# Low-level otoacoustic emissions may predict susceptibility to noise-induced hearing loss[a)]

Judi A. Lapsley Miller,[b)] Lynne Marshall, Laurie M. Heller,[c)] and Linda M. Hughes
*Naval Submarine Medical Research Laboratory, Groton, Connecticut 06349-5900*

(Received 6 October 2005; revised 11 April 2006; accepted 19 April 2006)

In a longitudinal study with 338 volunteers, audiometric thresholds and otoacoustic emissions were measured before and after 6 months of noise exposure on an aircraft carrier. While the average amplitudes of the otoacoustic emissions decreased significantly, the average audiometric thresholds did not change. Furthermore, there were no significant correlations between changes in audiometric thresholds and changes in otoacoustic emissions. Changes in transient-evoked otoacoustic emissions and distortion-product otoacoustic emissions were moderately correlated. Eighteen ears acquired permanent audiometric threshold shifts. Only one-third of those ears showed significant otoacoustic emission shifts that mirrored their permanent threshold shifts. A Bayesian analysis indicated that permanent threshold shift status following a deployment was predicted by baseline low-level or absent otoacoustic emissions. The best predictor was transient-evoked otoacoustic emission amplitude in the 4-kHz half-octave frequency band, with risk increasing more than sixfold from approximately 3% to 20% as the emission amplitude decreased. It is possible that the otoacoustic emissions indicated noise-induced changes in the inner ear, undetected by audiometric tests. Otoacoustic emissions may therefore be a diagnostic predictor for noise-induced-hearing-loss risk. [DOI: 10.1121/1.2204437]

PACS number(s): 43.64.Jb, 43.64.Wn [BLM] Pages: 280–296

## I. INTRODUCTION

Evoked otoacoustic emissions (OAEs) are sounds produced by the inner ear in response to acoustic stimulation (Kemp, 1978). These sounds can be measured in the ear canal with a low-noise microphone. OAEs are thought to be generated by the outer hair cells (OHCs), which are susceptible to noise damage (e.g., Liberman *et al.*, 1986; Nordmann *et al.*, 2000; Rask-Andersen *et al.*, 2000). Diminished OAE amplitudes may be an early warning sign of incipient noise-induced hearing loss (NIHL), and therefore they may have a role to play in hearing-conservation programs.

Cross-sectional studies have shown that OAEs are sensitive indicators of permanent noise-induced damage to the inner ear in groups of noise-exposed people, with lower OAE levels associated with higher audiometric thresholds (e.g., LePage and Murray, 1993; LePage *et al.*, 1993; LePage and Murray, 1998; Desai *et al.*, 1999; Mansfied *et al.*, 1999; Attias *et al.*, 2001). Furthermore, noise-exposed people tend to have lower OAEs than people with similar audiometric thresholds but no noise exposure (Bicciolo *et al.*, 1993; LePage and Murray, 1993; LePage *et al.*, 1993; Murray and LePage 1993; Attias *et al.*, 1995, 1998; Xu *et al.*, 1998; Desai *et al.*, 1999; Attias *et al.*, 2001).[1] This finding has led to the hypothesis that, in individuals, OAEs may decrease prior to changes in audiometric thresholds. However, a cross-sectional design means the purported progression of changes in OAEs due to noise exposure and the relationship to changes in audiometric thresholds cannot easily be demonstrated in individual ears.

Longitudinal studies have shown that permanent changes in OAEs and permanent changes in audiometric thresholds do not necessarily occur together, both for groups of noise-exposed people and for noise-exposed individuals (Engdahl *et al.*, 1996; Murray *et al.*, 1998; Murray and LePage, 2002; Lapsley Miller *et al.*, 2004; Konopka *et al.*, 2005; Seixas *et al.*, 2005a, b). Typically, group changes in OAEs are seen, but often there are no group changes in audiometric thresholds. Studies that have considered changes in individual ears have also found that permanent threshold shifts (PTSs) do not necessarily correlate with changes in OAEs (Murray *et al.*, 1998; Murray and LePage, 2002; Lapsley Miller *et al.*, 2004). The actual progression of OAE changes and hearing loss in individuals has not been documented to date, and the existing data are ambiguous and inconsistent, partly for methodological reasons. These reasons include (a) noise exposures that were not severe enough to permanently elevate audiometric thresholds; (b) study durations that were not long enough to measure slowly progressing hearing loss; (c) OAE stimulus levels that were too high to optimally detect OAE changes (Sutton *et al.*, 1994; Marshall and Heller, 1998; Marshall *et al.*, 2001); (d) difficulty getting volunteers who had not been recently exposed to noise [i.e., baseline measurements were contaminated by temporary threshold shifts (TTS)]; (e) difficulty getting an appropriate age-matched and sex-matched control group; (f) getting volunteers without previous noise exposure; (g) separating out the effects of aging and NIHL; and (h) achieving

[a)]Preliminary results have been reported at the Association for Research in Otolaryngology Midwinter Meeting, St. Petersburg Beach, FL, January 2002, and the International Military Noise Conference, Baltimore, MD, April 2001.
[b)]Electronic mail: judi@psychophysics.org
[c)]Now at the Department of Cognitive and Linguistic Sciences, Brown University, Providence, RI 02912.

sufficiently low test-retest variability (especially in field settings) to enable small changes in audiometric thresholds and OAEs to be detected.

It is unclear whether diminished OAEs are predictive of eventual NIHL, especially within an individual. The ideal study would follow a large number of volunteers, many of whom would eventually get PTS, over a period of years. Many subjects are needed because the incidence of NIHL is low in any one year, even in severely noise-exposed populations. To date, no one has amassed enough PTS cases to identify the best predictors. This question can be addressed in a more limited way by testing two points in time (before and after a particular noise exposure) to determine whether those with low amplitude OAEs on the preexposure test are more at risk for NIHL as measured postexposure. While this is not as desirable as a long-term multi-measurement longitudinal study because it does not provide information about why the OAE is at a low level, such a study can provide some information about which OAE parameters and properties seem to be the most predictive of PTS. In most populations, NIHL is a gradual process, and age can be a confounding factor, so studying this issue is more easily accomplished in a young population exposed to high levels of noise. One such population is the crew of an aircraft carrier.

An aircraft carrier, especially during flight operations, is one of the noisiest working environments known (Yankaskas, 1999; Yankaskas and Shaw, 1999). This environment puts sailors at risk for NIHL because even when using hearing protection as recommended, noise dosages still can exceed risk limits.[2] Naval hearing-conservation regulations mandate single hearing protection when noise levels exceed 84 dBA or impulse noise exceeds 140 dB pSPL, and double hearing protection (earplugs plus muffs or cranial helmets) when levels exceed 104 dBA (Navy Occupational Health and Safety Program, 1999). Double hearing protection ideally can provide attenuation up to 30 dB, but cannot provide sufficient attenuation to remove the risk of NIHL in the extreme noise levels present on an aircraft carrier. Furthermore, unlike many noise-hazardous industrial environments, there may be no truly quiet time for ears to recover from these shipboard exposures. This also implies that damage-risk criteria, which assume a daily quiet recovery time (Passchier-Vermeer, 1993), may not apply to this population. Poor hearing protection usage (Bjorn *et al.*, 2005), coupled with very high noise levels and with little quiet time for recovery, means sailors on aircraft carriers are at high risk for noise-induced hearing loss.

The main aim of the present study was to assess changes in audiometric thresholds and OAEs in sailors after 6 months of hazardous noise exposure on an aircraft carrier. A modified test battery was used, based on earlier studies (Sutton *et al.*, 1994; Kummer *et al.*, 1998; Marshall and Heller, 1998; Lapsley Miller *et al.*, 2004), which indicated that lower-level OAE stimuli were more sensitive to NIHL. The hypotheses were that (a) group average audiometric thresholds would increase (worsen), and group average OAE amplitudes would decrease (worsen); (b) individual cases of noise-induced PTS would be associated with significant emission shifts (SESs), but there would be more sailors with SESs than PTSs, and (c) ears with low-level or absent OAEs at predeployment testing would be more likely to show PTS at postdeployment testing.

## II. MATERIALS AND METHODS

### A. Volunteers

Audiometric thresholds and OAEs were measured in 338 sailors (35 women, median age 22 years, range 18 to 46 years; 303 men, median age 22 years, range 18 to 41 years) before and after 9 months on a Nimitz-class aircraft carrier, including 6 months at sea. Approximately 47% of the sailors were from the Air Department, who worked around aircraft and their launch and recovery mechanisms on or below the flight deck, as well as in the hangar bays; 19% were from the Engineering Department, who worked in various locations below deck; 32% were from the Reactor Department, who worked in the machinery spaces; and 2% were from other departments. Additionally, a control group of 28 volunteers (sailors and research staff; 8 women, median age 31 years, range 20 to 53 years; 20 men, median age 26 years, range 20 to 47 years) completed an identical protocol with no intervening noise exposure between pre- and posttesting. The posttest for the control group occurred 20 min to 2 days after the predeployment testing. A suitable age- and sex-matched control group that could be noise-free over 9 months was not available.

### B. Stimuli and equipment

Pure-tone audiograms were obtained at frequencies 0.5, 1, 2, 3, 4, and 6 kHz using a modified Hughson-Westlake procedure (with the usual 10-dB descending and 5-dB ascending steps). Four microprocessor-controlled audiometers were used (three Tremetrics RA400 and one RA500), all with TDH 39 earphones and MX-41/AR cushions, and one Beltone 120 manual audiometer, with TDH 50P earphones and MX-41/AR cushions. For the most part, the Tremetrics audiometers were used in automated mode. Middle-ear pressures were estimated from the peak of an immitance tympanogram with a 226-Hz tone using a Grason Stadler GSI 33 version 2 analyzer at a sweep speed of 12.5 daPa/s to minimize hysteresis.

OAEs were measured with the ILO292 Echoport system (Otodynamics Ltd., England), using the distortion-product OAE (DPOAE) probe. It was covered by an acoustic-immitance probe tip, which had been enlarged using a grinding tool, to allow better placement and manipulation in the ear canal.

### C. OAE test battery

Transient-evoked OAEs (TEOAEs) evoked with a 74 dB pSPL click (abbreviated herein to $TEOAE_{74}$) were measured in nonlinear mode, where responses to three clicks at one polarity and one click 9.5 dB higher with opposite polarity were added together to reduce linear artifact from the stimulus (Bray, 1989). TEOAEs were collected and averaged until 260 low-noise averages were obtained.[3] The results were windowed, filtered, and analyzed into half-octave bands [which is optimal according to Marshall and Heller (1996)]. At predeployment testing, every attempt was made to get a flat stimulus spectrum during calibration. At postdeployment testing, every attempt was made to get the same stimulus pattern during calibration as in predeployment testing.

In order of presentation, DPOAEs were measured with stimulus levels $L_1/L_2$=57/45, 59/50, 61/55, and 65/45 dB SPL (abbreviated herein to $DP_{57/45}$, $DP_{59/50}$, $DP_{61/55}$, and $DP_{65/45}$). The first three levels specified a DPOAE I/O function (Kummer *et al.*, 1998); the fourth level is sensitive to TTS (Marshall *et al.*, 2001). For all stimulus levels, the $f_2/f_1$ ratio was 1.22, with $f_2$=1.8, 2.0, 2.2, 2.5, 2.8, 3.2, 3.6, 4.0, and 4.5 kHz.[4]

Individual in-the-ear calibration was used for both TEOAEs and DPOAEs.

### D. Procedure

All OAE and audiometric testing occurred in single-walled sound-attenuating booths. OAE testing was done pierside, near the ship, in a mobile test van. Most audiometric testing was done in the medical department on the ship, which was docked at the pier, but some testing was done in the mobile van to expedite testing as many volunteers as possible. The left ear was tested first. At predeployment testing, volunteers were screened for clear ear canals (cerumen was removed if present), audiometric thresholds of ≤25 dB HL from 0.5 to 3 kHz and ≤30 dB HL at 4 kHz, and peak immitance within the range of ±50 daPa atmospheric pressure, with grossly normal amplitude, slope, and smoothness of the tympanogram. 85% of the ears had normal audiometric thresholds, using a strict criterion of ≤15 dB HL at 1 to 4 kHz. If the definition of normal is relaxed to include thresholds at 20 dB HL (which is often used for hearing screening using OAEs, e.g., Gorga *et al.*, 1993), 98% of the ears had normal thresholds. There was a greater incidence of slight hearing losses at higher frequencies. At 1 kHz, 98% had thresholds ≤15 dB HL and 2% had 20 dB HL thresholds. At 2 kHz, 97% had thresholds ≤15 dB HL and 2% had 20 dB HL thresholds. At 3 kHz, 94% had thresholds ≤15 dB HL, 5% had 20 dB HL thresholds, and 1% had 25 dB HL thresholds. At 4 kHz, 90% had thresholds ≤15 dB HL, 7% had 20 dB HL thresholds, 1% had 25 dB HL thresholds, and 1% had 30 dB HL thresholds. These percentages were similar for the group that got PTS during the deployment and the group that did not. Volunteers who did not meet screening criteria did not continue in the study.

At postdeployment testing, volunteers were asked to complete a detailed noise history covering the previous 9 months. They then underwent the same testing as for predeployment testing. At that time, they were screened only for peak immitance within ±50 daPa atmospheric pressure, and in all cases the tympanometric peak was within this range shortly before OAE testing.

During postdeployment data collection, the Navy hearing-conservation significant-threshold-shift (STS) criteria at that time of the study (a shift of at least 15 dB at 1, 2, 3, or 4 kHz, or an average shift of at least 10 dB at 2, 3, and 4 kHz) were used to detect changes in audiometric thresholds in individuals (Navy Occupational Health and Safety Program, 1999). STSs were confirmed with manual audiometry, immediately if possible, or as soon as possible thereafter (up to 9 days). If the volunteer had been noise-free and the STS was confirmed, it was considered a PTS. If the volunteer had recently been exposed to noise, they were asked to return for a 14-h noise-free follow-up to see if their STS was permanent or temporary.

### E. Data definitions, cleaning, and reduction

Because the testing was conducted in a military working environment, some of the data were unavoidably affected by background ambient and electrical noise, despite testing in a sound-attenuating booth calibrated to ANSI standards (ANSI, 1991) and using a power-line conditioner. Data-collection logistics meant that much of the OAE testing was done using a hook-up to the naval base's mains power supply rather than batteries. Once the problem was identified, the equipment was run on battery as much as possible. The short testing time available for each volunteer meant that it was not always possible to obtain clean data. Data points and/or test conditions contaminated with extreme stimulus levels, bad calibrations, high noise levels, large differences in noise level between tests, or many unexplained outliers were removed from the data set in an objective fashion, using the same elimination rules across the entire dataset of all volunteers.[5]

An OAE was considered present if, for $TEOAE_{74}$, the amplitude was greater than 0 dB SNR above the noise level, and, for DPOAEs, the amplitude was greater than the noise level, which was defined as two standard deviations above the noise floor.

For the remaining "good" frequencies and levels, the percentage of measurable OAEs was calculated (i.e., those OAE amplitudes with good SNR) and any frequencies where less than 70% of OAEs were measurable were dropped ($TEOAE_{74}$ at 0.7 and 5.7 kHz).

Although the actual criteria used were liberal at each stage of screening, a large amount of data was rendered unusable. Losing $DP_{61/55}$ and the two DPOAE frequencies (see footnote 5) meant that planned analyses involving DPOAE input-output functions and half-octave analyzed DPOAEs had to be dropped. The remaining test conditions were $TEOAE_{74}$, which was analyzed into half-octave bands centered at 1.0, 1.4, 2.0, 2.8, and 4.0 kHz, and $DP_{65/45}$, $DP_{59/50}$, and $DP_{57/45}$ at 1.8, 2.0, 2.5, 2.8, 3.2, 3.6, and 4.0 kHz.

For some cases, a predeployment OAE was measurable, but the postdeployment OAE was below the noise level. These postdeployment OAE amplitudes with bad SNR were substituted with the noise level in some circumstances [similarly to Lapsley Miller *et al.* (2004)]. This occurred only if the noise level was below the predeployment OAE amplitude (otherwise, a high noise level may masquerade as an increase in OAE amplitude). This enabled the use of more data, such as the important cases where a normal OAE at predeployment testing disappeared below the noise level by postde-

TABLE I. The number of ears in each group that contributed to each analysis, listed by the section number. The number in parentheses is the total number of volunteers in the group. The numbers varied at each test frequency, OAE level, and OAE type because all good data were used. The exception was for the ANOVAs where volunteers were required to have complete OAE data sets for both ears at 2, 3, and 4 kHz.

| Analyses | No. of ears (volunteers) | Group | Notes |
|---|---|---|---|
| III A ANOVA | 150 (75) | Noise | Ear was a factor in ANOVAs. |
| III B Correlations among changes in audiometric thresholds and changes in OAEs | 169–338 (338) | Noise | The left and right ears were in separate analyses. |
| III C 1 Forming STS and SES criteria | 33–56 (28) | Control | Ears were pooled across volunteers. |
| III C 2 Applying STS and SES criteria to noise-exposed and control groups | Noise: 473–675 (338) Control: 33–56 (28) | Noise and Control | Noise and control groups were analyzed separately. Within groups, ears were pooled across volunteers. |
| III C 3–4 Identifying and describing PTS cases | 18 (15) | Noise | Both men and women were analyzed but no woman got PTS. |
| III C 4 Correlations between SES status and PTS status | PTS: 10–17 non-PTS: 473–572 | Noise | PTS and non-PTS volunteers were in the same analysis. Ears were pooled across volunteers. |
| III C 4 Correlations among SESs for PTS ears | 15 (12) | Noise | Three ears were not included due to missing data. |
| III D Susceptibility | PTS: 16–18 non-PTS: 524–559 | Noise | Only data from the male volunteers were used. Ears were pooled. |

ployment testing, with the caveat that true decreases in OAE amplitude may have been underestimated. For TEOAE$_{74}$, 6% of postdeployment measurements were replaced with the noise level. For DPOAEs, 5% of postdeployment measurements were replaced with the noise level.

For the susceptibility analyses, it was of interest to know if low or absent OAEs at predeployment increased the chance of PTS at postdeployment. Some of the analyses required estimating amplitudes for missing predeployment OAEs. To do this, the noise level was substituted for missing OAEs, providing the noise floor was not high. A noise level was considered acceptably low if it was within the tenth percentile of the corresponding OAE amplitude (not the noise floor) based on the group. Again, this process is conservative because it overestimates the actual amplitude of the OAE.

This research was conducted in compliance with all applicable federal regulations governing the protection of human subjects in research.

## III. RESULTS

Table I provides a breakdown of the number of ears contributing to each analysis, and whether the ears were noise exposed or controls. The numbers varied at each test frequency, OAE level, and OAE type because all good data were used. The exception was for the ANOVAs, where data from 75 volunteers with complete data sets for both ears were used.

### A. Group OAE and audiometric thresholds before and after noise exposure

The primary interest was to see if there were any changes in audiometric thresholds or OAEs between pre- and postdeployment tests. Of secondary interest was whether these changes differed across frequency, stimulus level (for DPOAEs), or ears. There were not enough female volunteers to group by sex.

Separate, repeated-measures ANOVAs were conducted on audiometric threshold, TEOAE, and DPOAE data for the

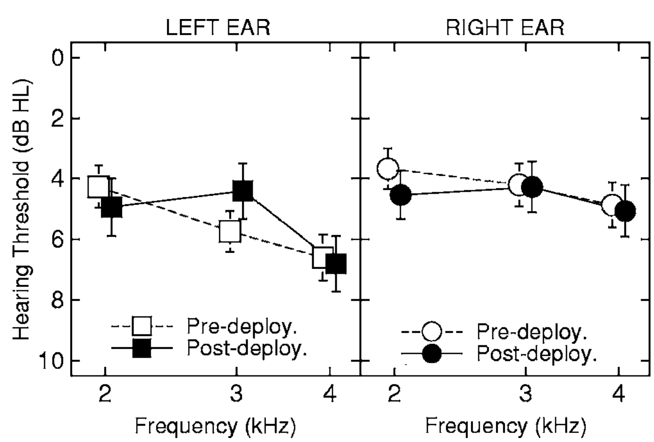


FIG. 1. Average group audiometric thresholds for left and right ears and pre- and postdeployment tests for the subgroup of 75 noise-exposed sailors with complete data sets used in the ANOVA. Error bars indicate one standard error of the mean. Frequency is plotted on a $\log_2$ scale. Data points are offset either side of the labeled frequency to aid interpretation.

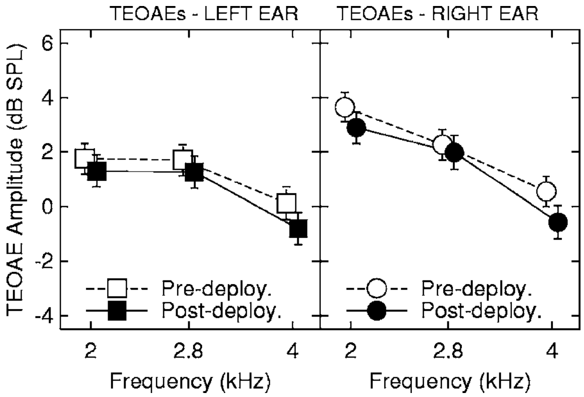


FIG. 2. Between pre- and postdeployment, average group TEOAE amplitudes significantly decreased by 1 dB at 4 kHz (combined over ears) for the 75 noise-exposed sailors with complete data sets used in the ANOVA. Left panel shows average group $TEOAE_{74}$ amplitudes for left ears; right panel shows average group $TEOAE_{74}$ amplitudes for right ears. Error bars indicate one standard error of the mean. Frequency is plotted on a $\log_2$ scale. Data points are offset either side of the labeled frequency to aid interpretation.

subgroup of 75 volunteers with "complete" data sets (8 women, median age 21 years; 67 men, median age 22 years). To maximize the number of volunteers with complete data, only the frequencies 2, 2.8 (or 3 for audiograms), and 4 kHz were included (i.e., there could be missing data at other frequencies). By selecting volunteers with complete data, a bias may have been introduced, because those volunteers with more missing data may have more noise-induced damage. However, by using complete data sets, comparisons across OAE stimulus types and frequencies could be more fairly made.

Figure 1 shows group average audiometric thresholds for left and right ears and pre- and postdeployment tests. A three-way, repeated-measures ANOVA was conducted for audiometric thresholds (test: pre- versus postdeployment; ear: left versus right; and frequency: 2, 3, and 4 kHz). There was no significant change in audiometric thresholds (main effect) between pre- and postdeployment tests ($F_{1,74}$=0.05, ns). There were, however, significant differences between ears ($F_{1,74}=5.82, p<0.05$) and across frequency ($F_{2,148}=3.68, p<0.05$). There was also a two-way interaction for test-by-frequency ($F_{2,148}=3.32, p<0.05$). Bonferroni *post hoc t*-test comparisons were used to establish which frequencies contributed to the test-by-frequency, two-way interaction. The familywise significance level was $p<0.05$, so, for three comparisons, $p<0.017$ was used. None were significant.

Figure 2 shows the group average TEOAE amplitudes for left and right ears and pre- and postdeployment tests. A three-way, repeated-measures ANOVA was conducted for $TEOAE_{74}$ amplitude (test: pre- versus postdeployment; ear: left versus right; and frequency: 2, 2.8, and 4 kHz). All three factors showed significant main effects. Particularly, there was a 0.66-dB decrease in $TEOAE_{74}$ amplitude between pre- and postdeployment testing ($F_{1,74}=12.3, p<0.05$). Ears also differed ($F_{1,74}=8.6, p<0.05$) as did frequency ($F_{2,148}=4.4, p<0.05$). There were two significant two-way interactions: test-by-frequency ($F_{2,148}=19.5, p<0.05$) and ear-by-frequency ($F_{2,148}=3.2, p<0.05$). Bonferroni *post hoc t*-test comparisons were used to establish which frequencies contributed to the test-by-frequency, two-way interaction. The familywise significance level was $p<0.05$, so, for three comparisons, $p<0.017$ was used. There was a significant 1.0-dB decrement in $TEOAE_{74}$ amplitude at 4 kHz.

Figure 3 shows the group average DPOAE amplitudes for each level, and pre- and postdeployment tests (ears combined). A four-way, repeated-measures ANOVA was conducted for DPOAE amplitude (test: pre- versus postdeployment; ear: left versus right, level: stimulus levels of 65/45, 59/50, and 57/45 dB SPL; and frequency: 2, 2.8, and 4 kHz). There was a 1.28-dB decrement in DPOAE amplitude between pre- and postdeployment testing ($F_{1,74}=27.4, p<0.05$). There were also main effects for level ($F_{2,148}=190.8, p<0.05$) and frequency ($F_{2,148}=24.2, p<0.05$) but not for ear ($F_{1,74}$=0.08,ns). There were three significant two-way interactions: test-by-level ($F_{2,148}=9.1, p<0.05$), ear-by-level ($F_{2,148}=10.1, p<0.05$), and level-by-frequency ($F_{4,296}=28.4, p<0.05$). Bonferroni *post hoc t*-test comparisons were used to establish which of the three levels contributed to the test-by-level, two-way interaction. The familywise significance level was $p<0.05$, so, for three comparisons, $p<0.017$ was used. Postdeployment DPOAE amplitudes for $DP_{59/50}$ and$DP_{57/45}$ were significantly lower than predeployment amplitudes (by 1.5 dB). None of the three- or four-way interactions were significant.

### B. Changes in OAEs and audiometric thresholds: Correlations

The relationship between *changes* in OAEs and *changes* in audiometric thresholds was assessed using Pearson corre-

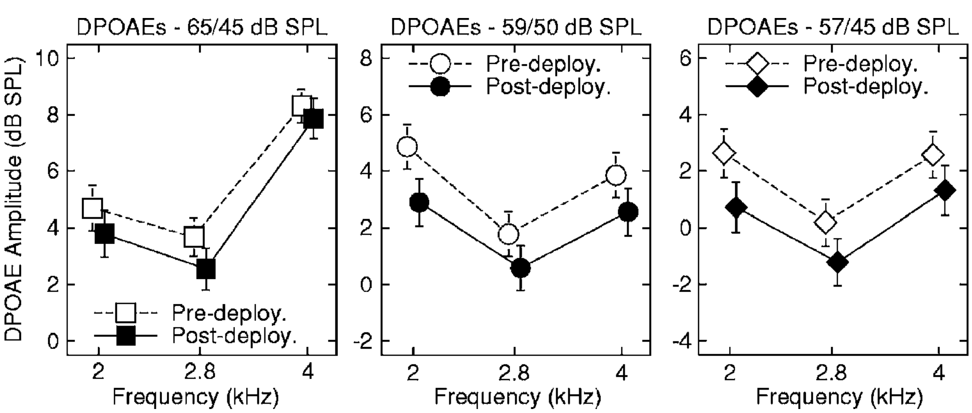


FIG. 3. Between pre- and postdeployment, average group DPOAE amplitudes for $DP_{59/50}$ and $DP_{57/45}$ significantly decreased by 1.5 dB for the 75 noise-exposed sailors with complete data sets used in the ANOVA. Left panel shows average group $DP_{65/45}$, middle panel shows average group $DP_{59/50}$, and right panel shows average group $DP_{57/45}$ amplitudes. Plots are averaged over ears. Error bars indicate one standard error of the mean. Frequency is plotted on a $\log_2$ scale. Data points are offset either side of the labeled frequency to aid interpretation.

lation coefficients for the arithmetic difference between pre- and postdeployment OAE amplitudes and audiometric thresholds, for all the valid data at every test, level, and frequency. The number of volunteers contributing to each correlation ranged from 169 to 338. Right and left ears were considered separately.

#### 1. Correlations between changes in audiometric thresholds and changes in TEOAEs

There was no correlation greater than 0.22 at any frequency, and most were not statistically significant at $p < 0.05$.

#### 2. Correlations between changes in audiometric thresholds and changes in DPOAEs

There was no correlation greater than 0.19 at any frequency or stimulus level, and most were not statistically significant at $p < 0.05$.

#### 3. Correlations between changes in DPOAEs and changes in TEOAEs

Generally, the strongest correlations were from 0.5 to 0.6 (statistically significant at $p < 0.05$), which occurred for same-frequency combinations and for $TEOAE_{74}$ frequencies 0.5 to 1 octave lower than the DPOAE frequency. This may reflect separate correlations with the two DPOAE components, with the reflection source originating from the $2f_1$-$f_2$ place and the distortion-source originating from near the $f_2$ place (both with a frequency of $2f_1$-$f_2$). Sometimes the correlation was highest when the $2f_1$-$f_2$ frequency was in the same $TEOAE_{74}$ half-octave and sometimes when the $f_2$ frequency was in the same $TEOAE_{74}$ half-octave, but the differences were not great, and many cases showed similar correlations across two or three $TEOAE_{74}$ half-octaves. Correlations between $TEOAE_{74}$ and DPOAEs showed no consistent pattern across DPOAE stimulus level. In general, $DP_{57/45}$ and $DP_{59/50}$ showed stronger correlations with $TEOAE_{74}$ than $DP_{65/45}$. The strongest correlations tended to be for $TEOAE_{74}$ at 1 or 1.4 kHz with DPOAEs at 1.8 kHz (the lowest DPOAE frequency), regardless of stimulus level.

There was no evidence that changes in OAEs were correlated with changes in audiometric thresholds. Although DPOAEs and TEOAEs did tend to shift together, the correlation was only moderate. Furthermore, there was little-to-no correlation between left and right ears for either audiometric thresholds or OAEs, even for the same test type and frequency. The lack of correlation between changes in audiometric threshold and changes in OAEs may be due to the small number of ears that actually had *significant* changes in audiometric threshold or OAEs—the larger number of non-PTS ears, where changes are just due to test-retest variability, may have swamped any effect. This was investigated further by considering the OAEs of the PTS ears.

### C. Association between changes in audiometric threshold and changes in OAEs in individuals

Note that a +STS and a −STS indicate a worsening of audiometric thresholds and OAE amplitude, respectively, whereas a −STS and a +STS indicate an improvement of audiometric thresholds and OAE amplitude, respectively. The plus or minus sign comes from subtracting the predeployment test result from the postdeployment test result.

TABLE II. Significant threshold shift (STS) criteria calculated from the control group (56 ears). Shown for each audiometric frequency and some averaged frequency combinations are the group average threshold shifts (postdeployment-predeployment), the standard error of measurement ($SE_{MEAS}$), the resulting STS criteria, and the Navy STS criteria, which was used to diagnose PTS onsite.

| Frequency (kHz) | Average shift (dB) | $\Delta SE_{MEAS}$ (dB) | STS (dB) | Navy STS (dB) |
|---|---|---|---|---|
| 1 | −1.3 | 2.8 | 15 | 15 |
| 2 | −0.2 | 2.1 | 15 | 15 |
| 3 | −1.6 | 3.4 | 15 | 15 |
| 4 | −0.2 | 3.8 | 20 | 15 |
| 6 | −1.1 | 5.5 | 25 | … |
| Mean 2 and 3 | −0.9 | 2.4 | 10 | … |
| Mean 3 and 4 | −0.5 | 2.8 | 12.5 | … |
| Mean 2, 3 and 4 | −0.5 | 2.2 | 8.3 | 10 |

#### 1. Standard error of measurement used to define individual significant shifts

Criteria based on the $SE_{MEAS}$ were used to detect significant audiometric thresholds and OAE shifts between pre- and postdeployment tests (similarly to Lapsley Miller and Marshall, 2001; Marshall *et al.* 2001, 2002; Lapsley Miller *et al.*, 2004),[6] for each frequency of interest from the group of 28 control volunteers, who had received no intervening noise exposure (see Table I; 56 ears were included for audiometric thresholds and between 29 to 43 ears for OAEs, because only OAEs with good SNR at both pre- and postdeployment were used).[7]

Tables II and III summarize $SE_{MEAS}$ and the resulting STS and SES criteria, respectively, for each frequency of interest.

#### 2. STS and SES cases identified using derived criteria

Table IV shows the percentage of STSs detected and Table V shows the percentage of SESs detected when applying the derived criteria to the data set of 338 volunteers. The percentages are relative to the amount of good data (i.e., after removing the cases with poor calibrations, etc., as described earlier). Virtually no STSs were detected in the control group, but more were detected in the noise-exposed group. Nearly as many −STSs (improvement of audiometric thresholds) were seen as +STSs (deterioration of audiometric thresholds), except for the averaged frequencies of 2 and 3 kHz, and 2, 3, and 4 kHz. This indicated that the test-retest variability was too great to reliably see noise-induced audiometric-threshold shifts at single frequencies, and it was only when a wider frequency band was examined that significant noise-induced changes were apparent. The STS criterion for 6 kHz was deemed too large (at 25 dB) to reliably detect shifts at this frequency and was therefore not used. For subsequent analyses, data for ears with STS were used only

TABLE III. $DP_{57/45}$, $DP_{59/50}$, $DP_{65/45}$, and $TEOAE_{74}$ significant emission shift (SES) criteria. Shown for each single DPOAE frequency and half-octave TEOAE band are the number of control-group ears going into the calculation, $SE_{MEAS}$, and the resulting SES criterion.

| OAE type | Frequency (kHz) | Ears | $SE_{MEAS}$ (dB) | SES (dB) |
|---|---|---|---|---|
| $DP_{57/45}$ | 1.8 | 41 | 2.3 | 6.9 |
| | 2.0 | 33 | 2.7 | 8.0 |
| | 2.5 | 39 | 1.9 | 5.7 |
| | 2.8 | 34 | 2.6 | 7.8 |
| | 3.2 | 37 | 2.1 | 6.3 |
| | 3.6 | 39 | 2.7 | 8.2 |
| | 4.0 | 38 | 1.7 | 5.2 |
| $DP_{59/50}$ | 1.8 | 40 | 2.2 | 6.5 |
| | 2.0 | 39 | 2.9 | 8.5 |
| | 2.5 | 38 | 2.3 | 7.0 |
| | 2.8 | 38 | 2.6 | 7.9 |
| | 3.2 | 36 | 2.5 | 7.5 |
| | 3.6 | 39 | 2.2 | 6.5 |
| | 4.0 | 35 | 2.0 | 6.1 |
| $DP_{65/45}$ | 1.8 | 42 | 2.0 | 6.1 |
| | 2.0 | 42 | 1.9 | 5.6 |
| | 2.5 | 39 | 1.5 | 4.6 |
| | 2.8 | 40 | 1.7 | 5.1 |
| | 3.2 | 43 | 1.8 | 5.4 |
| | 3.6 | 43 | 1.6 | 4.7 |
| | 4.0 | 42 | 1.9 | 5.7 |
| $TEOAE_{74}$ | 1.0 | 40 | 2.5 | 7.5 |
| | 1.4 | 46 | 2.0 | 6.1 |
| | 2.0 | 40 | 1.1 | 3.2 |
| | 2.8 | 39 | 1.3 | 3.8 |
| | 4.0 | 35 | 1.2 | 3.7 |

if the STS was confirmed to be PTS (with a repeat audiogram showing the STS was maintained, noise-free for at least 14 h prior to testing, and a noise history consistent with hazardous noise exposure). The data sets for ears with no STS were used for comparison with the PTS ears.

TABLE IV. Percentage of significant threshold shifts (STSs) detected with the derived criteria based on the $SE_{MEAS}$, for the noise-exposed group ($n=338$) and for the control group ($n=28$). Shown are the percentages of good data, the percentages of +STSs (deterioration in audiometric thresholds), relative to the good data, and the percentage of −STSs (improvement in audiometric thresholds), relative to the good data.

| | Noise-exposed ears | | | Control ears | | |
|---|---|---|---|---|---|---|
| Frequency (kHz) | Good data (%) | +STS (%) | −STS (%) | Good data (%) | +STS (%) | −STS (%) |
| 1.0 | 100 | 1 | 1 | 100 | 0 | 0 |
| 2.0 | 100 | 2 | 0 | 100 | 0 | 0 |
| 3.0 | 100 | 1 | 1 | 100 | 0 | 0 |
| 4.0 | 100 | 1 | 1 | 100 | 2 | 0 |
| 6.0 | 100 | 1 | 1 | 100 | 2 | 0 |
| Mean 2 and 3 | 100 | 3 | 1 | 100 | 0 | 0 |
| Mean 3 and 4 | 100 | 2 | 2 | 100 | 0 | 0 |
| Mean 2, 3, and 4 | 100 | 5 | 1 | 100 | 0 | 0 |

TABLE V. Percentage of significant emission shifts (SESs) for each OAE type detected with the derived criteria based on the $SE_{MEAS}$, for the noise-exposed group ($n=338$), and for the control group ($n=28$). Shown are the percentages of good data, the percentages of −SESs (decrease in OAE amplitude), relative to the good data, and the percentage of +SESs (increase in OAE amplitude), relative to the good data.

| | | Noise-exposed ears | | | Control ears | | |
|---|---|---|---|---|---|---|---|
| Test | Frequency (kHz) | Good data (%) | −SES (%) | +SES (%) | Good data (%) | −SES (%) | +SES (%) |
| $DP_{57/45}$ | 1.8 | 75 | 11 | 2 | 72 | 5 | 2 |
| | 2.0 | 72 | 6 | 2 | 64 | 3 | 3 |
| | 2.5 | 70 | 12 | 3 | 71 | 2 | 0 |
| | 2.8 | 73 | 5 | 2 | 62 | 3 | 6 |
| | 3.2 | 74 | 8 | 3 | 69 | 5 | 0 |
| | 3.6 | 76 | 3 | 2 | 72 | 2 | 0 |
| | 4.0 | 80 | 13 | 6 | 69 | 0 | 8 |
| $DP_{59/50}$ | 1.8 | 79 | 10 | 3 | 72 | 5 | 0 |
| | 2.0 | 79 | 4 | 2 | 74 | 0 | 2 |
| | 2.5 | 75 | 6 | 2 | 69 | 5 | 3 |
| | 2.8 | 77 | 5 | 1 | 69 | 0 | 3 |
| | 3.2 | 79 | 6 | 2 | 67 | 0 | 5 |
| | 3.6 | 80 | 7 | 4 | 69 | 5 | 3 |
| | 4.0 | 83 | 9 | 4 | 62 | 3 | 3 |
| $DP_{65/45}$ | 1.8 | 83 | 7 | 2 | 74 | 0 | 5 |
| | 2.0 | 83 | 8 | 4 | 74 | 2 | 0 |
| | 2.5 | 82 | 12 | 4 | 72 | 7 | 0 |
| | 2.8 | 83 | 11 | 3 | 69 | 0 | 3 |
| | 3.2 | 86 | 7 | 3 | 74 | 2 | 0 |
| | 3.6 | 86 | 10 | 6 | 76 | 2 | 0 |
| | 4.0 | 88 | 6 | 4 | 74 | 5 | 2 |
| $TEOAE_{74}$ | 1.0 | 85 | 5 | 1 | 75 | 5 | 2 |
| | 1.4 | 88 | 5 | 2 | 82 | 2 | 4 |
| | 2.0 | 81 | 12 | 7 | 75 | 5 | 0 |
| | 2.8 | 75 | 8 | 4 | 70 | 0 | 3 |
| | 4.0 | 71 | 12 | 3 | 68 | 0 | 5 |

OAEs, on the other hand, showed more evidence of noise-induced changes. Table V summarizes the percentage of SESs found for each OAE type, level, and frequency, for both the noise-exposed group and for the control group. Shown is the percentage of good data relative to all data (i.e., the percentage of pre- and postdeployment measurement pairs that could be used to calculate differences), the percentage of −SESs (deterioration of OAEs), and the percentage of +SESs (improvement of OAEs), both relative to the amount of good data. There was little difference between the percentages of +SESs for noise-exposed and control groups, indicating that many increments were just due to variability (even though it is theoretically possible for OAEs to increase in amplitude with noise damage). There were, however, more −SESs for the noise-exposed group, compared to the control group, indicating the effects of noise exposure on the OAEs. There does not appear to be any indication that higher frequencies showed more OAE changes, as might have been expected.

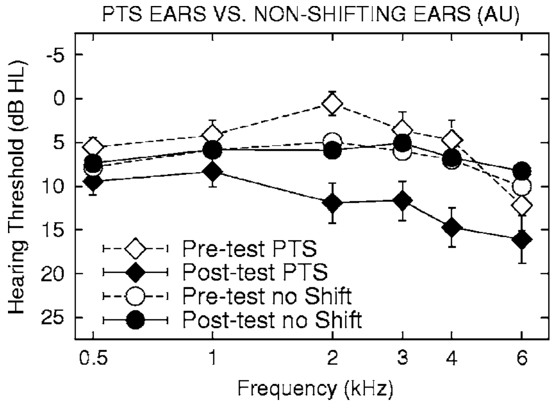


FIG. 4. Average audiometric thresholds for PTS ears ($n=18$) and non-PTS ears ($n=559$) for the volunteers with no shifts in either ear. Audiometric thresholds for PTS ears were, as expected, worse at postdeployment, with the biggest increases in threshold at 2 to 4 kHz. Error bars indicate one standard error of the mean and are smaller than the symbols for the non-PTS ears.

#### *3. Permanent threshold shifts*

Fifteen sailors (18 ears) were diagnosed with PTS (3 bilateral, 10 left ears, 2 right ears); this is 4.4% of sailors who were tested postdeployment. Their median age was 24 years (compared with 22 years for the group of sailors with no shifts), median length of service was 3.5 years (compared with 3 years for the group of sailors with no shifts), and all were male. Figure 4 shows the average audiograms for pre- and postdeployment for the 18 PTS ears and for the group of sailors with no shifts in either ear (599 ears). The PTS ears had slightly better audiometric thresholds at predeployment, but increased thresholds at postdeployment, especially in the 2- to 4-kHz range. Two volunteers with PTS at 4 kHz were diagnosed with an STS using the Navy criterion of 15 dB, which is less strict than the derived criterion of 20 dB. Both, however, had follow-up audiograms confirming the shift, so they were included in the PTS group. Most PTSs were on the order of 15 dB, with the largest being 25 dB.

From the noise histories of the volunteers with PTS, there were no particular commonalities among the noise exposures, other than there being proportionally more sailors with PTS from the reactor department (11%), compared with the air (4%) and engineering (3%) departments.

Thirteen of the 15 sailors with PTS had incomplete data sets, so only the data for two sailors with PTS contributed to the subset of 75 sailors used for the ANOVA reported in the previous section. The data sets tended to be incomplete due to the absence of OAEs rather than measurement problems. This tendency for absent OAEs among the PTS cases is examined later.

#### *4. Pattern of SESs for the 18 PTS ears*

For $TEOAE_{74}$, 11 ears showed no changes (though there were some missing data for seven ears), five ears showed changes in one half-octave band (though only one ear had no missing data), one ear showed changes across more than one band and also had some missing data, and one ear had essentially no measurable OAEs. For DPOAEs, generalizing over level, seven ears had no OAE changes at any level, four ears had OAE changes at all levels (though not necessarily at the same frequency), three ears had some OAE changes but not at all levels, one ear had essentially no OAEs, and three ears had no usable data at any frequency. There were many missing data (due to bad calibrations, etc.), so potentially some OAE shifts were not detected. Excluding the cases where all data were missing due to measurement problems or where OAEs were absent at most frequencies at predeployment, 31% of PTS ears showed at least one SES in TEOAEs, and 50% of PTS ears showed at least one SES in DPOAEs (across levels). However, many of these SESs were improvements in OAE amplitude, many were not in the same frequency band as the PTS, and there was only some consistency between changes in TEOAEs and changes in DPOAEs. Consistency among changes within DPOAE levels was also not high.

The nonparametric phi coefficient (Siegel, 1956) was used as a measure of association for the $2\times 2$ cross-tabulated tables of PTS ears versus non-PTS ears (at any frequency) versus ears with and without SESs (at any frequency) to determine whether PTSs and SESs tended to occur together (PTS ears: $n=10$ to 17; non-PTS ears: $n=473$ to 572). The phi coefficient can be interpreted similarly to a correlation coefficient and can be used for small data sets. Because of the small number of PTS cases and the large amount of missing data, an ear was considered to have an SES if there was an SES at any frequency within an OAE type and level. Ears with no measurable SESs (due to missing data) were excluded. There was no correlation between PTS status and SES status for any OAE type.

To investigate if there was an association between TEOAE SESs and DPOAE SESs in the PTS ears, each PTS ear (excluding the three cases with extensive missing data) was flagged as having either (a) no SESs at any frequency or (b) at least one SES at any frequency, for the conditions $TEOAE_{74}$, $DP_{57/45}$, $DP_{59/50}$, and $DP_{65/45}$. The phi coefficient was again used as a measure of association for the resulting six $2\times 2$ cross-tabulated tables. Phi was 0.58 for $TEOAE_{74}$ versus $DP_{57/45}$; 0.70 for $TEOAE_{74}$ versus $DP_{59/50}$; and 0.87 for $TEOAE_{74}$ versus $DP_{65/45}$. It is fair to say that when there was an SES for one OAE type, then there was often an SES for the other OAE type. Similarly, among the DPOAE levels, the association between levels $DP_{57/45}$ and $DP_{59/50}$ was 0.80, between $DP_{57/45}$ and $DP_{65/45}$ was 0.87, and between $DP_{59/50}$ and $DP_{65/45}$ was 0.93. DPOAE SESs and TEOAE SESs were associated with each other in the PTS ears, indicating that the SESs for the PTS ears were unlikely to be due to random fluctuations. However, this does not indicate that these SESs are related to the PTSs; it merely reinforces the finding from Sec. III B 3 that SESs across OAE type are related.

#### *5. Summary of changes in audiometric thresholds and concomitant changes in OAEs*

Although there is no compelling relationship between changes in audiometric thresholds and changes in OAEs, there is an association among changes in OAEs across OAE types, levels, and frequencies.

The number of PTS cases with low-level or absent OAEs was notable. If an OAE is already low level, it is

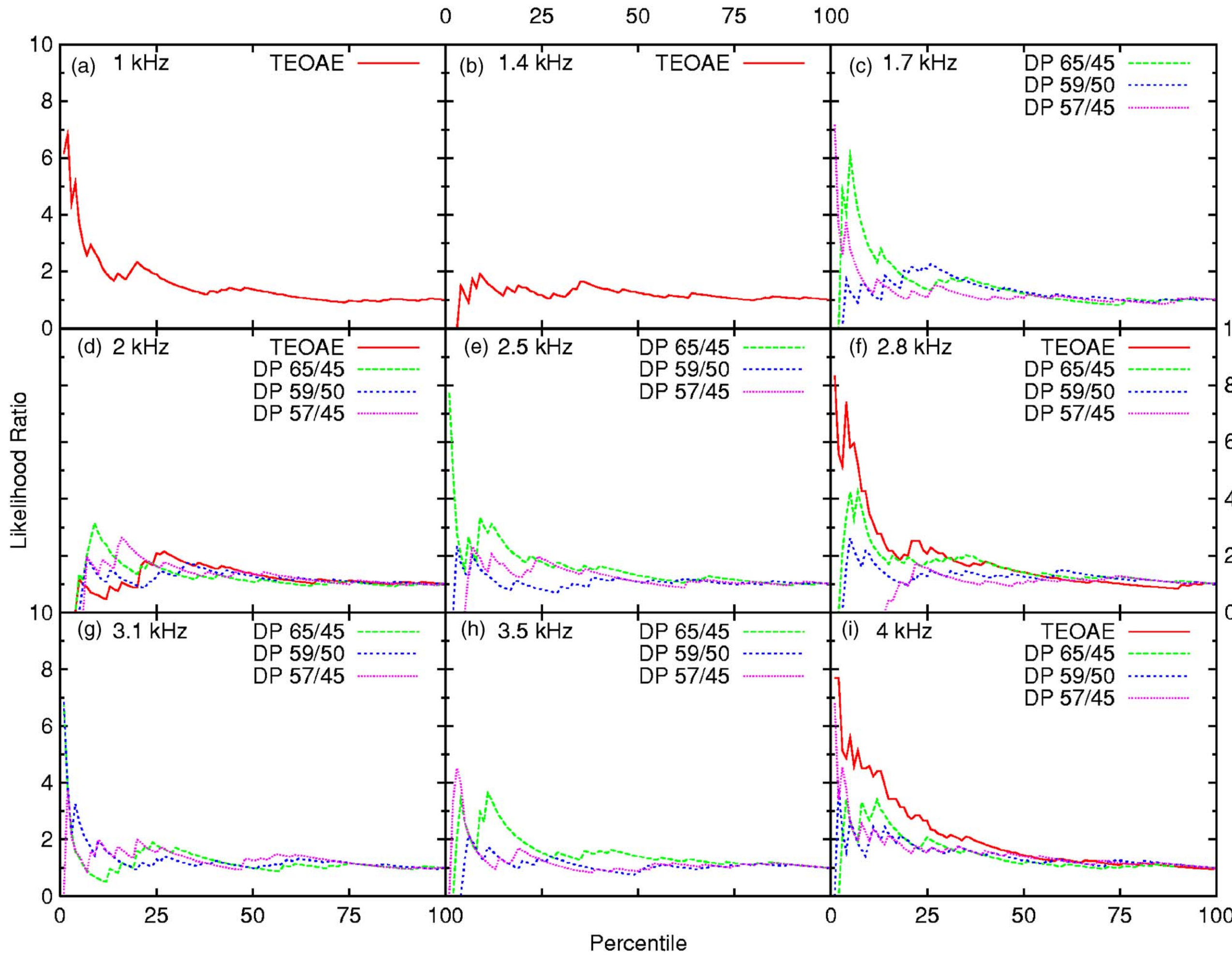


FIG. 5. (Color online) The likelihood ratio by percentile for each OAE test and frequency. For many OAE conditions, the likelihood of a PTS ear having an OAE level below the percentile criterion, compared with having an OAE level above the percentile criterion, increased as the percentile criterion (based on OAE amplitude) decreased. Not all OAE tests had measurements made at the same frequency; note that TEOAEs are half-octave band and DPOAEs are single-frequency measurements.

unlikely that further changes will be detected. A possible explanation is that noise damage prior to this study left many of these ears with subclinical damage, which makes these ears more likely to acquire hearing loss with further noise exposure. To investigate this theory, low-level and absent OAEs at predeployment were examined to see if they were predictive of PTS.

### D. Predictors of susceptibility to PTS

Earlier observations suggested that low-level or absent OAEs were more likely among the PTS ears. PTS risk (defined by likelihood ratios and positive predictive values) was estimated as a function of predeployment OAE level, with OAE levels converted into percentiles to enable comparison across OAE types.[8] By considering all possible percentiles, a prediction of PTS risk for any OAE amplitude is possible. Because no female sailors got PTS, this analysis was restricted to data from the male sailors. The greatest number of ears with good data was used for each condition, so the number of ears varied across conditions. As the total number of ears and the number of PTS ears were not constant across conditions, some caution must be taken in interpreting the results, particularly in comparing the advantage of various stimuli in predicting susceptibility.

In medical diagnostics, the likelihood ratio is a ratio of two probabilities: the probability of a particular test result among patients with a condition to the probability of that particular test result among patients without the condition (Zhou *et al.*, 2002). Here, the likelihood ratio indicates the relative probability that a predeployment OAE amplitude was below a given percentile in the group of ears that subsequently got a PTS, relative to the same result in the group of ears that did not subsequently get a PTS. For instance, a likelihood ratio of 1 would indicate that a particular predeployment test result was equally likely to occur for ears that subsequently got PTS and ears that did not get PTS. A likelihood ratio of 4 for a particular test result indicates the result was four times more likely among ears that got PTS than ears that did not get PTS. The likelihood ratio does not take the base rate (*a priori* probability) of PTS into account.

A cutoff defined by an OAE percentile can be applied as a diagnostic criterion for PTS risk. This cutoff is independent of the actual condition (presence or absence of PTS), and it is

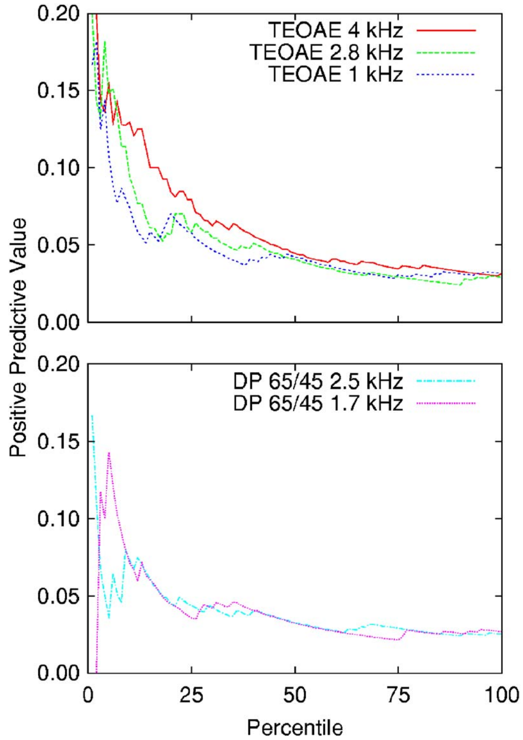


FIG. 6. (Color online) The positive predictive value (PPV) of a low-level OAE amplitude increased as the percentile criterion defining a low-level OAE amplitude decreased. PPV is the probability that a predeployment OAE amplitude was predictive of subsequent PTS. Shown are the PPVs, as a function of OAE percentile criterion, for the five OAE tests and frequencies that were most predictive of subsequent PTS. TEOAEs are shown in the top panel and DPOAEs in the bottom panel. These data should be viewed as indicative only; the jaggedness is caused by the small number of PTS cases contributing to the analysis. However, the general tendency is for TEOAEs to provide more predictive power than DPOAEs.

of interest to find an optimal cutoff value. For any specific percentile cutoff value, the likelihood ratio is defined as the ratio of the probability that a test result was below the cutoff given there was a PTS to the probability that a test result was below a percentile cutoff given there was not a PTS. For readers familiar with the theory of signal detectability and ROC analysis, this is equivalent to the ratio of the hit rate and false-alarm rate, though the data are transformed so that the PTS group is the "signal" and the non-PTS group is the "noise."

Figure 5 shows likelihood ratio as a function of percentile cutoff for each OAE and test frequency. For TEOAEs, there were 16 to 18 PTS ears and 524 to 559 non-PTS ears included in the analysis. For DPOAEs, there were 16 PTS ears and 546 to 550 non-PTS ears included in the analysis. For many cases, there was a clear trend of increasing risk with decreasing percentile cutoff. $TEOAE_{74}$ at 4, 2.8, and 1 kHz are the clearest cases—each shows that low-level TEOAE amplitudes were more likely among the ears that subsequently developed PTS in this population and noise environment. The risk starts to increase as the TEOAE amplitude moves below the 25th percentile. DPOAEs show a similar trend to TEOAEs, but they are not as consistent, nor do they reach as high a likelihood ratio.

The positive predictive value (PPV) (Zhou *et al.*, 2002), on the other hand, is the conditional probability of an ear

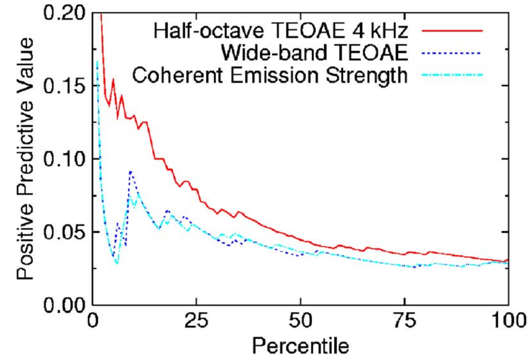


FIG. 7. (Color online) Positive predictive values (PPVs) as a function of percentile criterion for TEOAEs: half-octave 4 kHz, coherent emission strength, and the wideband TEOAE response. There is essentially no difference between CES and the wideband TEOAE, whereas the half-octave TEOAE at 4 kHz is a substantially better predictor of subsequent PTS.

from this population *getting* PTS within 9 months in this particular noise environment, *given* a test result of a low-level OAE. The PPV is also known as the *a posteriori* conditional probability: $P(\mathrm{PTS}|\mathrm{OAE}\leq\mathrm{cutoff})$, and it takes the base rate of PTS into account. PPVs are more useful than likelihood ratios for diagnosticians, because they can be used to estimate the probability of getting a PTS for a given population, timeframe, and noise environment.[9]

Figure 6 shows the best three $TEOAE_{74}$ frequencies and the best two DPOAE frequencies from Fig. 5, replotted as PPV. The base rate for an ear incurring PTS is approximately 3%.[10] For ears with results in the low percentiles, the risk of PTS increases to between 17% and 20% for the best TEOAE conditions and to between 14% and 17% for the best DPOAE conditions, depending on the percentile cutoff chosen.

LePage and Murray have also considered low-level TEOAEs as a predictor for hearing loss, based on their cross-sectional data set. They used an empirically derived TEOAE measure: coherent emission strength (CES, dB SPL), which represents the noise-free part of the TEOAE (LePage and Murray, 1993; LePage *et al.*, 1993; LePage, 1998). CES is a reweighting of the average TEOAE wide-band response by the square of the reproducibility (when linearly rescaled and transformed from $[-1:1]$ to $[0:1]$). For comparison, PPVs were calculated for the TEOAE data using CES and the TEOAE wide-band response, and compared to the best performing 4-kHz half-octave band (see Fig. 7). There were 563 non-PTS ears and 17 PTS ears contributing to the analysis. Data for eight ears were removed (including one PTS case) because the noise level was too high. The noise level was substituted for wide-band TEOAEs that were less than 0 dB SNR. No substitutions were made for CES, because this method allows for the use of TEOAEs with SNR less than 0 dB. Figure 7 shows PPVs for CES, wide-band TEOAEs, and 4-kHz half-octave TEOAEs. CES and wide-band TEOAEs were almost identical in their ability to predict PTS, and both were substantially worse than 4 kHz TEOAEs. The performance of CES and wide-band TEOAEs were similar to the best DPOAEs shown in Fig. 6. Focusing on the area most likely to be damaged by noise (4 kHz and above) increases the predictability for TEOAEs. The predictability for DPOAEs at 4 kHz, on the other hand, is slightly worse than CES and wide-band TEOAEs, especially at the lowest percentiles.

These predictors are based on a small number of PTS cases, and they should be treated as indicative only. The predictors are very much dependent on the specific population, elapsed time (only 9 months in this study), and noise environment studied. However, these data show promising signs that OAEs may be used as predictors for susceptibility to PTS.

### E. Summary of main findings

(1) Average audiometric thresholds did not change between pre- and postdeployment testing, but both average TEOAE and DPOAE amplitudes decreased for the group of 75 sailors with relatively complete data sets.

(2) There was no correlation between changes in audiometric thresholds and changes in OAEs for the entire group of noise-exposed sailors. There were, however, significant correlations between changes in OAEs across OAE types.

(3) Fifteen sailors (18 ears) were diagnosed with PTS. It was expected that significant changes in audiometric thresholds would be mirrored with significant changes in OAEs, but this was not the case in the majority of ears. Instead, the main observation was that PTS ears had many low-level or absent OAEs.

(4) There was no correlation between ears with PTS and without PTS and ears with SES or without SES. There was, however, a correlation among SESs across OAE types for the PTS ears, indicating that the SESs were probably not random fluctuations.

(5) Low-level and absent predeployment OAEs were predictive of postdeployment PTS.

(6) The best predictor of postdeployment PTS was predeployment TEOAE amplitude at 4 kHz, with lower amplitudes indicating increased risk.

## IV. DISCUSSION

### A. Why did PTS occur?

PTS occurred due to a combination of high noise levels and imperfect hearing protection usage. Fifteen sailors had a documented PTS in at least one ear after 6 months' deployment on an aircraft carrier, despite an active hearing-conservation program and the use of hearing protection. As reported earlier, the noise levels on aircraft carriers often exceed the maximum noise reduction ability of hearing protection. Furthermore, from the self-reported noise-exposure histories in the current study, the majority of sailors with PTS used hearing protection inconsistently. However, many sailors without PTS were also poor users of hearing protection. It was common for sailors to report using only single hearing protection in situations where double hearing protection was required. Sometimes no hearing protection was used when single hearing protection was required. It is also likely that many were not fitting hearing protection correctly. In a study across multiple platforms (aircraft carriers and amphibious assault ships), the vast majority of sailors omitted their earplugs some or all of the time, and did not insert them to a proper depth (Bjorn *et al.*, 2005). It is likely that these results generalize to the population of the current study, and it is therefore of no surprise that there was documented PTS.

### B. Derived STS and SES criteria

It was important to develop site-specific STS criteria in the current study, because the Navy STS criteria used at the time were not based on test-retest reliability measurements. The STS criteria were identical to the Navy criteria at 1, 2, and 3 kHz, but were larger at 4 kHz and smaller at the average of 2, 3, and 4 kHz. In comparison to previously published studies, the STS criteria were larger at 4 and 6 kHz than the criteria developed in Lapsley Miller *et al.* (2004). The TEOAE criteria were smaller than in this earlier study, possibly because the testers were more experienced, and possibly because every attempt was made to match the postdeployment stimulus waveform and spectrum to the predeployment waveform and spectrum by manipulating the angle and depth of the probe in the volunteer's ear. DPOAEs were not comparable because here they were based on measurements at individual frequencies, rather than averaged within half-octave bands. However, they were comparable to those reported elsewhere (Franklin *et al.*, 1992; Beattie and Bleech, 2000; Seixas *et al.*, 2005b). OAE reliability depends on the measurement paradigm and equipment, therefore the values from the current study may not generalize to other settings.

### C. Relationships between PTS and SES

As summarized in the introduction, cross-sectional human studies and longitudinal animal studies all indicate, for various reasons, that there should be a relationship between noise-induced PTS and SES. However, longitudinal human studies have yet to offer a clear-cut picture. Differences in noise exposures during the study, differences in prior noise exposures, hearing-protection usage, age, sex, and individual susceptibility all mean that each person is at a different stage in developing noise-induced inner-ear changes and noise-induced hearing loss. In the current study, there were sailors with no changes in OAEs or audiometric thresholds, changes in OAEs but not audiometric thresholds, changes in audiometric thresholds but not OAEs, and changes in OAEs and audiometric thresholds. Can all these scenarios be accounted for by current theories, or is it a sign that there is no relationship between OAEs, audiometric thresholds, and noise-induced change?

#### *1. No changes in OAEs or audiometric thresholds*

Out of the 338 noise-exposed sailors, only 12 (3.6%) had no measurable changes in audiometric thresholds or in any of their OAEs in either ear. A further 44 (13%) had no measurable changes, but also had some missing OAE data. All the other sailors had at least one significant shift in either audiometric thresholds or OAEs, though many of these are likely to be false positives. Why did some sailors have no changes? They may have been better users of hearing protection. They may have been lucky to not be as severely noise exposed, and so avoided any noise-induced damage. They may have had particularly tough ears (Cody and Robertson, 1983) or toughened ears [in laboratory rodents, intermittent noise exposure may increase resistance to noise damage (Henderson *et al.*, 1993)]. There are likely to be some undetected changes (false negatives), in part due to missing OAE data (from bad calibrations, high noise, etc.) and also due to test-retest variability. However, there were proportionally more ears in the control group with no changes across any of the measures, suggesting that the noise-exposed group did indeed have more changes due to noise exposure.

### 2. Changes in OAEs but not audiometric thresholds

More ears showed significant OAE shifts than permanent threshold shifts in the noise-exposed group. Furthermore, the ANOVAs indicated no group changes in audiometric thresholds, but small, significant, decreases in group OAE levels for the 75 volunteers with complete data sets.[11] This is mostly consistent with the other longitudinal studies in the literature where small group decreases in OAEs are often seen, but concomitant changes in group audiometric thresholds are not (Engdahl *et al.*, 1996; Seixas *et al.*, 2005a). Konopka *et al.* (2005) found an approximately 2-dB decrease in TEOAEs, but the only significant changes in audiometric thresholds were at 10 and 12 kHz (frequencies not normally measured). Lapsley Miller *et al.* (2004) also showed changes in audiometric thresholds along with changes in OAEs, with a standard audiometric-frequency range, but this result was not as clear-cut when considering individual PTS cases. In comparison, no consistent changes were found among an orchestra group over 5 and 9 years (Murray *et al.*, 1998; Murray and LePage, 2002), but there were issues with recent noise exposure and possible TTS at baseline, which would reduce the magnitude of any audiometric threshold shift.

There are at least four explanations for why there were more SESs than PTSs: sensitivity of the audiogram, high-frequency hearing loss, outer-hair-cell redundancy, and age-related changes. The parsimonious explanation is that OAE measurements have smaller test-retest variability than audiometry (as measured using a standard clinical protocol and audiometer), so smaller noise-induced changes to the inner ear can be detected. Although the audiometric reliability in the present study was not as low as possible (because of the requirement that the Navy's nonoptimal shipboard audiometers were used), even in the best of circumstances audiometric reliability is worse than OAE reliability when comparing the same frequency band (e.g., Lapsley Miller *et al.*, 2004).[12] Audiometric resolution is similar to OAE resolution if multiple audiometric test frequencies are combined, but at the price of a decrease in frequency specificity (11 out of 18 PTS cases had PTS over an average of two or three frequency bands). However, it would be expected that if audiometric resolution was much lower than OAE resolution, the PTS cases that were identified should show consistent SESs. This was not the case—only one-third of PTS cases also showed SESs (and these SESs were not necessarily consistent with the PTS). Differences in resolution cannot explain all the findings.

Only changes in audiometric thresholds up to 4 kHz were considered, because of the high test-retest variability at 6 kHz. It is possible that undetected high-frequency hearing loss at 6 kHz and above affected lower frequency OAEs. The mechanism by which this occurs is still being debated, but it could be due to intermodulation distortion of the OAE components (Avan *et al.*, 1995; Yates and Withnell, 1999; Withnell *et al.*, 2000). Recently, Konopka *et al.* (2005) reported group decreases in TEOAEs between 2 and 4 kHz, concomitant with group increases in high-frequency audiometric thresholds at 10 and 12 kHz, but no significant change in audiometric thresholds at the TEOAE frequencies.

An alternative theory is that there is outer-hair-cell (OHC) redundancy such that only some are required for normal hearing. Animal studies have shown that there can be extensive OHC loss without changes in hearing thresholds (e.g., Bohne and Clark, 1982; Hamernik *et al.*, 1989; Altschuler, 1992), and that OAEs can be more sensitive to the effects of noise damage to the inner ear than pure-tone thresholds (Hamernik *et al.*, 1996). LePage and Murray (1993) argue that because OAEs are a more direct measurement of OHC activity, the loss of some OHCs is more likely to show up as diminished OAEs levels rather than as hearing loss. OAEs would therefore show noise-induced changes prior to hearing loss (i.e., even if audiometric thresholds could be measured more sensitively, there would not be any increase in the amount of noise-induced PTS). This implies that OAE changes can indicate subclinical NIHL. LePage (1992) proposes that the ear is able to remap the cochlear place-to-frequency conversion to avoid gaps in frequency detection because of OHC loss, and hearing loss occurs only when this ability to remap is exceeded. In their large cross-sectional study, Le Page and Murray found that coherent emission strength (CES) may decrease by 80% before there is a change in audiometric thresholds (LePage *et al.*, 1993). They concluded that the "pure tone audiogram may not be a direct measure of cochlear damage so much as a measure of how much the cochlea can maintain normal performance despite ongoing damage" (LePage *et al.*, 1993). This is supported by the current findings that (a) although there were changes in audiometric thresholds and changes in OAEs, the two were not related contemporaneously, and (b) ears with low OAEs have less resistance to hazardous noise and are more likely to get PTS with continued noise exposure.

Another possibility is age-related OAE changes. Some studies have shown that OAEs decrease with age, even when audiometric thresholds are controlled (Dorn *et al.*, 1998; Cilento *et al.*, 2003, for women but not for men). Murray and LePage (1993) hypothesize that the OHC loss that occurs from birth throughout life causes the aging effect seen with OAEs, and that noise exposure accelerates the loss of OHCs. However, the present study was only 9 months long, and the volunteers were relatively young, so aging is unlikely to be a major factor.

These four possibilities—lower audiometric sensitivity, high-frequency hearing loss, OHC redundancy, and age-related changes—are not mutually exclusive, nor is it possible to easily tease out which, if any, is operating here. However, none of the explanations are contradictory—all can explain these findings to some extent.

### 3. Changes in audiometric threshold, but not OAEs

About two-thirds of the PTS ears did not have significant OAE changes consistent with their PTS. In most cases, this was because OAEs were already at low levels or were absent at predeployment testing. It is possible that OAEs had already decreased from earlier noise exposures (the large majority of volunteers had considerable military noise exposure prior to the study), and some sailors may also have had low-level OAEs due to genetics, illness, or environmental factors such as chemical exposure. Regardless of the cause, having low-level or absent OAEs was predictive of subsequent noise-induced hearing loss.

Massive, traumatic noise exposure can simultaneously affect both audiometric thresholds and OAEs, but here the largest PTS was only 25 dB, and most were only 15 dB. It is more likely that changes in OAEs preceded the changes in audiometric thresholds. By the time the audiometric thresholds were affected, the OAEs may have been sufficiently low or even absent. It is difficult to measure a change in a low-level OAE because measurements near or below the noise floor are not reliable. Even in the ears where PTSs and SESs occurred together, it is conceivable that the OAEs may have diminished before the audiometric thresholds, but both may have changed within the 9 months, so only the final outcome was observed.

Other explanations as to why some PTS ears showed no changes in OAEs include interaction of DPOAE sources and differences in intrasubject variability, though these factors are probably less influential than missing data. The DPOAE measured at frequency $2f_1$-$f_2$ actually consists of two frequency components—both with frequency $2f_1$-$f_2$, but with different magnitudes and phases. These two components—the reflection component and the distortion component—are thought to be generated from different parts of the cochlea (Shera and Guinan, 1999). If a DPOAE was dominated by the distortion component, then changes in the cochlea at the source of the reflection component (thought by Shera to be the more likely indicator of damage) may not show up (Shera, 2004). In the current study, the definition of a significant OAE shift was based on the group $\mathrm{SE}_{\mathrm{MEAS}}$. Some people exhibit little variation in OAEs over time—some show a great deal (Marshall and Heller, 1996). The criteria used in the current study may have been too strict for some people with very stable OAEs, therefore missing some SESs.

Finally, damage to structures other than OHCs may have caused the threshold shift. For example, stria vascularis may be affected by long-duration exposures (e.g., Bohne and Clark, 1982), and inner-hair cells may also be damaged by noise exposure, but usually not until greater amounts of hearing loss are observed (e.g., Hamernik *et al.*, 1989).

### 4. Both audiometric thresholds and OAEs change together

Only one-third of the PTS ears showed SESs; however, there was no strong or consistent pattern of SESs across ears, or across OAE types and frequencies. The ANOVA and correlational analyses gave scant evidence of audiometric thresholds and OAEs changing concomitantly. Sometimes OAEs improved when audiometric thresholds diminished or stayed the same. Although there were some positive SESs, they were at about the same rate as for the control group, so they most likely reflected random test-retest variability [although inner-ear damage can in some cases produce an increase in OAE amplitude (Withnell *et al.*, 2000)].

## D. Susceptibility

Ears with low-level or absent OAEs were more likely to get PTS. Low-level or absent OAEs may be a sign of genetic variability, or, more likely, a sign that these ears had already experienced subclinical damage from previous hazardous noise exposures. Although most volunteers had normal hearing at predeployment (some had a slight high-frequency hearing loss), most, if not all, had received considerable noise exposure prior to this study. The PTS ears had slightly better audiometric thresholds at predeployment compared with the non-PTS ears (see Fig. 4), so the initial audiometric thresholds themselves cannot explain the difference in OAE levels.

In the current study, TEOAEs were better PTS predictors than were DPOAEs. This may be because normal DPOAE microstructure has spectral nulls due to the interaction of the two DPOAE sources (e.g., Kalluri and Shera, 2001; Mauermann and Kollmeier, 2004). Measurements at or near nulls could easily show low-level DPOAE amplitudes. For instance, Shaffer *et al.* (2003) quite clearly showed how points in the DPOAE microstructure for a normal-hearing person can fall below norms. This would inflate the number of low-level DPOAEs seen in healthy ears. Such cases would not necessarily be predictive of hearing loss. Half-octave TEOAEs are less affected by fluctuations in microstructure because they represent the average level over a range of frequencies. A DPOAE design where frequencies were clustered closer together and the resulting amplitudes averaged may reduce the problem of nulls.

Dancer stated "There would be great interest in finding a test which predicts individual susceptibility to permanent threshold shift" (Dancer, 2000, p. 5-1). It looks promising that low-level or absent OAEs may indeed form the basis for such a test. Here, problems inherent with many previous susceptibility tests were bypassed in that the measurements were made on ears that developed PTS, rather than using TTS as a substitute (Ward, 1965; Humes, 1977). This is important since the mechanisms underlying TTS and PTS are different (e.g., Saunders *et al.*, 1985; Slepecky, 1986; Nordmann *et al.*, 2000).

Using low-level OAEs to measure PTS susceptibility is not like using a gene marker, which is stable and innate. Instead, an OAE reflects the current state of the inner ear which is likely a result of both genetic susceptibility and acquired susceptibility. Susceptibility to NIHL probably varies over time depending on factors such as noise exposure, illness, chemical exposure, and age. It will not be enough to take just one measurement at one point in time to determine susceptibility. Instead, personnel will require regular monitoring to see if their susceptibility is changing as they continue to accumulate OHC damage. With more data—especially longitudinal data over a longer timeframe and information about the outcome costs—optimal criteria for risk detection can be developed.

In the future, OAEs may also be used in other ways to gauge susceptibility to NIHL. For instance, the reflex strength of the auditory efferent medial olivocochlear system, whose fibers synapse primarily on the outer hair cells, may indicate NIHL susceptibility. One of the suggested physiological functions of the efferent system has been protection from acoustic overexposure (Cody and Johnstone, 1982; Reiter and Liberman, 1995; Maison and Liberman, 2000; Luebke and Foster, 2002). Maison and Liberman (2000) predicted susceptibility to NIHL from the strength of the auditory efferent reflex, as measured with ipsilateral DPOAE adaptation. Guinea pigs with a high reflex strength exhibited only small or no PTS whereas guinea pigs with a low reflex strength exhibited PTS. Muller *et al.* (2005) showed that measuring DPOAE amplitude changes in humans with a contralateral AER elicitor (which is easier to measure in humans) is also a suitable measure for determining AER strength. Others have suggested using TEOAEs (e.g., Berlin *et al.*, 1995) or stimulus-frequency otoacoustic emissions (e.g., Guinan *et al.*, 2003). No matter which OAE type is used, the challenge is to develop a clinical test that shows a large range of auditory efferent reflex strength across people, relative to the intrasubject test-retest reliability, to be able to validly distinguish people with large and small efferent reflex strength. Such a test must also be fast for use in clinical and field settings.

In the future, measuring both auditory efferent reflex strength and absolute OAE amplitude in normal-hearing ears may provide powerful indications of individual NIHL risk before significant hearing loss has occurred.

### E. Conclusions

When sailors are exposed to hazardous levels of shipboard noise, OAEs show the accumulated damage to the inner ear before hearing loss shows up in an audiogram. Moreover, diminished OAEs are predictive of subsequent hearing loss if the sailor remains in the noise-hazardous environment.

## ACKNOWLEDGMENTS

Thanks to Linda Westhusin, Michael McFadden, Joe Bertoline, Dan Pawluk, Doug Clingan, and John Correia for their assistance with study coordination, volunteer recruitment, and data collection. A special thanks to the crew of the *USS Dwight D. Eisenhower*, Bob Rogers, and the staff from Norfolk Naval Hospital Occupational Health Audiology Department. Thanks to the Commander in Chief, US Atlantic Fleet; the Commander, Naval Air Force, US Atlantic Fleet; and the Commander, Naval Sea Systems Command. Thanks to Vit Drga for comments and suggestions on data analyses; Wayne Horn, Jeff Gertner, Michael Qin, and Loring Crepeau for their comments on an earlier version of the manuscript; Kurt Yankaskas for comments on noise exposure levels; and many useful comments and suggestions from the anonymous reviewers.

This research was supported by grants from the Office of Naval Research (N0001400WX20195, N0001400WX20244, and 0601153N4508) and the US Army Medical Research and Materiel Command (DAMD17-01-IA-0679). The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, the Department of Defense, or the United States Government.

[1]These studies tend not to consider whether noise-exposed people with normal OAEs have similar audiometric thresholds to non-noise-exposed people with normal OAEs. Therefore, it is not entirely clear that OAEs are providing an advantage in detecting the early stages of NIHL.

[2]On a Nimitz-class aircraft carrier, flight-deck noise produced from aircraft launches ranges from 126 to 148 dBA peak depending on the proximity to the aircraft. The arresting gear and water brakes, as well as tools such as needle guns, grinders, and hydro-blasters, generate noise levels around 94 dBA peak in work and berthing areas. Sound levels in the hangar bay under the flight deck can exceed 120 dBA peak during flight operations. Other noisy areas include the main propulsion machinery space, machinery rooms, (work) shops, and the laundry, which is located above the propellers. Many of the berthing spaces are directly below the flight deck—some sailors even wear hearing protection while sleeping. Dosimetry data from a Nimitz-class aircraft carrier, reported by Rovig *et al.* (2004), showed 8-h time-weighted averages (using an 85 dBA damage-risk criterion with a 3 dB exchange rate) of 109 dBA (ranging from 96 to 120 dBA) for flight deck crew and 92 dBA (ranging from 79 to 98 dBA) for engineering crew. The average workday was 11.5 h. Unweighted peak noise levels were regularly clipped by the dosimeter's 150 dB SPL ceiling, so these noise levels are underestimated. Rovig *et al.* (2004) also found that in sailors with 4 or more years of service, 30% of flight deck crew and 37% of engineering crew had audiometric thresholds greater than 25 dB HL (at 1, 2, 3, or 4 kHz), compared with 5% of administrative crew. Many sailors reported not wearing double hearing protection because they felt it jeopardized speech communication.

[3]The noise-rejection level was set at 4 mPa by default, but was usually adjusted by the tester to approximately one standard deviation above the mean of the noise-level histogram, which was usually lower than 4 mPa.

[4]Each frequency was measured three times with 15 subaverages and then averaged. The frequencies where DPOAEs had a signal-to-noise ratio (SNR) less than 3 dB were automatically retested until the test time (50 s) expired. The noise-rejection level was set at 5 mPa.

[5]$\mathrm{TEOAE}_{74}$ stimulus levels were considered on-target if they were within ±4 dB of 74 dB pSPL (no data points were eliminated; 99.9% of data points were within ±3 dB of the target). DPOAE stimulus levels were considered on target if they were within ±6 dB of the target level, for both $L_1$ and $L_2$ (1.2% of data were eliminated across all levels and frequencies). Sometimes the ILO program could not obtain a good DPOAE calibration (usually due to standing waves). In these cases it automatically used an estimated level instead. The resulting DPOAEs produced many outliers—either unusually high or low, so all cases where levels were estimated were dropped from the DPOAE analyses (2.5% of all DPOAE data). Consideration of outliers showed many more outliers for $\mathrm{DP}_{61/55}$, compared to the other levels. The reason for this discrepancy could not be traced, so all data at this level were dropped. Furthermore, all DPOAE data at $f_2$=2.2 and 4.5 kHz were dropped because they were contaminated with a large, intermittent harmonic artifact at these frequencies. Sometimes the artifact elevated the noise level and sometimes it appeared to elevate the DPOAE amplitude, so it was not possible to identify the affected cases by just looking for high noise levels. When considering changes in OAEs, some outliers appeared to be due to large differences in the noise floor between pre- and postdeployment testing. Therefore, the cases where the absolute average difference between the pre- and postdeployment noise levels was larger than 3.5 dB (when averaged across 2.5 to 3.6 kHz for DPOAEs, and

when averaged across the half-octave bands centered at 2, 2.8, and 4 kHz for $TEOAE_{74}$) were removed. This affected 2.8% of $TEOAE_{74}$ measurements and 6.5% of DPOAE measurements.

[6]The $SE_{MEAS}$ can be used to specify the magnitude of a statistically significant change within an individual (Ghiselli, 1964) and is defined as $SE_{MEAS}=\sqrt{\frac{1}{2}(s_1^2+s_2^2)(1-r)}$ where $s_1^2$ and $s_2^2$ are the pre- and postdeployment variances, and $r$ is the correlation between pre- and postdeployment tests. Because the focus here is on the difference between pre- and postdeployment testing, $\Delta SE_{MEAS}$ is defined as $\sqrt{2}SE_{MEAS}$ (Beattie, 2003; Beattie *et al.*, 2003). Multiplying $\Delta SE_{MEAS}$ by an appropriate multiplier then gives the desired confidence interval. Here a multiplier of 2.12 is used, which gives a 98% confidence interval.

[7]The STS criteria were adjusted by adding the group average audiometric thresholds to account for a slight mean shift (most likely a practice effect). For individual ears, the resolution of the audiogram was 5 dB for single frequencies, and, due to averaging, 2.5 dB for the average of 2 and 3 kHz and the average of 3 and 4 kHz, and 1.66 dB for the average of 2, 3, and 4 kHz. In other words, the smallest change that can be detected is defined by the resolution. Each STS criterion was adjusted accordingly by rounding up to the next largest step. Since STS criteria are usually specified as inclusive (i.e., $\geqslant X$ dB, rather than $>X$ dB), another resolution step was added to give the STS criteria reported in Table II. For example, the $\Delta SE_{MEAS}$ for the average of 2, 3, and 4 kHz was 2.16 dB and the minimum resolution is 1.66 dB. Converting into an STS criterion for the 98% confidence interval gives $2.12\sqrt{2}\Delta\ SE_{MEAS}=6.48$ dB; rounding up to the next resolution step gives 6.66 dB, and then another step to 8.33 dB gives the inclusive criterion. Resolution was not an issue for SESs because the $SE_{MEAS}$ were orders of magnitude larger than the measurement resolution.

[8]The percentiles were calculated from the predeployment OAE data for the entire group of noise-exposed male sailors (606 ears, 303 volunteers), including the PTS ears and ears with absent OAEs, where noise levels were substituted to quantify absent OAEs, providing the noise level was low (as described earlier). Any ears with absent OAEs with noise levels higher than the cutoff were not included in these analyses. OAE amplitudes were converted to percentiles for each TEOAE and DPOAE level and frequency. Percentiles were calculated for left and right ears separately and for all ears combined. For the same percentile, OAE amplitude differed by up to 2.4 dB between the left and right ears.

[9]The PPV can be related to the likelihood ratio by using odds ratios where the *a posteriori* odds are equal to the *a priori* odds multiplied by the likelihood ratio (Zhou *et al.* 2002). In the current scenario, $PPV/(1-PPV)=\text{likelihood ratio}\times P(\text{PTS})/P(\text{no PTS})$. Other formulations and relationships may be derived using Bayes' theorem for conditional probabilities.

[10]This PTS rate is the percentage of ears with PTS and without high noise floors. It differs from the earlier PTS rate, which was the percentage of sailors with PTS.

[11]Because only 75 volunteers had complete data sets across 2 to 4 kHz, a further ANOVA was conducted for the group of 206 volunteers with complete data sets at just 4 kHz for TEOAEs and audiometric thresholds, to see if the larger group showed any changes between pre- and postdeployment. The two-way, repeated-measures ANOVA for TEOAE amplitude (test: pre- versus postdeployment; ear left versus right) showed a 0.95-dB decrement between pre- and postdeployment testing ($F_{1,205}=49.2, p<0.05$), but no significant difference between ears, and no significant interaction. The two-way, repeated-measures ANOVA for audiometric thresholds (test: pre- versus postdeployment; ear left versus right) showed no significant effect for test, but a significant difference between ears ($F_{1,205}=24.8, p<0.05$). The interaction was not significant. Even with the increased numbers, there was no significant change in audiometric thresholds.

[12]The ability to detect a PTS or SES is dependent on the test-retest reliability. Poorer reliability results in a larger criterion to decide that a significant shift has occurred. For example, compare the OAE and hearing-threshold criteria for 4 kHz from Tables II and III. To make this a fair comparison between OAEs and audiometric thresholds, the SES criteria are multiplied by 2.5 to convert them into "equivalent dB HLs" (Marshall and Heller, 1998) to give criteria of 9.3 dB for $TEOAE_{74}$, 14.0 dB for $DP_{57/45}$, 15.2 dB for $DP_{59/50}$, and 14.2 dB for $DP_{65/45}$. Both the TEOAE and DPOAEs are superior to single-frequency audiometric thresholds in their ability to detect a shift. However, when averaging the audiometric threshold over 2, 3, and 4 kHz, the criterion decreases to be comparable with TEOAEs at 4 kHz. A closer relationship between changes in OAEs and audiometric thresholds is found in the laboratory in animal studies and in human TTS studies (e.g., Marshall and Heller, 1998) when there is more precise information about both the audiometric changes and the noise that produced them. Differences in results between TTS studies and PTS studies could be attributed to the fact that the mechanisms underlying TTS and PTS differ (e.g., Saunders *et al.*, 1985; Slepecky, 1986; Nordmann *et al.*, 2000). However, this conclusion may not be warranted because the precision of the audiometric and noise measurements is higher for TTS studies.


Altschuler, R. A. (**1992**). "Acoustic stimulation and overstimulation in the cochlea: A comparison between basal and apical turns of the cochlea," in *Noise-Induced Hearing Loss*, edited by A. L. Dancer, D. Henderson, R. J. Salvi, and R. P. Hamernik (Mosby-Year Book, St. Louis), pp. 60–72.

ANSI (1991). "Maximum permissible ambient noise levels for audiometric test rooms (ANSI S3.1)" (American National Standards Institute, New York).

Attias, J., Horovitz, G., El-Hatib, N., and Nageris, B. (**2001**). "Detection and clinical diagnosis of noise-induced hearing loss by otoacoustic emissions," Noise Health **3**, 19–31.

Attias, J., Bresloff, I., Reshef, I., Horowitz, G., and Furman, V. (**1998**). "Evaluating noise induced hearing loss with distortion product otoacoustic emissions," Br. J. Audiol. **32**, 39–46.

Attias, J., Furst, M., Furman, V., Reshef, I., Horowitz, G., and Bresloff, I. (**1995**). "Noise-induced otoacoustic emission loss with or without hearing loss," Ear Hear. **16**, 612–618.

Avan, P., Bonfils, P., Loth, D., Elbez, M., and Erminy, M. (**1995**). "Transient-evoked otoacoustic emissions and high-frequency acoustic trauma in the guinea pig," J. Acoust. Soc. Am. **97**, 3012–3020.

Beattie, R. C. (**2003**). "Distortion product otoacoustic emissions: comparison of sequential versus simultaneous presentation of primary-tone pairs," J. Am. Acad. Audiol **14**, 471–484.

Beattie, R. C., and Bleech, J. (**2000**). "Effects of sample size on the reliability of noise floor and DPOAE," Br. J. Audiol. **34**, 305–309.

Beattie, R. C., Kenworthy, O. T., and Luna, C. A. (**2003**). "Immediate and short-term reliability of distortion-product otoacoustic emissions," Int. J. Audiol. **42**, 348–354.

Berlin, C. I., Hood, L. J., Hurley, A. E., Wen, H., and Kemp, D. T. (**1995**). "Binaural noise suppresses linear click-evoked otoacoustic emissions more than ipsilateral or contralateral noise," Hear. Res. **87**, 96–103.

Bicciolo, G., Ruscito, P., Rizzo, S., and Frenguelli, A. (**1993**). "Evoked otoacoustic emissions in noise-induced hearing loss," Acta Otorhinolaryngol. Ital. **13**, 505–515.

Bjorn, V. S., Albery, C. B., Shilling, R., and McKinley, R. L. (**2005**). "Navy Flight Deck Hearing Protection Use Trends: Survey Results," in *NATO Human Factors and Medicine Panel Symposium, New Directions for Improving Audio Effectiveness* (Amersfoort, The Netherlands).

Bohne, B., and Clark, W. W. (**1982**). "Growth of hearing loss and cochlear lesion with increasing duration of noise exposure," in *New Perspectives on Noise-induced Hearing Loss*, edited by R. P. Hamernik, D. Henderson, and R. Salvi (Raven, New York), pp. 283–301.

Bray, P. J. (1989). "Click evoked otoacoustic emissions and the development of a clinical otoacoustic hearing test instrument," unpublished Ph.D. thesis, Institute of Laryngology and Otology, Univ. College and Middlesex School of Medicine, London.

Cilento, B. W., Norton, S. J., and Gates, G. A. (**2003**). "The effects of aging and hearing loss on distortion product otoacoustic emissions," Otolaryngol.-Head Neck Surg. **129**, 382–389.

Cody, A. R., and Johnstone, B. M. (**1982**). "Temporary threshold shift modified by binaural acoustic stimulation," Hear. Res. **6**, 199–205.

Cody, A. R., and Robertson, D. (**1983**). "Variability of noise-induced damage in the guinea pig cochlea: electrophysiological and morphological correlates after strictly controlled exposures," Hear. Res. **9**, 55–70.

Dancer, A. (2000). "Individual susceptibility to NIHL and new perspective in treatment of acute noise trauma," in *RTO HFM Lecture Series on Damage Risk from Impulse Noise*, held in Maryland, USA, 5–6 June 2000 and Meppen, Germany, 15–16 June 2000, and published in RTO EN-11.

Desai, A., Reed, D., Cheyne, A., Richards, S., and Prasher, D. (**1999**). "Absence of otoacoustic emissions in subjects with normal audiometric thresholds implies exposure to noise," Noise Health **2**, 58–65.

Dorn, P. A., Piskorski, P., Keefe, D. H., Neely, S. T., and Gorga, M. P. (**1998**). "On the existence of an age/threshold/frequency interaction in distortion product otoacoustic emissions," J. Acoust. Soc. Am. **104**, 964–971.

Engdahl, B., Woxen, O., Arnesen, A. R., and Mair, I. W. (**1996**). "Transient evoked otoacoustic emissions as screening for hearing losses at the school for military training," Scand. Audiol. **25**, 71–78.

Franklin, D. J., McCoy, M. J., Martin, G. K., and Lonsbury-Martin, B. L. (**1992**). "Test/retest reliability of distortion-product and transiently evoked otoacoustic emissions," Ear Hear. **13**, 417–429.

Ghiselli, E. E. (**1964**). *Theory of Psychological Measurement* (McGraw-Hill, New York).

Gorga, M. P., Neely, S. T., Bergman, B. M., Beauchaine, K. L., Kaminski, J. R., Peters, J., Schulte, L., and Jesteadt, W. (**1993**). "A comparison of transient-evoked and distortion product otoacoustic emissions in normal-hearing and hearing-impaired subjects," J. Acoust. Soc. Am. **94**, 2639–2648.

Guinan, J. J., Backus, B. C., Lilaonitkul, W., and Aharonson, V. (**2003**). "Medial olivocochlear efferent reflex in humans: otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs," J. Assoc. Res. Otolaryngol. **4**, 521–540.

Hamernik, R. P., Ahroon, W. A., and Lei, S. F. (**1996**). "The cubic distortion product otoacoustic emissions from the normal and noise-damaged chinchilla cochlea," J. Acoust. Soc. Am. **100**, 1003–1012.

Hamernik, R. P., Patterson, J. H., Turrentine, G. A., and Ahroon, W. A. (**1989**). "The quantitative relation between sensory cell loss and hearing thresholds," Hear. Res. **38**, 199–211.

Henderson, D., Subramaniam, M., and Boettcher, F. A. (**1993**). "Individual susceptibility to noise-induced hearing loss: an old topic revisited," Ear Hear. **14**, 152–168.

Humes, L. E. (**1977**). "Review of four new indices of susceptibility to noise-induced hearing loss," J. Occup. Med. **19**, 116–118.

Kalluri, R., and Shera, C. A. (**2001**). "Distortion-product source unmixing: a test of the two-mechanism model for DPOAE generation," J. Acoust. Soc. Am. **109**, 622–637.

Kemp, D. T. (**1978**). "Stimulated acoustic emissions from within the human auditory system," J. Acoust. Soc. Am. **64**, 1386–1391.

Konopka, W., Pawlaczyk-Luszczynska, M., Sliwinska-Kowalska, M., Grzanka, A., and Zalewski, P. (**2005**). "Effects of impulse noise on transiently evoked otoacoustic emission in soldiers," Int. J. Audiol. **44**, 3–7.

Kummer, P., Janssen, T., and Arnold, W. (**1998**). "The level and growth behavior of the 2 f1-f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss," J. Acoust. Soc. Am. **103**, 3431–3444.

Lapsley Miller, J. A., and Marshall, L. (**2001**). "Monitoring the effects of noise with otoacoustic emissions," Semin. Hear. **22**, 393–403.

Lapsley Miller, J. A., Marshall, L., and Heller, L. M. (**2004**). "A longitudinal study of changes in evoked otoacoustic emissions and pure-tone thresholds as measured in a hearing conservation program," Int. J. Audiol. **43**, 307–322.

LePage, E. L. (**1992**). "Hysteresis in cochlear mechanics and a model for variability in noise-induced hearing loss," in *Noise-Induced Hearing Loss*, edited by A. L. Dancer, D. Henderson, R. J. Salvi, and R. P. Hamernik (Mosby-Year Book, St. Louis), pp. 106–115.

LePage, E. L. (**1998**). "Occupational noise-induced hearing loss: Origin, characterisation and prevention," Acoust. Aust. **26**, 57–61.

LePage, E. L., and Murray, N. M. (**1993**). "Click-evoked otoacoustic emissions: Comparing emission strengths with pure tone audiometric thresholds," Aust. J. Audiol. **15**, 9–22.

LePage, E. L., and Murray, N. M. (**1998**). "Latent cochlear damage in personal stereo users: a study based on click-evoked otoacoustic emissions," Med. J. Aust. **169**, 588–592.

LePage, E. L., Murray, N. M., Tran, K., and Harrap, M. J. (**1993**). "The ear as an acoustical generator: Otoacoustic emissions and their diagnostic potential," Acoust. Aust. **21**, 86–90.

Liberman, M. C., Dodds, L. W., and Learson, D. A. (**1986**). "Structure-function correlation in noise-damaged ears: A light and electron-microscopic study," in *Basic and Applied Aspects of Noise-induced Hearing Loss*, edited by R. J. Salvi, D. Henderson, R. P. Hamernik, and V. Colletti (Plenum, New York), pp. 163–177.

Luebke, A. E., and Foster, P. K. (**2002**). "Variation in inter-animal susceptibility to noise damage is associated with alpha 9 acetylcholine receptor subunit expression level," J. Neurosci. **22**, 4241–4247.

Maison, S. F., and Liberman, M. C. (**2000**). "Predicting vulnerability to acoustic injury with a noninvasive assay of olivocochlear reflex strength," J. Neurosci. **20**, 4701–4707.

Mansfield, J. D., Baghurst, P. A., and Newton, V. E. (**1999**). "Otoacoustic emissions in 28 young adults exposed to amplified music," Br. J. Audiol. **33**, 211–222.

Marshall, L., and Heller, L. M. (**1996**). "Reliability of transient-evoked otoacoustic emissions," Ear Hear. **17**, 237–254.

Marshall, L., and Heller, L. M. (**1998**). "Transient-evoked otoacoustic emissions as a measure of noise-induced threshold shift," J. Speech Lang. Hear. Res. **41**, 1319–1334.

Marshall, L., Lapsley Miller, J. A., and Heller, L. M. (**2001**). "Distortion-product otoacoustic emissions as a screening tool for noise-induced hearing loss," in *Noise Induced Hearing Loss Basic Mechanisms, Prevention and Control*, edited by D. Henderson, D. Prasher, R. D. Kopke, R. Salvi, and R. P. Hamernik (Noise Research Network, London), pp. 453–470.

Marshall, L., Lapsley Miller, J. A., Hughes, L. M., and Westhusin, L. J. (**2002**). "Changes in evoked otoacoustic emissions and hearing thresholds after a six-month deployment on an aircraft carrier," Assoc. Res. Otolaryngol. Abs., **25**, 203.

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

Muller, J., Janssen, T., Heppelmann, G., and Wagner, W. (**2005**). "Evidence for a bipolar change in distortion product otoacoustic emissions during contralateral acoustic stimulation in humans," J. Acoust. Soc. Am. **118**, 3747–3756.

Murray, N. M., and LePage, E. L. (**1993**). "Age dependence of otoacoustic emissions and apparent rates of ageing of the inner ear in an Australian population," Aust. J. Audiol. **15**, 59–70.

Murray, N. M., and LePage, E. L. (2002). "A nine-year longitudinal study of the hearing of orchestral musicians," in *International Auditory Congress*, Melbourne, Australia.

Murray, N. M., LePage, E. L., and Mikl, N. (**1998**). "Inner ear damage in an opera theatre orchestra as detected by otoacoustic emissions, pure tone audiometry and sound levels," Aust. J. Audiol. **20**, 67–78.

Navy Occupational Health and Safety Program (1999). OPNAVINST 5100.23E: Hearing conservation and noise abatement (Chief of Naval Operations, Washington DC).

Nordmann, A. S., Bohne, B. A., and Harding, G. W. (**2000**). "Histopathological differences between temporary and permanent threshold shift," Hear. Res. **139**, 13–30.

Passchier-Vermeer, W. (1993). Noise and Health (Publication No. A93/02E) (Health Council of The Netherlands, The Hague).

Rask-Andersen, H., Ekvall, L., Scholtz, A., and Schrott-Fischer, A. (**2000**). "Structural/audiometric correlations in a human inner ear with noise-induced hearing loss," Hear. Res. **141**, 129–139.

Reiter, E. R., and Liberman, M. C. (**1995**). "Efferent-mediated protection from acoustic overexposure: relation to slow effects of olivocochlear stimulation," J. Neurophysiol. **73**, 506–514.

Rovig, G. W., Bohnker, B. K., and Page, J. C. (**2004**). "Hearing health risk in a population of aircraft carrier flight deck personnel," Mil. Med. **169**, 429–432.

Saunders, J. C., Dear, S. P., and Schneider, M. E. (**1985**). "The anatomical consequences of acoustic injury: A review and tutorial," J. Acoust. Soc. Am. **78**, 833–860.

Seixas, N. S., Goldman, B., Sheppard, L., Neitzel, R., Norton, S. J., and Kujawa, S. G. (**2005a**). "Prospective noise induced changes to hearing among construction industry apprentices," Occup. Environ. Med. **62**, 309–317.

Seixas, N. S., Neitzel, R., Brower, S., Goldman, B., Somers, S., Sheppard, L., Kujawa, S. G., and Norton, S. (**2005b**). "Noise-related changes in hearing: a prospective study among construction workers," in 30th Annual NHCA National Hearing Conservation Conference, Tucson, AZ.

Shaffer, L. A., Withnell, R. H., Dhar, S., Lilly, D. J., Goodman, S. S., and Harmon, K. M. (**2003**). "Sources and mechanisms of DPOAE generation: implications for the prediction of auditory sensitivity," Ear Hear. **24**, 367–379.

Shera, C. A. (**2004**). "Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions," Ear Hear. **25**, 86–97.

Shera, C. A., and Guinan, J. J. (**1999**). "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs," J. Acoust. Soc. Am. **105**, 782–798.

Siegel, S. (**1956**). *Nonparametric Statistics for the Behavioral Sciences* (McGraw-Hill, New York).

Slepecky, N. (**1986**). "Overview of mechanical damage to the inner ear: noise as a tool to probe cochlear function," Hear. Res. **22**, 307–321.

Sutton, L. A., Lonsbury-Martin, B. L., Martin, G. K., and Whitehead, M. L. (**1994**). "Sensitivity of distortion-product otoacoustic emissions in humans to tonal over-exposure: time course of recovery and effects of lowering $L_2$," Hear. Res. **75**, 161–174.

Ward, W. D. (**1965**). "The concept of susceptibility to hearing loss," J. Occup. Med. **7**, 595–607.

Withnell, R. H., Yates, G. K., and Kirk, D. L. (**2000**). "Changes to low-frequency components of the TEOAE following acoustic trauma to the base of the cochlea," Hear. Res. **139**, 1–12.

Xu, Z. M., Van Cauwenberge, P., Vinck, B., and De Vel, E. (**1998**). "Sensitive detection of noise-induced damage in human subjects using transiently evoked otoacoustic emissions," Acta Otorhinolaryngol. Belg. **52**, 19–24.

Yankaskas, K. D. (1999). "Hearing conservation: The engineering part of the equation," Navy Med. Sept.-Oct., 21–25.

Yankaskas, K. D., and Shaw, M. F. (1999). "Landing on the roof: CVN noise," Nav. Eng. J. July, 23–34.

Yates, G. K., and Withnell, R. H. (**1999**). "The role of intermodulation distortion in transient-evoked otoacoustic emissions," Hear. Res. **136**, 49–64.

Zhou, X.-H., Obuchowski, N. A., and McClish, D. K. (**2002**). *Statistical Methods in Diagnostic Medicine* (Wiley-Interscience, New York).