin (Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race)? 2. (Pick One) Are you American Indian or Alaskan native; Asian; Black or African American; Native Hawaiian or Other Pacific Islander; White; More than one race; Other or Unknown (please specify if known). Once we realized the possibility of different outcomes in subjects from different racial backgrounds, all subjects answering “White” on question 2 were categorized as such, and all subjects choosing *any* other category on question 2 were categorized as Non-White, no matter what their response to question 1 (the Ns did not permit any finer categorizations). Two subjects answered yes on question 1, and did not answer question 2; they were categorized as Non-White (on the reasonable assumption—for Texas—that they also would have answered Hispanic on question 2 had that response been an option). The tallies for the various response categories are shown in Table I. Although the Non-White and White males were similar in number, there were twice as many Non-White females as White females. Note that we did not analyze the seemingly

large group of Asians separately because we knew from observation that this was a racially heterogeneous group of east Asians (Chinese, Japanese, Vietnamese, Koreans) and south Asians (Indians, Pakistanis, Sri Lankans) which we had no way to partition using the NIH questions.

B. Experimental procedures

Subjects were recruited using notices on bulletin boards around campus, advertisements in the student newspaper and the student job website, and word of mouth. Prior to hiring, potential subjects were screened audiometrically and given a full explanation of the purpose and logistics of the study. For the behavioral tests, only the right ear was tested; for DPOAEs, both ears were tested for males, only right ears for females.

1. Behavioral measures

Same-sex crews of four to eight listeners were tested simultaneously on the various behavioral tasks. They worked during the same 2 h time period each day of the workweek. They were seated in separate test booths inside a large, double-walled, sound-proofed room (Acoustic Systems, model 284221). Each booth had a set of TDH-39 headphones and identical arrays of small lamps that were lighted sequentially to indicate the progress of each listening trial. For the Greenwood task, the procedure was two-interval, two-alternative forced choice. On every trial of a block, there were two observation intervals, the signal was presented in one of those, and each listener pressed one of two response keys to indicate the interval he/she believed to contain the signal. Feedback as to the correct interval was provided immediately after the response interval using lamps dedicated to that purpose. The trial sequence was as follows: Warning interval and lamp (350 ms), pause (500 ms), first observation interval and lamp (300 ms), pause (500 ms), second observation interval and lamp (300 ms), response interval (1000 ms), feedback interval and lamp (350 ms), pause (~ 10 ms). Trials were presented in blocks of 50 trials each.

The level of the signal was varied adaptively for each listener individually. After three consecutive correct decisions, the signal was decreased by 2 dB, and after each error it was increased by 2 dB. This three-down/one-up procedure estimates the signal level necessary for 79% correct decisions (Levitt, 1971). The first two reversals of each block were discarded and the final even number of reversals was averaged to determine the estimate of sensitivity for that block. Only those blocks of trials having at least 46 responses, at least four reversals, and a standard deviation for those reversals smaller than 3.5 dB were used to estimate a subject’s sensitivity in each of the psychophysical conditions. For the Greenwood task, the males averaged ten blocks of trials satisfying these criteria for each condition of listening, and the females averaged five satisfactory blocks per condition during menses and eight satisfactory blocks per condition during the midluteal phase.

Typically, two to four different auditory tasks were tested per daily session, and the Greenwood task was revisited about every 4 days for male crews and about every 2 days for female crews. Between 21 and 24 blocks typically

TABLE I. Numbers of subjects in each of the categories included in NIH question 2 on ethnicity and race (bottom), plus those answering only question 1 (top).

Category	Females	Males
Question 1		
Hispanic (yes)	2 ^a	0 ^a
Question 2		
Native Alaskan	0	0
Asian	18	6
Black	3	3
Native Hawaiian/Pacific Islander	0	0
Other/Unknown	2	2
More than one race	4	1
Total Non-White	29	12
White	14	15

^aIncludes only those subjects who answered yes to question 1 and did not answer question 2.

were obtained per daily session. There were pauses of about 15 s between blocks, pauses of about 45 s after every three or six blocks (to allow the experimenter to change to a new listening condition), and longer breaks after about twelve blocks. A set of three to six consecutive blocks was collected on one task and/or set of parameters before moving to the next set. For the first block of a set, the signal level on the first trial was the same for all subjects; it was selected to be easily detectable, but the listeners were instructed to inform the experimenter immediately (via intercom) if he/she was unable to detect it. That block then would be restarted with a higher signal level on the first trial. For subsequent blocks of each set, the initial signal level was set individually for each subject; namely, 10 dB above the mean of his/her reversals for the previous block.

For the Greenwood task, the signal was a tone of 3.6 kHz, 300 ms in duration, including 20-ms rise–fall times. The primary masker was a 3.0-kHz tone of 62 dB SPL, which was present continuously during each trial. Thus, the signal and masker were simultaneously present. In addition, on some blocks of trials there was a second continuous masker: A narrow band of noise having shoulder frequencies of 2.0 and 2.8 kHz and extremely steep skirts, and presented concurrently with the tonal masker. The spectrum level of this narrowband noise (NBN) was either 17 or 27 dB SPL (overall levels of about 46 and 56 dB SPL, respectively), on separate blocks of trials. Thus, there were three basic Greenwood conditions: Tonal masker with no NBN masker and tonal masker with one of two levels of the NBN masker. When the NBN was absent, the detectability of the signal was substantially better (weaker signal required) than when the NBN was present, presumably because the NBN deprived the listener of the ability to listen to the weak CT generated at 2.4 kHz by the nonlinear interaction of the 3.0-kHz masker and the 3.6-kHz signal. The difference in signal level needed with and without the NBN (the Greenwood effect) was taken as an indicator of the strength of the CT for each listener (see Fig. 1). The primary statistics of interest are the correlations between the Greenwood conditions and the magnitudes of the DPOAEs measured in the same frequency region, in the same listeners, and in the same ear.

In addition, absolute sensitivity in the frequency region of interest was measured using a pure-tone stimulus of 3.0 kHz, a frequency that is midway between the CT at 2.4 kHz and the tonal signal at 3.6 kHz. Its duration was 300 ms, including 20-ms rise–fall times.

An acoustic coupler was used to measure the level of the CDT emitted by the TDH-39 earphones when activated by tones of 3.0 and 3.6 kHz. When both tones were at 70 dB SPL each, the level at 2.4 kHz was at the noise floor of our system, which was about 55 dB down from 70 dB. When the two tones were increased to 90 dB each, the level measured at 2.4 kHz remained at the noise floor, which was about 75 dB down from 90 dB. Thus, in the Greenwood conditions, when our subjects were presented with a 62-dB tonal masker and a tonal signal 15–25 dB weaker, there was essentially no energy present at the CT frequency in the output of the headphones.

FIG. 1. Diagram of stimuli used for behavioral tests. (Top) The tonal masker (3.0 kHz, 62 dB SPL) was present continuously throughout each trial. The level of the tonal signal (3.6 kHz, 300 ms in duration) was varied adaptively trial-by-trial (3 down/1 up) to estimate the level necessary for 79% correct decisions for each listener individually. When this tonal signal was added to this tonal masker, the combination tone (CT) generated at 2.4 kHz could have helped with detection of the signal. (Bottom) On some blocks of trials a second masker (a band of noise from 2.0 to 2.8 kHz) also was continuously present. If a listener were using information at the 2.4-kHz region for detection, then the narrowband masker should have worsened detectability (and it almost always did). The difference between signal detectability with and without the narrowband masker was taken as the magnitude of the Greenwood effect.

2. OAE measures

Two Etymotic ER-2 (Etymotic Research, Elk Grove, IL) earphones were used to evoke the DPOAEs and an Etymotic ER-10B+ microphone was used to record them. The procedure was basically the same as has been reported elsewhere (McFadden *et al.*, 2006a; *McFadden et al.*, 2006b). Briefly, pairs of equal-level tones having a ratio $f_{\text{higher}}/f_{\text{lower}}$ of approximately 1.21 were presented simultaneously (one to each earphone) to the right ear for about 4 s while the level at the frequency $2f_{\text{lower}} - f_{\text{higher}}$ (the DPOAE) was recorded. The basic stimulus configuration is shown in Fig. 2. Our procedure is a variant of CUBDIS. The same set of primary-tone frequencies was used for all subjects irrespective of the existence of SOAEs.

DPOAE strength was measured using multiple pairs of primary tones (see the following), but here we report only the data collected with a single pair—about 3.0762 and 3.7598 kHz, which give rise to a DPOAE at about 2.3926 kHz. Testing always began with these primary tones set to 50 dB SPL and proceeded to successively weaker levels (in 7-dB steps) until either the primaries were 34 dB or

DPOAE PROCEDURE

FIG. 2. Summary of procedure for estimating the strength of the DPOAE. The two primary tones always were presented at the same level and fixed in the ratio $f_{\text{higher}}/f_{\text{lower}} \sim 1.21$. The levels first were fixed at 50 dB SPL, and the level of the DPOAE at the frequency $2f_{\text{lower}} - f_{\text{higher}}$ (illustrated here as $2f_L - f_H$) was measured. Then the levels of the primaries were decreased and increased in increments of 7 dB, the DPOAEs measured again, and a gain function for the DPOAE thereby determined. For each subject, a straight line was fitted over the linear range of the gain function, and that fit was used to calculate both the strength of the primary tones necessary to produce a DPOAE of -5 dB SPL and the strength of the DPOAE for primary tones of 50 dB SPL. The primary frequencies actually used were about 3.0762 and 3.7598 kHz, producing a DPOAE at about 2.3926 kHz. (A segment of the ordinate was omitted because the DPOAEs typically are about 60 dB weaker than the level of the primary tones.)

the DPOAE level became -15 dB SPL or weaker (the noise floor for these measurements was about -25 dB SPL; see the following). The primary tones then were set to 57 dB, and progressively higher levels were tested (again in 7-dB steps) until either the maximum level of the primary tones (71 dB) was reached or the DPOAE level obtained was within 3 dB of that obtained for the previous (weaker) pair of primary tones. The gain function obtained was fitted with a straight line over its linear range and the resulting fit was used to estimate two related characterizations of the magnitude of the DPOAE for each subject.

One estimate of DPOAE magnitude was the strength of the DPOAE when the primary tones were 50 dB SPL. The second estimate was the level of the primary tones necessary to produce a DPOAE of -5 dB SPL.¹ Not surprisingly, these two, reciprocal characterizations of the strength of the DPOAE proved to be highly negatively correlated (see Sec. III D), but both were used here because some investigators think in terms of the primaries needed to produce a fixed level of DPOAE and some investigators think in terms of how strong a DPOAE is produced by a given pair of primary tones.

As noted, DPOAE data actually were collected using multiple pairs of primary tones, not just the single pair noted earlier (about 3.0762 and 3.7598 kHz). The reason is that the strength of DPOAEs can show considerable fluctuation even over narrow frequency ranges (e.g., Gaskill and Brown, 1990; Heitmann *et al.*, 1998; Talmadge *et al.*, 1999), and,

until writing this report, we always viewed this microstructure to be irrelevant to the experimental questions of interest. Accordingly, we developed a standard procedure that estimates DPOAE strength averaged across the microstructure (see McFadden *et al.*, 2006a; McFadden *et al.*, 2006b). The above-described procedures and decision rules are used to obtain not one gain function but six, one for each of six different pairs of primary tones. Here, all pairs of primary tones had an $f_{\text{lower}}/f_{\text{higher}}$ ratio of about 1.21, and the pairs were selected so that the DPOAE covered the range of about 2.246–2.734 kHz in increments of about 90 Hz. Those six gain functions were collapsed into one composite gain function by taking the median of the six DPOAE magnitudes for each level of the primary tones. The composite gain function calculated for each subject was fitted with a straight line over its linear range and the resulting fit was used to estimate the two related measures of DPOAE strength defined earlier. Thus, our standard procedure smoothed whatever microstructure existed across the approximately 0.5-kHz band surrounding the 2.4 kHz location on the basilar membrane, the frequency of the CT in the Greenwood behavioral task. When analyzing the data, we noted that smoothing the DPOAE microstructure led to weaker correlations between the physiological and behavioral data than were obtained from the single pair of primary tones closest in frequency to the tones used in the Greenwood task (about 3.0762 and 3.7598 kHz); accordingly, only the latter data are reported here. Note that no special treatment was given to the DPOAE data of subjects having SOAEs in the vicinity of the DPOAE frequencies.²

Also measured was the click-evoked otoacoustic emission (CEOAE), which is an echo-like response to a brief sound. The value of this measure here is as an indication of the *general* strength of the reflection-based mechanism that underlies some forms of OAE (Shera and Guinan, 1999). Again, the data-collection procedure was essentially identical to that used previously (e.g., McFadden *et al.*, 2006a; McFadden *et al.*, 2006b). The click stimuli were presented in sets of ten at a rate of about 10/s using one of the two ER-2 earphones; the sound in the ear canal for the 50 ms following each click was recorded; if the response met a collection of pre-established criteria, it was summed with all such previous responses during that block of presentations; and the presentations continued until 250 acceptable responses had been collected.³

The first 6 ms of the averaged response was discarded (because it contains stimulus artifact mixed with the cochlear response); the following 20.48 ms of the averaged response was bandpass filtered (1.0–5.0 kHz using an eighth-order elliptical digital filter); the rms amplitude of the filter output was calculated and transformed into decibels SPL; and that value was used to characterize the strength of the CEOAE response for that subject. Our CEOAE measure is “linear”; no differencing is involved. Two click levels were tested in separate blocks. Those click stimuli had peak amplitudes equal to the maximum amplitude of a 1.0-kHz tone of 75 dB or 69 dB SPL (peak-equivalent). The duration of the electrical transient delivered to the earphone was 100 μ s, creating a rarefaction pulse in the ear canal.⁴

3. Menstrual effects

One of the goals of the larger research project was to obtain behavioral and physiological measures during different phases of the menstrual cycle. Accordingly, only normal-cycling women were included in the study, and they and their female experimenters kept track of their cycles using diaries and other techniques. DPOAEs and CEOAEs were measured during as many menses and midluteal phases as possible in the 8–10 week time period of employment for the female crews. Data collected during the 5-day interval beginning with noticeable bleeding was defined as “menses.” The definition of midluteal was the interval 3–12 days prior to the onset of menses. The timing of the midluteal measurements was, necessarily, only a *prediction* at the time the DPOAE and CEOAE data were collected, but those physiological measurements were included only if the predicted timing was confirmed by the onset of that subject’s next menses; hence the differences in the Ns shown for the behavioral measures in our data tables.

Behavioral measures were obtained every week day, and the testing protocol ensured that every condition of listening was re-tested every 2.5 days for the females. For the analyses reported here, the blocks of trials collected in the first 5 days of the menstrual cycle were pooled, as were the blocks collected in the midluteal phase (and across multiple menses or midluteal phases of the same subject when possible). Again, the onset of the subsequent menses defined and confirmed the midluteal phase for each individual subject, so the actual position in the cycle occupied by each block of trials was determined retrospectively. Because of weekends and occasional holidays, the number of blocks contributing to the menses and midluteal estimates necessarily varied across subjects. In some instances no includable (see previous discussion) blocks were obtained for one phase; hence the differences in the Ns shown for the behavioral measures in our data tables.

C. Statistical procedures

Effect sizes are used to characterize the magnitudes of the differences between groups on the same variable, and correlations are used to characterize relationships between dependent variables. Here effect size was calculated as the difference between the means in the two conditions of interest divided by the square root of the weighted mean of the variances for the two conditions. By convention, effect sizes of 0.2, 0.5, and 0.8 are taken as small, medium, and large differences, respectively (Cohen, 1992). Of greatest interest here were the correlations between the magnitude of the Greenwood effect and the two measures of the strength of the DPOAE described previously; of lesser interest were the correlations of those measures both with absolute sensitivity and with the CEOAE. The correlations calculated were Pearson-product-moment correlations.

With the small Ns available here, especially after partitioning by race, extreme values had the ability to affect the summary statistics and correlations disproportionately. Accordingly, individual-subject values that were more than 2.5 SD from the relevant mean were excluded prior to

calculating the final means, effect sizes, and correlations reported here. This exclusion rule eliminated at most only one or (rarely) two subjects per dependent variable (not always the same subject) for any of the combinations of subject group and racial grouping. Subjects excluded from a pooled group were added back into consideration after partitioning by race on the grounds that they may not be extreme in the more select group. In one instance, a male was 3.65 SD below the mean in his pooled group (and was excluded) but became only 2.43 SD below the mean in his Non-White group (from which he was not excluded). Conversely, another male went from being 2.4 SD below the mean in his pooled group (and not excluded) to 5.6 SD below the mean in his Non-White group (and excluded). Of course, a subject excluded from one of the basic Greenwood conditions necessarily also was excluded from the calculation of the corresponding Greenwood effect. This same exclusion process was used with both the behavioral and physiological measures.

When a large number of measurements are made on the same subjects, and multiple pairwise comparisons are made on those measurements, it is difficult to know how to protect against false-positive errors. Here we used a resampling technique as a way of assessing how likely our outcomes were due to chance. For example, correlations were calculated for all the relevant comparisons between and among the behavioral and physiological measures; we will call these the *actually obtained correlations*. Then, for each pair of measures being correlated, the true value for one of the measures was retained for all subjects, but all the values for the second measure were replaced by values achieved by different subjects in that same group; those values were (re) assigned at random and without replacement. That “resampled” correlation was calculated and stored, and this process of random assignment for the second variable was repeated 20 000 times. For these 20 000 calculations, a tally was kept of how often the *absolute value* of the resampled correlation was equal to or larger than the absolute value of the corresponding *actually obtained correlation* for that comparison. These tallies provide estimates of how rare the actually obtained correlations were, given the data obtained from this sample of subjects. The tallies divided by 20 000 provide *implied significance* values for the actually obtained correlations.

A similar procedure was used to estimate how (un)likely the various group differences could be attributable to chance. For example, to evaluate the sex differences in DPOAEs and CEOAEs: (1) The data for the two sexes were pooled for each variable of interest; (2) a random sample equal to the number of females was drawn; (3) the remainder of the resample was taken to be “males”; (4) the effect sizes for all of the comparisons of interest were calculated; (5) the process was repeated 20 000 times; and, again, (6) a tally was kept of how often the (absolute values of the) resampled effect sizes were equal to or greater than the absolute values of the effect sizes actually obtained for these subjects. Those tallies then were used to calculate the proportions of the resamples that yielded group differences more extreme than actually obtained—the implied significance values for the actual outcomes. Note that because the effect sizes (or the

correlations) were calculated for all dependent variables for each resample, whatever dependencies might exist across the multiple dependent variables are preserved in our resampling procedure. We have used versions of this resampling procedure for effect sizes in the past (e.g., [McFadden et al., 2010a](#)). Because our tallies were based on the *absolute values* of the resampled correlations or effect sizes, our estimates of implied significance values should be viewed as both conservative and “two-tailed.” [As would be expected, when only those resamples that were both larger than the actually obtained value *and* had the same sign were tallied, the implied significance values were almost exactly half of those (i.e., more rare) obtained using the absolute value.]

1. Terminology

Because comparisons of different sorts are of interest here, having clear terminology is essential. Accordingly, we will use “*subject groups*” to denote the menses female, midluteal female, and male groups, and we will use “*race groupings*” to denote the Pooled, Non-White, and White groups of subjects. We will continue to use “*Greenwood effect*” to denote the difference in detection performance between the condition with no NBN noise added to the tonal masker and one of the conditions having the NBN noise added. The latter conditions are called the “*basic (or individual) Greenwood conditions*.”

III. RESULTS

We wish to acknowledge at the outset that the comparisons here between behavior and physiology are imperfect because the stimuli used to collect the two sets of data were not identical. For behavior, the masker was 3.0 kHz, 62 dB SPL, and continuously present, and the signal was 3.6 kHz, about 35–45 dB SPL, and 300 ms in duration. By contrast, for the DPOAE measurements, the primary tones were 3.0762 and 3.7598 kHz, always equal in level (ranging from about 36 to 71 dB SPL), and 4 s in duration. Thus, the CT was at 2.4 kHz for the behavioral tests, and the DPOAE was at 2.3926 kHz, a 7-Hz difference. In addition, some of the important behavioral conditions included a narrowband noise centered at 2.4 kHz, and the DPOAE conditions did not (although see Sec. III F for an exception). Also, different earphones were used for the behavioral and physiological measurements. This lack of identity was attributable primarily to the constraints of the larger study from which these data were extracted. Whether the use of more similar stimuli and the same transducer in the two domains would result in outcomes different from what we are about to report must be resolved by future research. Because the ultimate comparison here is correlation, the question is whether the *ordering* of the subjects is changed for one of the measurements by any of the dissimilarities of stimuli or equipment.

A. Basic Greenwood conditions

The mean signal levels (and SEs) for the three basic Greenwood conditions [numbered (1)–(3)] and the three race groupings are shown at the left-hand side of Table II.

(Subjects were allowed to contribute to each of these nine means separately; usable data did not have to be available for all three conditions, or for both menstrual phases, in order to contribute to one.) For conditions (1)–(3), the same basic patterns of results are evident in the Pooled data (bold type) as in the data for the Non-White and White groups. Namely, males were slightly worse than females at detecting the tonal signal when only the tonal masker was present [condition (1)], but they were somewhat better than the females when the additional NBN masker was added [conditions (2) and (3)]. Second, note that increasing the NBN masker from 17 to 27 dB had little additional effect on signal detectability [condition (2) vs condition (3)]. [Although we did not measure this, the NBN maskers themselves were presumed to be too weak to produce any masking on the 3.6-kHz signal; see [Greenwood \(1971\)](#).] Third, note that, for all three race groupings, the differences in signal detectability for females in the two phases of the menstrual cycle generally were small for all three basic Greenwood conditions. On the right-hand side of Table II are the effect sizes for the sex difference for each of the three basic Greenwood conditions for each race grouping; with few exceptions, these sex differences were small to medium in size ([Cohen, 1992](#)) for conditions (1)–(3), and they generally did not achieve implied significance during resampling. That is, there was little by way of sex or race differences in mean behavioral performance on the three basic Greenwood conditions.

Not shown in Table II are the results for detecting a 3.0-kHz tone in quiet. Mean detectability for the pooled males (8.4 dB) was intermediate to that for the pooled menses females (7.5 dB) and the pooled midluteal females (8.7 dB). The corresponding effect sizes (females minus males) were -0.16 and $+0.04$ for the menses and midluteal comparisons, respectively. When the data were partitioned by race, however, markedly different patterns of results emerged. For the Non-Whites, all three subject groups (menses females, midluteal females, and males) had mean sensitivities within 0.25 dB of about 9.6 dB SPL. For the Whites, the menses females had the best sensitivity (5.6 dB SPL), the midluteal females were intermediate (6.7 dB SPL), and the males had the worst sensitivity (8.3 dB SPL). The White results are generally in accord with the literature on sex differences in hearing sensitivity, which also suggests that females are more sensitive than males (reviewed in [McFadden, 1998](#)). The effect sizes for the various pairwise comparisons within the White group ranged between 0.25 and 0.55, and none achieved implied significance during resampling.

B. The Greenwood effect

Of greatest interest here are the *differences* in detection performance with and without the NBN masker—the Greenwood effect—because they are indications of the strength of the CT. The Greenwood effect exhibited large, unexpected race differences. The magnitudes of the Greenwood effect for both levels of the NBN masker are shown on the left-hand side of Table II [conditions (4) and (5)]. [These means are not simple differences between the values in the rows above them because they were calculated on a subject-by-

TABLE II. Means and SEs (in parentheses) for performance in the three basic Greenwood conditions and for the two Greenwood effects, and the effect sizes for the various comparisons plus the resampling outcomes, all shown separately for the three subject groups and the three race groupings.

Condition ^b	Groups ^c	Mean ^a detection performance (dB SPL)			Effect size (sex diff.)	
		Females		Males	Menses minus males	Midluteal minus males
		Menses	Midluteal			
(1) No NBN masker	Pooled	35.5 (0.85)	34.8 (0.81)	36.0 (1.29)	−0.09	−0.21
	Non-White	35.2 (1.28)	35.0 (1.06)	34.6 (1.81)	0.09	0.06
	White	34.9 (1.30)	34.6 (1.24)	36.9 (1.92)	−0.31	−0.37
(2) 17-dB NBN masker	Pooled	44.5 (1.02)	44.5 (0.87)	42.5 (1.27)	0.31	0.33
	Non-White	42.9 (1.58)	44.4 (1.16)	40.2 (0.90)	0.43	0.83 ^e
	White	45.6 (1.45)	45.6 (0.94)	42.5 (1.83)	0.48	0.54
(3) 27-dB NBN masker	Pooled	45.0 (1.01)	44.8 (0.92)	42.4 (1.24)	0.41	0.39
	Non-White	44.1 (1.28)	44.6 (1.21)	41.1 (1.73)	0.46	0.57
	White	47.2 (1.46)	45.1 (1.43)	42.8 (1.77)	0.73 ^d	0.39
(4) Greenwood effect for 17-dB masker	Pooled	9.4 (1.02)	9.4 (0.83)	6.5 (0.93)	0.54^d	0.58^e
	Non-White	8.0 (1.25)	9.1 (1.12)	6.2 (1.55)	0.33	0.53
	White	11.4 (1.53)	10.8 (1.01)	5.6 (1.09)	1.21 ^f	1.31 ^f
(5) Greenwood effect for 27-dB masker	Pooled	10.4 (0.91)	9.9 (0.79)	6.4 (0.81)	0.80^f	0.76^f
	Non-White	9.4 (1.19)	9.6 (1.06)	6.5 (1.22)	0.56	0.62
	White	12.2 (1.25)	10.6 (1.14)	5.9 (1.11)	1.47 ^g	1.10 ^f

^aSmaller means in Conditions 1–3 correspond to better detection.

^bNo NBN masker = 3.0-kHz tonal masker and 3.6-kHz signal; NBN masker = condition (1) plus a narrowband noise centered over the combination tone at 2.4 kHz; Greenwood effect = difference between conditions (2) and (1) or between (3) and (1), but only for those subjects having usable data for both conditions contributing to the difference (see the text).

^cNs: 21–27 Non-White menses; 24–26 Non-White midluteal; 10–11 Non-White males; 12–13 White menses; 13–14 White midluteal; 15 White males.

^dImplied significance, from resampling: $0.10 > p > 0.05$.

^eImplied significance, from resampling: $0.05 > p > 0.01$.

^fImplied significance, from resampling: $0.01 > p > 0.001$.

^gImplied significance, from resampling: $0.001 > p > 0.0001$.

subject basis; only subjects having a score for both conditions being differenced were included in conditions (4) and (5).] Note that, for all the race groupings, the Greenwood effects for the females were larger than those for the males (about 9.5 and 6.5 dB for the pooled females and males, respectively), and that was true for both levels of the NBN masker. Because of these sex differences, the female and male data were not pooled for the correlations between behavior and physiology in what follows. Also note that, again, there were no clear patterns of difference in effect for the two phases of the menstrual cycle.

The effect sizes for the sex difference in the Greenwood effect are shown at the bottom right-hand side of Table II, calculated separately for the two NBN masker conditions and the two menstrual phases. When all subjects were pooled, the four effect sizes for sex difference were moderate to large (Cohen, 1992), and they did achieve implied significance during resampling. However, when the subjects were dichotomized by race, a marked dissimilarity became evident; the large, significant sex differences existed only for the White group. That is, for the White group, the two sexes were much more different in their Greenwood effects than was true for the Non-White group. (Note that the sex differences for the Greenwood effect were larger than the sex differences for the basic Greenwood conditions.)

Because the presence of the NBN masker was so detrimental to detection performance, there can be no doubt that

the detection of the 3.6-kHz signal, when masked only by the 3.0-kHz tonal masker, was largely determined by cochlear activity in the vicinity of the CT that arises at 2.4 kHz because of the nonlinear interaction of the tonal signal and masker. In Sec. III D we examine whether the individual differences in performance in the Greenwood task are correlated with the individual differences in DPOAE strength.

C. DPOAE strength

As noted, two reciprocal measures of DPOAE strength were extracted from the data collected: The level of the primary tones necessary to produce a DPOAE of -5 dB SPL, and the level of the DPOAE produced by primary tones of 50 dB. Those values, along with the mean strength of the CEOAE responses for two click levels, are shown in Table III for all three race groupings and the three subject groups. (Recall that the DPOAE data in Table III are those collected with the single pair of primary tones 3.0762 and 3.7598 kHz.) Past research reported small, or no, sex difference for DPOAEs (see McFadden *et al.*, 2009; but compare Valero *et al.*, 2008), and that was confirmed here. The sex differences for DPOAE conditions (1) and (2) in Table III were small to medium for all race groupings, and none achieved implied significance during resampling. In addition, although there was a hint of a menstrual effect on DPOAE strength for White females (midluteal weaker), none of the effect sizes

TABLE III. Means and SEs (in parentheses) for DPOAE and CEOAE measurements, and the effect sizes for the various comparisons plus the resampling outcomes, all shown separately for the three subject groups and the three race groupings.

Condition ^b	Groups ^c	Mean ^a strength (dB SPL)			Effect size (sex diff.)	
		Females		Males	Menses minus males	Midluteal minus males
		Menses	Midluteal			
(1) Primary level for DPOAE = −5 dB SPL	Pooled	53.0 (1.39)	51.9 (1.32)	54.1 (1.59)	−0.12	−0.27
	Non-White	54.1 (1.96)	52.8 (1.73)	50.7 (5.03)	0.27	0.18
	White	52.3 (1.83)	50.2 (1.98)	53.7 (1.77)	−0.21	−0.51
(2) DPOAE level for 50-dB primaries	Pooled	−7.7 (1.03)	−7.0 (1.04)	−9.9 (1.73)	0.29	0.38
	Non-White	−7.9 (1.45)	−8.3 (1.29)	−8.9 (3.44)	0.12	0.07
	White	−7.4 (1.40)	−4.7 (1.63)	−7.8 (1.55)	0.06	0.52
(3) CEOAE level for 75-dB click	Pooled	12.8 (0.53)	12.2 (0.56)	11.5 (0.69)	0.40	0.20
	Non-White	12.6 (0.64)	11.8 (0.68)	11.7 (1.40)	0.24	0.02
	White	13.4 (0.95)	12.9 (0.96)	11.2 (0.74)	0.70 ^d	0.56
(4) CEOAE level for 69-dB click	Pooled	9.9 (0.59)	9.8 (0.68)	8.1 (0.86)	0.44^d	0.40
	Non-White	9.6 (0.69)	9.5 (0.84)	8.8 (1.78)	0.18	0.15
	White	10.5 (1.13)	10.4 (1.18)	7.5 (0.88)	0.82 ^e	0.80 ^e

^aLarge means in Condition 1 and small means in Conditions 2–4 correspond to weaker OAEs.

^bDPOAE = distortion-product otoacoustic emission at 2.393 kHz (primary tones of about 3.076 and 3.760 kHz); CEOAE = click-evoked otoacoustic emission; dB SPL = decibels sound-pressure level.

^cNs: 26–29 Non-White menses; 24–26 Non-White midluteal; 11 Non-White males; 13–14 White menses; 12–13 White midluteal; 14–15 White males.

^dImplied significance, from resampling: $0.10 > p > 0.05$.

^eImplied significance, from resampling: $0.05 > p > 0.01$.

for sex difference in DPOAEs achieved implied significance for any race grouping (for menstrual effects on other types of OAE, see Haggerty *et al.*, 1993; Yellin and Stillman, 1999; McFadden, 1998). Here, DPOAE strength was about 60 dB weaker than the primary tones [condition (2), Table III], which is about typical for humans (see McFadden *et al.*, 2009); the DPOAEs of other mammalian species tend to be considerably stronger than those in humans.

For CEOAEs, the sex differences seen in Table III were generally in accord with the literature (McFadden, 2009, 2011), but with a twist. As expected, the CEOAEs of females were stronger than those of males [conditions (3) and (4)], and this was true for both phases of the menstrual cycle. However, the sex differences for the White group were substantially larger than those for the Non-White group. Three of the four effect sizes for Whites achieved implied significance during resampling, but none did for the Non-Whites. The effect sizes for Whites are comparable to those previously reported by this lab for samples that were not partitioned by race (e.g., McFadden and Pasanen, 1998; McFadden *et al.*, 2009).

D. Correlations between behavioral and physiological measures of the CT

One way to test the idea that people with strong DPOAEs perceive strong CTs is to correlate the measured magnitudes of the DPOAEs with the measured magnitudes of the Greenwood effect. This was done and the results are shown in conditions (1)–(4) in Table IV. Note first that, for the females, there is nothing in the data to suggest a substantial relationship between our behavioral and physiological measures. The only correlation that catches the eye is for

Non-White menses females in condition (2), and even that was only marginally significant under resampling. For the males, however, there is evidence of a relationship between DPOAE strength and the Greenwood effect (top right-hand side of Table IV). Furthermore, this relationship clearly is attributable to the Non-White males. For each of the four DPOAE conditions at the top of Table IV, the correlation for Non-White males was either significant or marginally significant under resampling. We have no explanation for why this relationship should exist only for the Non-White males.

As a check on whether the behavioral data were related to the strength of the cochlear-amplifier system *in general*, correlations were calculated for CEOAE magnitude against the Greenwood effect. The values for the 69-dB click are shown in conditions (5) and (6) of Table IV; the outcomes for the 75-dB click were highly similar, so they are not shown. For both sexes and for all race groupings, the correlations between CEOAE strength and the Greenwood effect all were small, and none achieved implied significance during resampling. Accordingly, it seems safe to conclude that individual differences in the general strength of the reflection mechanism underlying OAEs are not strongly related to the individual differences in the Greenwood effect.

Not shown in Table IV is that the two measures of DPOAE strength were highly negatively correlated with each other: about −0.92 and −0.89 for the pooled menses and pooled midluteal females, respectively, and about −0.93 for the pooled males. This directionality is expected and the magnitude is gratifying. Also, the magnitudes of the Greenwood effects measured with the 17- and 27-dB NBN maskers were highly positively correlated with each other—about 0.95 for both pooled females and pooled males.

TABLE IV. Correlations between the relevant behavioral and physiological measures, shown separately for the three subject groups and the three race groupings and with implied significance indicated.

Conditions compared ^a	Groups ^b	Females		Males
		Menses	Midluteal	
(1) Greenwood effect with 17-dB NBN masker vs primary level for DPOAE = -5 dB SPL	Pooled	-0.19	+0.02	-0.41^c
	Non-White	-0.36	-0.01	-0.67 ^c
	White	-0.02	+0.33	-0.16
(2) Greenwood effect with 17-dB NBN masker vs DPOAE strength with 50-dB primary tones	Pooled	+0.31	-0.12	+0.46^d
	Non-White	+0.48 ^a	-0.13	+0.69 ^c
	White	+0.17	-0.46	-0.00
(3) Greenwood effect with 27-dB NBN masker vs primary level for DPOAE = -5 dB SPL	Pooled	-0.12	-0.06	-0.28
	Non-White	-0.26	-0.14	-0.77 ^d
	White	-0.01	+0.13	-0.05
(4) Greenwood effect with 27-dB NBN masker vs DPOAE strength with 50-dB primary tones	Pooled	+0.14	-0.01	+0.37^c
	Non-White	+0.26	+0.02	+0.74 ^d
	White	+0.13	-0.11	-0.07
(5) Greenwood effect with 17-dB NBN masker vs CEOAE strength with 69-dB click	Pooled	+0.21	-0.17	+0.06
	Non-White	+0.06	-0.20	-0.02
	White	+0.38	-0.11	-0.11
(6) Greenwood effect with 27-dB NBN masker vs CEOAE strength with 69-dB click	Pooled	+0.03	-0.15	+0.11
	Non-White	-0.06	-0.16	+0.38
	White	+0.10	-0.18	-0.26

^aGreenwood effect = difference between No-NBN-Masker and With-NBN (of 17 or 27 dB spectrum level); NBN = narrowband noise (2.0–2.8 kHz wide); DPOAE = distortion-product otoacoustic emission at 2.393 kHz (primary tones of about 3.076 and 3.760 kHz); dB SPL = decibels sound-pressure level.

^bNs: 20–24 Non-White menses; 20–24 Non-White midluteal; 10–11 Non-White males; 11–12 White menses; 11–12 White midluteal; 14–15 White males.

^cImplied significance, from resampling: $0.10 > p > 0.05$.

^dImplied significance, from resampling: $0.05 > p > 0.01$.

The Greenwood effects are difference scores; might it be that one or more of the basic Greenwood conditions used to calculate those differences *was* strongly correlated with DPOAE strength for the females or White males, but that fact was masked by the differencing operation? The condition with no NBN masker present is the most interesting in this context because it is intuitive that the DPOAE ought to be best correlated with behavior when the CT is unmasked (the presence of the NBN ought to make listeners more similar). The correlations between DPOAE strength and the basic Greenwood conditions are shown in Table V for the three subject groups and the three race groups.

Except for the Non-White males, the correlations in Table V are generally larger than those for the Greenwood effects in Table IV. Also, the White females appear to have a moderate relationship between DPOAE strength and performance in the individual Greenwood conditions; see in particular the conditions with no NBN masker [conditions (1) and (2) in Table V]. Partly because of the White females, many of the correlations for the pooled female group also achieved implied significance during resampling. Note that

TABLE V. Correlations between the relevant behavioral and physiological measures, shown separately for the three subject groups and the three race groupings and with implied significance indicated.

Conditions compared ^a	Groups ^b	Females		Males
		Menses	Midluteal	
(1) Tonal masker plus no NBN masker vs primary level for DPOAE = -5 dB SPL	Pooled	-0.40^d	-0.39^d	-0.02
	Non-White	-0.15	-0.36	+0.37
	White	-0.48 ^c	-0.43	-0.20
(2) Tonal masker plus no NBN masker vs DPOAE strength with 50-dB primary tones	Pooled	+0.35^c	+0.42^d	+0.00
	Non-White	+0.11	+0.35	-0.40
	White	+0.50 ^c	+0.56 ^d	+0.27
(3) Tonal masker plus 17-dB NBN masker vs primary level for DPOAE = -5 dB SPL	Pooled	-0.52^c	-0.37^c	-0.30
	Non-White	-0.40	-0.38	-0.09
	White	-0.40	-0.29	-0.30
(4) Tonal masker plus 17-dB NBN masker vs DPOAE strength with 50-dB primary tones	Pooled	+0.57^c	+0.29	+0.33
	Non-White	+0.46 ^c	+0.22	-0.13
	White	+0.57 ^d	+0.35	+0.30
(5) Tonal masker plus 27-dB NBN masker vs primary level for DPOAE = -5 dB SPL	Pooled	-0.39^d	-0.40^d	-0.19
	Non-White	-0.34	-0.44 ^c	-0.16
	White	-0.47	-0.29	-0.25
(6) Tonal masker plus 27-dB NBN masker vs DPOAE strength with 50-dB primary tones	Pooled	+0.35^c	+0.37^c	+0.23
	Non-White	+0.29	+0.33	+0.11
	White	+0.60 ^d	+0.42	+0.26

^aDPOAE = distortion-product otoacoustic emission at 2.393 kHz (primary tones of about 3.076 and 3.760 kHz); NBN = narrowband noise (2.0–2.8 kHz wide); dB SPL = decibels sound-pressure level.

^bNs: 21–27 Non-White menses; 20–22 Non-White midluteal; 10–11 Non-White males; 12–13 White menses; 12–13 White midluteal; 14–15 White males.

^cImplied significance, from resampling: $0.10 > p > 0.05$.

^dImplied significance, from resampling: $0.05 > p > 0.01$.

^eImplied significance, from resampling: $0.01 > p > 0.001$.

the Non-White males, who had such high correlations in Table IV for the Greenwood effect, did not exhibit any significant correlations with the individual Greenwood conditions in Table V. Nor did the White males.

When CEOAE strength (our proxy for general strength of the cochlear-amplifier system) was compared with the individual Greenwood conditions shown in Table V, the correlations for the pooled subjects all were smaller than 0.24, half of them were smaller than 0.10, and none even approached implied significance during resampling. With minor exceptions, the same lack of significance and pattern was true of these correlations within race groups. Accordingly, none of those correlations are shown. Apparently the individual differences in the general strength of the reflection mechanism underlying OAEs (Shera and Guinan, 1999) also are not strongly related to the individual differences in performance in the basic Greenwood conditions.

Also not shown in Tables IV or V is an indication of the reliability of those correlations. For the pooled males, the data collected on even-numbered dates were compared with those collected on odd-numbered dates for each of the five Greenwood conditions. The weakest of those correlations was 0.86 (for the Greenwood effect with the 27-dB NBN

masker), and the rest ranged from 0.90 to 0.98. For the pooled females, it was possible to compare data collected during menses with those collected during the midluteal phase for both the behavioral and the physiological measures. The weakest correlation for the five Greenwood conditions was 0.84 (for the Greenwood effect with the 17-dB NBN masker) and the rest ranged from 0.92 to 0.97. Thus, by these tests, the reliabilities for the behavioral conditions were high for both sexes (when the races were pooled). When the menses and midluteal data for the two physiological measures of DPOAE strength were compared, the correlations were 0.82 and 0.72, and when similar comparisons were made for the two CEOAE measures, the correlations were 0.84 and 0.94. So, the reliabilities were reasonably high for both the behavioral and physiological measures. Of course, after partitioning the data by race, the Ns for some combinations of race and sex (or menstrual phase) became rather small, so those correlations need to be interpreted with caution.

E. Correlations with absolute sensitivity

As noted, the detectability of a 3.0-kHz tone (300 ms duration) in the quiet also was measured as part of the larger study. The purpose was to obtain a general proxy for absolute sensitivity in this frequency region, even though we knew that this measure would be imperfect because of the microstructure present in the audiogram. In Table VI are the correlations of absolute sensitivity with the basic Greenwood conditions, with the Greenwood effect, and with the two measures of DPOAE strength. Superscripts indicate those correlations that achieved implied significance during resampling.

Intuition again suggested that the strongest correlations with absolute sensitivity would be for the comparison with the Greenwood condition having no NBN masker, and again this intuition proved *wrong*. For nearly all the combinations of subject group and race grouping, the correlations were low between absolute sensitivity and performance in the Greenwood condition having no NBN present [condition (1) in Table VI]. However, the correlations between absolute sensitivity and the three other basic Greenwood conditions, and the two Greenwood effects were moderately large, and many achieved some level of implied significance under resampling [see conditions (2)–(5) in Table VI]. Furthermore, some of the correlations with absolute sensitivity seemed to differ for the Non-White and White female groups [see conditions (4) and (5)]. Some of the comparisons in Table VI show the most similarity between subject groups and race groupings of all the comparisons reported in this paper. By comparison, the correlations between absolute sensitivity and the two measures of DPOAE strength [conditions (6) and (7) in Table VI] exhibited no apparent pattern across subject groups or race groupings.

Not shown in Table VI are the correlations of absolute sensitivity with CEOAE strength; those were universally in the range of 0.22–0.36 for the pooled subjects (negative for the pooled males and positive for the pooled females), and resampling revealed that they were not rare outcomes. When

TABLE VI. Correlations between absolute sensitivity, Greenwood conditions, and DPOAE strength, all shown separately for the three subject groups and the three race groupings and with implied significance indicated.

Conditions compared ^a	Groups ^b	Females		Males
		Menses	Midluteal	
(1) Absolute sensitivity vs tonal masker plus no NBN masker	Pooled	+0.04	−0.17	−0.17
	Non-White	−0.08	−0.14	−0.12
	White	−0.55 ^c	−0.31	−0.16
(2) Absolute sensitivity vs tonal masker plus 17-dB NBN masker	Pooled	−0.39^c	−0.55^c	−0.47^d
	Non-White	−0.45 ^c	−0.55 ^d	−0.56
	White	−0.48 ^c	−0.37	−0.43
(3) Absolute sensitivity vs tonal masker plus 27-dB NBN masker	Pooled	−0.54^c	−0.56^f	−0.47^d
	Non-White	−0.54 ^d	−0.57 ^e	−0.47
	White	−0.61 ^d	−0.55 ^d	−0.38
(4) Absolute sensitivity vs Greenwood effect with 17-dB NBN masker	Pooled	−0.39^c	−0.41^d	−0.40^c
	Non-White	−0.46 ^c	−0.43 ^c	−0.24
	White	−0.02	+0.01	−0.44
(5) Absolute sensitivity vs Greenwood effect with 27-dB NBN masker	Pooled	−0.55^c	−0.48^e	−0.44^d
	Non-White	−0.55 ^d	−0.51 ^d	−0.49
	White	−0.09	−0.36	−0.34
(6) Absolute sensitivity vs primary level for DPOAE = −5 dB SPL	Pooled	+0.26	+0.32^c	+0.30
	Non-White	+0.21	+0.39	+0.31
	White	+0.06	+0.02	+0.51 ^c
(7) Absolute sensitivity vs DPOAE strength with 50 dB primary tones	Pooled	−0.10	−0.22	−0.25
	Non-White	−0.11	−0.26	−0.18
	White	−0.06	+0.04	−0.32

^aTonal signal was 3.0 kHz, 300 ms in duration; DPOAE = distortion-product otoacoustic emission at 2.393 kHz (primary tones of about 3.076 and 3.760 kHz); NBN = narrowband noise (2.0–2.8 kHz wide); dB SPL = decibels sound-pressure level.

^bNs: 20–26 Non-White menses; 23–26 Non-White midluteal; 10–11 Non-White males; 12–14 White menses; 13–14 White midluteal; 14–15 White males.

^cImplied significance, from resampling: $0.10 > p > 0.05$.

^dImplied significance, from resampling: $0.05 > p > 0.01$.

^eImplied significance, from resampling: $0.01 > p > 0.001$.

^fImplied significance, from resampling: $0.001 > p > 0.0001$.

those same correlations were calculated within race, the values for all conditions were slightly higher for the Non-Whites than for the Whites, but the values all were 0.41 or smaller, and not statistically rare.

F. DPOAEs and narrowband noise

For the Greenwood task, measurements were obtained both with and without the narrowband noise present, whereas all our DPOAE measurements were made without the narrowband noise. This procedural difference raises the question of how our DPOAEs might have been affected by the presence of that noise band. Accordingly, supplementary physiological measurements were made on six subjects, some of whom also had served in the main study reported here.

The noise band used for these supplementary measurements extended from 2.0 to 2.8 kHz, just as in the Greenwood task; however, this noise band had a 75-Hz spectral notch centered at 2.4 kHz so that the DPOAE could be observed. The procedure was to fix the level of the lower frequency primary tone at 60 dB SPL and to vary the level of

the higher frequency primary from 22 to 71 dB SPL in 7-dB increments. This was done both without the noise band present and with it present at spectrum levels ranging from 17 to 32 dB SPL in 5-dB increments. For each combination of these parameters, the strength of the DPOAE was measured.

For the three subjects having strong DPOAE responses, three different outcomes were observed. For one subject, the obtained functions were basically the same for all values of the noise band. For another subject, the strength of the DPOAE diminished as the level of the narrowband noise was increased. For the third subject, the strength of the DPOAE *increased* as the noise level was increased. For the three subjects having weak DPOAE responses, two showed little change across the various levels of the narrowband noise, and the third showed a marked decrease in DPOAE strength with increasing noise level. In the ear showing increased DPOAE strength with increasing noise level, the increases were 3–5 dB at different levels of the higher frequency primary tone. In the ears showing decreasing DPOAE strength with increasing noise level, the decreases were 4–10 dB.

The implication is that, when no noise was present, the various constituent components of the DPOAE were out of phase by differing amounts in the different ears. Thus, there were varying degrees of constructive and destructive interference in the DPOAE waveforms reaching the ear canals. Then, when the narrowband noise diminished the DPOAE component(s) coming from the 2.4-kHz region of the cochlea, the overall DPOAE in the ear canal accordingly weakened, strengthened, or remained the same. In the current context, it is natural to wonder if the correlations between DPOAE strength and the Greenwood effect would differ across subsets of subjects showing these differences in DPOAE response patterns. To answer this question, additional selected subjects would have to be tested.

IV. SUMMARY AND DISCUSSION

A. Summary

We summarize our results as follows.

- (1) The NBN was highly effective at reducing detection performance (Table II), suggesting strongly that information extracted from the 2.4-kHz region of the cochlea was contributing to our listeners' decisions, along with whatever information was being extracted from the 3.6-kHz region activated directly by the signal.
- (2) Both the behavioral and physiological measures differed between the sexes and between the two broad categories of racial background we were able to identify. The sex and race differences interacted for both the behavioral and physiological measures.
- (3) The primary finding was that, for most subjects, the individual differences in the Greenwood effect were only moderately correlated with the strength of a 2.4-kHz DPOAE measured in those same cochleas (Tables IV and V). The Non-White males were an exception to this generalization; their correlations between Greenwood effect and DPOAE magnitude reached about 0.75, which

achieved implied significance during resampling. We have no explanation for why this single group of subjects was different from the others in this important regard.

- (4) The correlations between DPOAE strength and the individual Greenwood conditions were nearly always larger for the females than the males (Table V), but those correlations generally were modest.
- (5) For all race groupings, females exhibited larger Greenwood effects than males, and this sex difference was much larger for the White subjects than for the Non-White subjects (Table II). For the White subjects, the effect sizes for sex difference in the Greenwood effect all were >1.0 , and they did achieve implied significance during resampling, whereas for the Non-White subjects these effect sizes were moderate and did not achieve implied significance.
- (6) There were no evident sex or race differences in detection performance in the condition having only the tonal masker with no NBN masker. Sex and race differences became more evident for the other two basic Greenwood conditions, and they were most evident for the Greenwood effect (Table II).
- (7) Absolute sensitivity was poorly correlated with detection performance in the condition with no NBN masker, but it was moderately correlated with performance in the conditions having the NBN masker, as well as with the two measures of DPOAE (Table VI). Had absolute sensitivity been measured at 2.4 kHz instead of 3.0 kHz, perhaps the latter relationships would have been stronger.
- (8) In accord with the literature (see [McFadden et al., 2009](#)), the sex differences in OAE strength were generally smaller for DPOAEs than CEOAEs (Table III).
- (9) CEOAE strength, the proxy measure for general strength of the reflection mechanism underlying some forms of OAE, also was not well-correlated with any of the behavioral measures (Table IV).
- (10) For CEOAEs, the White subjects exhibited substantial sex differences but the sex differences were essentially zero for the Non-White subjects. If replication confirms this outcome, it will raise interesting questions about the origins of sex differences in OAEs (see [McFadden, 1998, 2009, 2011](#)). It is likely that the past reports on this topic were based on largely White samples, but that is impossible to verify from the published details provided.
- (11) There were essentially no menstrual effects for either the behavioral or physiological measures (Tables II and III). This was true within race groups as well as for the pooled subjects.

B. Discussion

Previous investigators have compared OAEs with behavioral measures of CTs, auditory sensitivity, frequency selectivity, gap detection, psychophysical tuning curves, intensity discrimination, overshoot, and other psychoacoustical measures, with varying degrees of success (e.g., see [Wilson, 1980](#);

Shannon and Houtgast, 1980; Long and Tubis, 1988; Zwicker and Harris, 1990; McFadden and Mishra, 1993; Micheyl and Collett, 1994, 1996; Micheyl *et al.*, 1997; McFadden and Pasanen, 1994; Probst and Harris, 1996; Neumann *et al.*, 1997; Smurzynski *et al.*, 1999; Smurzynski *et al.*, 2001; Furst *et al.*, 1988; Shera *et al.*, 2002; Purcell *et al.*, 2007; Keefe *et al.*, 2009; Goodman *et al.*, 2009; McFadden *et al.*, 2010b; Walsh *et al.*, 2010; Rodriguez *et al.*, 2010). Additional attempts are inevitable and desirable because comparisons of this sort carry considerable potential to illuminate which aspects of auditory perception are attributable to mechanisms operating in the cochlea and which to later, neural mechanisms. [In this regard, note that nonlinearities continue to be evident at stages of processing beyond the cochlea; for example, Rose (1980) demonstrated that the temporal patterns of firing of auditory neurons contain periodicities associated with various CT frequencies.]

Although we did not design the original study with race differences in mind, finding them is not totally unexpected. Several lines of evidence suggest that skin pigmentation (and, by implication, intra-cochlear melanocytes) can affect auditory measures such as hearing sensitivity, number of SOAEs, susceptibility to noise-induced hearing loss, and even auditory brainstem responses (see McFadden and Wightman, 1983; Whitehead *et al.*, 1993; McFadden and Loehlin, 1995; Lin *et al.*, 2012; McFadden *et al.*, 2012). Possible race differences in middle-ear characteristics also have been suggested (Shahnaz and Davies, 2006; Shahnaz and Bork, 2006). To our knowledge, ours is among the first demonstrations of race differences in a mainstream psychoacoustical task.

The two categories of race used here are clearly imperfect for anyone interested in the mechanisms underlying the differences reported, but it was the best we could do under the circumstances. All recruitment and selection of subjects was done without regard to race, the only race information collected was the two NIH items, and the Ns were stretched thin after simply dichotomizing by race. Targeted recruitment would lead to more, more homogeneous, and larger groups and thereby to a better chance of establishing the precise relationships between pigmentation and auditory performance. While the race differences reported here are interesting to contemplate, they must be viewed only as preliminary findings at this point.

As noted, we followed our typical procedure for measuring DPOAEs by collecting gain functions for six pairs of primary tones covering a range of about 0.8 kHz (so the resulting DPOAEs cover a range of about 0.5 kHz) and then averaging across them. The purpose of this procedure is to cancel out the across-frequency fluctuations in DPOAE strength called its microstructure (e.g., Gaskill and Brown, 1990; Heitmann *et al.*, 1998). However, all the analyses reported here were based on the data collected using just one of those six pairs of primary frequencies, the pair closest in frequency to the masker and signal frequencies used in the Greenwood task. This decision was made while we still were working exclusively with data pooled across race; the majority of the correlations between behavior and physiology for the pooled subjects were larger when using the data from the

single pair of primary frequencies. After the data were partitioned by race, the appropriateness of the decision was further reinforced. For example, 72% of the 36 correlations in Table V for the Non-Whites and Whites were larger when using the data from the single pair of primary frequencies than when using the composite data; the average increase was about 0.10. When only the Greenwood conditions having both the NBN masker and the tonal masker present were considered [items (3)–(6) in Table V], 88% of the 24 correlations within race were larger for the single pair of primaries. Also worth noting is that when the DPOAE data obtained using the single pair of primary frequencies were used, the signs of all six correlations in Table V for the White males were the same as the signs for the females, whereas the signs of *all* of those male correlations were reversed when the data from the composite gain functions were used. The implication is that the relationships between our physiological and behavioral measures are highly dependent upon local characteristics of the auditory periphery; simply smoothing out the DPOAE microstructure that is present in the general frequency region of the CT of interest had the effect of clouding those (generally weak-to-moderate) relationships. Using composite data has its place, but apparently this is not one. In regard to microstructure, recall that there was a 7-Hz difference between the 2.4-kHz CT in the Greenwood task and the 2.3926-kHz DPOAE measured here; we doubt that this small difference contributed substantially to the weak correlations observed, but future research may reveal otherwise.

Because the absence of substantial correlations between the Greenwood effect and DPOAE strength is somewhat counterintuitive, there is value in trying to identify factors that may have contributed to those modest correlations. We described our study in detail here as an aid to others looking for any conceptual or procedural errors we may have made. Also, one or more of the following may prove to be relevant; these points are not intended to be mutually exclusive.

- (1) The DPOAE measured in the ear canal is known to be the sum of a combination of sources originating from multiple sites in the cochlea (e.g., Kim, 1980; Shera and Guinan, 1999; Konrad-Martin *et al.*, 2001; Martin *et al.*, 2009), and there is evidence for substantial individual differences in the phases at which the multiple sources add to produce the DPOAE in the ear canal (e.g., Mauer-mann and Kollmeier, 2004; Sec. III F of the present paper). Thus, the correlations between behavior and physiology could be diminished by the uncontrolled patterns of reinforcement or cancellation of the sources in individual ears.
- (2) Let us assume that the subject “knows” about the additional cue for detection in the Greenwood no-NBN-masker condition because of activity at the CT region of the cochlea (here 2.4 kHz). Further assume that the NBN is most effective at disrupting the DPOAE source arising from this cochlear place, and less effective at disrupting the other cochlear source(s). Finally assume that these other source(s) contribute more strongly to the sound recorded in the ear canal than does the source located at the place of the CT frequency (see Konrad-Martin *et al.*,

2001; but compare Talmadge *et al.*, 1999). If true, then our correlations (Tables IV and V) compared DPOAEs originating primarily from sites *other than* the cochlear place of the CT frequency with behavioral performance based primarily on the activity *at* the CT place, and, as a result, the correlations were small. Note that, even with these assumptions, our correlations did not *have* to be low (and they were not for the Non-White males). Had the various DPOAE sources themselves been correlated, the overall DPOAE in the ear canal could have been highly correlated with the Greenwood effect.

- (3) When measuring DPOAEs, a range of primary levels was used for each frequency pair, but for the Greenwood conditions, the tonal masker always was fixed at 62 dB SPL. Often, the gain functions for DPOAEs begin to saturate in the region of about 60–65 dB SPL. Thus, in the listening condition having no NBN, there also may have been a compression of the range of possible magnitudes of the CT cue, which would have diminished the magnitudes of the correlations between physiology and behavior. This suggests that, had a weaker *tonal* masker been used, the Greenwood effects might have been smaller, but the individual differences in those Greenwood effects might have correlated more strongly with the individual differences in our DPOAE measures.
- (4) As noted previously, the DPOAE measurements were made using equal-level primary tones, but the behavioral conditions depended upon an interaction between a tonal masker of 62 dB SPL and a tonal signal 20–30 dB weaker than that masker to produce the CT. Had we measured the DPOAEs using primary tones not equal in level (see Whitehead *et al.*, 1995; Dhar *et al.*, 1998; Konrad-Martin *et al.*, 2001), it is possible that the resulting individual differences would correlate better with the individual differences in the Greenwood effect than was true here. (Contradicting this logical possibility is the observation that using primary tones of different levels changes the strength of the DPOAE only slightly; see Whitehead *et al.*, 1995.)
- (5) Our listeners never were instructed to listen in the frequency region below the tonal masker; the implied emphasis always was on the 3.6-kHz region of the signal. Clearly, most or all of the listeners did discover (consciously or not) that they could detect better by using the off-frequency information (as evidenced by the effectiveness of the NBN maskers), but perhaps if they had been *instructed* to listen at the 2.4-kHz region, the relationship between behavioral and physiological measures might have been stronger. One might expect this factor to be most relevant for the condition with no NBN masker.
- (6) Perhaps some individuals are simply not proficient at hearing (and using) the CT that is generated in our Greenwood conditions, and that “cognitive” factor acts to diminish the correlations of interest. Concretely, perhaps post-cochlear frequency resolution strongly affects performance in the Greenwood conditions. If the individual differences in frequency resolution are independent of the mechanism responsible for the magnitude of the

CT, then the process of frequency resolution could reduce the correlations between the behavioral and physiological measures of the CT. That is, perhaps the individual differences in DPOAE strength *do* affect the magnitude of the CT information available in the cochlea, but that information then is filtered prior to the detection process in a way that eliminates the original individual differences in the cochlea, and thereby reduces the correlations of interest here.

- (7) An anonymous reviewer of this paper pointed out that individual differences in the inner- and middle-ear systems have the potential to reduce the correlations between the behavioral and physiological measures used here. The reason is that the DPOAE measures involved two passes through the inner- and middle-ear systems (the input and output paths), while the Greenwood task presumably involved only the first pass. The reviewer suggested that the second pass has the potential to decorrelate (slightly) the behavioral and physiological measures.

No matter what the ultimate explanation for the generally weak correlations between the physiological and behavioral measures reported here, interest in the relationships between OAEs and behavior surely will continue because OAEs appear to carry the potential to provide additional valuable insights into the cochlear contributions to auditory experience and performance.

Students of human perception have known for centuries that individual people perceive the world slightly differently. Some of these individual differences are inborn and some acquired. For example, some males experience a different palette of hues from other males; some astronomers have slower reaction times to stellar transits than others; some people hear different frequencies as specific, nameable pitch entities while others experience pitch only in a relative way; some people have high-frequency hearing loss that alters their experience of certain speech sounds. Together these examples illustrate that all of us live inside machines, machines that differ slightly in their ability to transduce energy at the surface of the body, transform it via various neural codes, and present it to a sensorium that extracts meaning from the information and builds up a mental representation of the external world. In recent years, evidence has begun to accumulate that some of these individual differences are shared within groups of subjects. For example, there appear to be certain sex differences in audition, and there may be some perceptual differences across the menstrual cycle (McFadden, 1998, 2011). To this still-short list of group differences we now can add race differences in audition (also see McFadden *et al.*, 2012). For now, these additions only can be penciled on to the list until verification is obtained from future research, but if confirmed, race differences have the potential to serve as an additional source of information about the basic functioning of the auditory system and to provide additional insight into the origins of some individual differences in auditory physiology and behavior. In the future, making universal statements about human hearing may require evidence from more diverse samples. With regret, we must leave future work on these topics to others.

ACKNOWLEDGMENTS

This work was supported by a research grant awarded to D.M. by the National Institute on Deafness and other Communication Disorders (NIDCD; RO1 DC000153). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDCD or the National Institutes of Health. We gratefully acknowledge the assistance of K. P. Walsh. Helpful comments were made on a draft of this paper by J. C. Loehlin and C. A. Champlin. Three anonymous reviewers also provided excellent comments.

¹In past work with non-humans (e.g., [McFadden et al., 2006a](#); [McFadden et al., 2006b](#)), we estimated the level of the primary tones necessary to produce a DPOAE of 0 dB SPL, not -5 dB. However, the DPOAEs of humans are weaker than those of non-humans (see [McFadden et al., 2009](#)). As a consequence, the range of primary-tone levels used here sometimes did not produce a DPOAE as strong as 0 dB, meaning that the estimates of the 0-dB point often were extrapolations rather than interpolations. Using a target of -5 dB instead of 0 dB resolved this inelegance. In the few remaining instances (about 5 out of 181 total test sessions) for which the estimate of the primary levels necessary for a DPOAE of -5 dB SPL was an extrapolation beyond the linear range of the gain function, both that measure and the estimate of DPOAE magnitude for primary tones of 50 dB were excluded. That is, subjects with extremely weak DPOAEs did not contribute to the means and correlations reported here.

²None of the six pairs of primary tones had the same frequencies as the tones used in the Greenwood task because we initially intended to use the composite functions to obtain the DPOAE measures for comparison with the behavioral data. Also, in accord with our standard procedure, the $f_{\text{higher}}/f_{\text{lower}}$ ratio was 1.21. Even so, the DPOAE frequency for the specific pair of frequencies used was only about 7 Hz different from the 2.4-kHz CT in the Greenwood task.

³To minimize the inclusion of noisy responses, two criteria were established for each ear immediately prior to collection of CEOAE data. First, clicks were presented and responses recorded for 20 s, and the maximum peak amplitude in each interval beginning 6 ms post-click was calculated. The median and standard deviation of this set of amplitudes were determined. Second, the ambient level in the ear canal was recorded during a 20-s silent period, and the median and standard deviation of those levels were calculated. Then, immediately before presenting a set of ten clicks for CEOAE data collection, the noise level in the ear canal was checked, and if necessary, click presentation was suspended until that level did not exceed 0.25 standard deviation above the median level measured in the quiet.

⁴For DPOAE measurements, the noise floor of the system, as measured in 24-Hz-wide frequency bins flanking the DPOAE frequency, typically was -25 dB SPL or lower. For CEOAEs, the noise floor was estimated by analyzing the 20.48-ms segment of the averaged response beginning 20 ms after click onset. For most ears having no spontaneous otoacoustic emissions (SOAEs), the cochlear response to the click will have decayed completely by that time. For ears having no SOAEs, the CEOAE noise floor so measured was about -5 dB SPL in the band from 1.0 to 5.0 kHz. Both noise-floor measurements correspond to a spectrum level of about -40 dB SPL.

Cohen, J. (1992). "A power primer," *Psychol. Bull.* **112**, 155–159.

Dhar, S., Long, G. R., and Culpepper, N. B. (1998). "The dependence of the distortion product $2f_1-f_2$ on primary levels in non-impaired human ears," *J. Speech Hear. Res.* **41**, 1307–1318.

Dorn, P. A., Piskorski, P., Gorga, M. P., Neely, S. T., and Keefe, D. H. (1999). "Predicting audiometric status from distortion product otoacoustic emissions using multivariate analyses," *Ear. Hear.* **20**, 149–163.

Furst, M., Rabinowitz, W. M., and Zurek, P. M. (1988). "Ear canal acoustic distortion at $2f_1-f_2$ from human ears: Relation to other emissions and perceived combination tones," *J. Acoust. Soc. Am.* **84**, 215–221.

Gaskill, S. M., and Brown, A. M. (1990). "The behavior of the acoustic distortion product, $2f_1-f_2$, from the human ear and its relation to auditory sensitivity," *J. Acoust. Soc. Am.* **88**, 821–839.

Goodman, S. S., Fitzpatrick, D. F., Ellison, J. C., Jesteadt, W., and Keefe, D. H. (2009). "High frequency click-evoked otoacoustic emissions and behavioral thresholds in humans," *J. Acoust. Soc. Am.* **125**, 1014–1032.

Greenwood, D. D. (1971). "Aural combination tones and auditory masking," *J. Acoust. Soc. Am.* **50**, 502–543.

Haggerty, H. S., Lusted, H. S., and Morton, S. C. (1993). "Statistical quantification of 24 hour and monthly variabilities of spontaneous otoacoustic emission frequency in humans," *Hear. Res.* **70**, 31–49.

Hall, J. L. (1972). "Monaural phase effect: Cancellation and reinforcement of distortion products f_2-f_1 and $2f_1-f_2$," *J. Acoust. Soc. Am.* **51**, 1872–1881.

Hall, J. L. (1975). "Nonmonotonic behavior of distortion product $2f_1-f_2$: Psychophysical observations," *J. Acoust. Soc. Am.* **58**, 1046–1050.

Heitmann, J., Waldmann, B., Schnitzler, H. U., Plinkert, P. K., and Zenner, H. P. (1998). "Suppression of distortion product otoacoustic emissions (DPOAE) near $2f_1-f_2$ removes DP-gram fine structure—evidence for a secondary generator," *J. Acoust. Soc. Am.* **102**, 1527–1531.

Keefe, D. H., Schairer, K. S., Ellison, J. C., Fitzpatrick, D. F., and Jesteadt, W. (2009). "Use of stimulus-frequency otoacoustic emissions to investigate efferent and cochlear contributions to temporal overshoot," *J. Acoust. Soc. Am.* **125**, 1595–1604.

Kemp, D. T. (1979). "Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea," *Arch. Otorhinolaryngol.* **224**, 37–45.

Kim, D. O. (1980). "Cochlear mechanics: Implications of electrophysiological and acoustical observations," *Hear. Res.* **2**, 297–317.

Konrad-Martin, D., Neely, S. T., Keefe, D. H., Dorn, P. A., and Gorga, M. P. (2001). "Sources of distortion product otoacoustic emissions revealed by suppression experiments and inverse fast Fourier transforms in normal ears," *J. Acoust. Soc. Am.* **109**, 2862–2879.

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

Licklider, J. C. R. (1951). "Basic correlates of the auditory stimulus," in *Handbook of Experimental Psychology*, edited by S. S. Stevens (Wiley, New York), pp. 985–1039.

Lin, F. R., Maas, P., Chen, W., Carey, J. P., Ferrucci, L., and Thorpe, R. (2012). "Association of skin color, race/ethnicity, and hearing loss among adults in the USA," *J. Assoc. Res. Otolaryngol.* **13**, 109–117.

Long, G. R., and Tubis, A. (1988). "Modification of spontaneous and evoked otoacoustic emissions and associated psychoacoustic microstructure by aspirin consumption," *J. Acoust. Soc. Am.* **84**, 1343–1353.

Martin, G. K., Stagner, B. B., Fahey, P. F., and Lonsbury-Martin, B. L. (2009). "Steep and shallow phase gradient distortion product otoacoustic emissions arising basal to the primary tones," *J. Acoust. Soc. Am.* **125**, EL85–EL92.

Mauermann, M., and Kollmeier, B. (2004). "Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source," *J. Acoust. Soc. Am.* **116**, 2199–2212.

McFadden, D. (1998). "Sex differences in the auditory system," *Dev. Neuropsychol.* **14**, 261–298.

McFadden, D. (2009). "Masculinization of the mammalian cochlea," *Hear. Res.* **252**, 37–48.

McFadden, D. (2011). "Sexual orientation and the auditory system," *Front. Neuroendocrinol.* **32**, 210–213.

McFadden, D., Garcia-Sierra, A., Hsieh, M. D., Maloney, M. M., Champlin, C. A., and Pasanen, E. G. (2012). "Relationships between otoacoustic emissions and a proxy measure of cochlear length derived from the auditory brainstem response," *Hear. Res.* **289**, 63–73.

McFadden, D., Hsieh, M. D., Garcia-Sierra, A., and Champlin, C. A. (2010a). "Differences by sex, ear, and sexual orientation in the time intervals between successive peaks in auditory evoked potentials," *Hear. Res.* **270**, 56–64.

McFadden, D., and Loehlin, J. C. (1995). "On the heritability of spontaneous otoacoustic emissions: A twins study," *Hear. Res.* **85**, 181–198.

McFadden, D., Martin, G. K., Stagner, B. B., and Maloney, M. M. (2009). "Sex differences in distortion-product and transient-evoked otoacoustic emissions compared," *J. Acoust. Soc. Am.* **125**, 239–246.

McFadden, D., and Mishra, R. (1993). "On the relation between hearing sensitivity and otoacoustic emissions," *Hear. Res.* **71**, 208–213.

McFadden, D., and Pasanen, E. G. (1994). "Otoacoustic emissions and quinine sulfate," *J. Acoust. Soc. Am.* **95**, 3460–3474.

McFadden, D., and Pasanen, E. G. (1998). "Comparison of the auditory systems of heterosexuals and homosexuals: Click-evoked otoacoustic emissions," *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2709–2713.

- McFadden, D., and Pasanen, E. G. (1999). "Spontaneous otoacoustic emissions in heterosexuals, homosexuals, and bisexuals," *J. Acoust. Soc. Am.* **105**, 2403–2413.
- McFadden, D., Pasanen, E. G., Raper, J., Lange, H. S., and Wallen, K. (2006a). "Sex differences in otoacoustic emissions measured in rhesus monkeys (*Macaca mulatta*)," *Horm. Behav.* **50**, 274–284.
- McFadden, D., Pasanen, E. G., Weldele, M. L., Glickman, S. E., and Place, N. J. (2006b). "Masculinized otoacoustic emissions in female spotted hyenas (*Crocuta crocuta*)," *Horm. Behav.* **50**, 285–292.
- McFadden, D., Walsh, K. P., Pasanen, E. G., and Grenwelge, E. M. (2010b). "Overshoot using very short signal delays," *J. Acoust. Soc. Am.* **128**, 1915–1921.
- McFadden, D., and Wightman, F. L. (1983). "Audition: Some relations between normal and pathological hearing," *Annu. Rev. Psychol.* **34**, 95–128.
- Micheyl, C., and Collet, L. (1994). "Interrelations between psychoacoustical tuning curves and spontaneous and evoked otoacoustic emissions," *Scand. Audiol.* **23**, 171–178.
- Micheyl, C., and Collet, L. (1996). "Involvement of the olivocochlear bundle in the detection of tones in noise," *J. Acoust. Soc. Am.* **99**, 1604–1610.
- Micheyl, C., Perrot, X., and Collet, L. (1997). "Relationship between auditory intensity discrimination in noise and olivocochlear efferent system activity in humans," *Behav. Neurosci.* **111**, 801–807.
- Neumann, J., Uppenkamp, S., and Kollmeier, B. (1997). "Relations between notched noise suppressed TEOAE and the psychoacoustical critical bandwidth," *J. Acoust. Soc. Am.* **101**, 2778–2788.
- Probst, R., and Harris, F. P. (1996). "Effect of otoacoustic emissions on just-noticeable differences for intensity in normally hearing subjects," *J. Acoust. Soc. Am.* **100**, 504–510.
- Purcell, D. W., Ross, B., Picton, T. W., and Pantev, C. (2007). "Cortical responses to the $2f_1$ - f_2 combination tone measured indirectly using magnetoencephalography," *J. Acoust. Soc. Am.* **122**, 992–1003.
- Rodriguez, J., Neely, S. T., Patra, H., Kopun, J., Jesteadt, W., Tan, H., and Gorga, M. (2010). "The role of suppression in psychophysical tone-on-tone masking," *J. Acoust. Soc. Am.* **127**, 361–369.
- Rose, J. E. (1980). "Neural correlates of some psychoacoustic experiences," in *Neural Mechanisms in Behavior*, edited by D. McFadden (Springer, New York), pp. 1–37.
- Shahnaz, N., and Bork, K. (2006). "Wideband reflectance norms for Caucasian and Chinese young adults," *Ear Hear.* **27**, 774–788.
- Shahnaz, N., and Davies, D. (2006). "Standard and multifrequency tympanometric norms for Caucasian and Chinese young adults," *Ear Hear.* **27**, 75–90.
- Shannon, R. V., and Houtgast, T. (1980). "Psychophysical measurements relating suppression and combination tones," *J. Acoust. Soc. Am.* **68**, 825–829.
- Shera, C. A., and Guinan, J. J., Jr. (1999). "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs," *J. Acoust. Soc. Am.* **105**, 782–798.
- Shera, C. A., Guinan, J. J., Jr., and Oxenham, A. J. (2002). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3318–3323.
- Smooenburg, G. F. (1972a). "Auditory region of combination tones," *J. Acoust. Soc. Am.* **52**, 603–614.
- Smooenburg, G. F. (1972b). "Combination tones and their origin," *J. Acoust. Soc. Am.* **52**, 615–632.
- Smooenburg, G. F. (1974). "On the mechanisms of combination tone generation and lateral inhibition in hearing," in *Psychological Models and Physiological Facts in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin), pp. 332–343.
- Smurzynski, J., Harris, F. P., and Probst, R. (2001). "Otoacoustic emissions in relation to auditory perception: Beyond the pure-tone audiogram," *Semin. Hear.* **22**, 361–375.
- Smurzynski, J., Probst, R., and Grandori, F. (1999). "Special issue: Otoacoustic emissions and psychoacoustics," *Audiology* **38**, 241–280.
- Talmadge, C. L., Long, G. R., Tubis, A., and Dhar, S. (1999). "Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions," *J. Acoust. Soc. Am.* **105**, 275–292.
- Valero, M. D., Pasanen, E. G., McFadden, D., and Ratnam, R. (2008). "Distortion product otoacoustic emissions in the common marmoset (*Calithrix jacchus*): Parameter optimization," *Hear. Res.* **243**, 57–68.
- Walsh, K. P., Pasanen, E. G., and McFadden, D. (2010). "Overshoot measured physiologically and psychophysically in the same human ears," *Hear. Res.* **268**, 22–37.
- Wegel, R. L., and Lane, C. E. (1924). "The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear," *Phys. Rev.* **23**, 266–285.
- Whitehead, M. L., Kamal, N., Lonsbury-Martin, B. L., and Martin, G. K. (1993). "Spontaneous otoacoustic emissions in different racial groups," *Scand. Audiol.* **22**, 3–10.
- Whitehead, M. L., McCoy, M. J., Lonsbury-Martin, B. L., and Martin, G. K. (1995). "Dependence of distortion-product otoacoustic emissions on primary levels in normal and impaired ears. I. Effects of decreasing L_2 below L_1 ," *J. Acoust. Soc. Am.* **97**, 2346–2358.
- Wilson, J. P. (1980). "The combination tone, $2f_1$ - f_2 , in psychophysics and ear-canal recording," in *Psychophysical, Physiological and Behavioral Studies in Hearing*, edited by G. Van den Brink and F. A. Bilsen (Delft University Press, Delft, The Netherlands), pp. 43–50.
- Wit, H. P., Langevoort, J. C., and Ritsma, R. J. (1981). "Frequency spectra of cochlear acoustic emissions ('Kemp-echoes')," *J. Acoust. Soc. Am.* **70**, 437–445.
- Yellin, M. W., and Stillman, R. D. (1999). "Otoacoustic emissions in normal-cycling females," *J. Am. Acad. Audiol.* **10**, 400–408.
- Zwicker, E., and Harris, F. P. (1990). "Psychoacoustical and ear canal cancellation of ($2f_1$ - f_2)-distortion products," *J. Acoust. Soc. Am.* **87**, 2583–2591.