J Am Acad Audiol 24:486–504 (2013)

# Computer-Automated Tinnitus Assessment: Noise-Band Matching, Maskability, and Residual Inhibition

DOI: 10.3766/jaaa.24.6.5

James A. Henry*†
Larry E. Roberts‡
Roger M. Ellingson*
Emily J. Thielman*

**Abstract**

**Background:** Psychoacoustic measures of tinnitus typically include loudness and pitch match, minimum masking level (MML), and residual inhibition (RI). We previously developed and documented a computer-automated tinnitus evaluation system (TES) capable of subject-guided loudness and pitch matching. The TES was further developed to conduct computer-aided, subject-guided testing for noise-band matching (NBM), MML, and RI.

**Purpose:** The purpose of the present study was to document the capability of the upgraded TES to obtain measures of NBM, MML, and RI, and to determine the test-retest reliability of the responses obtained.

**Research Design:** Three subject-guided, computer-automated testing protocols were developed to conduct NBM. For MML and RI testing, a 2–12 kHz band of noise was used. All testing was repeated during a second session.

**Study Sample:** Subjects meeting study criteria were selected from those who had previously been tested for loudness and pitch matching in our laboratory. A total of 21 subjects completed testing, including seven females and 14 males.

**Results:** The upgraded TES was found to be fairly time efficient. Subjects were generally reliable, both within and between sessions, with respect to the type of stimulus they chose as the best match to their tinnitus. Matching to bandwidth was more variable between measurements, with greater consistency seen for subjects reporting tonal tinnitus or wide-band noisy tinnitus than intermediate types. Between-session repeated MMLs were within 10 dB of each other for all but three of the subjects. Subjects who experienced RI during Session 1 tended to be those who experienced it during Session 2.

**Conclusions:** This study may represent the first time that NBM, MML, and RI audiometric testing results have been obtained entirely through a self-contained, computer-automated system designed specifically for use in the clinic. Future plans include refinements to achieve greater testing efficiency.

**Key Words:** Compensation, hearing disorders, loudness matching, loudness perception, pitch matching, pitch perception, rehabilitation, reliability of results, tinnitus, tinnitus diagnosis

**Abbreviations:** LM = loudness match; MML = minimum masking level; NBM = noise-band matching; PM = pitch match; RI = residual inhibition

Three decades have elapsed since the Ciba Foundation in London (Evered and Lawrenson, 1981) and the National Academy of Sciences (McFadden 1982) made recommendations to standardize tinnitus evaluation procedures. Four measures of tinnitus psychoacoustic characteristics were recommended: loudness, pitch, maskability, and residual inhibition (RI). Specific procedures for obtaining these indirect measures were provided by Vernon and Meikle (1981). The techniques described, however, required specialized instrumentation.

*VA RR&D National Center for Rehabilitative Auditory Research, VA Medical Center, Portland, OR; †Department of Otolaryngology, Oregon Health and Science University, Portland, OR; ‡Department of Psychology, Neuroscience, and Behaviour, McMaster University, Hamilton, Ontario, Canada

James A. Henry, Ph.D., VA Medical Center (NCRAR), PO Box 1034, Portland, OR 97207; E-mail: james.henry@va.gov

Funding for this study provided by Veterans Affairs Rehabilitation Research and Development (RR&D) Service (C2300R and F7070S).

Most audiologists who perform tinnitus measurement do not have access to such equipment, which may in large part explain why these procedures have never attained universal usage. Audiologists generally utilize a clinical audiometer in some fashion to obtain some or all of the measures.

Chronic neurophysiologic tinnitus is the result of pathology somewhere within the auditory system, and has been observed to be associated with almost every form of hearing loss (Nuttall et al, 2004). There are myriad perceptual manifestations of tinnitus, but little progress has been made to associate tinnitus perceptual characteristics with different forms of hearing loss. Subcategories of tinnitus undoubtedly exist, but progress identifying them has been hindered by the nonuniformity in procedures used to measure tinnitus psychoacoustic parameters. A valid classification scheme of these parameters could provide a basis on which to diagnose tinnitus etiology and to specify the optimal form of treatment. Examples of how this information has already been used include these: (1) The perceived pitch of tinnitus tends to be in the range of 2–8 kHz for noise-induced (or presbycusic) tinnitus, 250–2000 Hz for middle ear tinnitus, and 125–250 Hz for Ménière's disease (Douek and Reid, 1968). (2) There is a strong correlation between the perceived tinnitus pitch and the area of steepest decline in the audiograms of impaired ears (Meikle 1991). (3) Subjectively rated loudness of tinnitus (using either a numeric rating scale or visual analog scale of tinnitus loudness) correlates well with the level of tinnitus handicap; however, tinnitus loudness matches do not correspond with perceived loudness, nor with tinnitus handicap (Henry and Meikle, 2000). (4) Minimum masking levels (MMLs) can be used as predictors of the potential efficacy of treatment with tinnitus masking (Vernon et al, 1990). (5) It has been reported that sound therapy for tinnitus results in reduced MMLs (Jastreboff et al, 1994; Davis et al, 2007, 2008). (6) Residual inhibition (partial or complete reduction in tinnitus loudness following certain sound stimulation) occurs in 80–90% of patients with tinnitus (Henry and Meikle, 2000, Vernon and Meikle, 2000).

We have conducted research since 1995 to develop computer-automated techniques for obtaining tinnitus loudness matches (LMs) and pitch matches (PMs). These efforts have resulted in numerous studies documenting the reliability of the techniques (e.g., Henry et al 1996, 2000, 2001b; Henry, Flick, et al, 2004). A series of studies established these measures could be obtained reliably when patients were given control over level and frequency of the test stimuli when performing, respectively, LM and PM testing (Henry et al, 2006, 2009; Henry, Rheinsburg, et al, 2004). Patient control of these parameters allowed testing to be conducted rapidly, which is important for clinical application of the technique. The system was subsequently expanded to perform noise-band matching (NBM) and to obtain measures of MML and RI. The purpose of the present study was to document the system capability of performing these new tests and to determine the test-retest reliability of the responses obtained.

## METHOD

### Subjects

The limited output of the noise stimuli for NBM dictated that subjects must have tinnitus LMs within 70 dB SPL and PMs no higher than 8 kHz. Therefore, subjects meeting these criteria were selected from those who had previously been tested for LM and PM in our laboratory. A total of 21 subjects completed the new testing protocol, including seven females and 14 males (mean age = 57 yr; SD = 12; range = 28–78). Each subject attended two test sessions. Test sessions were separated by an average of 6.2 days (SD = 6.4; range = 1–25).

Ten of the subjects were U.S. military veterans. Only three of the subjects had normal hearing sensitivity (all hearing thresholds 20 dB HL or better). Figure 1 shows mean hearing thresholds for the 21 subjects. Nineteen of the subjects had binaural tinnitus, and two had unilateral tinnitus. Table 1 lists the types of external sounds subjects could pick to describe the sound of their tinnitus, as well as the percentages of subjects who chose each type of sound. Although subjects could choose more than one sound, in general the tinnitus was described as "tonal" for 16 of the subjects and "nontonal" for the remaining five. All but two of the subjects had experienced tinnitus for at least 3 yr (Fig. 2).

### Hardware and Software

Details of the system hardware and software used for this study have been described (Henry, Rheinsburg, et al, 2004). The major change to the system that could have affected patient responses was the custom-built "knob" device that enabled direct patient control over auditory stimuli during testing. Knob control of testing stimuli was developed in an effort to shorten testing time for clinical application. Use of the knob provided no cues to the patient; that is, the knob rotated smoothly (without detents) and continuously (without stops or labels) in either direction. In the aforementioned study, LM and PM testing was performed repeatedly by each subject to document that test-retest response reliability of these measures was unaffected by using the knob. In addition to the knob device, the system used a slate-type (tablet-style) computer (Aqcess Technologies Qbe Personal Computing Tablet) with a pen-touch-enabled monitor enabling patients to advance the instruction screens and respond to the auditory tests using a pen-type pointing instrument. Graphical instruction screens were developed

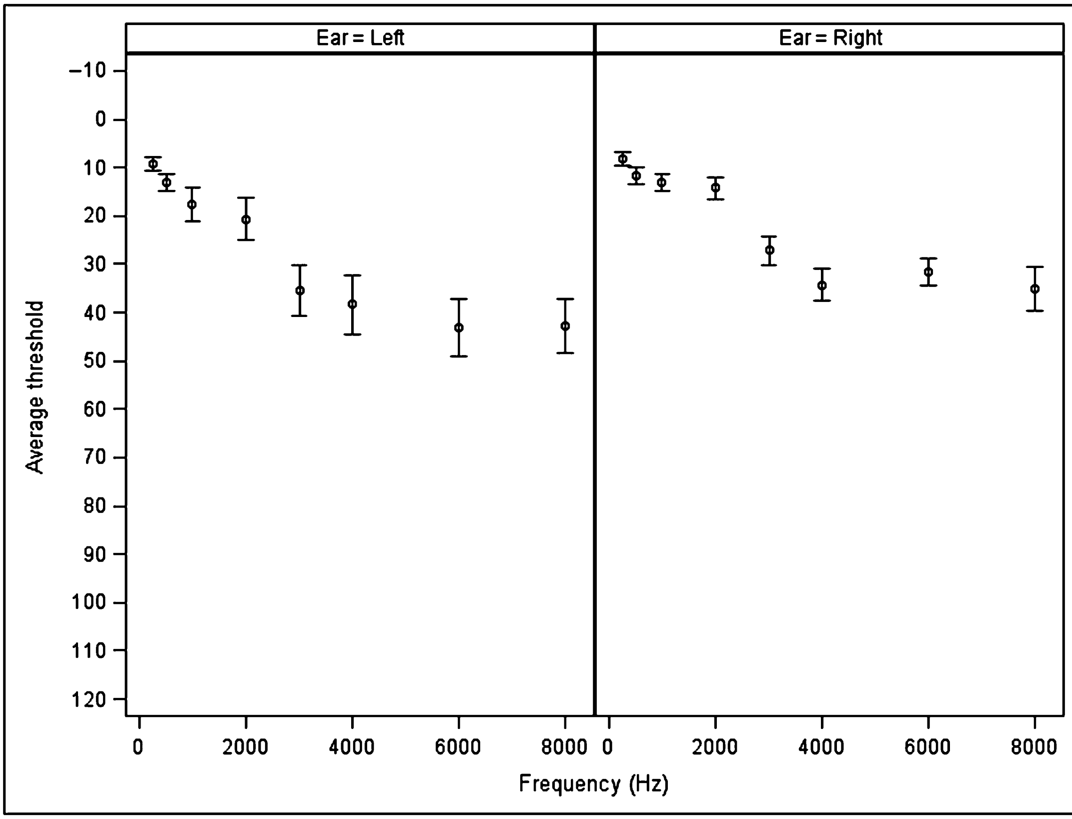


**Figure 1.** Mean hearing thresholds (dB HL) for the 21 subjects. Error bars reflect standard error.

using Microsoft PowerPoint software. Generation and control of all stimulus parameters was accomplished by the custom-built Programmable Auditory Laboratory (PAL 3000) (Flick et al, 2000). Earphones were ER-4B Canal Phone™ insert earphones (Henry et al, 2001a, 2003). System calibration was performed automatically by special computer programming as previously described (Henry et al, 1999).

For the present study, the automated system underwent further programming to incorporate the additional tinnitus measurement tests that included NBM (three methods), and measurement of MML and RI. These tests were added to the previous tests (hearing threshold, LM, PM). Additional, but similar, instruction and response screens were developed to allow test stimuli to continue to be controlled by the patient for the new tests. The new testing paradigms required the addition of new system hardware components to support stimulus generation using direct digital waveform synthesis techniques.

## Hardware

To accommodate the additional testing protocols, new hardware was integrated into the system stimulus generation path (see Fig. 3). The Creative Extigy Sound Blaster USB sound card was added to support near real-time, software generated, stimulus waveform generation and manipulation of various noise bands. In addition, the slate-type touch-screen computer was

**Table 1. Frequency of Responses for Sounds That Most Closely Resemble Tinnitus (data obtained from initial survey completed by subjects)**

| | "No" | | "Yes" | |
|---|---|---|---|---|
| Type of Sound | Frequency | Percent | Frequency | Percent |
| Ringing | 13 | 61.9 | 8 | 38.1 |
| Clear tone | 13 | 61.9 | 8 | 38.1 |
| More than one tone | 15 | 71.4 | 6 | 28.6 |
| Whistle | 20 | 95.2 | 1 | 4.8 |
| Hissing | 16 | 76.2 | 5 | 23.8 |
| Buzzing | 15 | 71.4 | 6 | 28.6 |
| Hum | 19 | 90.5 | 2 | 9.5 |
| Music | 21 | 100 | 0 | 0.0 |
| Sizzling | 19 | 90.5 | 2 | 9.5 |
| Transformer noise | 20 | 95.2 | 1 | 4.8 |
| High tension wire | 15 | 71.4 | 6 | 28.6 |
| Crickets, insects | 19 | 90.5 | 2 | 9.5 |
| Pulsating | 20 | 95.2 | 1 | 4.8 |
| Pounding | 21 | 100 | 0 | 0.0 |
| Ocean roar | 20 | 95.2 | 1 | 4.8 |
| Clicking | 21 | 100 | 0 | 0.0 |
| Other | 20 | 95.2 | 1 | 4.8 |

*Note:* Subjects chose "yes" or "no" to indicate whether the type of sound resembled their tinnitus sound.

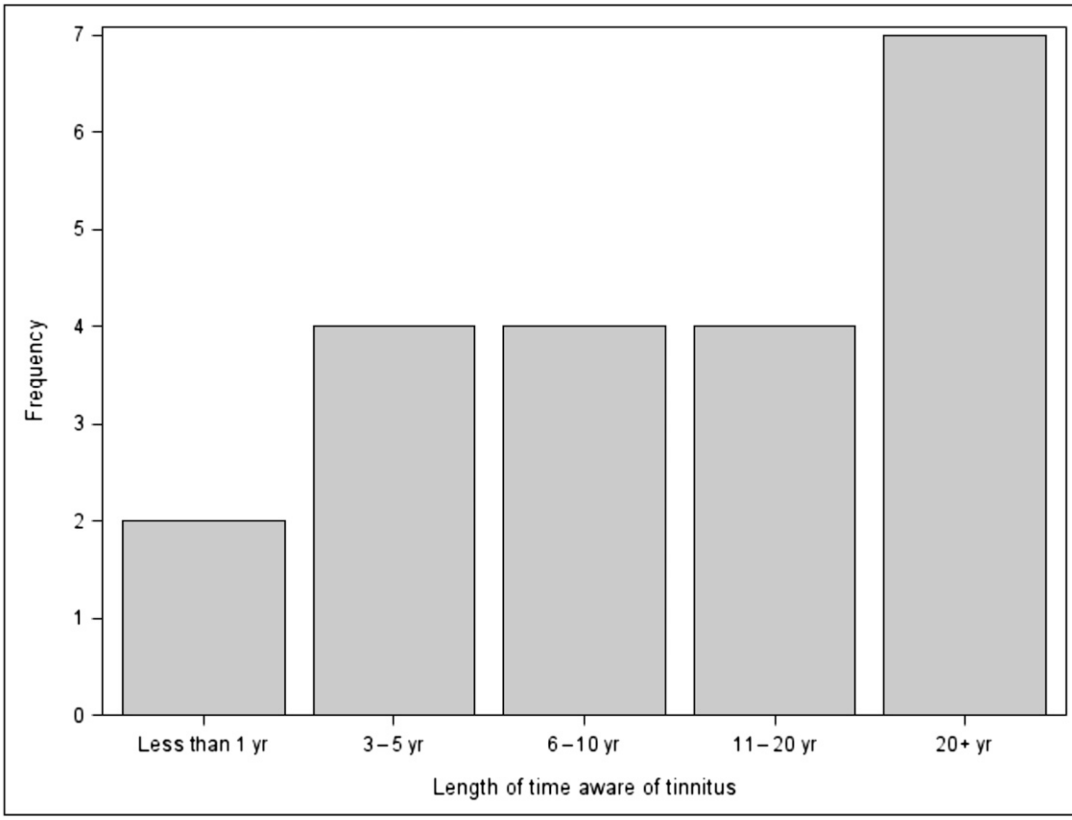


**Figure 2.** Numbers of subjects reporting tinnitus of different durations ("length of time aware of tinnitus").

replaced by an ASUS P4B-533-E Pentium IV 2.5-GHZ IBM PC. This was done to provide the required power needed to efficiently run the upgraded stimulus generation software and keep the user controlled acoustic stimulus changing highly responsive. As compared to a pure-tone sine wave, an extremely narrow bandwidth noise stimulus has very limited acoustic power. Therefore the previously used tail-end attenuator on the output of the PAL3000 headphone amplifier was removed in order to achieve adequate output levels when testing with narrow-band noise stimuli. Finally, the standard three-wire ER-4B Canal Phone™ insert earphone cabling was modified to a four-wire cable setup to decrease channel crosstalk during binaural testing.

## PowerPoint Platform Software Interface

Microsoft PowerPoint screens for subject instruction were developed for the new testing procedures. These instruction screens provided subjects with the guidance necessary for performing tinnitus band-match testing, and measurement of MML and RI. The complete subject interface was created using the "slide-show" type paradigm designed and created by user-interface designers on the project team. This stand-alone "storyboarding" of the user-interface experience greatly simplified the creation of an effective and attractive set of subject instruction and response displays.

## Custom Software

Extensive software development was necessary to support the new "more patient-friendly" PowerPoint slide-based user interface. A shared system code library was developed to exchange control and response messages between the tinnitus test system application and the user interface running as a PowerPoint application slide show.

The PAL3000 was used for all pure-tone testing as before, including pure tones used during sound-match testing. The previously used digital Noise Stimulus generator was upgraded considerably to support real-time, variable-bandwidth, variable center-frequency noise bands. Full bandwidth, stereo, acoustic transducer equalization was added using inline real-time digital filters to support acoustically calibrated variable-bandwidth stimuli. As before, a continuously varying, real-time random white noise generator was used as the foundation for noise stimuli waveform data points with the Stimulus Generator being implemented on top of the Microsoft Windows DirectX audio platform (Ellingson et al, 2004).

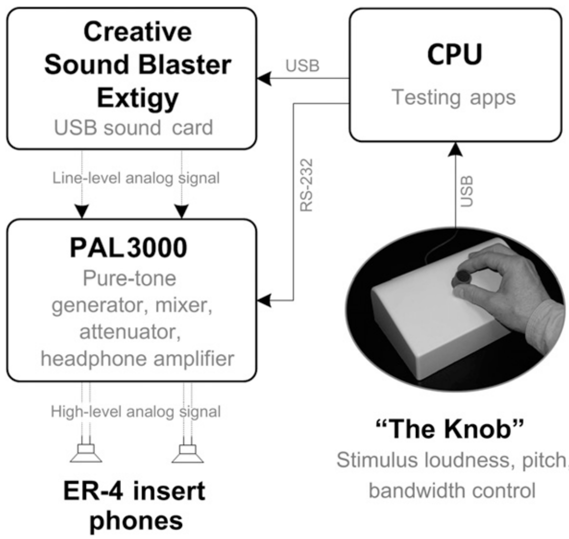


**Figure 3.** Block diagram of tinnitus testing system as modified for this study. The CPU is a standard PC-type computer running software applications to: (1) display instruction, testing, and help instructions to the subject; (2) generate digital stimulus waveforms; (3) monitor "the knob" and adjust digital stimulus parameters based on subject inputs; (4) control sound card–based digital-to-analog stimulus generator via a USB interface; and (5) control the PAL3000 signal generation, mixing, and attenuation headphone driving device via an RS-232 interface. Acoustic signals were presented to the subject via ER-4B insert-type transducers connected to the PAL3000 via a four-wire interface.

The variable-bandwidth noise stimulus generation was performed using multiple buffer segments, each with independent filter state storage. The actual noise-band shaping was implemented using stereo, digital elliptic-type bandpass filters with 96 dB attenuation, and 1 dB ripple specification. For stability reasons, the filter order varied from 9 for bandwidth greater than 6999 Hz to 3 for bandwidth less 11 Hz. A stereo, 55 tap finite impulse response (FIR) filter was used for acoustic equalization. The complete 44,100 Hz, stereo, real-time, digitally generated, acoustically equalized, noise-band stimulus generation chain was computationally intensive. For software development efficiency reasons, a fairly high-powered desktop computer was used for the testing platform. It is possible that the less powerful, previously used 400 MHz slate-style computers could be used for future systems, but such benchmark testing has not been carried out as part of this project due to time constraints.

## System Calibration

A computer controlled calibration routine was used to calibrate the tinnitus system. The PAL3000, Bruel & Kjaer 2231 sound level meter, ZI 9101 interface module, 4157 acoustic coupler, and Etymotic ER-4B Canal Phone™ transducer were used to conduct calibration procedures. Calibration values were stored in a database for online real-time access, for system documentation, and as documentation and for future calibration reference starting points.

For each stimulus channel, two separate "virtual" instrument signal paths were calibrated, one for PAL3000 generated pure-tone stimuli and one for the digitally generated noise stimuli. PAL3000 pure tones were calibrated at each test frequency for each ear using traditional methods as previously described (Henry et al, 1999). Calibration of the noise generator was quite different. The first step was to record the acoustic transducer frequency response characteristics across the full system test bandwidth. Next step was to generate the digital equalization filter coefficients necessary to flatten the transducer acoustic response to support variable-bandwidth calibrated noise band stimuli. The system was then operated with the noise generator running at full bandwidth, and the sound power output was measured using the linear scale that became the singular "cal factor" stored for the noise stimulus channel. The acoustical equalization of the actual stimulus output level was accomplished using a mathematical transformation to adjust the full bandwidth cal factor value to cal factors for other requested bandwidths.

## Testing Algorithms

The same testing protocols for determining hearing threshold, LM, and PM that were used in the previous study (Henry, Rheinsburg, et al, 2004) were used for this portion of the study. Briefly, subjects were seated in front of the computer and instructed in how to respond to the testing. One ER-4B earphone was inserted in the "stimulus ear" by the examiner. (The stimulus ear was contralateral to the predominant tinnitus if tinnitus was asymmetric. If tinnitus was symmetric, then the stimulus ear was the ear with better hearing, or it was selected randomly if hearing was also symmetric.) The computer provided all further instructions, and the examiner sat close by to be available to answer questions. The computer then obtained hearing thresholds and tinnitus LMs at 1/3-octave frequencies from 1 to 8 kHz. For pitch matching, subjects rotated the dial, resulting in the presentation of adjacent test frequencies at the loudness-matched levels. Test frequencies were presented in a saw-tooth fashion—when the highest test frequency (8 kHz) was reached, consecutively lower frequencies were presented; when the lowest test frequency (1 kHz) was reached, consecutively higher frequencies were presented. When subjects decided which frequency best matched their tinnitus, they touched "Go" on the computer screen to record this frequency and proceed to the next measurement. The pitch-matching procedure was repeated five times, resulting in five PMs. The following new testing protocols began immediately upon completion of the five repeated PMs.

### *Noise-Band Matching*

Three testing paradigms (Methods 1, 2, and 3) were developed to conduct NBM. Technical limitations with the system allowed for testing only from 1 through 8 kHz.

*Method 1.* For each subject, the mean of the five repeated PMs was determined, which identified the "center frequency" for Method 1 of the NBM testing. Subjects were instructed to "turn the knob until the sound is the closest match to your tinnitus." (Fig. 4 shows the screens that were presented to subjects to provide instructions and to guide responses during testing.)

Noise-band matching was repeated for five consecutive trials. For each trial, testing started with presentation of the center-frequency pure tone. The subject turned the knob to change the bandwidth around the center frequency (within the frequency limits). As the knob was turned either clockwise or counterclockwise, the frequency bandwidth of the tone widened exponentially, that is, the first band was 2 Hz wide, the second 4 Hz wide, the third 8 Hz wide, and so on. The noise bands increased in this manner until they reached their maximum frequency range at either end of the frequency spectrum (1 and 8 kHz). When the maximum bandwidth was achieved, further turning of the knob in the same direction caused the noise band to progressively narrow to the pure-tone starting point. Continued turning in the same direction would repeat the entire progression, whereas reversing the direction of turning would reverse the bandwidth widening/narrowing progression (in a saw-tooth fashion, similar to PM testing).

Each noise band was presented at the loudness-matched level of the tinnitus at the test frequency closest to the mean of the five PMs. Regardless of the width of the noise band, a constant sound pressure level was maintained.

Following selection of a noise bandwidth as a tinnitus match, subjects were instructed to listen again to the chosen stimulus and to confirm the accuracy of the match. Five response choices for this "Tinnitus Sound Match Check" were provided, ranging from "exact match" to "not a match," as shown in the last panel of Figure 4.

*Method 2.* For Method 2 testing, the program first calculated the mean of subjects' five bandwidths obtained during Method 1. The program then selected a center frequency at random from the test set of frequencies (1/3-octave frequencies between 1 and 8 kHz). The first noise band presented thus had a bandwidth as averaged from Method 1, and a center frequency selected at random between 1 and 8 kHz.

Subjects followed the same instructions as for Method 1 (Fig. 4) to turn the knob and select the closest match from the different noise bands. As the knob was turned, the same frequency width was maintained while the center frequency shifted across the frequency range. As for Method 1 subjects had five opportunities to select a tinnitus sound match, after which the "closeness of the match" was queried. The knob responded in a continuously variable manner with no stops as before. For each of the five trials, the initial center frequency was selected by the computer at random.

*Method 3.* For the final stage of tinnitus NBM the system averaged the results from the five sound matches selected in Method 2 to assign a new center frequency. As for Method 1, subjects then again had the opportunity to increase or decrease the stimulus frequency bandwidth from pure-tone to full bandwidth in exponential frequency steps to find the best tinnitus match. The best match and associated confirmation accuracy were completed five times.

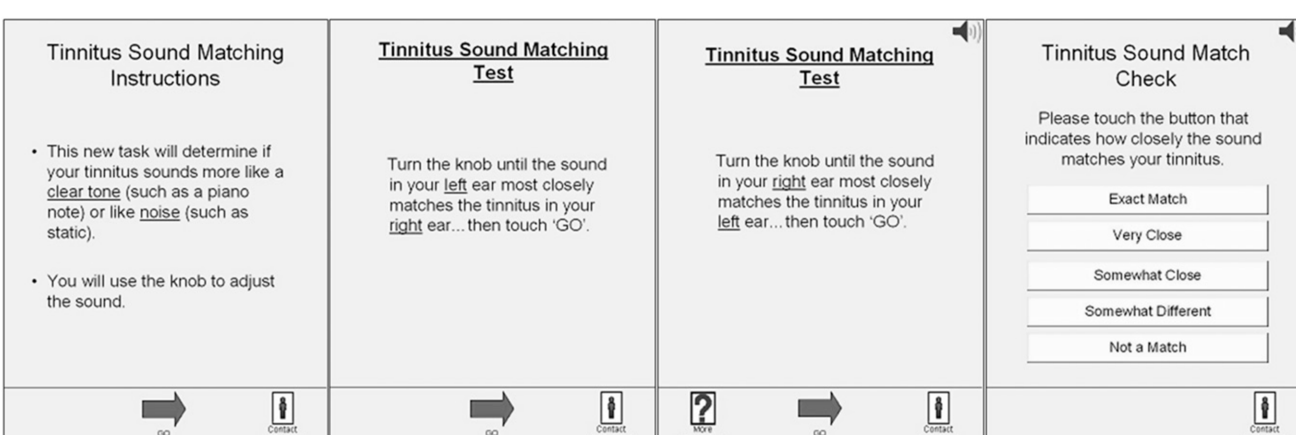


**Figure 4.** Instruction and testing screens for performing noise-band matching (NBM) with the automated system. Note that instructions are shown for both right-ear and left-ear testing, although testing was performed in only one ear (the test ear, i.e., the ear contralateral to the predominant tinnitus).

### *Minimum Masking Level*

Testing for MML used a 2–12 kHz band of noise, which is the standard stimulus for clinical measurement of MML as described by Vernon et al (1990). As this was the first test that used binaural presentation of test tones, the examiner inserted a second earphone into the "tinnitus ear" (contralateral to "stimulus ear") of subjects just prior to conducting the test. Threshold testing for the noise occurred in each ear separately, using the automated hearing-threshold technique (Henry, Rheinsburg, et al, 2004). (Fig. 5 shows the screens that were presented to subjects to provide instructions and to perform threshold testing.) Once both thresholds were obtained, the program randomly selected a noise level, to the nearest 1 dB, between 5 and 20 dB above the noise thresholds. For each ear, the noise was presented at the same sensation level (dB SL, i.e., level above threshold). For example, if the noise thresholds were 0 dB SPL for the right ear and 10 dB SPL for the left ear and the program randomly selected 8 dB SL as the starting level, then the binaural presentation of noise would start at 8 dB SPL in the right ear and 18 dB SPL in the left ear. Subjects were instructed to "turn the knob until the noise just covers the tinnitus in both ears." The noise was raised and lowered in 1 dB increments when the knob was rotated clockwise and counterclockwise, respectively. Two responses were obtained, and the average of the two responses was the final MML result. It should be noted that the duration of the stimulus at any level was variable according to the amount of time subjects chose to listen to the stimulus at each level.

### *Residual Inhibition*

*RI* refers to the common phenomenon of a temporary reduction in loudness of tinnitus (or temporary silencing of tinnitus) as a result of stimulation with (typically) broad-band noise. As with MML testing, a 2–12 kHz band of noise was used for clinical testing of RI (Vernon et al, 1990). The intensity level of the sound was presented binaurally at 10 dB above the MML value that was established for each ear. Subjects were instructed to listen to the noise for 1 min, at which time the noise signal was terminated and they could report the loudness of their tinnitus. (Fig. 6 shows the screens that were presented to subjects to provide instructions and to perform testing.) If the stimulus was too loud during the test, subjects could touch the "too loud" button on the screen. Each time this was done the stimulus decreased by 5 dB. During the poststimulus time, subjects repeatedly reported the "percentage loudness" of their tinnitus until the loudness returned to "100 percent of its usual loudness" or until 5 min (from when the stimulus stopped) had elapsed.

### Time of Testing

Time of testing to complete the battery of tests (pure-tone thresholds, LM, PM, NBM [three methods], MML, and RI) averaged 47 min for Session 1 (SD = 24; range = 21–110), and 45 min for Session 2 (SD = 25; range = 20–110). These times reflect actual testing time and do not include time required to prepare subjects for testing (i.e., otoscopy, explaining overall function of the system, and inserting earphones).

## RESULTS

### NBM: Method 1

Table 2 shows an example of subject individual responses for Method 1 of NBM. The frequencies shown in the table indicate the range of frequencies subjects chose as a match to their tinnitus. Table 2 shows only results from the first of five consecutive matches during the first session (each subject completed an additional four matches during Session 1 and five additional matches during Session 2). The column under the heading "Bandwidth

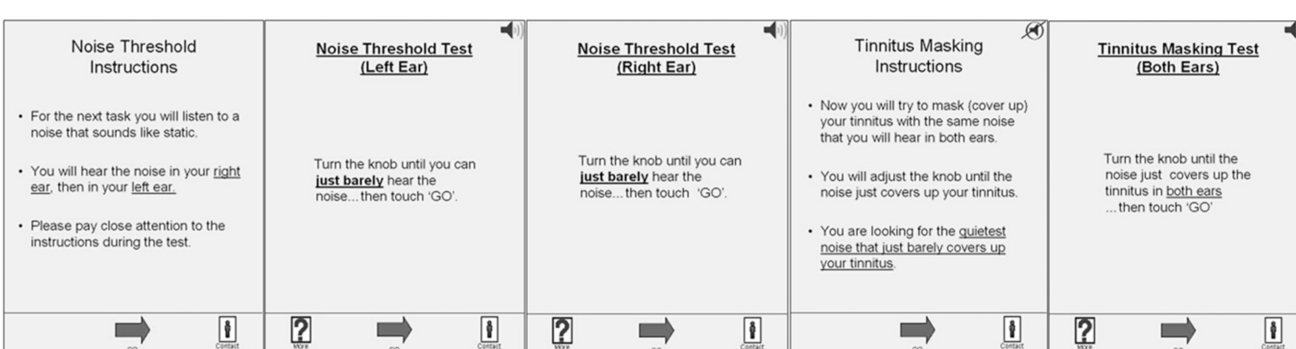


**Figure 5.** Instruction and testing screens for obtaining tinnitus minimum masking levels (MMLs) with the automated system. This was the first test that used earphones in both ears. These screens indicate monaural presentation of stimuli for "noise thresholds" in each ear separately, followed by binaural presentation of stimuli to determine the MML.

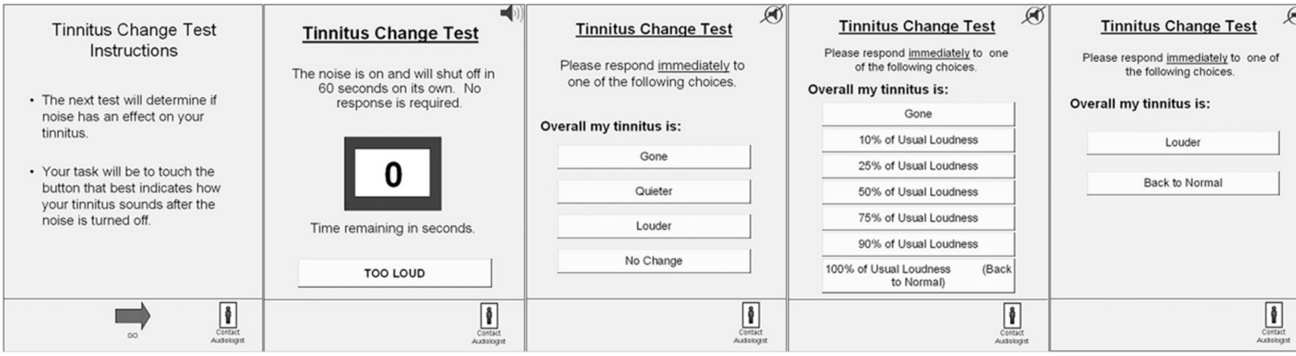


**Figure 6.** Instruction and testing screens for measuring tinnitus residual inhibition (RI) with the automated system. RI testing was done with binaural presentation of 2–12 kHz bandwidth noise.

of Match" shows that subjects chose bandwidths ranging between 0 and 2048 Hz. (A bandwidth of 0 Hz indicates the subject chose a pure tone as a closer match than any of the bandwidth choices.) Eleven of the subjects chose a pure tone as the closest match for this first trial. Only three subjects chose a bandwidth exceeding 32 Hz. When asked to confirm how close the match was, only three subjects indicated their choice was "not a match." Other choices included "somewhat close," "very close," and "exact match."

Tables 3 and 4 show, for Sessions 1 and 2 respectively, all of the subjects' bandwidth choices for Method 1 of NBM. It can be seen that subjects were generally consistent with respect to the bandwidths of their repeated matches. That is, subjects who chose 0 Hz tended to choose 0 Hz or very narrow bands (2, 4, or 8 Hz) repeatedly. Subjects who chose wider bands tended to be consistent in choosing wide bands of noise. There were of course exceptions to these trends (note subjects 6 and 8 in Table 3). Subjects were also generally consistent between sessions, although there were again some subjects who were very inconsistent between sessions (note subjects 11, 13, 15, and 16 between Tables 3 and 4). Notwithstanding these inconsistencies a significant test-retest correlation was found ($r = 0.563$, $p = 0.0078$). Inspection of this relationship in Figure 7 and Tables 3 and 4 shows that 16 subjects chose bandwidths of

**Table 2. Results of the First of Five Repeated Noise-Band Matches during Session 1, for Noise-Band Matching with Method 1**

| Subject | Lower Stop Frequency (Hz) | Center Frequency (Hz) | Upper Stop Frequency (Hz) | Bandwidth of Match (Hz) | Subject's Judgment of Match |
|---|---|---|---|---|---|
| 1 | 8000 | 8000 | 8000 | 0 | Very close |
| 2 | 6349 | 6350 | 6351 | 2 | Very close |
| 3 | 5037 | 5040 | 5041 | 4 | Somewhat close |
| 4 | 8000 | 8000 | 8000 | 0 | Very close |
| 5 | 6350 | 6350 | 6350 | 0 | Somewhat close |
| 6 | 4784 | 5040 | 5296 | 512 | Not a match |
| 7 | 8000 | 8000 | 8000 | 0 | Very close |
| 8 | 2000 | 2000 | 2000 | 0 | Not a match |
| 9 | 6334 | 6350 | 6366 | 32 | Very close |
| 10 | 3175 | 3175 | 3175 | 0 | Very close |
| 11 | 3159 | 3175 | 3191 | 32 | Very close |
| 12 | 3175 | 3175 | 3175 | 0 | Very close |
| 13 | 8000 | 8000 | 8000 | 0 | Exact match |
| 14 | 3174 | 3175 | 3176 | 2 | Very close |
| 15 | 5040 | 5040 | 5040 | 0 | Very close |
| 16 | 8000 | 8000 | 8000 | 0 | Somewhat close |
| 17 | 5326 | 6350 | 7374 | 2048 | Somewhat close |
| 18 | 2519 | 2520 | 2521 | 2 | Very close |
| 19 | 5032 | 5040 | 5048 | 16 | Somewhat close |
| 20 | 7744 | 8000 | 8000 | 256 | Not a match |
| 21 | 8000 | 8000 | 8000 | 0 | Very close |
| **Mean** | 5581 | 5656 | 5719 | 138 | |
| **SD** | 2104 | 2118 | 2142 | 454 | |

**Table 3. Bandwidths (Hz) of Noise Matches for Noise-Band Matching, Method 1 (Session 1)**

| | Bandwidth (Hz) of Noise-Band Match | | | | | | |
|---|---|---|---|---|---|---|---|
| Subject | Trial 1 | Trial 2 | Trial 3 | Trial 4 | Trial 5 | Mean | SD |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 2 | 2 | 64 | 0 | 0 | 14 | 28 |
| 3 | 4 | 2 | 2 | 2 | 2 | 2 | 1 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 512 | 4096 | 0 | 4096 | 0 | 1741 | 2160 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 7000 | 4 | 4 | 16 | 1405 | 3128 |
| 9 | 32 | 4 | 136 | 64 | 37 | 55 | 50 |
| 10 | 0 | 0 | 2 | 0 | 2 | 1 | 1 |
| 11 | 32 | 0 | 0 | 0 | 0 | 6 | 14 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 2 | 8 | 4 | 0 | 1 | 3 | 3 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 2048 | 5746 | 384 | 2048 | 2122 | 2470 | 1972 |
| 18 | 2 | 4 | 2 | 2 | 16 | 5 | 6 |
| 19 | 16 | 8 | 2 | 2 | 0 | 6 | 7 |
| 20 | 256 | 256 | 128 | 256 | 128 | 205 | 70 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Mean** | 138 | 816 | 35 | 308 | 111 | 282 | 354 |
| **SD** | 454 | 2060 | 90 | 976 | 462 | 689 | 886 |

**Table 4. Bandwidths (Hz) of Noise Matches for Noise-Band Matching, Method 1 (Session 2)**

| | Bandwidth (Hz) of Noise-Band Match | | | | | | |
|---|---|---|---|---|---|---|---|
| Subject | Trial 1 | Trial 2 | Trial 3 | Trial 4 | Trial 5 | Mean | SD |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 2 | 0 | 256 | 2 | 0 | 52 | 114 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 512 | 0 | 0 | 0 | 0 | 102 | 229 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 16 | 8 | 0 | 4 | 8 | 7 | 6 |
| 9 | 32 | 32 | 128 | 16 | 64 | 54 | 45 |
| 10 | 0 | 0 | 2 | 0 | 0 | 0 | 1 |
| 11 | 128 | 0 | 256 | 128 | 128 | 128 | 91 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 512 | 64 | 16 | 16 | 2 | 122 | 219 |
| 14 | 2 | 0 | 0 | 0 | 0 | 0 | 1 |
| 15 | 512 | 256 | 0 | 4096 | 128 | 998 | 1742 |
| 16 | 128 | 1024 | 512 | 128 | 1024 | 563 | 449 |
| 17 | 2048 | 2048 | 2048 | 1024 | 1024 | 1638 | 561 |
| 18 | 0 | 2 | 2 | 4 | 8 | 3 | 3 |
| 19 | 0 | 0 | 4 | 4 | 8 | 3 | 3 |
| 20 | 128 | 128 | 128 | 256 | 512 | 230 | 167 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Mean** | 191 | 179 | 160 | 270 | 138 | 186 | 173 |
| **SD** | 462 | 486 | 452 | 905 | 316 | 411 | 392 |

0 Hz or close to 0 on both tests; one subject chose bandwidths exceeding 1500 Hz on both tests; and the remaining four gave discrepant results between sessions. Comparing the grand means and means of standard deviations of the bandwidths between sessions, these numbers became smaller during the second session, indicating that, for the second session, subjects as a group chose narrower bands of noise and deviated less in response choices from trial to trial. Paired $t$-tests, however, determined that the between-sessions differences in means were not significant ($p > .05$), nor were the between-sessions differences in mean standard deviations ($p > .05$).

Table 5 provides a summary of how subjects responded following each noise-band match when asked to indicate how closely the stimulus they just selected matched their tinnitus. For the first session, responses were fairly normally distributed across the different response choices. For the second session, responses were more skewed toward responses indicating closer subjective matches ("very close" and "exact match"). Notably, across all trials, the response option "somewhat different" was selected only once.

## NBM: Method 2

Method 2 maintained a constant bandwidth, and turning the knob changed the center frequency of the stimulus. Tables 6 and 7 show, for Sessions 1 and 2 respectively, the center frequencies that were selected by each subject during each of the five trials for Method 2.

If, for a single subject, the average bandwidth as determined from Method 1, Session 1 consisted of a tone

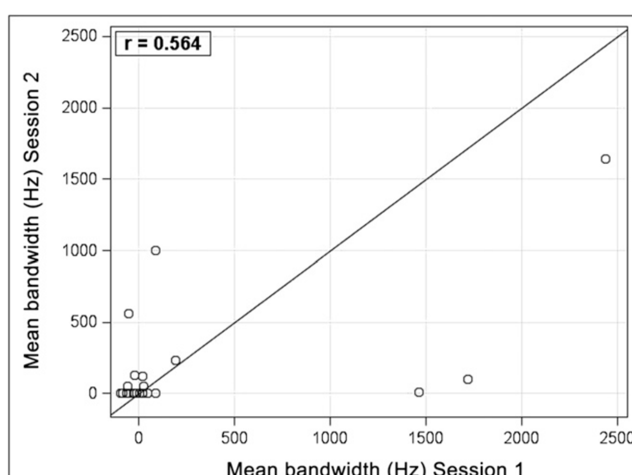


**Figure 7.** Scatterplot showing correlations of bandwidth responses between Sessions 1 and 2 for Method 1. Note that the overlapping circles represent multiple points that all occur at 0,0 on the graph. The diagonal line is the line of unity, that is, the line of best fit if Session 1 and 2 values were identical for all subjects.

**Table 5. Numbers of Subjects, for Each of the 10 Individual Noise-Band Matches from Method 1, Choosing the Different Response Options When Asked If Their Previous Noise-Band Selection Matched Their Tinnitus**

| | Session 1 | | | | | Session 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Subject's Impression of Match | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 |
| Not a match | 3 | 3 | 2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 |
| Somewhat different | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Somewhat close | 5 | 6 | 7 | 4 | 5 | 4 | 4 | 5 | 4 | 4 |
| Very close | 12 | 9 | 7 | 8 | 8 | 12 | 11 | 9 | 9 | 8 |
| Exact match | 1 | 3 | 5 | 6 | 6 | 4 | 4 | 6 | 7 | 8 |

(indicated by a "0" for the "Mean" bandwidth in Table 3), then for Method 2 turning the knob changed the frequency of the tone (equivalent to pitch matching with pure tones). In Tables 6 and 7, subject numbers have an asterisk if these subjects were responding to tones during Method 2.

A strong linear correlation was observed between sessions for the mean noise-band matches ($r = 0.831, p < 0.0001$). Figure 8 shows that subjects choosing center frequencies at or near 8 kHz appeared to be as consistent between sessions as subjects choosing lower frequencies as their best tinnitus match. The within-subjects reliability of the repeated noise-band choices during Method 2 within a session would be indicated by the standard deviations (last column of Tables 6 and 7). These standard deviations averaged across subjects resulted in a mean standard deviation of 1019 Hz for Session 1 and 641 Hz for Session 2. While this difference might suggest an improvement in within-subjects test-retest reliability from Session 1 to Session 2, a paired $t$-test revealed that the difference was not significant ($p > .05$).

The means and standard deviations across measures within sessions for the nine subjects who listened to tones during Method 2 (indicated by asterisks in Tables 6 and 7) were calculated and compared to those of the 12 subjects who listened to bands of noise during Method 2. For Session 1, the mean standard deviation was 668 Hz for the subjects who listened to tones and 1282 Hz for the subjects who listened to bands of noise. Similarly, for Session 2, the mean standard deviation was 333 Hz for the nine subjects who listened to tones and 871 Hz for the subjects who listened to bands of noise. Thus, for both Sessions 1 and 2, the standard deviations

**Table 6. Center Frequency (Hz) of Noise Matches for Noise-Band Matching, Method 2 (Session 1)**

| | Center Frequency (Hz) of Noise-Band Match | | | | | | |
|---|---|---|---|---|---|---|---|
| Subject | Trial 1 | Trial 2 | Trial 3 | Trial 4 | Trial 5 | Mean | SD |
| 1* | 8000 | 2000 | 8000 | 8000 | 8000 | 6800 | 2683 |
| 2 | 3175 | 5040 | 6350 | 5040 | 6350 | 5191 | 1303 |
| 3 | 8000 | 5040 | 6350 | 5040 | 8000 | 6486 | 1482 |
| 4* | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 5* | 6350 | 8000 | 6350 | 6350 | 6350 | 6680 | 738 |
| 6 | 1260 | 5040 | 5040 | 1585 | 5040 | 3593 | 1985 |
| 7* | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 8 | 5040 | 4000 | 6350 | 3175 | 2520 | 4217 | 1520 |
| 9 | 6350 | 8000 | 8000 | 8000 | 8000 | 7670 | 738 |
| 10 | 5040 | 5040 | 1585 | 5040 | 1000 | 3541 | 2063 |
| 11 | 3175 | 2520 | 3175 | 4000 | 8000 | 4174 | 2202 |
| 12* | 3175 | 3175 | 2520 | 2520 | 2520 | 2782 | 359 |
| 13* | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 14 | 2520 | 3175 | 3175 | 2520 | 3175 | 2913 | 359 |
| 15* | 5040 | 5040 | 2520 | 2520 | 2000 | 3424 | 1490 |
| 16* | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 17 | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 18 | 4000 | 1000 | 5040 | 5040 | 5040 | 4024 | 1749 |
| 19 | 4000 | 6350 | 1000 | 4000 | 2520 | 3574 | 1988 |
| 20 | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 21* | 8000 | 6350 | 8000 | 8000 | 8000 | 7670 | 738 |
| **Mean** | 5768 | 5608 | 5784 | 5659 | 5929 | 5749 | 1019 |
| **SD** | 2285 | 2310 | 2470 | 2336 | 2562 | 2061 | 888 |

*Subjects who responded to pure tones during Method 2 (all other subjects listened to a band of noise as averaged from Method 1 results).

**Table 7. Center Frequency (Hz) of Noise Matches for Noise-Band Matching, Method 2 (Session 2)**

| Subject | Center Frequency (Hz) of Noise-Band Match | | | | | | |
|---|---|---|---|---|---|---|---|
| | Trial 1 | Trial 2 | Trial 3 | Trial 4 | Trial 5 | Mean | SD |
| 1* | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 2 | 8000 | 8000 | 6350 | 6350 | 8000 | 7340 | 904 |
| 3 | 4000 | 4000 | 8000 | 4000 | 5040 | 5008 | 1732 |
| 4* | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 5* | 6350 | 6350 | 6350 | 6350 | 6350 | 6350 | 0 |
| 6 | 6350 | 6350 | 6350 | 6350 | 6350 | 6350 | 0 |
| 7* | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 8 | 5040 | 5040 | 5040 | 2520 | 3175 | 4163 | 1223 |
| 9 | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 10 | 5040 | 3175 | 3175 | 3175 | 6350 | 4183 | 1456 |
| 11 | 4000 | 5040 | 5040 | 6350 | 6350 | 5356 | 1002 |
| 12* | 2520 | 2520 | 3175 | 2520 | 2520 | 2651 | 293 |
| 13* | 8000 | 6350 | 6350 | 8000 | 8000 | 7340 | 904 |
| 14 | 2520 | 2520 | 2520 | 2520 | 1585 | 2333 | 418 |
| 15* | 4000 | 4000 | 5040 | 2520 | 4000 | 3912 | 899 |
| 16* | 8000 | 8000 | 6350 | 6350 | 6350 | 7010 | 904 |
| 17 | 8000 | 8000 | 8000 | 8000 | 8000 | 8000 | 0 |
| 18 | 1585 | 1000 | 1585 | 2520 | 1000 | 1538 | 622 |
| 19 | 5040 | 1000 | 4000 | 5040 | 5040 | 4024 | 1749 |
| 20 | 6350 | 8000 | 5040 | 8000 | 8000 | 7078 | 1345 |
| 21* | 6350 | 6350 | 6350 | 6350 | 6350 | 6350 | 0 |
| **Mean** | 5864 | 5605 | 5748 | 5663 | 5927 | 5761 | 641 |
| **SD** | 2128 | 2462 | 1983 | 2234 | 2270 | 2072 | 624 |

*Subjects who responded to pure tones during Method 2 (all other subjects listened to a band of noise as averaged from Method 1 results).

were almost twice as large when subjects listened to bands of noise versus those who listened to pure tones. Unpaired *t*-tests revealed, however, that these within-session differences were not significant ($p > .05$). The test-retest reliability between subjects was higher for subjects matching with pure tones ($r = 0.93$, $p < 0.0003$) than with noise bands ($r = 0.74$, $p < 0.006$), although the difference between these correlations was not significant.

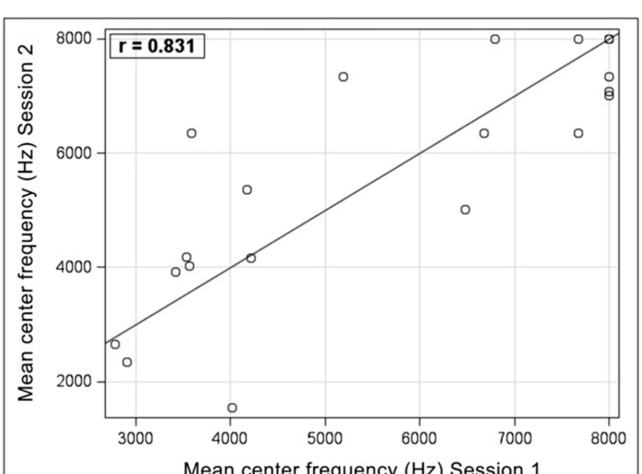


**Figure 8.** Scatterplot showing correlations of mean center-frequency responses between Sessions 1 and 2 for Method 2. The diagonal line is the line of unity, that is, the line of best fit if Session 1 and 2 values were identical for all subjects.

## NBM: Method 3

For Method 3, center frequencies obtained from the Method 2, Session 1 repeated tests were averaged to determine the center frequency for Method 3. Turning the knob widened or narrowed the bandwidth around the center frequency as done in Method 1.

Tables 8 and 9 show, for Sessions 1 and 2 respectively, all individual bandwidth choices for Method 3 of noise-band matching. As for Method 1, subjects were generally consistent with respect to the bandwidths of their repeated matches. Subjects 6 and 8 were again exceptions to these otherwise consistent trends. Subjects were also generally reliable between sessions; although there were some subjects who were very unreliable between sessions for Method 1 (subjects 11, 13, 15, and 16), these same subjects were not seen to be particularly unreliable for Method 3.

When comparing the grand mean and mean standard deviations between sessions for Method 3, these numbers became smaller during the second session. This trend, also seen for Method 1, would reflect that, during the second session, subjects as a group chose narrower

**Table 8. Bandwidths (Hz) of Noise Matches for Noise-Band Matching, Method 3 (Session 1)**

| Subject | Bandwidth (Hz) of Noise-Band Match | | | | | | |
|---|---|---|---|---|---|---|---|
| | Trial 1 | Trial 2 | Trial 3 | Trial 4 | Trial 5 | Mean | SD |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 512 | 0 | 0 | 0 | 4096 | 922 | 1788 |
| 3 | 2 | 0 | 0 | 0 | 2 | 1 | 1 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 256 | 4096 | 0 | 0 | 870 | 1807 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 110 | 1024 | 32 | 512 | 256 | 387 | 400 |
| 9 | 4 | 2 | 8 | 128 | 4 | 29 | 55 |
| 10 | 0 | 2 | 0 | 2 | 0 | 1 | 1 |
| 11 | 425 | 2 | 0 | 0 | 0 | 85 | 190 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 1 | 4 | 0 | 0 | 32 | 7 | 14 |
| 15 | 256 | 384 | 256 | 128 | 64 | 218 | 125 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 1024 | 1024 | 1024 | 1024 | 1024 | 1024 | 0 |
| 18 | 16 | 136 | 2 | 4 | 8 | 33 | 58 |
| 19 | 16 | 2 | 5 | 16 | 16 | 11 | 7 |
| 20 | 512 | 256 | 128 | 128 | 128 | 230 | 167 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Mean** | 137 | 147 | 264 | 92 | 268 | 182 | 220 |
| **SD** | 268 | 311 | 907 | 243 | 906 | 333 | 534 |

**Table 9. Bandwidths (Hz) of Noise Matches for Noise-Band Matching, Method 3 (Session 2)**

| | Bandwidth (Hz) of Noise-Band Match | | | | | | |
|---|---|---|---|---|---|---|---|
| Subject | Trial 1 | Trial 2 | Trial 3 | Trial 4 | Trial 5 | Mean | SD |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 4 | 2 | 2 | 2 | 0 | 2 | 1 |
| 9 | 32 | 32 | 16 | 32 | 14 | 25 | 9 |
| 10 | 0 | 0 | 16 | 4 | 0 | 4 | 7 |
| 11 | 8 | 2 | 8 | 2 | 16 | 7 | 6 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 2 | 0 | 2 | 0 | 1 | 1 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 512 | 128 | 512 | 512 | 1024 | 538 | 319 |
| 16 | 512 | 512 | 512 | 32 | 1024 | 518 | 351 |
| 17 | 1024 | 1024 | 1024 | 1024 | 1024 | 1024 | 0 |
| 18 | 0 | 2 | 2 | 2 | 2 | 2 | 1 |
| 19 | 0 | 2 | 0 | 0 | 2 | 1 | 1 |
| 20 | 64 | 256 | 256 | 64 | 256 | 179 | 105 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Mean** | 103 | 93 | 112 | 80 | 160 | 110 | 38 |
| **SD** | 260 | 246 | 262 | 243 | 366 | 263 | 101 |

bands of noise, and they were more consistent in their response choices from test to test. Paired $t$-tests, however, determined that the between-sessions differences in means were not significant ($p > .05$), nor were the between-sessions differences in standard deviations ($p > .05$). The between-subject correlation obtained for Method 3 ($r = 0.451$, $p = 0.04$; Fig. 9) was lower

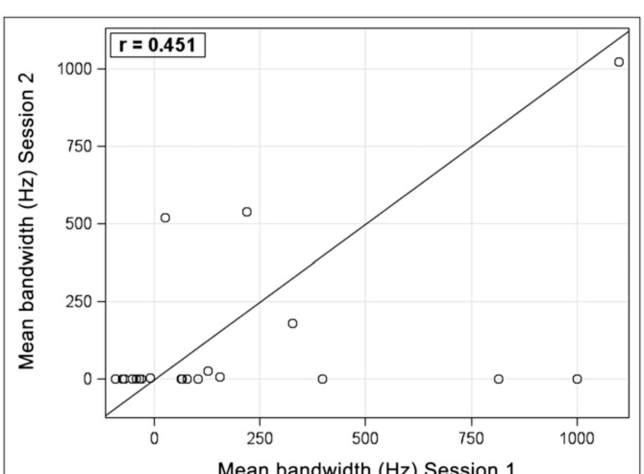


**Figure 9.** Scatterplot showing correlations of mean bandwidth responses between Sessions 1 and 2 for Method 3. Note that the overlapping circles represent multiple points that all occur at 0,0 on the graph. The diagonal line is the line of unity, that is, the line of best fit if Session 1 and 2 values were identical for all subjects.

but not notably different from that reported above for Method 1 ($r = 0.565$, $p = 0.0078$; Fig. 7).

Further inspection of the results for Methods 1 and 3 pointed to sources of consistency and inconsistency in assessing the bandwidth of tinnitus. From Figure 7 it can be seen that the correlation between the two sessions of Method 1 was strongly influenced by the single subject (#17) who gave the largest bandwidth matches in the two sessions, exceeding 1500 Hz in both sessions (see Tables 3 and 4). This subject also gave the largest bandwidth matches in Method 3, exceeding 1000 Hz in both sessions of this method (Tables 8 and 9). The subject was therefore consistent in reporting a tinnitus of comparatively wide bandwidth in all four measurements. Figure 7 also reveals a cluster of subjects (n = 10) who gave bandwidth matches averaging 0 or 1 Hz in the first session of Method 1. Seven of these 10 subjects repeated their ratings across the four bandwidth measurements of Methods 1 and 3. These results suggest that subjects experiencing either a narrow tonal tinnitus or a noisy tinnitus can report consistently on the bandwidth of their tinnitus. Subjects reporting intermediate bandwidth matches were less consistent between the sessions with each method.

## Minimum Masking Levels

Prior to testing for MML, threshold testing of the 2–12 kHz noise band was done monaurally in each ear. Table 10 shows results of this threshold testing. These results indicate there was no particular trend for subjects to provide either higher or lower noise thresholds when testing was repeated across sessions. The absolute values of the differences show, on average, that differences in noise thresholds between sessions averaged 4.0 dB in the left ears and 4.3 dB in the right ears.

Table 11 shows results of binaural MML testing. Any negative MML would indicate that the MML was below the threshold level. These negative values only occurred for two subjects during Session 1 and for one subject during Session 2. The negative values for MML were small and are thus explainable as normal variability for both thresholds and for MML (these subjects probably had an MML very close to their noise threshold). All but three of the subjects provided repeated MMLs within 10 dB of each other across sessions. The three more variable subjects differed by 14, 29, and 34 dB between sessions. The results in Table 11 indicate a significant test-retest reliability of $r = 0.808$, $p = 0.0001$ (Fig. 10).

MML measures have been used by some investigators as a surrogate for tinnitus loudness measures. We were able to test this relationship by evaluating the correlation between the MMLs (from Session 1) and the subjective ratings of tinnitus loudness (obtained

**Table 10. Monaural Thresholds for 2–12 kHz Noise, and Differences in Thresholds between Sessions**

| | Left Ear | | | | Right Ear | | | |
|---|---|---|---|---|---|---|---|---|
| Subject | Session 1 | Session 2 | Between-Session Difference | Absolute Value of Difference | Session 1 | Session 2 | Between-Session Difference | Absolute Value of Difference |
| 1 | 36 | 36 | 0 | 0 | 31 | 33 | 2 | 2 |
| 2 | 46 | 44 | −2 | 2 | 54 | 48 | −6 | 6 |
| 3 | 44 | 44 | 0 | 0 | 54 | 52 | −2 | 2 |
| 4 | 34 | 33 | −1 | 1 | 40 | 44 | 4 | 4 |
| 5 | 68 | 74 | 6 | 6 | 54 | 58 | 4 | 4 |
| 6 | 48 | 48 | 0 | 0 | 48 | 46 | −2 | 2 |
| 7 | 40 | 43 | 3 | 3 | 40 | 44 | 4 | 4 |
| 8* | 101 | 101 | 0 | 0 | 24 | 25 | 1 | 1 |
| 9 | 22 | 24 | 2 | 2 | 38 | 34 | −4 | 4 |
| 10 | 30 | 29 | −1 | 1 | 46 | 57 | 11 | 11 |
| 11 | 70 | 64 | −6 | 6 | 58 | 62 | 4 | 4 |
| 12 | 52 | 49 | −3 | 3 | 50 | 47 | −3 | 3 |
| 13 | 49 | 46 | −3 | 3 | 47 | 46 | −1 | 1 |
| 14 | 54 | 36 | −18 | 18 | 54 | 54 | 0 | 0 |
| 15 | 52 | 50 | −2 | 2 | 63 | 56 | −7 | 7 |
| 16 | 64 | 56 | −8 | 8 | 36 | 34 | −2 | 2 |
| 17 | 90 | 90 | 0 | 0 | 58 | 55 | −3 | 3 |
| 18 | 40 | 56 | 16 | 16 | 48 | 68 | 20 | 20 |
| 19 | 31 | 24 | −7 | 7 | 47 | 43 | −4 | 4 |
| 20 | 26 | 24 | −2 | 2 | 24 | 22 | −2 | 2 |
| 21 | 64 | 68 | 4 | 4 | 57 | 61 | 4 | 4 |
| **Mean** | 51 | 49 | −1.0 | 4.0 | 46 | 47 | 0.9 | 4.3 |
| **SD** | 20 | 21 | 6.3 | 5.0 | 11 | 12 | 6.1 | 4.3 |

*This subject could not hear the noise in her left ear; 101 is the code for "no response."

on the initial survey). Results revealed a Pearson product-moment correlation of 0.19 ($p > .05$), which suggested at best a weak relationship between the two measures.

## Residual Inhibition

Tables 12 and 13 summarize results of RI testing for Sessions 1 and 2, respectively. Of the 21 subjects, nine did not experience any RI during the first session, and eight did not experience any during the second session. Two of the subjects who did not experience any RI during the first session (#5 and #8) did experience it during the second session. One subject who experienced RI during the first session (#14) did not have any RI during the second session.

Of the 12 subjects who experienced RI during the first session, eight experienced partial RI (tinnitus reduced in loudness) and four experienced complete RI (tinnitus inaudible). Of the 13 subjects who experienced RI during the second session, seven experienced partial RI and six experienced complete RI.

Including all subjects, whether they experienced RI, the mean duration of RI was 79 sec for Session 1 and 113 sec for Session 2. A fairly strong linear correlation was observed between sessions for the mean duration of RI ($r = 0.706$, $p = 0.003$) (Fig. 11). Recalculating the mean RI duration only for those who experienced any RI, the RI lasted an average of 137 sec for Session 1 and 182 sec for Session 2.

During the first session, 10 of the subjects opted to reduce the output of the noise during the 1 min of RI testing, due to the stimulus being perceived as uncomfortably loud. Six subjects reduced the output during the second session, which included five of those who reduced the output during the first session.

## DISCUSSION

A group of 21 subjects completed a tinnitus test battery over two sessions with the automated tinnitus assessment system. Each of the subjects was able to complete the testing protocol with minimal intervention from the attending research audiologist who sat in the adjoining control room. The current tools differ from automated methods devised to measure tinnitus attributes in the laboratory (Noreña et al, 2002, Roberts et al, 2008) in that NBM, MML, and RI were obtained through automation using a self-contained system designed specifically for use in clinical settings. We discuss the results for each tinnitus attribute separately, considering relation to previous results and factors that may contribute to reliability.

**Table 11. Binaural Minimum Masking Levels (MMLs), in dB Sensation Level (SL)**

| Subject | MML (dB SL) Session 1 | MML (dB SL) Session 2 | Between-Session Difference | Absolute Value of Difference |
|---|---|---|---|---|
| 1 | 29 | 32 | 3 | 3 |
| 2 | 9 | 8 | −1 | 1 |
| 3 | 2 | 9 | 7 | 7 |
| 4 | 12 | 22 | 10 | 10 |
| 5 | 2 | 5 | 3 | 3 |
| 6 | 29 | 28 | −1 | 1 |
| 7 | 24 | −5 | −29 | 29 |
| 8 | 74 | 40 | −34 | 34 |
| 9 | 15 | 10 | −5 | 5 |
| 10 | 0 | 3 | 3 | 3 |
| 11 | −3 | 11 | 14 | 14 |
| 12 | −2 | 1 | 3 | 3 |
| 13 | 32 | 28 | −4 | 4 |
| 14 | 10 | 18 | 8 | 8 |
| 15 | 18 | 24 | 6 | 6 |
| 16 | 24 | 32 | 8 | 8 |
| 17 | 27 | 24 | −3 | 3 |
| 18 | 16 | 20 | 4 | 4 |
| 19 | 7 | 10 | 3 | 3 |
| 20 | 58 | 56 | −2 | 2 |
| 21 | 4 | 7 | 3 | 3 |
| **Mean** | 18 | 18 | 0 | 7 |
| **SD** | 19 | 15 | 11 | 9 |

## Center Frequency

Subjects in the current study were selected from a larger sample tested previously to have tinnitus loudness matches equal to or less than 70 dB SPL and pitch matches not exceeding 8 kHz. The familiarity of the subjects with their tinnitus, and their prior experience with the graphical user interface, may have enhanced the reliability of pitch matching between the two sessions of Method 2. Reliability was reflected by the significant correlation between sessions (Fig. 8) and by a trend toward diminished within-subject variability across repeated trials within each session. Although the sounds used for pitch determination could not exceed 8000 Hz, the frequencies obtained covered the range of 2800–8000 Hz suggesting that reliability was not constrained or inflated by ceiling effects. Perceptual judgments were likely based only on the perceived properties of tinnitus, because cues for pitch matching were not available from the featureless dial that was used for pitch determinations.

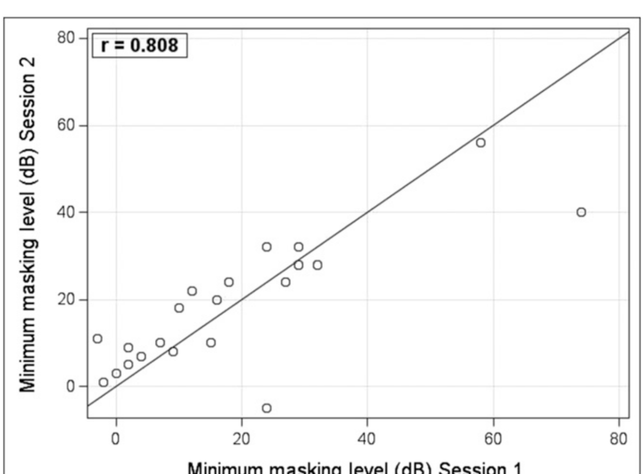


**Figure 10.** Scatterplot showing correlations of minimum masking levels (dB SL) between Sessions 1 and 2. The diagonal line is the line of unity, that is, the line of best fit if Session 1 and 2 values were identical for all subjects.

It may be noteworthy that between-subject reliability of NBM for Method 2 was especially high for subjects choosing pure tones as the best match for their tinnitus ($r = 0.93$). This relationship accounted for 87% of the between-subject variance in pitch matching for these subjects compared to 55% for subjects selecting a band of noise to represent their tinnitus. While this difference did not reach significance in our sample (nine subjects performed pitch matches with pure tones rather than narrow-band noise), it is provocative in supporting the concept of a pure tonal tinnitus in a substantial number of tinnitus sufferers. It is possible that off-frequency listening (Florentine and Houtsma, 1983; Huss and Moore, 2005) may have been a source of error for subjects matching a band of noise to their tinnitus. However, hearing thresholds did not correlate with the absolute difference in center frequency between the two repeated sessions and were not different between subjects matching with pure tones or a band of noise. This suggests that subjects with a noisy tinnitus simply had a more difficult time matching to the dominant frequency in the percept.

The current approach, which identifies a single center frequency to represent tinnitus, differs from the approaches taken by Noreña et al (2002) and Roberts et al (2008) in which subjects rate a number of frequencies for their resemblance or "likeness" to their tinnitus sound. The result is a "tinnitus spectrum" that has been found to cover the region of threshold shift (or "slope") in the audiogram, highlighting the role of hearing impairment in the generation of tinnitus. Tinnitus spectra peak at significantly higher likeness ratings in the hearing loss region for subjects reporting a "tonal" (pure tone) tinnitus rather a tinnitus of wider bandwidth (Roberts et al, 2008), but there is nevertheless a substantial spread of ratings across frequencies in tonal cases that could reflect not the presence of multiple frequencies in the tinnitus but proximity to a single dominant pitch. Significant test-retest reliability was found when likeness ratings were correlated between sessions for each tested frequency (Roberts et al, 2008) and when

**Table 12. Session 1 Results of Residual Inhibition Testing**

| Subject | Left Initial Presentation Level (dB SPL) | Right Initial Presentation Level (dB SPL) | Level Reduced by Subject | Occurrence of Residual Inhibition | Total Seconds of Residual Inhibition |
|---|---|---|---|---|---|
| 1 | 75 | 70 | No | None | 0 |
| 2 | 65 | 73 | No | None | 0 |
| 3 | 56 | 66 | No | None | 0 |
| 4 | 56 | 62 | No | None | 0 |
| 5 | 80 | 66 | Yes | None | 0 |
| 6 | 87 | 87 | Yes | Partial | 87 |
| 7 | 74 | 74 | Yes | None | 0 |
| 8 | 104 | 108 | Yes | None | 0 |
| 9 | 47 | 63 | No | None | 0 |
| 10 | 40 | 56 | No | Partial | 199 |
| 11 | 77 | 65 | Yes | Partial | 51 |
| 12 | 60 | 58 | Yes | Partial | 58 |
| 13 | 91 | 89 | Yes | Partial | 38 |
| 14 | 75 | 75 | Yes | Partial | 87 |
| 15 | 80 | 91 | No | Complete | 141 |
| 16 | 98 | 70 | Yes | Complete | 276 |
| 17 | 127 | 95 | Yes | Partial | 195 |
| 18 | 66 | 74 | No | None | 0 |
| 19 | 48 | 64 | No | Complete | 186 |
| 20 | 94 | 92 | No | Partial | 205 |
| 21 | 78 | 71 | No | Complete | 115 |
| **Mean** | 75 | 75 | | | 78 |
| **SD** | 21 | 14 | | | 89 |

likeness ratings were correlated across frequencies within a subset of individual subjects (Zhou et al, 2011). However, the relation of the dominant likeness rating to pitch determinations made by the current procedure is presently unknown. Both procedures (each automated and employing featureless response dials)

**Table 13. Session 2 Results of Residual Inhibition Testing**

| Subject | Left Initial Presentation Level (dB SPL) | Right Initial Presentation Level (dB SPL) | Level Reduced by Subject | Occurrence of Residual Inhibition | Total Seconds of Residual Inhibition |
|---|---|---|---|---|---|
| 1 | 78 | 75 | No | None | 0 |
| 2 | 62 | 66 | Yes | None | 0 |
| 3 | 63 | 71 | No | None | 0 |
| 4 | 65 | 76 | No | None | 0 |
| 5 | 89 | 73 | No | Complete | 161 |
| 6 | 48 | 46 | Yes | Partial | 85 |
| 7 | 48 | 49 | No | None | 0 |
| 8 | 110 | 75 | Yes | Partial | 212 |
| 9 | 44 | 54 | No | None | 0 |
| 10 | 42 | 70 | No | Partial | 67 |
| 11 | 85 | 83 | No | Partial | 95 |
| 12 | 60 | 58 | Yes | Complete | 53 |
| 13 | 84 | 84 | No | Partial | 54 |
| 14 | 64 | 82 | Yes | None | 0 |
| 15 | 84 | 90 | No | Complete | 281 |
| 16 | 98 | 76 | No | Complete | 511 |
| 17 | 124 | 89 | Yes | Partial | 162 |
| 18 | 86 | 98 | No | None | 0 |
| 19 | 24 | 46 | No | Complete | 158 |
| 20 | 90 | 88 | No | Partial | 234 |
| 21 | 85 | 78 | No | Complete | 297 |
| **Mean** | 73 | 73 | | | 113 |
| **SD** | 24 | 15 | | | 135 |

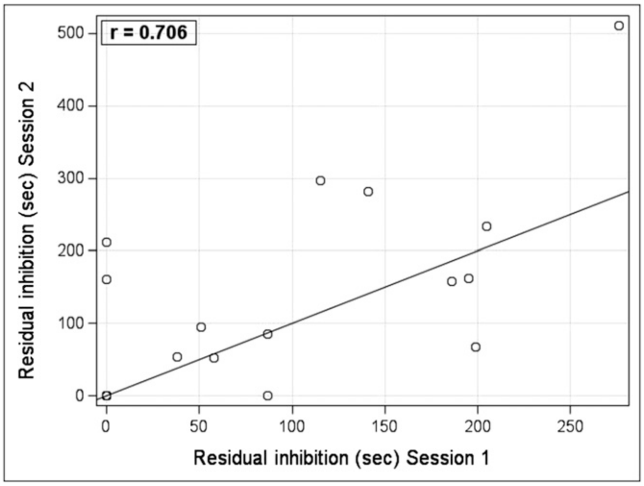


**Figure 11.** Scatterplot showing correlations of residual inhibition responses (in seconds) between Sessions 1 and 2. The diagonal line is the line of unity, that is, the line of best fit if Session 1 and 2 values were identical for all subjects.

have an advantage over two-alternative forced choice methods for determining tinnitus pitch in that subjects are not required to choose between two sounds both of which resemble tinnitus or neither of which do.

## Bandwidth

When assessed with Method 1 most subjects were generally consistent within and between sessions with respect to the stimuli they chose as the best match to their tinnitus (Tables 3 and 4). There was also a trend for the bandwidths that subjects chose as the best match to become narrower from Session 1 to Session 2, although the difference was not significant. Similarly, the bandwidths that subjects chose as the best match tended to become more consistent from Session 1 to Session 2, although this difference was also not significant. Of the 21 subjects tested, nine chose a pure tone as their bandwidth match in each session (six of the nine chose a pure tone for both sessions). When asked to judge the quality of their matches, for each of the 10 trials across the two sessions, only three of the 21 subjects judged their selection as "not a match" (all other subjects indicated that their selection was "somewhat close," "very close," or an "exact match").

The second assessment of bandwidth (Method 3) essentially duplicated Method 1 but using a center frequency that was determined by the average of the frequencies selected during the preceding measurement of tinnitus pitch (Method 2). One might have expected an improvement in reliability with Method 2 since the bandwidth adjustment was now performed around what was expected to be an improved frequency match. However, the results replicated those of the first method with no improvement seen, although more subjects chose zero as their bandwidth on the second session compared to any other bandwidth test.

When the results of Methods 1 and 3 are considered together, subjects reporting either a narrow-band tonal tinnitus or a wide-band noisy tinnitus were consistent across their bandwidth matches. Other subjects choosing an intermediate noise sound as their preferred bandwidth showed large discrepancies between their repeated matches. These latter subjects reduced the test-retest reliability of bandwidth matching to a level (Fig. 8) considerably below that found for determining the dominant tinnitus pitch (Fig. 7). Similarly, Roberts et al (2008) found that 70% of individuals choosing either a 5 kHz pure tone as the best match to their tinnitus ("tonal" tinnitus), or a wide band of noise (center frequency 5 kHz ±15% at −10 dB) as the best match to their tinnitus ("hissing" tinnitus), repeated their choices on a second test compared to 43% of subjects choosing an intermediate noise band as their best tinnitus match (center frequency 5 kHz ±5% at −10 dB, "ringing" tinnitus). Consistent bandwidth matching by subjects reporting a narrow-band tonal tinnitus or a wide-band noisy tinnitus appears to give a strong indication of the "spectral content" of a patient's tinnitus percept. The presence of a band of frequencies in a tinnitus percept may help to explain why some patients have so much difficulty performing pitch matches with pure tones—that is, if their tinnitus consists of a broad spectrum of frequencies, repeated tonal pitch matches would be expected to span the range of frequencies.

## Minimum Masking Levels

Hearing thresholds, using the 2–12 kHz noise stimulus, varied around a central point, i.e., repeated thresholds did not show any trends of becoming higher or lower upon repeated testing (Table 10). When the between-sessions noise threshold differences were converted to absolute values, the mean differences averaged only about 4 dB. This variability is well within clinical standards that have been established for test-retest reliability of hearing thresholds using pure tones (Witting and Hughson, 1940; Corso and Cohen, 1958; High et al, 1961; Atherley and Dingwall-Fordyce, 1963; Hickling, 1964).

Between-sessions repeated MMLs were within 10 dB of each other for all but three of the subjects (Table 11). Even including the latter three subjects, the average magnitude of the between-sessions differences was only 7 dB. These results supported a test-retest correlation between the two measurement sessions of $r = 0.807$ ($p < 0.05$) with only two of the 21 subjects falling notably off the line of unity (Fig. 10). Considering the subjective nature of tinnitus, and the potential for tinnitus fluctuations between sessions, this result suggests that MMLs can be obtained reliably. Results of repeated

MMLs have not been reported in the literature, so there is no external basis for comparison. It is noteworthy that some studies of sound therapy for tinnitus have reported improvements (reductions) in MML after treatment (Jastreboff et al, 1994; Davis et al, 2007, 2008), which implies reliable variance in the measure.

The age range for the 21 subjects was 28–78 yr. This fairly large range might raise the question of test-retest reliability of the measures as a function of age. It was possible to address this question with the MML data. Table 11 shows the differences in MMLs between Session 1 and Session 2 for all of the subjects. These differences were evaluated as a function of age, revealing little relationship between age and test-retest reliability in the MML data. The Pearson product-moment correlation between these two variables was 0.14.

### Residual Inhibition

The literature shows that about 80–90% of tinnitus patients will experience RI when tested with the clinical procedure developed by Vernon and Meikle in 1981. Our results show that 12 of 21 subjects (57%) reported RI during the first session, and 13 of 21 (62%) reported it during the second session. These percentages are smaller than the clinical average, which might be partially explainable because subjects were allowed to reduce the volume of the test stimulus. This is not normally done for clinical RI testing but was included as an option for this study to ensure that subjects were not forced to listen to noise that was uncomfortably loud. Ten subjects reduced the output during the first session, and six did during the second session.

The percentage of subjects experiencing RI in the present study is, however, in broad agreement with results reported by Roberts et al (2008). In their sample of 59 subjects with bilateral tinnitus, 69.5% reported at least some degree of RI (defined as a suppression of 20% of scale or greater averaged over three trials) after listening to 11 narrow-band noise maskers one at a time, with center frequencies differing between 500 and 12,000 Hz. The mean duration of RI was on the order of 25 sec for subjects who reported RI, compared to 137 sec for such subjects in the first session of the present study increasing nonsignificantly to 182 sec in the second session. The difference between these two studies with respect to the duration of RI is likely a consequence of the fact Roberts et al (2008) used a masking sound of only 30 sec duration compared to 1 min here, so that a "residual inhibition function" relating RI depth to the center frequency of the masking sounds could be determined. The RI function paralleled the tinnitus spectrum, both increasing above the edge of normal hearing in proportion to the degree of threshold shift up to the limit of hearing. Although the proportion of subjects reporting RI in the current study appears to be less than that reported for patients assessed in clinical practice, the mean durations of RI found in the present study are typical of what is described for clinical settings.

A striking feature of our RI results was that subjects who reported RI tended to do so consistently. No subject reporting RI in the first session failed to report it in the second, and RI duration showed significant test-retest reliability ($r = 0.706$; Fig. 11). These findings suggest that RI can be a stable individual trait. Why some tinnitus sufferers experience RI and others do not is a puzzle waiting to be understood. Roberts et al (2008) noted that subjects younger than 50 yr old who had comparatively good hearing to 12 kHz were resistant to RI. RI resistance may occur when the bandwidth of the masking sound is not well matched to the pattern of cochlear pathology present in such cases.

### Conclusions, Limitations, and Future Directions

Testing capabilities enabled by the automated system include measurements of hearing thresholds, tinnitus LM and PM, NBM, MML, and RI in a controlled test environment. The assessment could comprise a complete clinical test battery for psychoacoustic quantification of tinnitus. The total time of testing for the measurement reported here averaged 46 min to complete the battery of tests. Our objective is to conduct an entire tinnitus test battery in an average of 30 min. We plan to refine some of the testing for greater testing efficiency. All of the measures are repeated at least twice, and we will analyze the responses carefully to determine the minimum number of responses necessary for each test without compromising test-retest reliability.

Some examples of potential time-saving changes include the following:

1. Obtaining hearing threshold responses only once at each frequency. The computer could be programmed to analyze each pair of responses "on the fly" for a given patient, and to switch to single-response testing only if the patient achieves a criterion level of reliability.
2. Obtaining LM only once at each frequency. As with hearing thresholds, ongoing computer analyses of responses could determine if a single response would be sufficient.
3. Obtaining PM three instead of five times. If a patient selects the same frequency three times in succession, the computer could be programmed to omit the last two PM trials and advance to the next test.
4. Performing only one noise-band test. Analysis of the present data from the three NBM methods will assist in determining a single test to perform NBM. It is also possible that a single trial is sufficient to

determine that the patient's PM with a pure tone is a better match than any noise band, which would obviate NBM in that case.

5. Obtaining an NBM fewer than five times. As for PMs, if a patient selects the same noise-band three times in a row, further matches could be unnecessary.

Each of these potential changes is being considered as we attempt to improve testing efficiency. These changes will be incorporated into a complete redesign of the system. We fully anticipate the development of a clinic-ready tinnitus measurement system as we complete current testing. The availability of such a system will promote, for the first time, interclinic uniformity of tinnitus assessment. This capability is greatly needed within the Veterans Affairs health-care system to address the rapidly increasing numbers of veterans who are claiming tinnitus as a service-connected disability and who are appearing in clinics to receive assessment and treatment.

Notable within-subject variability in repeated pitch matching within sessions was observed in Tables 6 and 7, even though the test-retest reliability of mean determinations between sessions was highly significant (Fig. 8). We intend to make a modification that could improve pitch match reliability within and between sessions. When conducting pitch matching using manual methods, it has been commonly reported that patients confuse the octave of their pitch match (Henry and Meikle, 2000). That is, when selecting a match, it is thought that patients may select a frequency that is one octave higher or one octave lower in frequency than the actual perceived pitch of the tinnitus. To remedy this problem, examiners present tones one octave above and one octave below the patient's pitch match to determine if the patient has made an "octave confusion." For our pitch match protocol we (perhaps wrongly) assumed that octave confusion testing was not necessary because patients have the capability of raising or lowering the frequency of the test tone at will. This capability was thought to preclude the need for octave confusion testing. However, we intend to evaluate the benefit of modifying the pitch match test algorithm such that, when a pitch match is made, the computer will present tones an octave above and an octave below (if such frequencies are available in the test set) the selected frequency and query the patient as to which frequency sounds "most like the tinnitus."

Another concern that could affect test-retest reliability of the measures is the inherent variability that may exist in the tinnitus itself for some patients. Within-session fluctuations in tinnitus would be highly unusual. However, between-session fluctuations could occur for some patients. The subjects in this study responded to an initial survey question about whether their tinnitus perception fluctuated over time ("Does the loudness of your tinnitus tend to fluctuate up and down?") as well as a follow-up question ("If your tinnitus shows loudness fluctuations, how large are the changes usually?"). We examined the subjects who reported greater fluctuations to determine if they showed higher variability in their between-session repeated measures. No difference was found in this measure between subjects who reported greater fluctuations and subjects who reported more stable tinnitus percepts. Future investigations of test-retest reliability of tinnitus measures should query subjects as to the stability of their tinnitus during and between test sessions.

A measurement system designed for the clinic will also facilitate comparison of data with alternative automated methods for measuring tinnitus spectra (Noreña et al, 2002; Roberts et al 2008) and RI (Roberts et al, 2008). The latter procedures were designed as research tools giving measurements of the frequencies that comprise tinnitus, its bandwidth, loudness, and RI functions. The procedure of Roberts et al (2008) included a brief, automated front-end training program to acquaint the subject with the graphical user interface and introduce the concepts of loudness and pitch. A limitation of these procedures at present is the time taken for completion, which can range up to 1 hr or more exclusive of generating an RI function, and the need for specialized sound delivery and calibration equipment in the clinic. All approaches to automated measurement including the methods described in this article will require substantial baseline data if patients are to understand their tinnitus relative to population norms.

In their comprehensive review of psychoacoustic methods, Henry and Meikle (2000) noted that the measurement of tinnitus sounds is complicated by the variability that many patients report in the quality of their tinnitus. Examples include complex spectrotemporal tinnitus sounds, variations over hours or days, and different percepts in the two ears. The subjects studied here and by Roberts et al (2008) and Noreña et al (2002) were long-standing cases with reasonably stable tinnitus percepts. While capturing the full range of tinnitus may remain a challenge, methods for characterizing stable tinnitus are likely to represent an advance and may be suitable for the majority of individuals afflicted with this condition.

**Acknowledgments.** The authors wish to thank Betsy Rheinsburg, M.S., for collecting data for this study and Garnett McMillan, Ph.D., for his contributions to the statistical analyses and creation of figures.

## REFERENCES

Atherley GR, Dingwall-Fordyce I. (1963) The reliability of repeated auditory threshold determination. *Br J Ind Med* 20:231–235.

Corso JF, Cohen A. (1958) Methodological aspects of auditory threshold measurements. *J Exp Psychol* 55(1):8–12.

Davis PB, Paki B, Hanley PJ. (2007) Neuromonics Tinnitus Treatment: third clinical trial. *Ear Hear* 28(2):242–259.

Davis PB, Wilde RA, Steed LG, Hanley PJ. (2008) Treatment of tinnitus with a customized acoustic neural stimulus: a controlled clinical study. *Ear Nose Throat J* 87(6):330–339.

Douek E, Reid J. (1968) The diagnostic value of tinnitus pitch. *J Laryngol Otol* 82(11):1039–1042.

Ellingson RM, Lilly DJ, Fausti SA. (2004) Method to generate near real-time variable bandwidth and center frequency noise stimuli. In: Ueha S, ed. *Proceedings of the 18th International Congress on Acoustics*. Kyoto: Acoustical Society of Japan, IV-2999–IV-3002.

Evered D, Lawrenson G, eds. (1981) *Tinnitus. Ciba Foundation Symposium 85*. London: Pitman Books.

Flick CL, Fausti SA, Henry JA, Ellingson RM, Silaski GC. (2000) Development of a computerized system for conducting audiometric research. In: Ouslander JG, ed. *Second National Department of Veterans Affairs Rehabilitation Research and Development Conference Proceedings*. Washington, DC: VA Rehab R&D Service, 72.

Florentine M, Houtsma AJ. (1983) Tuning curves and pitch matches in a listener with a unilateral, low-frequency hearing loss. *J Acoust Soc Am* 73(3):961–965.

Henry JA, Fausti SA, Mitchell CR, Flick CL, Helt WJ. (1996) An automated technique for tinnitus evaluation. In: Reich GE, Vernon JA, eds. *Proceedings of the 5th International Tinnitus Seminar 1995*. Portland: American Tinnitus Association, 325–326.

Henry JA, Fausti SA, Flick CL, Helt WJ, Ellingson RM. (2000) Computer-automated clinical technique for tinnitus quantification. *Am J Audiol* 9(1):36–49.

Henry JA, Flick CL, Gilbert AM, Ellingson RM, Fausti SA. (1999) Reliability of tinnitus loudness matches under procedural variation. *J Am Acad Audiol* 10(9):502–520.

Henry JA, Flick CL, Gilbert AM, Ellingson RM, Fausti SA. (2001a) Reliability of hearing thresholds: computer-automated testing with ER-4B Canal Phone earphones. *J Rehabil Res Dev* 38(5): 567–581.

Henry JA, Flick CL, Gilbert AM, Ellingson RM, Fausti SA. (2001b) Comparison of two computer-automated procedures for tinnitus pitch matching. *J Rehabil Res Dev* 38(5):557–566.

Henry JA, Flick CL, Gilbert AM, Ellingson RM, Fausti SA. (2003) Reliability of computer-automated hearing thresholds in cochlear-impaired listeners using ER-4B Canal Phone earphones. *J Rehabil Res Dev* 40(3):253–264.

Henry JA, Flick CL, Gilbert AM, Ellingson RM, Fausti SA. (2004) Comparison of manual and computer-automated procedures for tinnitus pitch-matching. *J Rehabil Res Dev* 41(2):121–138.

Henry JA, James KE, Owens K, Zaugg T, Porsov E, Silaski G. (2009) Auditory test result characteristics of subjects with and without tinnitus. *J Rehabil Res Dev* 46(5):619–632.

Henry JA, Meikle MB. (2000) Psychoacoustic measures of tinnitus. *J Am Acad Audiol* 11(3):138–155.

Henry JA, Rheinsburg BE, Ellingson RM. (2004) Computer-automated tinnitus assessment using patient control of stimulus parameters. *J Rehabil Res Dev* 41(6A):871–888.

Henry JA, Rheinsburg B, Owens KK, Ellingson RM. (2006) New instrumentation for automated tinnitus psychoacoustic assessment. *Acta Otolaryngol Suppl* 126(556):34–38.

Hickling S. (1964) The validity and reliability of pure tone clinical audiometry. *N Z Med J* 63:379–382.

High WS, Glorig A, Nixon J. (1961) Estimating the reliability of auditory threshold measurements. *J Aud Res* 1:247–262.

Huss M, Moore BC. (2005) Dead regions and pitch perception. *J Acoust Soc Am* 117(6):3841–3852.

Jastreboff PJ, Hazell JWP, Graham RL. (1994) Neurophysiological model of tinnitus: dependence of the minimal masking level on treatment outcome. *Hear Res* 80(2):216–232.

McFadden D. (1982) *Tinnitus—Facts, Theories and Treatments*. Washington, DC: National Academy Press.

Meikle MB. (1991) How tinnitus is related to hearing impairment. *Assessment and Re-Habilitation: Papers Presented at 3rd Bi-Annual Workshop, 15th-17th March 1991*. Turramurra, New South Wales, Australia: Australian Association of Audiologists in Private Practice, 4–6.

Noreña A, Micheyl C, Chéry-Croze S, Collet L. (2002) Psychoacoustic characterization of the tinnitus spectrum: implications for the underlying mechanisms of tinnitus. *Audiol Neurootol* 7 (6):358–369.

Nuttall AL, Meikle MB, Trune DR. (2004) Peripheral processes involved in tinnitus. In: Snow JB, ed. *Tinnitus: Theory and Management*. Lewiston, NY: BC Decker, 52–68.

Roberts LE, Moffat G, Baumann M, Ward LM, Bosnyak DJ. (2008) Residual inhibition functions overlap tinnitus spectra and the region of auditory threshold shift. *J Assoc Res Otolaryngol* 9(4): 417–435.

Vernon J, Griest S, Press L. (1990) Attributes of tinnitus and the acceptance of masking. *Am J Otolaryngol* 11(1):44–50.

Vernon JA, Meikle MB. (1981) Tinnitus masking: unresolved problems. In: Evered D, Lawrenson G, eds. *Tinnitus. Ciba Foundation Symposium 85.*London: Pitman Books, 239–256.

Vernon JA, Meikle MB. (2000) Tinnitus masking. In: Tyler RS, ed. *Tinnitus Handbook*. San Diego: Singular Publishing Group, 313–356.

Witting EG, Hughson W. (1940) Inherent accuracy at a series of repeated clinical audiograms. *Laryngoscope* 50:259–269.

Zhou X, Henin S, Long GR, Parra LC. (2011) Impaired cochlear function correlates with the presence of tinnitus and its estimated spectral profile. *Hear Res* 277(1–2):107–116.