J Am Acad Audiol 11: 156-161 (2000)

A Re-evaluation of Tinnitus Reliability Testing

Gary P. Jacobson*
Jennifer A. Henderson†
Devin L. McCaslin‡

Abstract

The purposes of these investigations were to (1) evaluate tinnitus loudness measures of unskilled normal listeners asked to imagine that they are experiencing a constant monaural tonal tinnitus and (2) compare the performance of these listeners to that of a sample of patients with tinnitus. Subgroups of 24 patients participated in two investigations. Results suggest that (1) normal subjects asked to imagine a high-pitched, tonal tinnitus show significantly greater tinnitus loudness matching levels (a) when they do not have an external reference, (b) after a 1-week interval, and (c) for low-frequency matching tones; (2) patients with high-pitched tinnitus did not demonstrate statistically significant differences in tinnitus loudness judgments within or between sessions or frequencies; (3) tinnitus patients do not differ significantly from normals feigning tinnitus in the variability of tinnitus loudness matching levels within a single session (two measures) or after a 1-week interval (one measure); and (4) normals feigning tinnitus tended to choose significantly greater loudness matching levels than did tinnitus subjects.

Key Words: Audiometry, medicolegal evaluation, tinnitus

Abbreviations: SL = sensation level, THI= Tinnitus Handicap Inventory

At present, there is no objective method to validate a patient's contention that he or she experiences tinnitus. Despite this observation, tinnitus continues to be a compensable disability affecting, for example, an estimated 80,000 veteran patients evaluated by the Department of Veterans Affairs. Further, it is not uncommon for otologists and audiologists to be called upon to serve as expert witnesses for plaintiffs claiming tinnitus as a primary injury following trauma secondary to automobile accidents, falls, and assaults. A subjective technique has been proposed by Johnson (1987, as cited in Shulman et al, 1991; Vernon and Meikle, 1988) to help substantiate a claim that tinnitus exists. This evaluation technique has been referred to as Tinnitus Reliability Testing. In Tinnitus Reliability Testing, multiple within-session behavioral estimates of tinnitus loudness (in dB SL) are obtained and a standard deviation calculated. It has been proposed that a standard deviation of no more than 3 dB may suggest that the patient has an internal reference (i.e., tinnitus) against which he or she is comparing the test tone. What are not known, however, are the loudness matching characteristics of individuals without tinnitus who claim that they do have tinnitus. Accordingly, the purposes of the present preliminary investigations were to (1) evaluate tinnitus loudness measures of unskilled normal listeners asked to imagine they are experiencing constant tonal tinnitus in one ear and (2) to compare the performance of unskilled, normal-hearing listeners to a sample of tinnitus patients.

METHOD

Investigation 1

Subjects were 13 normal-hearing, otologically and neurologically normal, unskilled listeners (mean age = 32 years [SD = 7.5 years], one male). Normal hearing was defined as auditory thresholds ≤ 20 dB HL at octave and interoctave frequencies from 250 through and including 8000 Hz.

*Division of Audiology, Henry Ford Hospital, Detroit, Michigan; †Arizona State University, Tempe, Arizona; ‡Ohio State University, Columbus, Ohio

Reprint requests: Gary P. Jacobson, Division of Audiology, Henry Ford Hospital, 2799 West Grand Blvd., Detroit, MI 48202

Table 1 Conditions and Tasks Requested of Normal Subjects in Investigation 1

Condition	*Reference*	*Instructions*	*Stimuli*
1: No reference†	None	Example: "Tell me to increase/decrease the loudness of the tone I am putting in your left ear until it matches the loudness of the tinnitus in your right ear."	Continuous pure-tone stimuli ascending in frequency from 250–8000 Hz. Stimulus intensity began 3 dB below audiometric threshold and increased at a rate of 1 dB/5 sec. When instructed by subject, the intensity was increased (or decreased) at a rate of 1 dB/5 sec.‡
2: Remembered reference†	Continuous 4000 Hz, 6 dB SL pure tone presented to "tinnitus ear" for 5 minutes, then switched off for 5 minutes prior to loudness measures	Same as above	Same as above
3: Continuous reference†	Continuous 4000 Hz, 6 dB SL pure tone presented to "tinnitus ear" during loudness measures	Same as above	Same as above

*Loudness matching stimuli always were presented to the contralateral ear (i.e., ear opposite tinnitus ear).
†Measures were conducted two times on day 1 (with a 15-minute break between measures) and once on day 2 (i.e., 1 week later).
‡Equal loudness judgments were conducted three times at each frequency before the value was accepted.

Subjects were seated in a quiet room for 5 minutes to ensure that they were tinnitus free. ER3A insert earphones were placed in their ears and auditory thresholds were determined in 1-dB increments using a bracketing technique for 10 octave and interoctave frequencies from 250 to 8000 Hz. Following threshold measurements, subjects were assigned one ear (counterbalanced across the sample) that became their "tinnitus" ear. That is, they were requested to imagine that they had a steady tonal tinnitus in that ear. The other ear was used for "tinnitus" loudness matching measures.

The three listening conditions are shown in Table 1. In condition 1 (no reference), subjects were told to imagine that they had a tone of constant loudness and pitch in one ear (i.e., the ear designated by the examiner) and that this imagined sound would serve as their tinnitus or "ringing in the ears." Tinnitus loudness judgments were obtained by having subjects listen to continuous tones in the opposite ear (i.e., 10 frequencies consisting of octave and interoctave frequencies ascending from 250–8000 Hz). For each frequency, the tone was increased manually in intensity in 1-dB increments by the examiner until threshold was reached. Subjects then instructed the examiner to either manually increase or decrease the intensity of the pure tone until the perceived loudness of the tone equaled the loudness of the subject's "tinnitus" in the contralateral ear. Equal loudness estimates were calculated in dB SL (e.g., "tinnitus" loudness match in dB HL minus auditory threshold dB HL at that frequency). For condition 2 (remembered reference), a continuous 4000-Hz, 6 dB SL tone was presented for 5 minutes to the "tinnitus" ear. The tone then was switched off for 5 minutes and the tinnitus loudness measures were repeated. In condition 3 (continuous reference), the 6 dB SL, 4000-Hz pure tone was presented continuously in the "tinnitus" ear during the "tinnitus" loudness judgments.

"Tinnitus" loudness matching measures were conducted twice in one recording session with a 15-minute rest period between recording blocks in an attempt to evaluate group short-term, test–retest reliability. The three conditions were conducted one time a week later in an attempt to evaluate group long-term test–retest reliability.

Investigation 2

Subjects were 11 patients seen for tinnitus management at Henry Ford Hospital or at the Veterans Affairs Medical Center in Dayton, Ohio (mean age = 47.54 years [SD = 11.57 yrs], 10 males). Group audiometric thresholds for this sample are shown in Table 2.

Tinnitus was (by report) bilateral in all patients. Tinnitus was bilateral but predominating in the right ear for six patients and bilateral but predominating in the left ear for two patients. All subjects pitch matched their tinnitus to pure-tone frequencies exceeding 3000 Hz. The mean duration of bothersome tinnitus for this sample was 11.72 years (SD = 14.36 years). Mean score on the Tinnitus Handicap Inventory (THI; Newman et al, 1998) for this sample was 50.90 points (SD = 28.28 points), representing moderate self-perceiving tinnitus handicap (Newman et al, 1996).

Methods for obtaining tinnitus loudness matching judgments were identical to those shown in condition 1 for Investigation 1. The tinnitus ear evaluated was the right ear for nine patients and the left ear for the remaining two patients (i.e., the ear used to present the loudness matching tones was the left ear for nine patients and the right ear for two patients).

RESULTS

Investigation 1

The data from the normal subjects were subjected to an analysis of variance (ANOVA), with tinnitus "loudness" (in dB SL) as the dependent variable and session (i.e., day 1, 15 minutes after the first set of measurements on day 1 and 1 week later), condition (i.e., no reference, remembered reference, and continuous reference), and frequency (i.e., 250–8000 Hz) serving as independent variables.

Table 3 Group Means and Standard Deviations of "Tinnitus" Loudness Matching Judgments for the Three Sessions (i.e., Collapsed across Frequencies and Conditions) for the Normal Group

	Session 1	*Session 2 (+ 15 Minutes)*	*Session 3 (+1 Week)*
Mean (dB SL)	12.05	11.13	13.79
SD (dB SL)	7.03	6.92	11.69

Results revealed significant session ($F = 8.81$, $df = 2$, $p < .001$; Table 3), condition ($F = 60.01$, $df = 2$, $p < .001$; Table 4), and frequency ($F = 2.57$, $df = 9$, $p = .006$; Table 5) main effects. Tinnitus loudness matching SLs were greater for the last session (+ 1 week), for the first condition (no reference), and for tinnitus matching tones that were less than 4000 Hz. There was a statistically significant session × condition interaction ($F = 3.64$, $df = 4$, $p = .006$; Table 6). Results of Bonferroni-corrected post hoc testing (Tukey) revealed that tinnitus loudness matching SLs in condition 1 (no reference) were significantly greater than those in conditions 2 and 3 and tinnitus loudness matching SLs for session 3 were significantly greater than for sessions 1 and 2.

Investigation 2

Data from the sample of tinnitus patients were subjected to an ANOVA with tinnitus loudness estimates as the dependent variable and session and frequency serving as independent variables. There were no statistically significant main effects or session × frequency interactions. That is, although the magnitude of the tinnitus patient's loudness judgments tended to increase over time, the differences were not statistically significant within or between recording sessions (see bottom half of Table 7).

Table 2 Mean (SD) Audiometric Thresholds for the Tinnitus Sample (N = 11)

	Frequency (Hz)									
	250	*500*	*750*	*1000*	*1500*	*2000*	*3000*	*4000*	*6000*	*8000*
Right ear mean	8.90	10.60	12.90	11.60	15.20	18.30	26.50	31.10	38.40	37.30
	(11.22)	(10.14)	(8.41)	(9.56)	(12.40)	(23.22)	(25.56)	(25.70)	(25.48)	(23.29)
Left ear mean	5.91	9.55	9.64	9.91	15.00	20.64	30.55	31.64	37.18	35.27
	(3.50)	(4.63)	(5.98)	(3.76)	(7.58)	(13.04)	(21.57)	(21.17)	(25.68)	(25.03)

Table 4 Group Means and Standard Deviations of "Tinnitus" Loudness Matching Judgments for the Three Conditions (i.e., Collapsed across Frequencies and Sessions) for the Normal Group

	Condition 1 (No Reference)	*Condition 2 (Remembered Reference)*	*Condition 3 (Continuous Reference)*
Mean (dB SL)	15.97	10.14	10.51
SD (dB SL)	10.1	7.1	7.0

Finally, data from the normal subjects for condition 1 (i.e., no reference) were compared to analogous data obtained from the tinnitus subjects. The comparisons were made separately for each of the 10 frequencies used for tinnitus loudness matching (i.e., ANOVA with tinnitus loudness match in dB SL as the dependent variable, and subject class [tinnitus versus normal] and session as independent variables). With the exception of 750 Hz and 8000 Hz ($p > .05$), the results revealed statistically significant main effects for subject class. That is, normal subjects loudness matched their "tinnitus" to continuous pure tones at significantly greater sensation levels than did tinnitus patients (i.e., 250 Hz, $F = 4.32$, $df = 1$, $p = .042$; 500 Hz, $F = 10.55$, $df = 1$, $p = .002$; 1000 Hz, $F = 6.396$, $df = 1$, $p = .014$; 1500 Hz, $F = 6.995$, $df = 1$, $p = .010$; 2000 Hz, $F = 4.14$, $df = 1$, $p = .046$; 3000 Hz, $F = 12.24$, $df = 1$, $p = .001$; 4000 Hz, $F = 8.79$, $df = 1$, $p = .004$; 6000 Hz, $F = 6.346$, $df = 1$, $p = .014$). Group mean data illustrating this effect are shown in Table 7. There were no statistically significant group differences in tinnitus loudness matching SLs within or between sessions.

DISCUSSION

An unexpected finding in this investigation was that normal subjects were consistent in their loudness judgments of their feigned tinnitus. The loudness judgments of normals were of greater magnitude after a 7-day interval and when they did not have either a remembered or continuous reference. However, the group comparisons were not statistically significant (i.e., normals feigning tinnitus versus patients with tinnitus). In this regard, there was a trend for even our tinnitus subjects to demonstrate an increase in tinnitus loudness judgments over the 7-day interval.

It was our expectation that loudness judgments of subjects who did not have a continuous internal referent would vary greatly within and between sessions; this did not occur. It is possible that an explanation for this phenomenon can be found in the methods we used to obtain the loudness judgments. An ascending method was employed for presenting the continuous comparison tones (i.e., we increased tone intensity from subthreshold to suprathreshold) in an attempt to obviate the effect of residual inhibition. Thus, it might have been possible for subjects to use as a "benchmark" when the comparison tone first became audible. Different results might have occurred had we compared loudness judgments obtained using both ascending and descending tone presentation methods. We would recommend that future investigations evaluate this hypothesis. An alternate explanation is that it is possible for even unsophisticated listeners to create a long-lived sensory memory of sound loudness.

Another finding, that normals demonstrated significantly greater sensation levels for their feigned tinnitus than did patients with tinnitus, probably occurred because normals matched their feigned tinnitus to tones that were clearly audible. In this regard, tinnitus patients in the present investigation loudness matched their tinnitus to tones at low sensation levels. It is a well-accepted observation that when given an external tone to match to their tinnitus, patients seldom choose levels exceeding 15 to 16 dB SL (Meikle and Walsh, 1983). Further, in carefully controlled psychoacoustic investigations (i.e., where some training has occurred prior to mea-

Table 5 Group Means and Standard Deviations for "Tinnitus" Loudness Matching Judgments for 10 Frequencies (i.e., Collapsed across Sessions and Conditions) for the Normal Group

	Frequency (Hz)									
	250	*500*	*750*	*1000*	*1500*	*2000*	*3000*	*4000*	*6000*	*8000*
Mean (dB SL)	13.44	13.23	12.08	13.19	12.52	13.02	12.84	11.6	10.1	10.03
SD (dB SL)	7.98	8.33	9.04	8.76	7.88	8.95	8.1	7.83	9.06	9.49

Table 6 Means (SD) of "Tinnitus" Loudness Matching Judgments (in dB SL) for Session X Conditions (i.e., Collapsed across Frequencies) for the Normal Group

Condition	*Session 1*	*Session 2 (+ 15 Minutes)*	*Session 3 (+ 1 Week)*
1	14.63 (5.83)	14.44 (8.86)	19.68 (14.28)
2	10.56 (7.59)	9.29 (4.66)	10.70 (8.83)
3	10.98 (6.89)	9.68 (5.24)	10.99 (8.86)

surements), most match their tinnitus to tones of 6 dB SL or less (Meikle and Walsh, 1983).

In general, it was observed that for both normals feigning tinnitus and for tinnitus patients that tinnitus loudness matching values (in dB SL) tended to be of a greater magnitude for low-frequency matching tones. This phenomenon can be explained for patients with tinnitus that is high in pitch. We know from prior research that, for example, the dB SL required for equal loudness judgments to a high-frequency tinnitus can decrease as a function of frequency as the tinnitus matching tone approximates the pitch of the patient's tinnitus (i.e., equivalent to "convergence" type masking functions reported previously by Feldman, 1969).

There are some limitations to the present investigation. The data were collected manually as opposed to using a system under computer control (Henry et al, 1999). It always is possible that examiner bias can contaminate data collected in this manner. In the loudness matching trials, tone intensities were presented using an ascending technique like that described by Vernon and Meikle (1988). It is possible that a greater disparity in between-group (i.e., normals versus patients with tinnitus) tinnitus loudness judgments might have been obtained by calculating differences in tinnitus loudness judgments using ascending and descending tone intensity presentation methods. Thus, it might be expected that the variability in tinnitus loudness judgments would be greater, as are auditory threshold measures, using ascending versus descending assessment techniques for normals feigning tinnitus. Despite the methodological limitations, the within-group variability in tinnitus loudness matching measures was modest.

The results of this investigation support the contention that Tinnitus Reliability Testing using an ascending tinnitus loudness matching technique may not adequately discriminate between patients feigning tinnitus and those who actually have tinnitus. Further, partial validation of a patient's complaint that he or she experiences tinnitus may be found not in the variability of loudness matching measures but instead in the magnitude of their loudness matching levels.

Table 7 Means (SD) of Tinnitus Loudness Matching Judgments for 10 Frequencies and Three Sessions for the Normal and Tinnitus Groups*

	Frequency (Hz)									
Session	*250**	*500***	*750*	*1000**	*1500***	*2000**	*3000***	*4000***	*6000**	*8000*
Normal										
1	18.30 (5.45)	16.23 (5.63)	14.77 (6.30)	16.15 (6.09)	14.23 (5.70)	14.61 (5.33)	15.54 (4.64)	13.92 (5.87)	11.38 (5.14)	11.15 (6.16)
2	15.07 (6.06)	15.15 (7.58)	13.85 (9.04)	14.61 (9.60)	14.69 (7.45)	15.46 (8.98)	16.00 (8.31)	14.92 (9.29)	12.77 (10.40)	11.92 (12.70)
3	20.00 (12.86)	21.20 (15.11)	20.30 (17.56)	20.40 (17.09)	20.30 (14.43)	20.00 (15.13)	19.90 (12.94)	17.80 (10.76)	19.20 (15.77)	17.70 (16.31)
Tinnitus										
1	13.45 (7.65)	9.72 (5.90)	11.18 (5.01)	10.45 (4.41)	9.82 (6.03)	11.73 (7.85)	10.09 (5.20)	9.36 (4.59)	8.54 (3.61)	9.91 (5.37)
2	14.54 (7.06)	10.09 (5.49)	12.91 (3.94)	12.45 (4.86)	10.63 (5.77)	12.91 (7.12)	11.18 (5.88)	10.54 (6.14)	9.73 (4.69)	10.27 (4.49)
3	13.73 (6.23)	13.64 (6.14)	12.64 (6.01)	11.73 (7.07)	11.64 (7.05)	12.27 (7.18)	11.27 (5.27)	11.73 (6.15)	10.27 (5.69)	11.00 (6.15)

Notice the group differences in sensation levels at each frequency. The SLs are greater overall for the normal group.
*Group difference (i.e., normal versus tinnitus collapsed across sessions) with a probability estimate of $p \leq .05$.
**Group difference with an associated probability estimate of $p \leq .01$.

Acknowledgment. This paper was presented at the 22nd meeting of the Association for Research in Otolaryngology (ARO), St. Petersburg Beach, FL, February 13–18, 1999.

REFERENCES

Feldmann H. (1969). Studies on the masking of subjective tinnitus—a contribution to the pathphysiology of tinnitus. *Laryngol Rhinol Otol* 48:528–545.

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

Johnson RM. (1987). Tinnitus and disability—a method of evaluation. In: Shulman A, Aran J-M, Tonndorf J, Feldmann H, Vernon JA, eds. *Tinnitus: Diagnosis / Treatment.* Malvern, PA: Lea & Febiger, 444.

Meikle MB, Walsh ET. (1983). Characteristics of tinnitus and related observations in over 1800 tinnitus clinic patients. *J Laryngol Otol Suppl* 9:17–21.

Newman CW, Jacobson GP, Spitzer JB. (1996). Development of the Tinnitus Handicap Inventory. *Arch Otolaryngol Head Neck Surg* 122:143–148.

Newman CW, Sandridge SA, Jacobson GP. (1998). Psychometric adequacy of the tinnitus handicap inventory (THI) for evaluating treatment outcome. *J Am Acad Audiol* 9:153–160.

Shulman A, Aran J-M, Tonndorf J, Feldmann H, Vernon JA. (1991). *Tinnitus: Diagnosis / Treatment.* Philadelphia: Lea & Febiger.

Vernon JA, Meikle MB. (1988). Measurement of tinnitus: an update. In: Kitahara M, ed. *Tinnitus: Pathophysiology and Management.* New York: Igaku-Shoin, 36–52.