

Differences in common psychoacoustical tasks by sex, menstrual cycle, and race

Dennis McFadden,^{a)} Edward G. Pasanen, and Mindy M. Maloney

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

Erin M. Leshikar and Michelle H. Pho

Department of Communication Sciences and Disorders, University of Texas, 2504-A Whitis Avenue, A1100, Austin, Texas 78712-0114, USA

(Received 16 November 2017; revised 13 March 2018; accepted 24 March 2018; published online 24 April 2018)

The psychoacoustical literature contains multiple reports about small differences in performance depending upon the sex and phase of the menstrual cycle of the subjects. In an attempt to verify these past reports, a large-scale study was implemented. After extensive training, the performance of about 75 listeners was measured on seven common psychoacoustical tasks. For most tasks, the signal was a 3.0-kHz tone. The initial data analyses failed to confirm some past outcomes. Additional analyses, incorporating the limited information available about the racial background of the listeners, did confirm some of the past reports, with the direction and magnitude of the differences often diverging for the White and Non-White listeners. Sex differences and race differences interacted for six of the seven tasks studied. These interactions suggest that racial background needs to be considered when making generalizations about human auditory performance, and when considering failures of reproducibility across studies. Menstrual differences were small, but generally larger for Whites than Non-Whites. Hormonal effects may be responsible for the sex and cycle differences that do exist, and differences in intra-cochlear melanocytes may account for the race differences. © 2018 Acoustical Society of America. <https://doi.org/10.1121/1.5030998>

[BLM]

Pages: 2338–2354

I. INTRODUCTION

In the nearly 40 years since the discovery of otoacoustic emissions (OAEs; Kemp, 1978, 1979), relatively little has been learned about how OAEs relate to everyday hearing. Both OAEs and performance in common psychoacoustical tasks exhibit large individual differences in people with normal hearing. Thus, a reasonable question is: Might some of the individual differences in cochlear physiology co-vary with the individual differences in performance on certain psychoacoustical tasks? To address that question, a large-scale study was implemented. Multiple crews of normal-hearing listeners were tested physiologically and behaviorally for several weeks each on seven psychoacoustical tasks, and their OAEs were measured. Also of interest were any differences by sex or menstrual cycle for the psychoacoustical tasks (McFadden, 1998). It is the latter findings that are reported here; the correlations between OAEs and psychoacoustical performance are reported in a companion article (McFadden *et al.*, 2018). A third report is in preparation on psychoacoustical performance and auditory evoked potentials (AEPs), which also were measured in this study.

Sex differences in psychoacoustical tasks were of interest in part because they are so well-established for OAEs. Females are known to have more spontaneous otoacoustic emissions (SOAEs; see Bilger *et al.*, 1990; Burns *et al.*, 1992;

Talmadge *et al.*, 1993; McFadden and Pasanen, 1999) and stronger click-evoked otoacoustic emissions (CEOAEs; McFadden and Pasanen, 1998) than males. Also, sex differences are known to be smaller for distortion-product otoacoustic emission (DPOAEs) than for other OAEs (McFadden *et al.*, 2009a). Because sex differences exist in the OAEs of newborns (Strickland *et al.*, 1985; Burns *et al.*, 1992) as well as in adults, and because OAEs appear to be relatively stable through life (Burns, 2017), it appears that OAEs are permanently influenced by prenatal events, possibly hormonal events (reviewed by McFadden, 2002, 2008, 2011). Bolstering this idea are results from various non-human species that were exposed to atypical hormone levels prenatally, and the results from studies on various special populations of humans also are of interest in this context (all reviewed by McFadden *et al.*, 2009b). For sex and hormonal differences in other species, also see the book by Bass *et al.* (2016).

Unlike OAEs, sex differences in auditory performance have not been systematically studied, but evidence does exist for some tasks (reviewed by McFadden, 1998). For example, females are reported to have better hearing sensitivity in the quiet (Chung *et al.*, 1983; Agrawal *et al.*, 2008), more overshoot (Wright, 1994), less two-tone suppression (Wright, 1994), and less noise-induced hearing loss (NIHL) (Royster *et al.*, 1980; Agrawal *et al.*, 2008) than males. (Overshoot is the difference in detectability of a brief tonal signal when its onset occurs soon after the onset of a wideband masker versus later in its time course; two-tone suppression is the

^{a)}Electronic mail: mcfadden@utexas.edu

difference in forward masking for a brief tone masked only by another tone of the same frequency versus masked by that tone plus another tone slightly higher in frequency.) By contrast, males are reported to be better than females at frequency discrimination and loudness discrimination (Rammsayer and Troche, 2012), at sound localization (reviewed by McFadden, 1998; Zündorf *et al.*, 2011), at using an off-frequency cue in a simultaneous tone-on-tone masking task (the Greenwood effect; see McFadden *et al.*, 2012b), and at detecting a tonal signal in the presence of a wideband masker whose frequency content varies from presentation to presentation (called profile analysis; see Neff *et al.*, 1996). It is not clear what factor(s) links these psychoacoustical tasks, but some of these previously studied tasks are among the seven studied for this report. More detail about the tasks studied is provided below.

Effects of the menstrual cycle on the auditory system have been reported occasionally. The existence and magnitude of these effects were of interest to us because of certain other, circumstantial evidence that hormones can affect the auditory system (reviewed by McFadden, 1998, 2000). The literature suggests that during menses, hearing sensitivity in the quiet is worse and NIHL is less than during other phases of the cycle (Davis and Ahroon, 1982; Swanson and Dengerink, 1988; reviewed by McFadden, 1998, 2000). Physiologically, SOAEs were reported to shift in frequency across the cycle (Bell, 1992; Haggerty *et al.*, 1993), and Wave V of the AEPs to fluctuate in latency (Elkind-Hirsch *et al.*, 1992). There is some evidence that oral contraceptives moderate these effects (McFadden, 2000). Also, there is evidence of OAEs fluctuating in rhesus and human males with seasonal changes in androgen levels (McFadden *et al.*, 2006; Snihur and Hampson, 2012). Viewed as a whole, these behavioral and physiological studies suggested that the cochlea, the auditory brain, and psychoacoustical performance all might vary with the changes in hormone levels associated with the menstrual cycle. If true, this relationship needed further documentation and extension to other psychoacoustical tasks; accordingly, this study was designed to permit the collection of both psychoacoustical and physiological measures during different phases of the cycle.

After all testing was completed and data analysis begun, we realized that some results differed depending upon the race/ethnicity of the subjects, and the analyses were changed to explore this variable. Race/ethnic differences have been reported for hearing sensitivity and NIHL (e.g., Royster *et al.*, 1980; Agrawal *et al.*, 2008; DaCosta *et al.*, 2008; Lin *et al.*, 2012), but otherwise race generally has been ignored by mainstream psychoacoustics. In the reports that do exist, the direction of effect was better hearing sensitivity and less NIHL in people with dark skins than in less-pigmented people. Previously we reported an interaction of sex difference with race for the Greenwood effect (McFadden *et al.*, 2012b), further discussed below.

Some investigators of auditory physiology also had noticed and studied race differences. The typical direction of effect was that people with dark skin or dark eyes had more SOAEs (Russell, 1992; Whitehead *et al.*, 1993; McFadden and Loehlin, 1995), stronger CEOAEs (Shahnaz, 2008), and less bothersome tinnitus (Shargorodsky *et al.*, 2010). A common

interpretation is that intracochlear melanocytes appear to be involved in the mechanisms underlying OAEs and in the mechanisms underlying cochlear protection and/or restoration after noise exposure (Lin *et al.*, 2012). Even with this knowledge of the auditory literature, the current study was not initially designed to explore race differences; we did not recruit with race in mind. Consequently, we are not able now to make some obvious comparisons that otherwise might have been possible, and all outcomes relevant to race must be interpreted with caution.

Although this report is about sex and race differences, there must be no uncertainty about the views of the authors about these findings. *Differences are not deficiencies*, and no difference reported here should be interpreted as a deficiency. Some of the differences we found may exist because of some direct evolutionary pressure(s) placed upon the hearing of early humans. However, most of the differences we observed were small, probably too small to affect everyday listening. That suggests that the observed differences are likely to be simply incidental by-products of responses to other evolutionary pressures that may well have been critical to the survival of individual members of the species. In Sec. IV we make the important point that race categories are a poor proxy for the likely actual basis for these apparent “race” differences: individual differences in pigmentation.

II. METHODS

One motivation for this large-scale study was to document the relationships between certain physiological measures of the auditory system (OAEs and AEPs) and performance on certain psychoacoustical tasks. Those relationships are reported elsewhere (McFadden *et al.*, 2018). Here we describe the seven psychoacoustical tasks studied and the results obtained for them. From the outset, the study was designed to measure sex differences in both the psychoacoustical tasks and the physiological measures, and, for the female subjects, also to measure any differences across the menstrual cycle. As noted above, after all the data were collected, we noticed some unanticipated race/ethnic effects (McFadden *et al.*, 2012b), so that variable is included in the analyses here. Because racial background was not systematically considered during recruitment, all outcomes relevant to that variable must be viewed as highly tentative.

All aspects of this study were approved by the Institutional Review Board of The University of Texas, Austin.

A. Subjects

Subjects were primarily students at this university. They were recruited using employment websites, flyers posted at various places on campus, and word of mouth. The acceptable age range was 18–33; in the end, the average age was 21 yrs for both females and males. Prior to hiring, all subjects were screened to have audiometrically normal hearing sensitivity (≤ 15 dB Hearing Level) for the standard audiometric frequencies between 250 and 8000 Hz, and normal middle-ear function as measured by a clinical audiometric screening device (Auto Tympanometry 38, Grason-Stadler, Eden Prairie, MN). All subjects gave informed consent prior to

being hired, and all hired subjects completed an extensive questionnaire with items about physical characteristics, past exposures to intense sound, sexual experiences and fantasies (for categorizing heterosexuals and non-heterosexuals), drug use, and race/ethnicity.

Because one of the variables of interest was the effect of the menstrual cycle on auditory performance, and because previous research suggested that oral contraceptives could affect the OAEs important to the larger study (McFadden, 2000), only non-users of oral or other hormonal contraceptives (“naturally-cycling” women) were recruited and hired. Other work revealed that sexual orientation also could affect OAEs (McFadden and Pasanen, 1998, 1999). Non-heterosexuals were not excluded from the study, but their data were excluded from all analyses reported here; specifically, the data for seven females and five males were excluded for sexual orientation. Orientation was determined using the standard Kinsey items on fantasies and experience from our questionnaire.

Because race/ethnicity emerged as a relevant variable only after data collection was complete, we tested a racially heterogeneous group that reflected the makeup of the student body at this university (and probably at most universities in the U.S.). Then, the best we could do for our analyses was use the standard demographic items required by our funding agency¹ to partition our subjects into two categories—White and Non-White.

Data were collected for 140 subjects, 61 females and 79 males. Of this total, 11 females and six males quit the study prior to the completion of testing, and they were excluded from the analyses reported here. Also excluded were the non-heterosexual subjects (7 females, 5 males) mentioned above. As explained below, the data from 41 male subjects could not be included with the other male subjects and they are reported separately; for 12 of those males no information was available about race/ethnicity because of an oversight with their questionnaires. Some other subjects did not provide answers to the race/ethnicity items, and for various technical reasons, not every subject contributed to every psychoacoustical and physiological measure. Accordingly, the N 's vary slightly across conditions and sometimes the N 's for the subjects partitioned by race do not sum to the N for the subjects pooled over race.

B. Procedures for psychoacoustical tasks

Same-sex crews of 4–8 listeners were hired to work approximately 2 h/day, 5 days/week. The female crews typically worked for 8–10 weeks and the male crews for 6–8 weeks. Subjects were paid an hourly wage plus a bonus upon completion of all testing. All members of a crew were tested simultaneously on the psychoacoustical tasks. Each sat in a separate test booth inside the same large, double-walled test room (Acoustic Systems, model 284221, Hillside, IL); subjects could not see each other during the tests. Each booth had a set of TDH-39 headphones in circumaural cushions, although stimuli were presented only to the right ear. A Macintosh G4 computer (Apple Computer, Cupertino, CA) running software written by E.G.P. (LabVIEW[®] from National Instruments,

Austin, TX) controlled the trial sequence and warning lights, presented the stimuli, and collected the responses for all subjects simultaneously. Stimuli were generated digitally at a sampling rate of 50 kHz with 16-bit resolution; the digital/analog conversion was performed by an interface board (model PCI-MIO-16XE-10 from National Instruments, Austin, TX) installed on the internal bus of the G4 computer. For all psychoacoustical tasks, the procedure was two-interval, two-alternative forced choice. On every trial of a 50-trial 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 contained the signal. Immediately after the response interval, one of two lights was lit to indicate which interval contained the signal. The trial-timing sequence for all tasks was: 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), and pause (~10 ms).

For all listening conditions, the level of the signal was varied adaptively for each subject 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 47 responses, at least four reversals, and a standard deviation for those reversals smaller than 3.5 dB were kept as eligible for analysis (other exclusion rules are described below).

For the first block of a set, the signal level on the first trial was selected to be easily detectable for all subjects. For subsequent blocks of each set, the initial signal level was determined individually for each subject; namely, 10 dB above the mean of his/her reversals for the previous block.

Seven psychoacoustical tasks or abilities were studied. One involved no masking: detection of a 3.0-kHz tone in the quiet. Four involved simultaneous masking, and two involved temporal (forward) masking. Most of these tasks required multiple conditions of listening (types of blocks). In all, there were 16 different conditions of listening, and multiple blocks of trials were collected for each. With the exception of the Greenwood task (see below), the signal to be detected always was a 3.0-kHz tone. Cosine-squared gating envelopes were used for all rise/decay times.

The subjects in the (all-female) crews kept daily diaries on their menstrual cycles and the (all-female) experimenters used this diary information to partition each subject's psychoacoustical data into three categories: Menses, Midluteal, and the remainder (nominally Ovulatory). The partitioning was done for each subject individually, *after* all data collection was complete. The onset of menses was used for partitioning (the onset of menses was defined as the first day of bleeding, not just spotting). All blocks of trials collected during the first 5 days of a cycle were defined as Menses, all blocks collected between 3 and 12 days prior to the onset of menses were defined as Midluteal, and all blocks collected

in between were defined as Ovulatory. Note that this procedure accommodated individual differences in cycle duration. For most subjects, at least some data were collected for two occurrences of each of the three cycle phases, but because of week-ends and occasional absences, the number of blocks of trials collected during each cycle phase varied considerably. [Female subjects also kept the (female) experimenters orally informed about each onset of menses so that OAEs (and AEPs) could be measured during both the Menses and Midluteal phases; for details see [McFadden et al. \(2018\)](#).]

Because the menstrual cycles of the females necessarily were not synchronized and the phase of each subject was defined by a future event (the onset of her next menses), it was desirable to collect data on as many psychoacoustical conditions as possible per daily session (dense sampling). Thus, for female crews, each daily session typically consisted of 3 blocks for each of 7 conditions of listening (7 “sets”). In this way, data were collected on all 16 conditions every 2–3 sessions for the female crews. For male crews, 6 blocks of trials for 4 different psychoacoustical conditions (4 sets) typically were collected per daily session, meaning that about 4 sessions were required to collect data for every condition of listening. For both sexes, there were breaks of approximately 15 s between blocks, approximately 45 s between sets, and approximately 10 min about half way through the 2 h session. The levels of the stimuli were checked for each subject booth at each transition between tasks.

C. VM malfunction

Midway through this study, the true root-mean-square voltmeter (VM) used for setting the levels of the stimuli for the psychoacoustical tasks malfunctioned, and this was not detected until after data collection was complete. The result was that, after the malfunction, all the acoustic stimuli used were 8.1 dB weaker than intended (after the malfunction, the VM was reliable, but wrong by a constant). Numerous lines of evidence revealed that the malfunction occurred between two crews of listeners, just prior to our beginning to test female crews; only male subjects were tested prior to the malfunction.

Because performance on many of the listening conditions varies with level, data obtained prior to the malfunction of the VM cannot easily be compared with data obtained after the malfunction. Thus, only the data from the 27 males tested after the malfunction are compared here with the female data. The data from the initial 42 males (for only some of whom we had race/ethnicity information) also are shown for completeness, but are not included in the detailed analyses. The two groups of male subjects are labeled Pre-VMm and Post-VMm for pre- and post-VM malfunction, respectively. To summarize, some all-male crews were run prior to the malfunction of the VM (the Pre-VMm males); then all the females and some additional males (the Post-VMm males) were tested.

All sound-pressure levels (SPLs) used here to describe the stimuli and the results were the actual levels used, not the initially intended levels. Unless otherwise indicated, the levels given were those used for the females and Post-VMm males.

D. Training and culling of blocks

Because the subjects were tested intensively for several weeks, many blocks of trials were collected for every listening condition, meaning we often were able to exclude individual blocks for various reasons and still have adequate data on which to base estimates of performance. Recall that to be eligible for analysis, a block of 50 trials needed to have at least 47 responses, at least four reversals, and a standard deviation of those reversals smaller than 3.5 dB.

In order to satisfy the goals of the larger study, all subjects needed to be fully experienced with each task prior to collection of the data used for the analyses. Accordingly, the first six blocks collected for each listening condition were considered to be practice and were discarded.

After data collection was complete, all blocks were carefully examined and (rarely) some were excluded for cause. For example, sometimes the data were atypical for an individual subject (or an entire crew) for one or more conditions during a test session. Our assumption was that the acoustic stimuli were miscalibrated by the experimenter, and those blocks of trials were excluded from subsequent analyses. For each subject individually, every remaining block of trials was compared with that subject’s mean performance on that condition of listening and if it was 10 dB different from (typically worse than) that mean, it was excluded as being aberrant for some unknown reason.

Finally, in an attempt to obtain the cleanest estimates of performance possible, means were calculated for each male subject for each condition of listening, and the four blocks most distant from each mean were excluded from all subsequent analyses. The same was done for the individual female subjects, except the *six* most-extreme blocks were excluded (more blocks were available for the female subjects because they were employed about 2 weeks longer than the male subjects).

In general, final performance for each condition of listening was based on between 12 and 27 blocks of trials for the females, distributed across the menstrual cycle, and between 11 and 20 blocks for the post-VMm males, all highly practiced subjects. For most of the seven psychoacoustical tasks studied, estimating performance involved comparing two or more conditions of listening (e.g., differences between with and without the second masker), each estimated using the above-described exclusion procedures. For comparisons across the menstrual cycle, only those female subjects having usable data for at least one menses phase and one midluteal phase were included, but those females excluded from menstrual comparisons were included for other comparisons.

E. Analyses

The primary comparisons of interest here are differences by sex, menstrual cycle, and race. These comparisons are summarized using effect size (a d' -like measure), which is defined here as the difference between the means of the two conditions of interest divided by the square root of the weighted means of the variances for those 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). We also report Pearson product-moment correlations between the performances on various pairs of psychoacoustical tasks. [Elsewhere we report the correlations between each of the psychoacoustical tasks and the OAE (or AEP) measures also obtained on these subjects (McFadden *et al.*, 2018).]

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. As noted, effect sizes were calculated for the various comparisons between the sexes, menstrual phases, and the races; these are called the *actually obtained effect sizes*. To assess the significance of each obtained effect size, (1) *all* of the individual values for the two groups or conditions of interest were pooled; (2) a sample the size of one of the two groups was drawn at random, and without replacement, from the pool; (3) the remaining values were taken to represent the second group; (4) an effect size was calculated for those two “groups” and saved; (5) the process was repeated 20 000 times, and a tally was kept of how often the resampled effect sizes exceeded the actually obtained effect size for that comparison. That tally was divided by 20 000 to yield a proportion that provides an estimate of how rare the actually obtained effect size was, given the actual data obtained from this sample of subjects. We call those proportions the *implied significance value* for the actually obtained effect size.

For resampling a correlation, the obtained values for the first variable were retained for all subjects, but all the values for the second measure were replaced at random by values achieved by other subjects in that same group; that correlation was calculated and saved (called a resampled correlation), and the process was repeated 20 000 times. The tally of the number of times the resampled correlations exceeded the actually obtained correlation (divided by 20 000) was taken as the implied significance of the actually obtained correlation.

When making our tallies we used the *absolute value* of the resampled correlations and effect sizes, meaning that our estimates of implied significance should be viewed as conservative (“two-tailed”). Here we use the term “negligible” to denote correlations or effect sizes that did not achieve implied significance values of 0.10 (marginally significant) or smaller. We have used versions of this resampling procedure in the past (e.g., [McFadden *et al.*, 2012b](#)).

III. PSYCHOACOUSTICAL TASKS: DESCRIPTIONS, RESULTS, COMMENTS

The results from the seven psychoacoustical tasks fell roughly into three categories, and that determined the order in which the tasks are considered below. Four tasks exhibited little or no sex difference when the subjects were pooled across race, but showed a moderate-to-large sex difference

for the White group when the subjects were partitioned by race. Two tasks exhibited a moderate-to-large sex difference when the subjects were pooled across race, but the sex difference was noticeably larger for one race group when the subjects were partitioned by race. One task exhibited about the same (small) sex difference whether the subjects were partitioned or not.²

A. Tonal detection in the quiet

Probably the most commonly mentioned sex difference for audition is for hearing sensitivity in the quiet. Various databases show that females are slightly more sensitive in the quiet than are males, especially at middle frequencies and above (e.g., [McFadden, 1993, 1998](#); [Agrawal *et al.*, 2008](#)). It always has been difficult to obtain data on this topic that are not contaminated by NIHL, and that problem has only become greater in recent times, even for college-age populations. Various lines of evidence confirm the casual observation that males typically spend more time in noisy environments than do females. Because this is a life-long characteristic, the potential for NIHL is higher in males than females. Accordingly, with increasing age, males become increasingly less sensitive to weak sounds than do females ([Agrawal *et al.*, 2008](#)), but this difference may be attributable solely to differences in NIHL.

The literature also contains evidence of race/ethnic differences in hearing sensitivity, overlaid on any existing sex differences. Industrial databases and research studies point to greater hearing sensitivity (and less NIHL) in people with highly pigmented skin. For example, Non-White males appear to be more sensitive than White males (e.g., [Royster *et al.*, 1980](#); [Agrawal *et al.*, 2008](#); [Lin *et al.*, 2012](#)). In parallel findings, more and stronger SOAEs and stronger CEOAEs have been reported for people with highly pigmented skin or eyes ([Russell, 1992](#); [Whitehead *et al.*, 1993](#); [McFadden and Loehlin, 1995](#)). A parsimonious interpretation is that the concentration of intra-cochlear melanin has some effect on the cochlear amplifiers. The issue of the melanin connection was been discussed by [Lin *et al.* \(2012\)](#) and [McFadden and Wightman \(1983\)](#).

Here, the detectability of a 3.0-kHz tone in the quiet was measured using a duration of 300 ms (including 20-ms rise/decay times). The results are summarized in [Fig. 1](#) and [Table I](#), column (1). When the subjects were pooled across race, the expected better sensitivity for females over males was not evident ([Fig. 1](#), top panel). However, when the subjects were partitioned by race ([Fig. 1](#), bottom panel), the expected result was present, but only for the White subjects. That is, there was an interaction between sex and race differences. White females were about 2.0 dB more sensitive than White males, which is similar in direction and magnitude to the sex difference reported by [Chung *et al.* \(1983\)](#) and [McFadden and Mishra \(1993\)](#). By contrast, Non-White females and males differed only by fractions of a decibel. [Table I](#), column (1), shows the effect sizes for the sex, race, and menstrual differences, essentially none of which achieved implied significance.

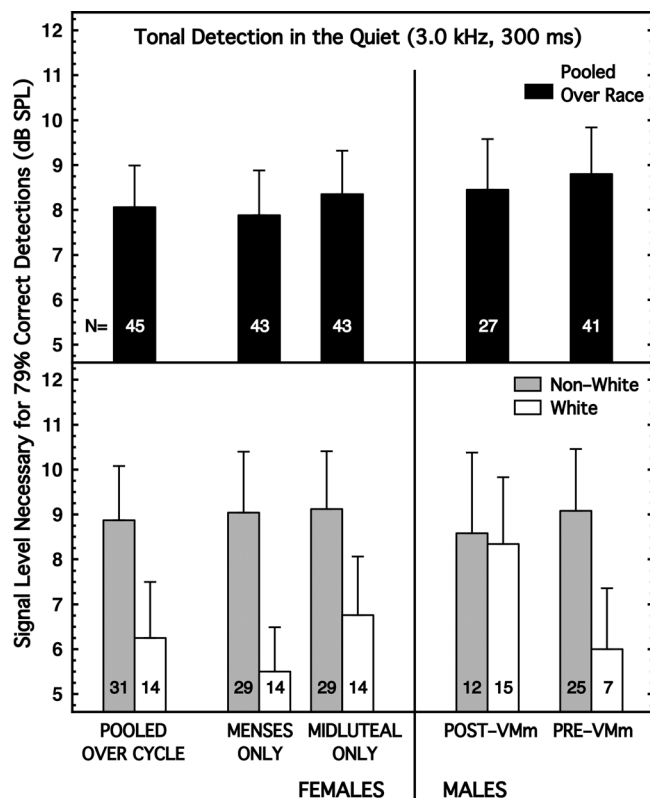


FIG. 1. Detection of a 3.0-kHz tone in the quiet (duration 300 ms). (Top) Subjects are pooled across race. (Bottom) Subjects are partitioned by race. For all figures shown, the error bars denote the standard error of the mean and footnote 2 is relevant. Discrepancies between the N 's for the pooled and partitioned conditions are attributable to the absence of responses to the questionnaire items on race/ethnicity or to anomalies in the data collected for individual subjects.

Contrary to the past literature, the Non-White females were *less* sensitive than the White females (Fig. 1, bottom panel), and the males exhibited no race difference. Because more-pigmented people have had better hearing sensitivity than less-pigmented people in so many reports (e.g., Agrawal *et al.*, 2008; Lin *et al.*, 2012), we presume the contradictory results here are attributable to the specific racial/ethnic makeup of our samples (probably to the heterogeneity of the Non-White group) compared to past samples.

Also contrary to the expectation from past literature, White females were *more* sensitive during menses than during the midluteal phase, although the difference was only about 1.0 dB, and the effect size was small and not significant [Table I, column (1)]. For Non-White females, the cycle difference was negligible, but the directionality was in accord with past reports (menses less sensitive). (Recall that, within race, these are the same subjects, just tested at different times in their cycles.)

We note that the Pre- and Post-VMm males in the Non-White category were quite similar in sensitivity (after the 8.1-dB correction because of the malfunction of the VM), but not so for the White males. We suspect that this difference also is attributable to the mix of individual racial/ethnic differences in our samples.

Two of our other psychoacoustical tasks (forward masking and two-tone suppression), required that we also measure

detectability in the quiet for a 3.0-kHz tone, but using a short duration (20 ms, including 5-ms rise/decay times). We know of no past survey results using short signals, so we had no strong expectations. No results are shown for this measure because there were no sex or menstrual-cycle differences whether or not the subjects were pooled or partitioned by race. For the females, the race differences were in the same direction as for the 300-ms signal, but the effect sizes were about half those shown (Fig. 1, bottom panel). Apparently, whatever mechanisms are responsible for (the small) sex and race differences in absolute sensitivity they require signal durations longer than 20 ms. Even with this duration effect, the correlation between absolute sensitivity measured with the long and short signals was +0.85 or higher for all subject groups (implied significance <0.001 from resampling), evidence of high reliability in our measures.

B. Overshoot

The detectability of a brief signal can differ considerably depending upon when it is presented during the time course of a gated noise masker. Consider a wideband masking noise of about 300 ms duration. When a 10-ms tonal signal is presented soon after the onset of the gated noise, it can be 10–20 dB less detectable than when presented later in the burst of noise. Here, this difference in detectability is called overshoot (e.g., Zwicker, 1965; Bacon, 2004; Walsh *et al.*, 2010; McFadden *et al.*, 2010); other labels used for this difference are temporal decline of masking and signal enhancement (Wright, 1996b). Overshoot is known to be greatest when the signal is high in frequency and shorter than about 10 ms in duration, and when the spectrum level of the masker is only moderately intense (about 20–25 dB/Hz; Bacon, 2004).

Several lines of evidence link overshoot to the mechanisms associated with cochlear amplification, making the correlations between overshoot and OAEs interesting. For example, when the cochlear-amplifier mechanism is inactivated, overshoot is diminished or abolished (Champlin and McFadden, 1989; McFadden and Champlin, 1990; Hicks and Bacon, 1999). Overshoot shrinks because sensitivity for the short-delay signal paradoxically *improves* with hearing loss. In addition, Wright (1994) found a significant sex difference in overshoot, with 20 female subjects having more overshoot than 20 male subjects.

We measured overshoot using a masking noise having a duration of 300 ms, a bandwidth of 0.1 to 6.0 kHz, and a spectrum level of approximately 17 dB/Hz (about 55 dB overall) (the masker was 8.1 dB stronger for the Pre-VMm males). The signal was a tone at 3.0 kHz with duration of 10 ms (5-ms rise/decay times, no steady-state segment). The signal was presented either 2 or 225 ms after noise onset (short and long delay, respectively), and the decibel difference in detectability for those two conditions (short minus long) was our measure of overshoot.

As with detection in the quiet, when subjects were pooled across race, the sex difference in overshoot was negligible; the difference in detectability between signals having short and long delays was about 11 dB for both females and males (Fig. 2, top panel). However, when subjects were

TABLE I. Effect sizes for various pairwise comparisons, shown separately for seven psychoacoustical tasks.

Comparison	Numerator for calculation	Race	Effect size						
			(1) Detection in quiet	(2) Overshoot	(3) Auditory Filter (ERB)	(4) Forward Masking	(5) Greenwood Effect	(6) Profile Analysis	(7) Two-tone Suppression
Sex difference	Females ignoring cycle minus	Pooled	-0.07	-0.07	-0.34	0.16	0.84^d	0.67^c	-0.31
		Non-White	0.04	0.13	-0.35	-0.10	0.60	0.75^b	-0.34
	post-VMm ^a males	White	-0.40	-0.34	-0.40	0.48	1.19^d	0.45	-0.47
		Pooled	-0.09	-0.23	-0.29	0.18	0.88^c	0.69^c	-0.40
	post-VMm males	Non-White	0.06	-0.02	-0.33	0.02	0.53	0.87^c	-0.51
		White	-0.58	-0.51	-0.27	0.41	1.53^e	0.24	-0.19
	Midluteal minus post-VMm males	Pooled	-0.02	-0.20	-0.31	0.23	0.76^d	0.63^c	-0.40
Non-White		0.08	-0.06	-0.22	0.02	0.51	0.76^b	-0.44	
	White	-0.30	-0.49	-0.52	0.54	1.14^d	0.40	-0.42	
Race difference	Non-White minus White	ALL females, ignoring cycle	0.43	0.30	0.22	-0.05	-0.16	0.41	0.39
		Menses females	0.56	0.33	0.07	0.11	-0.53	0.76^b	-0.21
	Midluteal females	0.37	0.12	0.51	0.04	-0.24	0.38	0.06	
	Males, post-VMm	0.04	-0.17	0.13	0.43	0.26	0.00	0.25	
	Males, pre-VMm	0.49	-0.04	0.46	-0.96^b	0.07	-0.52	0.27	
Cycle difference	Menses minus midluteal	Pooled	-0.05	0.06	0.06	-0.03	0.09	0.08	0.03
		Non-White	0.02	0.09	-0.04	0.00	0.00	0.21	-0.04
		White	-0.29	-0.05	0.22	-0.11	0.31	-0.20	0.20

^aVMm = malfunction of the voltmeter.

^b0.05 < *p* < 0.10 (all implied significance from resampling).

^c0.01 < *p* < 0.05.

^d0.001 < *p* < 0.01.

^e0.0001 < *p* < 0.001.

partitioned by race, an interaction between sex and race differences did emerge (Fig. 2, bottom panel). Non-White females generally exhibited about 1 dB *more* overshoot than Non-White males whereas White females generally had 1–2 dB *less* overshoot than White males. These differences across race were evident in the signs of the effect sizes for sex difference, but none of those effect sizes achieved implied significance [Table I, column (2)].

The direction of the sex difference reported by Wright (1994) was in accord only with the result here for Non-White subjects; more overshoot for females. Overshoot is known to be strongly level-dependent (Bacon, 2004), and because of the malfunction of the VM, the spectrum level of our masker was below the optimum range. So there is a temptation to attribute this reversal of the anticipated sex difference to that factor (somehow). However, our Pre-VMm males [who were tested with our intended (higher) spectrum level (25 dB/Hz)] did exhibit *smaller* overshoot than our Post-VMm males (tested with a spectrum level of 17 dB/Hz). The Wright (1994) masker also was 25 dB/Hz. It is logically possible that the two sexes do not respond exactly the same to changes in masker level so that the direction of the difference in overshoot between the sexes reverses as a function of level. But more likely, the difference in direction of effect for the Wright study and this one is attributable to one or more of the procedural, training, culling, or subject differences between the two studies.

No matter how the subjects were sorted by race, there were no significant differences in overshoot magnitude by menstrual cycle [Table I, column (2)].

Within sex, there was some evidence for a race difference for the females but not for the males [Table I, column (2)].

C. Bandwidth of the auditory filter (ERB)

The procedure of Patterson and Moore (1986) was used to determine the bandwidth of the auditory filter, specifically the equivalent rectangular bandwidth (ERB). The masker consisted of two simultaneous, equal-level bands of noise, each 2000 Hz in width, non-overlapping in frequency, and 340 ms in duration. The signal was a 3.0-kHz tone of 300 ms duration. Signal onset followed masker onset by 20 ms, and both signal and masker had 20-ms rise/decay times. On some blocks of trials, the two masker bands were spectrally contiguous, producing a continuous spectrum from 1.0 to 5.0 kHz. On other blocks of trials, the two bands were shifted in opposite directions such that they were separated by a notch centered at 3.0 kHz. On different blocks, the width of this notch was about 600, 900, 1200, or 1800 Hz (20%, 30%, 40%, or 60% of the center frequency). Because the masker bands were not filtered but synthesized with infinitely steep frequency slopes, the levels in the notches were at the noise floor of our electronics. For all notch widths the spectrum levels of the two masking bands were constant at approximately 17 dB/Hz (about 53 dB overall, and 8.1 dB higher for the Pre-VMm males).

The measure of interest was the level of the signal necessary for 79% correct detections for each of the notch bandwidths. When the notch width was small, that signal level

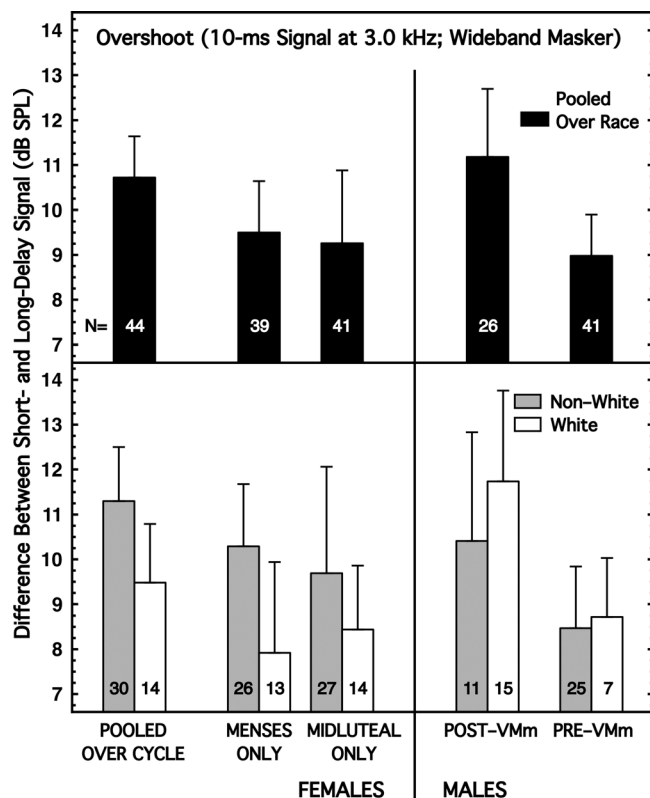


FIG. 2. Overshoot. The masker was a wideband noise, 300 ms in duration and spectrum level of 17 dB/Hz (25 dB/Hz for Pre-VMm males). The signal was a 3.0-kHz tone of 10 ms duration and presented either 2 or 225 ms after the onset of the masker. Decibel difference in performance for the two delays (short minus long) is the magnitude of overshoot. (Top) Subjects are pooled across race. (Bottom) Subjects are partitioned by race.

should be similar to the level with no notch, and as the notch width increased, the signal level should approach that for detection in the quiet. [Patterson and Moore \(1986\)](#) suggested fitting the obtained data with a function called Roex (for rounded exponential) as a way to extract an estimate of the bandwidth of the auditory filter being used to detect the 3.0-kHz tone. We implemented their Roex function with a LabVIEW[®] program written by E.G.P., typically fitting five data points per subject. A small number of subjects were excluded from the analyses because the fits to their data were poor. We extracted three measures of auditory-filter bandwidth: -3 dB, -10 dB, and the ERB. Within groups, the correlations for the various pairwise comparisons of these measures typically were 1.0, so only the ERB values are presented here. The results are summarized in Fig. 3 and Table I, column (3).

Similar to the results for detection in the quiet and overshoot, there were no evident sex (or menstrual-cycle) differences (Fig. 3, top panel) when the subjects were pooled across race, but when the subjects were partitioned by race (Fig. 3, bottom panel), sex differences emerged. Specifically, the ERBs were 25–30 Hz narrower for females than for males, and this was true for both Non-White and White subjects. The effect sizes for sex difference generally were somewhat larger for the Whites than the Non-Whites, but none achieved implied significance [Table I, column (3)].

Within race, there was no convincing evidence of menstrual-cycle differences. Note, however, that all female

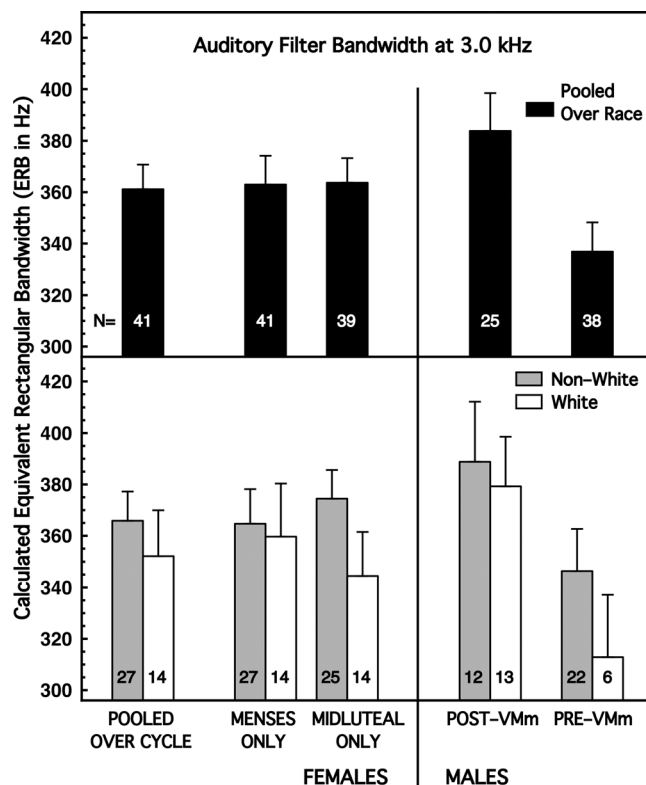


FIG. 3. Bandwidth of the auditory filter (ERB) determined using a pair of noise bands adjusted to have different frequency separations (notch widths) on different blocks of trials. Noise was 17 dB/Hz (25 dB/Hz for Pre-VMm males). Signal was a 3.0-kHz tone; both signal and noise were 300 ms, presented simultaneously. (Top) Subjects are pooled across race. (Bottom) Subjects are partitioned by race.

groups exhibited slightly wider (more male-like) ERBs during menses, when estrogen and progesterone levels were lowest, than during the midluteal phase.

Within sex, the ERBs for the White subjects always were smaller than those for the Non-White subjects, but those race differences generally were small, and none achieved implied significance [Table I, column (3)].

[Patterson and Moore \(1986\)](#) proposed a rough generalization that the average ERB is about 11% of the signal frequency in young adults, and [Wright \(1996a\)](#) reported an average ERB of about 13% for a 2.0-kHz signal. The ERB values here were in the range of 11%–13% (about 330–380 Hz) across groups even though they were determined with weaker-than-typical masking noises.

D. Tone-on-tone forward masking

Most of the masking tasks studied involved the simultaneous presentation of masker and signal. Two of the tasks were different because they involved non-simultaneous masking—forward masking and two-tone suppression. One reason for being interested in temporal masking was that [Hicks and Bacon \(1999\)](#) had reported covariations both between auditory-filter width and forward masking and between auditory-filter width and two-tone suppression, and [Moore et al. \(1999\)](#) also had reported the former. Also, some aspects of temporal masking have been linked to the strength of the cochlear amplifiers (reviewed by [Bacon, 2004](#)).

To measure forward masking, one condition of listening involved a tonal masker and a brief tonal signal. The masker was a 3.0-kHz tone of 42 dB SPL (re 20 micropascal), with a duration of 300 ms (including 5-ms rise/decay times). The signal also was 3.0 kHz with a duration of 20 ms (including 5-ms rise/decay). The signal was presented 5 ms following the complete offset of the masker. In addition, the detectability of the 20-ms signal was measured in the quiet. The measure of forward masking was the decibel difference between detection in the quiet and detection following the tonal masker.

Once again, when subjects were pooled over race, there was a small sex difference of about 1 dB for tone-on-tone forward masking (Fig. 4, top panel), with females exhibiting more forward masking than males. That effect size [Table I, column (4)] was small and non-significant. When the subjects were partitioned by race, however, the sex difference was about 1 dB for the Non-Whites and about 2 dB for the Whites (Fig. 4, bottom panel), with both groups of females exhibiting greater forward masking than the males. The effect sizes for sex difference all were larger for Whites than for Non-Whites, but none achieved implied significance [Table I, column (4)].

Within sex, none of the race differences of interest achieved implied significance for forward masking [Table I, column (4)].

Whether partitioned by race or not, there was no evidence for a difference by menstrual cycle for forward masking [Table I, column (4)].

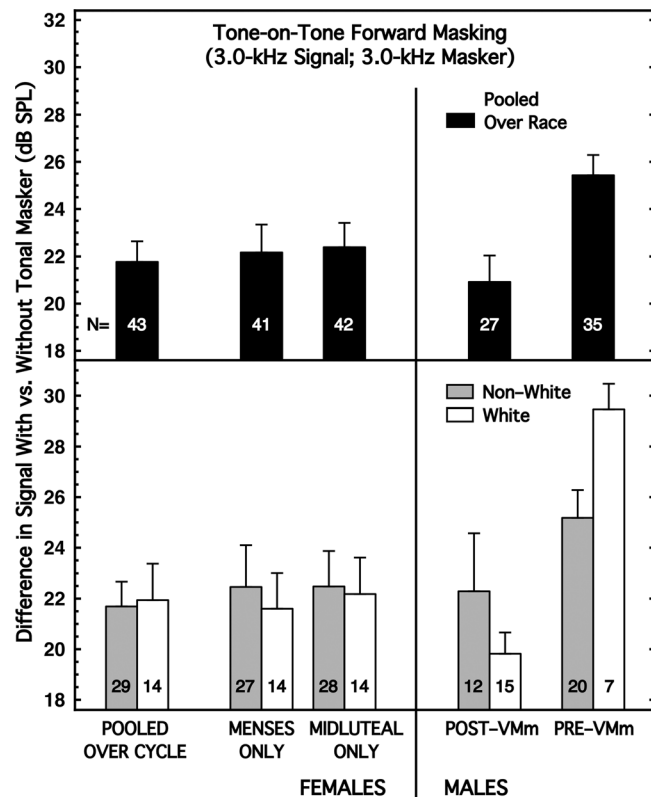


FIG. 4. Forward masking. Masker was a 3.0-kHz tone of 300 ms duration and 42 dB SPL (50 dB SPL for Pre-VMm males). The signal was the same frequency but 20 ms in duration and presented 5 ms after masker offset. Decibel difference in detection with and without tonal masker is amount of forward masking. (Top) Subjects are pooled across race. (Bottom) Subjects are partitioned by race.

E. Tone-on-tone masking with a distortion product: Greenwood task

When a tonal signal and a simultaneous tonal masker have the appropriate frequency ratio, a distortion product can be generated within the cochlea at the frequency of $2f_{\text{lower}} - f_{\text{higher}}$. This distortion product, or combination tone (CT), lies below the masking pattern of the tonal masker and thus can be an unintended cue for detection of the signal. Greenwood (1971) demonstrated the strong contribution of the CT to detection by adding a narrowband noise (NBN) as a second masker centered at the frequency of the CT. That second masker worsened performance by as much as 15–20 dB. We have called the decibel difference in performance with and without the second masker the Greenwood effect. Although we already have reported on the sex and race differences in the Greenwood task (McFadden *et al.*, 2012b), we repeat some of those results here for completeness and for consistency of presentation. Also, the data for the Pre-VMm males have not been reported previously.

Our tonal masker was 3.0 kHz, 62 dB SPL, and continuously present. The signal was a 3.6-kHz tone of 300 ms duration (including 20-ms rise/decay times). Thus, the CT was at 2.4 kHz. On some blocks of trials, a second masker also was presented; it was a continuous NBN 800 Hz in width centered at 2.4 kHz and fixed in level within a block (17 or 27 dB spectrum level on different blocks; approximately 46 and 56 dB overall level, respectively). The patterns of results for the 17- and 27-dB conditions were highly similar, although the magnitudes of the various differences generally were greater for the 27-dB condition. The 17- and 27-dB Greenwood conditions were positively correlated with each other (0.74–0.95) for all groups of subjects (implied significance $0.001 < p < 0.01$). Accordingly, we show the results only for the 27-dB condition.

The Greenwood task fell into the second category of results; namely, sex differences with subjects pooled over race, but a larger sex difference for one of the race groups. The top panel of Fig. 5 shows the magnitude of the sex difference for the Greenwood task when the subjects were pooled across race. The Greenwood effect was about 4 dB smaller for the males pooled across race than for the females pooled across race. As column (5) of Table I reveals, the effect size for this difference was 0.84, and its implied significance was about 0.004.

When the subjects were partitioned by race, the sex differences remained for both races (Fig. 5, bottom panel), but they were substantially larger for the White subjects than for the Non-White subjects [Table I, column (5)]. Two factors contributed to this difference across race; the White males had a smaller Greenwood effect than the Non-White males, and the White females had a larger Greenwood effect than the Non-White females (Fig. 5, bottom panel). The effect sizes for sex difference for the White subjects were the largest observed in this study [Table I, column (5)], and all were highly significant.

Within sex, the race differences were small and non-significant [Table I, column (5)].

When the subjects were partitioned by race, a small, and non-significant, difference by menstrual cycle was evident for the White females only [Table I, column (5)].

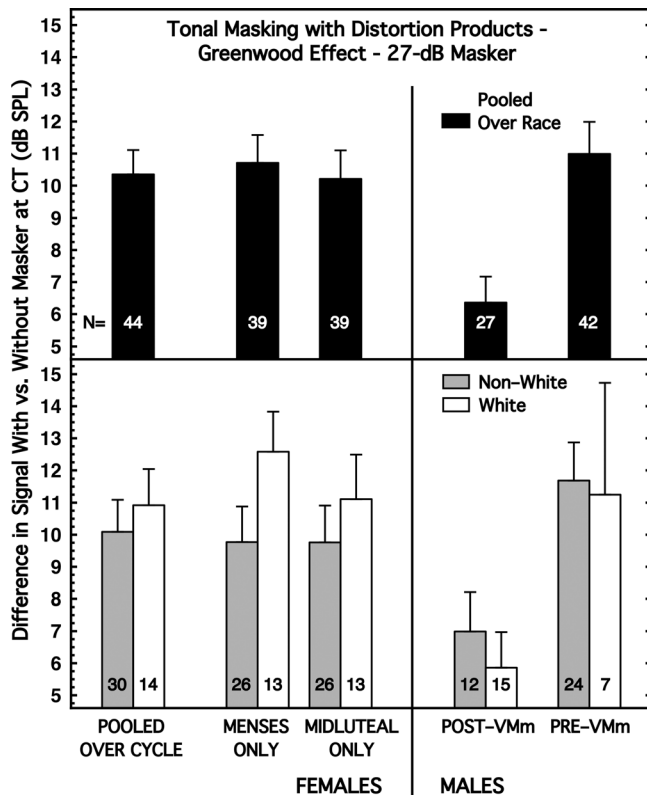


FIG. 5. Greenwood task. The signal was a 3.6-kHz tone of 300 ms duration. One masker was a 3.0-kHz tone of 62 dB SPL (70 dB SPL for Pre-VMm males). The second masker was a narrowband of noise centered over the $2f_{\text{lower}} - f_{\text{higher}}$ CT at 2.4 kHz and having a spectrum level of 27 dB/Hz (35 dB/Hz for Pre-VMm males). Both maskers were continuously present. Decibel difference in performance with and without the narrowband masker is the Greenwood effect. (Top) Subjects are pooled across race. (Bottom) Subjects are partitioned by race.

The Greenwood measure is a difference between detection with and without the narrowband masker located at the $2f_{\text{lower}} - f_{\text{higher}}$ frequency. Because the sex differences for the Greenwood task were so large, there is value in examining the origins of those difference scores. Ignoring race, with no narrowband masker present, females were about 2 dB better than males; with the narrowband masker present, males were about 3 dB better than females. That is, the large sex difference is attributable to a smaller change in detection for the males than the females when the narrowband masker was added. The implication is that the males were less sensitive to the low-frequency cue and thus suffered less when it was masked.

F. Uncertain complex masker (profile analysis)

Green (1988) and his colleagues demonstrated that perception of tonal signals is considerably more difficult when the spectral makeup of the masker is uncertain than when the masker is well-known. The effect came to be called profile analysis, a form of informational masking. Neff *et al.* (1996) studied a version of this effect and demonstrated that males were 7–8 dB better than females on their task, an effect size of about 0.73, perhaps the largest psychoacoustical sex difference previously in the literature. Neff *et al.* (1996) used a tonal signal of 1.0 kHz, 200 ms in duration. Their masker

was a 10-tone complex having an overall level of 60 dB SPL and a duration of 200 ms. The frequency components of their masker ranged from 0.3 to 3.0 kHz and were different for each observation interval of each trial.

Our profile task was similar to that of Neff *et al.* (1996). The masker consisting of 10 equal-level tones of 300 ms duration; the signal was a 3.0-kHz tone of 300 ms duration; both of those durations included 5 ms of rise/decay times; and masker and signal were gated simultaneously. The specific frequencies of the 10 masker tones were selected pseudorandomly from a master set of 199 tones created to be equally spaced in log frequency over the range of 0.5 to 6.0 kHz. Ten tones were sampled from that set for every observation interval, with two constraints: no tones could be between 2.8 and 3.2 kHz (the region of the signal) and at least 3 but no more than 7 of the 10 tones could lie above 3.2 kHz. The starting phase of each of the ten selected tones was random within and across observation intervals. In accord with Neff *et al.* (1996), the unmasked signal was presented on every trial during the warning interval preceding the two observation intervals; this cue was presented at a level that was easily detectable for all subjects. On different blocks of trials, each of the 10 tones in the masker had a level of either 17 or 42 dB (8.1 dB higher for the Pre-VMm males). The patterns of results for the 17- and 42-dB conditions were highly similar—although the magnitudes of the various differences generally were greater for the 42-dB condition. The 17- and 42-dB profile conditions were positively correlated with each other (0.74–0.88) for all groups of subjects (implied significance $0.001 < p < 0.01$). Accordingly, we show the results only for the 42-dB condition.

Prior to collecting data with the stimulus parameters described, subjects received several blocks of practice using 100-ms signals temporally centered in 300-ms maskers (100 ms of forward and backward fringe). This experience appeared to speed learning of the profile task.

Profile analysis was like the Greenwood task in that it exhibited a significant sex difference when the subjects were pooled across race, but an even larger sex difference for one of the race groups. The top panel of Fig. 6 shows that when the males were pooled across race they were about 7 dB more sensitive at profile analysis than were the females pooled across race; that effect size was about 0.67 and had an implied significance of about 0.02 [Table I, column (6)]. This direction and magnitude of effect was similar to the results of Neff *et al.* (1996) who used a lower signal frequency.

When the subjects were partitioned by race, the sex differences remained for both races (Fig. 6, bottom panel), but they were substantially larger for the Non-White subjects than for the White subjects [Table I, column (6)]. The reason was that the Non-White females were 4–8 dB less sensitive than the White females and the Post-VMm males did not differ by race. As a result, the effect sizes for sex difference achieved implied significance for the Non-White subjects but not for the White subjects. This was the only psychoacoustical task showing this direction of effect.

Within sex, the race differences were highly variable and only one achieved (marginal) significance [Table I, column (6)].

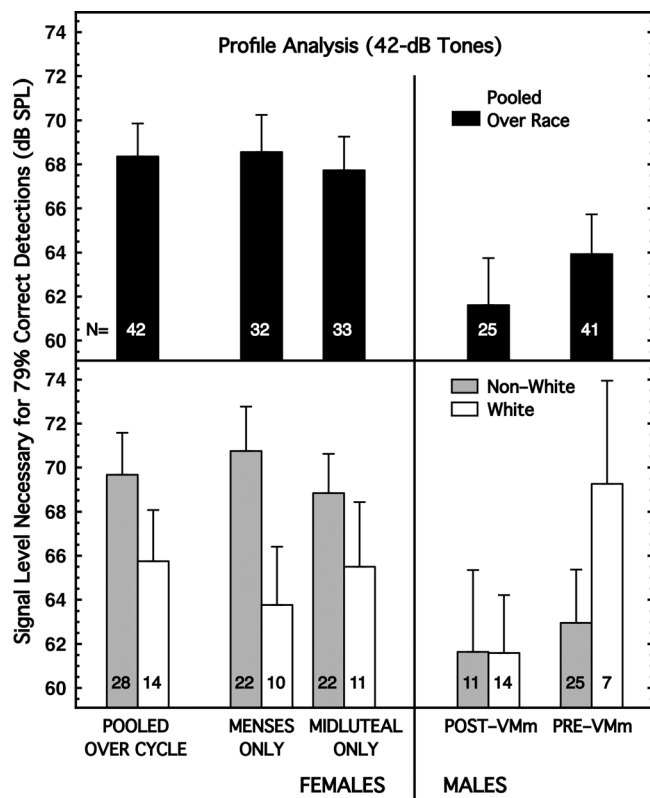


FIG. 6. Complex, uncertain masker (profile analysis). The masker was 10 tones pseudorandomly selected for each observation interval over the range of 0.5 to 6.0 kHz; each tone was 42 dB SPL in level (50 dB SPL for Pre-VMm males). The signal was a 3.0-kHz tone of 300 ms duration, simultaneously present with the masker. (Top) Subjects are pooled across race. (Bottom) Subjects are partitioned by race.

For the Non-White females, there was a small effect of the menstrual cycle, but it did not approach implied significance [Table I, column (6)].

A note in passing about the rate of growth of informational masking: The 25-dB difference in our two CT maskers led to a difference in required signal level of about 30 dB for females (pooled over race and cycle) and about 28 dB for Post-VMm males (pooled over race). For the Pre-VMm males (tested with more intense stimuli), that difference was about 22 dB. We are not aware of other studies on this point, but the work in Alexander and Lutfi (2004) is related.

G. Two-tone suppression

The term suppression has been used to denote several different auditory effects, often studied with different procedures. Here the interest is in what commonly is called two-tone suppression (Shannon, 1976). Historically, it has been compared to some versions of lateral suppression studied in vision.

To measure two-tone suppression, two forward-masking conditions were compared. In the first condition, the masker was a single tone of the same frequency as the tonal signal (one of our forward-masking conditions). In the second condition, the masker was a two-tone complex in which one tone was again at the frequency of the signal and the other was higher in frequency and more intense than the first.

Counter-intuitively, the two-masker condition typically leads to *less* masking than the single-masker condition (Shannon, 1976). The now-standard explanation is that the second, higher frequency masker is producing a suppression of the on-frequency masker tone, thereby causing it to be less effective as a forward masker. The greatest suppression occurs when the higher frequency masker is 1.15 to 1.2 times the lower frequency masker (Shannon, 1976).

There were several reasons for being interested in two-tone suppression for this study. First, Wright (1994) reported significantly less suppression in 20 female subjects than in 20 male subjects. Also, Hicks and Bacon (1999) found that the magnitudes of certain nonlinear effects measured psychophysically were positively correlated with two-tone suppression. Mills (1982) reported that exposure to moderately intense sounds produced a decline in two-tone suppression before NIHL was measurable. Taken together, the data suggested that two-tone suppression could be linked to some of the mechanisms that also are responsible for OAEs, and if so, the correlations between the individual differences in two-tone suppression and the individual differences in some forms of OAEs ought to be high (that issue is examined in McFadden *et al.*, 2018).

We measured two-tone suppression using one tonal masker at 3.0 kHz and 42 dB SPL, and a second tonal masker (the suppressor) at 3.45 kHz and 52 dB SPL (the levels were 8.1 dB higher for the Pre-VMm males). The durations of both maskers were 300 ms (including 5-ms rise/decay times). The signal was a tone of 3.0 kHz with a duration of 20 ms (5-ms rise/decay times), and a delay between masker offset and signal onset of 5 ms. The measure of two-tone suppression was the decibel difference in performance with and without the higher-frequency masker present (the latter being one of the conditions for our forward-masking task).

The two-tone-suppression task fell into the third category of effect. There was a small sex difference when the subjects were pooled across race (less suppression in females than in males), and that small sex difference persisted for both races when the subjects were partitioned by race. There was no interaction between sex and race differences. The data are shown in Fig. 7. Less suppression in females is the same direction of effect as reported by Wright (1994), but the difference here was less than 1 dB and never achieved implied significance [Table I, column (7)].

Within sex, the race differences were negligible and did not achieve implied significance [Table I, column (7)].

When pooled across race, there was no difference by phase of the menstrual cycle, and when the subjects were partitioned by race, the cycle differences remained small, and non-significant, for both races [Table I, column (7)].

Wright (1996b) reported that only about half of her 82 subjects exhibited suppression with the stimulus parameters she used; the others showed additional masking (negative suppression). A few of our subjects also showed additional masking when the suppressor tone was present, but they comprised less than 13% of our subjects, with no obvious differences across subject group. Also, the magnitude of the additional masking was less than 1.0 dB in the majority of cases. The racial mix of Wright's subjects was not reported.

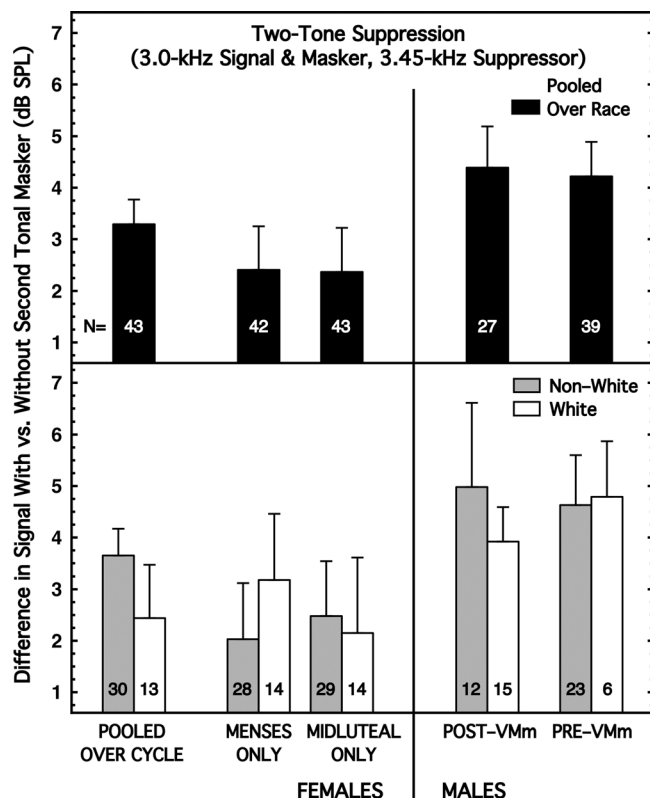


FIG. 7. Two-tone suppression. One masker was a 3.0-kHz tone of 42 SPL and the second masker was a 3.45-kHz tone of 52 SPL (both levels were 8 dB higher for Pre-VMm males); both were 300 ms in duration. The signal was a 3.0-kHz tone of 20 ms duration presented 5 ms after masker offset (forward masking). Decibel difference in detection with and without the second masker is the amount of two-tone suppression. (Top) Subjects are pooled across race. (Bottom) Subjects are partitioned by race.

H. Correlations between psychoacoustical tasks

Correlations between the various psychoacoustical tasks are shown in Table II, separately for females (bold font, top right, above the negative diagonal) and males (bottom left, below the diagonal). For the females, the data used were pooled across the menstrual cycle. The results of resampling are indicated by the superscripts.

In the spirit of providing the reader with some assistance when trying to encode the patterns of results in Table II, we offer the following general observations. (1) More pairings of tasks achieved implied significance for the males than for the females. (2) The outcomes for the two sexes were most consistent for the pairings involving detection in the quiet. That is, for both sexes, detection in the quiet was highly correlated with forward masking, width of the auditory filter, and the Greenwood effect, and weakly correlated with overshoot, two-tone suppression, and profile analysis. (3) There were a number of pairings of tasks that achieved implied significance for both sexes when the subjects were pooled across race, but not when they were partitioned by race. (4) There were some additional, individual pairings that achieved implied significance for only one sex for only one race group. (5) A number of pairings of tasks exhibited small correlations for all combinations of sex and race. We suggest that these five categories should represent a hierarchy of interest for future research.

Over the years a number of studies have examined the correlations between pairs of different auditory tasks (e.g., Elliott *et al.*, 1966; Festen and Plomp, 1981; Dreschler and Plomp, 1985; Van Rooij and Plomp, 1990; Kidd *et al.*, 2007). In most cases, the psychophysical tasks tested were chosen because of their known or apparent relationship to speech perception, which was not the interest here. Like here, Moore *et al.* (1999) reported a high correlation between the width of the auditory filter and a measure of forward masking (see Table II). Hicks and Bacon (1999) reported covariation between auditory filter width, two-tone suppression (not observed here), and growth of forward masking. Neumann *et al.* (1997) compared auditory filter widths and an OAE-based filter width extracted from transient-evoked OAEs, and got a correlation of 0.48, but the OAE-based filter widths evidenced some peculiarities. Unlike here, Wright (1996b) found strong correlations between overshoot and two-tone suppression.

The profile-analysis task is unusual in our collection of tasks in that the masking involved is informational rather than energetic (Alexander and Lutfi, 2004). That is, it is more of a perceptual/cognitive task than a sensory one. Even so, profile analysis often was highly correlated with other (energetic-masking) tasks.

IV. SUMMARY AND DISCUSSION

A large-scale, long-term study was implemented to examine the relationships between OAEs, AEPs, and seven common psychoacoustical tasks. Reported here were the sex and menstrual-cycle differences observed for those psychoacoustical tasks, along with their unanticipated interaction with race/ethnicity. Elsewhere we report the correlations between psychoacoustical performance and OAEs (McFadden *et al.*, 2018) or AEPs (report in preparation).

We did replicate several previous reports of sex differences in psychoacoustical performance, and we have added to the lists of tasks exhibiting and not exhibiting sex differences. Those replications are complicated by the interaction of sex and race differences. When we did replicate past sex differences, it was primarily for the White subjects (a similar pattern was observed in McFadden *et al.*, 2012a, 2012b). In historical perspective, this is not particularly surprising; most past studies in the United States surely involved primarily White subjects because, historically, that was the makeup of the student bodies of the large universities in this country, where most of the past studies were conducted. That sampling bias was not intentional; investigators studied who was available to them, samples of convenience, just as in this study. Besides, past investigators had little reason to believe that race/ethnicity would be relevant to the measurements they intended to make.

The interactions between sex and race in this study can be organized as follows: (1) four tasks exhibited little or no sex difference when the subjects were pooled over race, but when subjects were partitioned by race a moderate sex difference did exist for the White group. Those four tasks were: detection in the quiet at 3.0 kHz, overshoot, forward masking, and bandwidth of the auditory filter at 3.0 kHz. (2) Two

TABLE II. Correlations between psychoacoustical tasks (females^a above the negative diagonal; Post-VMm males below the diagonal).

Task	Race	Detection in quiet	Overshoot	Width of Aud. filter	Forward Masking	Greenwood	Profile Analysis	Two-tone Suppression
Detection in quiet	Pooled	—	-0.02	0.62^d	-0.67^d	-0.46^c	0.16	-0.14
	Non-White	—	-0.06	0.51	-0.63^c	-0.51	0.20	-0.25
	White	—	0.02	0.82^e	-0.86^c	-0.23	-0.16	-0.04
Overshoot	Pooled	-0.32	—	-0.09	0.23	-0.12	0.49^c	0.10
	Non-White	-0.33	—	-0.12	0.31	-0.02	0.56^b	0.18
	White	-0.30	—	-0.10	0.04	-0.43	0.22	-0.15
Width of auditory filter (ERB)	Pooled	0.74 ^c	-0.33	—	-0.46^c	-0.35	-0.06	0.11
	Non-White	0.79 ^c	-0.23	—	-0.32	-0.38	0.08	-0.03
	White	0.68 ^d	-0.41	—	-0.70^d	-0.30	-0.43	0.26
Forward masking	Pooled	-0.60 ^d	0.39 ^b	-0.74 ^e	—	0.38^b	0.16	0.24
	Non-White	-0.74 ^c	0.44	-0.82 ^d	—	0.46	0.26	0.23
	White	-0.51 ^c	0.46 ^b	-0.80 ^d	—	0.18	-0.04	0.28
Greenwood	Pooled	-0.44 ^c	0.40 ^b	-0.31	0.58 ^d	—	-0.08	0.24
	Non-White	-0.59	0.64 ^b	-0.50	0.78 ^c	—	-0.07	0.43
	White	-0.34	0.24	-0.12	0.33	—	-0.02	-0.12
Profile analysis	Pooled	-0.10	0.52 ^c	-0.19	0.40 ^b	0.33	—	-0.00
	Non-White	0.16	0.32	-0.03	0.31	0.28	—	0.28
	White	-0.35	0.74 ^d	-0.44	0.68 ^d	0.38	—	-0.62^e
Two-tone suppression	Pooled	0.16	-0.21	0.29	-0.37 ^b	-0.12	-0.07	—
	Non-White	0.10	-0.37	0.21	-0.39	-0.35	-0.06	—
	White	0.31	0.02	0.49 ^b	-0.49 ^b	0.19	-0.11	—

^aFemales were pooled over menstrual cycle for these correlations.

^b0.05 < *p* < 0.10 (all implied significance from resampling).

^c0.01 < *p* < 0.05.

^d0.001 < *p* < 0.01.

^e0.0001 < *p* < 0.001.

tasks did exhibit moderate-to-large sex differences when the subjects were pooled over race, but when the subjects were partitioned by race, the sex difference was greater for one race group than the other. Those tasks were: the Greenwood effect (larger for the White subjects) and profile analysis (larger for the Non-White subjects). Those two tasks also exhibited the largest effect sizes for sex of all the tasks studied. (3) One task exhibited a small sex difference whether the subjects were pooled over, or partitioned by, race: two-tone suppression. It is not clear why these seemingly similar tasks varied in this way, but apparently slightly different underlying mechanisms were involved across tasks. Future investigators interested in the physiological underpinnings of sex or race differences in the auditory system should consider the pattern of effect sizes in Table I when selecting psychoacoustical tasks to study.

The strengths of this study include: larger *N*'s and far more extensively trained subjects than in many past studies, multiple psychoacoustical tasks measured on the same subjects over a limited time span, the use of adaptive forced-choice methods to acquire the data, rigorous culling of atypical blocks of trials, statistical significance estimated using resampling techniques, precise definitions of the menstrual phases, and measurements of OAEs and AEPs made for all subjects over that same time span. In addition, we are able to report correlations both between psychoacoustical tasks (here) and between the physiological and psychoacoustical measures (McFadden *et al.*, 2018). Given the large number of pairwise comparisons made here, the resampling

procedures used provided protection against false-positive outcomes. Because subjects were extensively trained, we are confident that performance was essentially asymptotic for all subjects. As indirect evidence of the reliability of our measurements, correlations between the two Greenwood conditions with different masker levels and between the two profile-analysis conditions with different masker levels were in the range of 0.75–0.95 across tasks and groups. (Only one Greenwood condition and one profile-analysis condition were reported here to conserve space.)

Although the initial goals of this long-term study were somewhat compromised by the discovery of marked race/ethnicity differences, we hope our experience will help raise awareness in auditory science about this potentially relevant variable. Because this study was not designed initially to study race/ethnicity, we were not able to make some obvious comparisons that might have been possible had we recruited subjects with race in mind. So, many uncertainties remain. However, there is little doubt that pooling subjects over race did obscure underlying differences for some of our psychoacoustical tasks. The decibel differences admittedly were small, but the existence of race differences does deserve note. For one reason, the interactions between sex and race revealed here have the potential to explain past and future failures to replicate previously published outcomes. Failure to replicate could be simply a matter of not having similar mixes of race/ethnicity in the subjects tested. Indeed, any attempt to replicate the *present* findings is likely to succeed in finding small race differences for some tasks, but is likely

to fail to obtain the specific results reported here. The reason is both of our race groups were highly heterogeneous, and we are unable to provide the precise information that would permit a detailed replication. We regret that our partitioning could not have been more rigorous. We remind the reader that because subjects were not recruited with race in mind, all conclusions about possible race effects must be viewed as tentative.

We also attempted to replicate some past reports of psychoacoustical performance being affected by the *menstrual cycle* (cf. [McFadden, 1998, 2000](#)). All we found were negligible or small differences, at least one of which was opposite in direction from previously published reports. That said, the effect sizes for cycle typically were largest for the White subjects, although even the largest effect sizes for cycle difference were only about 0.2–0.3 (Table I). Perhaps these small effects are attributable in part to the frequency region tested, the extensive training our subjects received, and/or some other parameter of this study. We believe that our procedure for determining phase of the cycle was superior to that in many studies. We collected data continually through more than one cycle and defined the two phases after the fact for each subject individually, using her daily diary.

Correlations between the various psychoacoustical tasks generally were weak and not consistent across the sex and race groups (Table II). Large correlations were most common when detection in the quiet was one of the tasks, and large correlations were more common for males than females. The most consistent correlations were between detection in the quiet and forward masking, detection in the quiet and width of the auditory filter, and width of the auditory filter and forward masking.

The correlations between psychoacoustical tasks and OAEs (reported in [McFadden et al., 2018](#)) were generally weak for these subjects, suggesting that the individual differences existing in our psychoacoustical tasks are more attributable to post-cochlear (neural) mechanisms than to cochlear (“mechanical”) mechanisms. Nevertheless, there may be value in mentioning cochlear mechanisms of possible relevance to race differences in human physiological and psychoacoustical measurements. [In this regard, note that we also observed an interaction between sex and race for an AEP-derived measure of cochlear length ([McFadden et al., 2012a](#)).]

It is logically possible, perhaps even likely, that the same cochlear mechanism(s) involved in producing the sex and ear differences in OAEs also are involved in producing the interactions with race. “All” that need be speculated is that (for some reason) some genetic, congenital, or hormonal process operates on the cochlear amplifiers slightly differently across race. Unfortunately, we do not know the specific mechanism, but the sex and ear differences in the OAEs of newborns suggest that the cochlear amplifiers are weakened in males prenatally, presumably by the action of androgens (reviewed by [McFadden, 2002, 2008, 2011](#)). For race differences, a possible role for melanin also needs to be considered ([McFadden and Wightman, 1983](#); [Lin et al., 2012](#)). In addition to their presence in the epidermis, melanocytes are located at multiple sites in the cochlea ([LaFerriere et al.,](#)

[1974](#)), and there is a line of evidence suggesting that they play an important role in cochlear function. For example, people with darker skin or eyes have more SOAEs ([Russell, 1992](#); [Whitehead et al., 1993](#); [McFadden and Loehlin, 1995](#)), better hearing sensitivity (although not here), less NIHL, and less tinnitus than people with lighter skin ([Agrawal et al., 2008](#); [Shargorodsky et al., 2010](#); [Lin et al., 2012](#)). Also, humans with albinism have more NIHL than people with darker skin ([Garber et al., 1982](#)). If we assume that the activity of melanocytes in the cochlea is positively correlated with the melanocyte activity in the skin (in accord with [Bonaccorsi, 1965](#); but compare [Bartels et al., 2001](#)), then these facts about skin color suggest that melanocytes are somehow involved in the functioning of the cochlear-amplifier mechanism and in the protection of the cochlea from intense sounds (and/or in its recovery after such exposure). An additional fact is that many drugs that are ototoxic bind to the cochlear melanocytes ([Lindquist, 1973](#)), further suggesting that having functioning melanocytes is crucial to cochlear homeostasis.

The authors believe that the various race differences that have been reported for auditory physiology are in fact attributable to the individual differences that exist in degree of pigmentation inside the cochlea. Although the standard race categories establish groups of people that are ordered by degree of pigmentation, that ordering obviously is rough. Within categories, the variation in pigmentation can be large, meaning that the standard race categories are at best rough substitutes for the relevant characteristic: individual pigment concentration. [Lin et al. \(2012\)](#) provided evidence for this belief; they partitioned the subjects of a single race/ethnic category (Hispanic) into two groups based on skin color and found that the darker-skinned subjects had significantly better hearing sensitivity than the lighter-skinned subjects. That is, the heterogeneity of the overall category was simplified by further partitioning of the subjects by degree of pigmentation. Logic and intuition suggest that degree of pigmentation rightly should be viewed as a difference across individuals, not a difference across groups. When it comes to the auditory system, the standard categories of race are at best poor proxies for the relevant factor(s).

A related point is that auditory race differences inevitably will come to be less relevant with the passage of time. The obvious reason is that humans now are mating across racial/ethnic lines at rates far higher than in the past ([Livingston and Brown, 2017](#)), so distinctions based on ancestry will become less meaningful with time. However, individual differences in pigment concentration will remain, and should continue to affect audition if only slightly.

Some sex, race, and other group differences are the natural result of the wide array of individual differences that exist in humans because of the complex processes and interactions that occur during meiosis and after fertilization. Individual differences are the atoms of human diversity, and group differences, such as sex and race differences, are the molecules. Most branches of medicine now tacitly recognize the importance of individual differences when explaining the origins of various disorders and the varying effectiveness of certain treatments. At each stage of advancing knowledge,

people who are similar on a particular medical dimension will, of necessity, be grouped together and treatment will be tailored to that group. With the passage of time, additional knowledge will allow the partitioning of those groups into yet smaller groups, with different explanations about etiology whether or not there are differences in diagnosis or treatment. Ultimately, treatment for many disorders will be based on highly specific factors present in unique collections of individual people. Although the specific race (and sex) differences reported here are unlikely ever to be of importance medically, understanding the mechanisms underlying certain other race differences may prove valuable. Namely, knowledge of the mechanisms underlying the race differences in NIHL and hearing sensitivity (Agrawal *et al.*, 2008) might lead to schemes for activating those mechanisms on demand to help protect people from cochlear insults.

The sex differences in OAEs that are seen in adults also exist in newborns (e.g., Strickland *et al.*, 1985), so the differences in adults cannot be attributed solely to the result of post-natal events or life experiences. By comparison, the adult sex differences reported here for psychoacoustical tasks might be attributable to some combination of subtle differences in biology/physiology at some stage of life, and differential life experiences by the two sexes or race groups. Some auditory tasks do show complex patterns of acquisition (Huyck and Wright, 2013), and these patterns might differ across sex or race. Recall that, for each of the psychoacoustical tasks described here, the subjects received considerable practice and then atypical blocks were culled before estimates of sensitivity were calculated. The intent was to reduce if not eliminate any pre-experimental differences in experience, but there is no way to know how effective that attempt was. If it was effective, then the sex and race differences reported here do have some inherent biological/physiological underpinning. If not, then the differences observed “only” reflect brain differences acquired through life—some what less interesting, but not completely uninteresting.

There exist no obvious everyday issues or problems that would be solved or better addressed by additional detailed knowledge about the mechanisms underlying the small race, sex, or menstrual-cycle differences reported here. Knowing more about the underlying mechanisms surely would be satisfying to pure scientists interested in the fundamental mechanisms of hearing, but practical advances based on that knowledge seem remote indeed. Accordingly, the differences reported here are likely to be primarily of historic interest. That said, we have done our best to report how race differences affected the outcomes in our study. Because this lab now is closed due to retirements, we must leave any future investigations to others.

ACKNOWLEDGMENTS

This work was supported by a research grant awarded to D.M. by the NIDCD (Grant No. 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 and J. T. Tran during data

collection. We thank J. Smurzynski and an anonymous reviewer for many helpful comments. With the publication of this article and its companion, D.M. has published in this journal since 1966.

¹The National Institute on Deafness and Other Communication Disorders (NIDCD) required that these two items be asked of all subjects. Item 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)?” Item 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).” All subjects answering “White” on item 2 were categorized as such, and all subjects choosing *any* other category on item 2 were categorized as Non-White, no matter what their response to question 1 (the *N*'s did not permit any finer categorizations). Two subjects answered yes on item 1, and did not answer item 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). Subjects failing to answer both items were excluded from our analyses. Note how inadequate these items are for a systematic study of race effects; as one example, the category “Asian” pools people from China and Japan with people from India and Pakistan. For the females in this study, approximately 45% self-identified as Asian, 8% as Black, 36% as White, and 11% as “More Than One.” For the males, approximately 27% self-identified as Asian, 9% as Black, 2% as Pacific Islander, 46% as White, and 16% as “More Than One.”

²For some psychoacoustical tasks, the results for the Pre- and Post-VMm groups were somewhat different. For all such cases, the reader is reminded that the acoustic stimuli used for the Pre-VMm males were 8.1 dB stronger than those for the Post-VMm males; also, there were only 7 White males in the Pre-VMm group. All females were tested after the malfunction of the VM.

- Agrawal, Y., Platz, E. A., and Niparko, J. K. (2008). “Prevalence of hearing loss and differences by demographic characteristics among US adults,” *Arch. Intern. Med.* **168**, 1522–1530.
- Alexander, J. M., and Lutfi, R. A. (2004). “Informational masking in hearing-impaired and normal-hearing listeners: Sensation level and decision weights,” *J. Acoust. Soc. Am.* **116**, 2234–2247.
- Bacon, S. P. (2004). “Overview of auditory compression,” in *Compression: From Cochlea to Cochlear Implants*, edited by S. P. Bacon, R. R. Fay, and A. N. Popper (Springer, New York), pp. 1–17.
- Bartels, S., Ito, K., Trune, D. R., and Nuttall, A. L. (2001). “Noise-induced hearing loss: The effect of melanin in the stria vascularis,” *Hear. Res.* **154**, 116–123.
- Bass, A. H., Sisneros, J. A., Popper, A. N., and Fay, R. R. (eds.) (2016). *Hearing and Hormones* (Springer, New York), 209 pp.
- Bell, A. (1992). “Circadian and menstrual rhythms in frequency variations of spontaneous otoacoustic emissions from human ears,” *Hear. Res.* **58**, 91–100.
- Bilger, R., Matthies, M. L., Hammel, D. R., and Demorest, M. E. (1990). “Genetic implications of gender differences in the prevalence of spontaneous otoacoustic emissions,” *J. Speech Hear. Res.* **33**, 418–432.
- Bonaccorsi, P. (1965). “Il colore dell’iride come ‘test’ di valutazione quantitativa nell’uomo, della concentrazione di melanina nella stria vascolare” (“The color of the iris as a ‘test’ in the quantitative estimation, in man, of the melanin concentration in the stria vascularis”), *Ann. Laringol. Otol. Rinol. Faringol.* **64**, 725–738.
- Burns, E. M. (2017). “Even-longer-term stability of spontaneous otoacoustic emissions,” *J. Acoust. Soc. Am.* **142**, 1828–1831.
- Burns, E. M., Arehart, K. H., and Campbell, S. L. (1992). “Prevalence of spontaneous otoacoustic emissions in neonates,” *J. Acoust. Soc. Am.* **91**, 1571–1575.
- Champlin, C. A., and McFadden, D. (1989). “Reductions in overshoot following intense sound exposures,” *J. Acoust. Soc. Am.* **85**, 2005–2011.
- Chung, D. Y., Mason, K., Gannon, R. P., and Willson, G. N. (1983). “The ear effect as a function of age and hearing loss,” *J. Acoust. Soc. Am.* **73**, 1277–1282.
- Cohen, J. (1992). “A power primer,” *Psychol. Bull.* **112**, 155–159.

- DaCosta, E. A., Castro, J. C., and Macedo, M. E. G. (2008). "Iris pigmentation and susceptibility to noise-induced hearing loss," *Int. J. Audiol.* **47**, 115–118.
- Davis, M. J., and Ahroon, W. A. (1982). "Fluctuations in susceptibility to noise-induced temporary threshold shift as influenced by the menstrual cycle," *J. Aud. Res.* **22**, 173–187.
- Dreschler, W. A., and Plomp, R. (1985). "Relations between psychophysical data and speech perception for hearing-impaired subjects. II," *J. Acoust. Soc. Am.* **78**, 1261–1270.
- Elkind-Hirsch, K. E., Stoner, W. R., Stach, B. A., and Jerger, J. F. (1992). "Estrogen influences auditory brainstem responses during the normal menstrual cycle," *Hear. Res.* **60**, 143–148.
- Elliott, D. N., Riach, W. D., Sheposh, J. P., and Trahiotis, C. (1966). "Discrimination performance of high school sophomores on a battery of auditory tests," *Acta Oto-Laryngol., Suppl.* **216**, 1–59.
- Festen, J. M., and Plomp, R. (1981). "Relations between auditory functions in normal hearing," *J. Acoust. Soc. Am.* **70**, 356–369.
- Garber, S. R., Turner, C. W., Creel, D., and Witkop, C. J., Jr. (1982). "Auditory system abnormalities in human albinos," *Ear Hear.* **3**, 207–210.
- Green, D. M. (1988). *Profile Analysis: Auditory Intensity Discrimination* (Oxford University Press, New York), 138 pp.
- 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.
- Hicks, M. L., and Bacon, S. P. (1999). "Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing," *J. Acoust. Soc. Am.* **105**, 326–338.
- Huyck, J. J., and Wright, B. A. (2013). "Learning, worsening, and generalization in response to auditory perceptual training during adolescence," *J. Acoust. Soc. Am.* **134**, 1172–1182.
- Kemp, D. T. (1978). "Stimulated acoustic emissions from within the human auditory system," *J. Acoust. Soc. Am.* **64**, 1386–1391.
- Kemp, D. T. (1979). "Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea," *Arch. Otorhinolaryngol.* **224**, 37–45.
- Kidd, G. R., Watson, C. S., and Gygi, B. (2007). "Individual differences in auditory abilities," *J. Acoust. Soc. Am.* **122**, 418–435.
- LaFerriere, K. A., Arenberg, I. K., Hawkins, J. E., and Johnsson, L.-G. (1974). "Melanocytes of the vestibular labyrinth and their relationship to the microvasculature," *Ann. Otol. Rhinol. Laryngol.* **83**, 685–694.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Lin, F. R., Maas, P., Chien, 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.
- Lindquist, N. G. (1973). "Accumulation of drugs on melanin," *Acta Radiol. Diagn. (Stockh.)* **325**, 1–92.
- Livingston, G., and Brown, A. (2017). "Intermarriage in the U.S. 50 years after Loving v. Virginia," Pew Research Center report, pp. 1–35.
- McFadden, D. (1993). "A speculation about the parallel ear asymmetries and sex differences in hearing sensitivity and otoacoustic emissions," *Hear. Res.* **68**, 143–151.
- McFadden, D. (1998). "Sex differences in the auditory system," *Devel. Neuropsych.* **14**, 261–298.
- McFadden, D. (2000). "Masculinizing effects on otoacoustic emissions and auditory evoked potentials in women using oral contraceptives," *Hear. Res.* **142**, 23–33.
- McFadden, D. (2002). "Masculinization effects in the auditory system," *Arch. Sex. Behav.* **31**, 99–111.
- McFadden, D. (2008). "What do sex, twins, spotted hyenas, ADHD, and sexual orientation have in common?," *Perspect. Psychol. Sci.* **3**, 309–323.
- McFadden, D. (2011). "Sexual orientation and the auditory system," *Front. Neuroendocrinol.* **32**, 201–213.
- McFadden, D., and Champlin, C. A. (1990). "Reductions in overshoot during aspirin use," *J. Acoust. Soc. Am.* **87**, 2634–2642.
- McFadden, D., Garcia-Sierra, A., Hsieh, M. D., Maloney, M. M., Champlin, C. A., and Pasanen, E. G. (2012a). "Relationships between otoacoustic emissions and a proxy measure of cochlear length derived from the auditory brainstem response," *Hear. Res.* **289**, 63–73.
- 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. (2009a). "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. (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., Leshikar, E. M., Hsieh, M. D., and Maloney, M. M. (2012b). "Comparing behavioral and physiological measures of combination tones: Sex and race differences," *J. Acoust. Soc. Am.* **132**, 968–983.
- McFadden, D., Pasanen, E. G., Maloney, M. M., Leshikar, E. M., and Pho, M. H. (2018). "Correlations between otoacoustic emissions and performance in common psychoacoustical tasks," *J. Acoust. Soc. Am.* **143**, 2355–2367.
- McFadden, D., Pasanen, E. G., Raper, J., Lange, H. S., and Wallen, K. (2006). "Sex in otoacoustic emissions measured in rhesus monkeys (*Macaca mulatta*)," *Horm. Behav.* **50**, 274–284.
- McFadden, D., Pasanen, E. G., Valero, M. D., Roberts, E. K., and Lee, T. M. (2009b). "Effect of prenatal androgens on click-evoked otoacoustic emissions in male and female sheep (*Ovis aries*)," *Horm. Behav.* **55**, 98–105.
- McFadden, D., Walsh, K. P., Pasanen, E. G., and Grenwelge, E. M. (2010). "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," *Ann. Rev. Psych.* **34**, 95–128.
- Mills, J. H. (1982). "Effects of noise on auditory sensitivity, psychophysical tuning curves, and suppression," in *New Perspectives on Noise-Induced Hearing Loss*, edited by R. P. Hamernik, D. Henderson, and R. Salvi (Raven, New York), pp. 249–263.
- Moore, B. C. J., Vickers, D. A., Plack, C. J., and Oxenham, A. J. (1999). "Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism," *J. Acoust. Soc. Am.* **106**, 2761–2778.
- Neff, D. L., Kessler, C. J., and Dethlefs, T. M. (1996). "Sex differences in simultaneous masking with random-frequency maskers," *J. Acoust. Soc. Am.* **100**, 2547–2550.
- 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.
- Patterson, R. D., and Moore, B. C. J. (1986). "Auditory filters and excitation patterns as representations of frequency resolution," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, New York), pp. 123–178.
- Rammesayer, T. H., and Troche, S. J. (2012). "On sex-related differences in auditory and visual sensory functioning," *Arch. Sex. Behav.* **41**, 583–590.
- Royster, L. H., Royster, J. D., and Thomas, W. G. (1980). "Representative hearing levels by race and sex in North Carolina industry," *J. Acoust. Soc. Am.* **68**, 551–566.
- Russell, A. F. (1992). "Heritability of spontaneous otoacoustic emissions," Dissertation, University of Illinois at Urbana-Champaign, available at <https://www.ideals.illinois.edu/handle/2142/72384>.
- Shahnaz, N. (2008). "Transient evoked otoacoustic emissions (TEOAEs) in Caucasian and Chinese young adults," *Int. J. Audiol.* **47**, 76–83.
- Shannon, R. V. (1976). "Two-tone unmasking and suppression in a forward-masking situation," *J. Acoust. Soc. Am.* **88**, 741–744.
- Shargorodsky, J., Curhan, G. C., and Farwell, W. R. (2010). "Prevalence and characteristics of tinnitus among US adults," *Am. J. Med.* **123**, 711–718.
- Snihur, A. W. K., and Hampson, E. (2012). "Click-evoked otoacoustic emissions: Response amplitude is associated with circulating testosterone levels in men," *Behav. Neurosci.* **126**, 325–331.
- Strickland, E. A., Burns, E. M., and Tubis, A. (1985). "Incidence of spontaneous otoacoustic emissions in children and infants," *J. Acoust. Soc. Am.* **78**, 931–935.
- Swanson, S. J., and Dengerink, H. A. (1988). "Changes in pure-tone thresholds and temporary threshold shifts as a function of menstrual cycle and oral contraceptives," *J. Speech. Hear. Res.* **31**, 569–574.

- Talmadge, C. L., Long, G. R., Murphy, W. J., and Tubis, A. (1993). "New off-line method for detecting spontaneous otoacoustic emissions in human subjects," *Hear. Res.* **71**, 170–182.
- Van Rooij, J. C. G. M., and Plomp, R. (1990). "Auditive and cognitive factors in speech perception by elderly listeners," *J. Acoust. Soc. Am.* **88**, 2611–2624.
- 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.
- 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.
- Wright, B. A. (1994). "Individual, sex, and ear differences in measures of overshoot and psychophysical two-tone suppression," *J. Acoust. Soc. Am.* **95**(Suppl. 1), 2942–2943.
- Wright, B. A. (1996a). "Auditory filter asymmetry at 2000 Hz in 80 normal-hearing ears," *J. Acoust. Soc. Am.* **100**, 1717–1721.
- Wright, B. A. (1996b). "Correlated individual differences in conditions used to measure psychophysical suppression and signal enhancement," *J. Acoust. Soc. Am.* **100**, 3295–3303.
- Zündorf, I. C., Karnath, H.-O., and Lewald, J. (2011). "Male advantage in sound localization at cocktail parties," *Cortex* **47**, 741–749.
- Zwicker, E. (1965). "Temporal effects in simultaneous masking by white-noise bursts," *J. Acoust. Soc. Am.* **37**, 653–656.