J Am Acad Audiol 16:85–100 (2005)

Estimates of Loudness, Loudness Discomfort, and the Auditory Dynamic Range: Normative Estimates, Comparison of Procedures, and Test-Retest Reliability

LaGuinn P. Sherlock*
Craig Formby*

Abstract

The purpose of this series of experiments was to establish normative reference values for absolute and relative judgements of loudness discomfort and for the auditory dynamic range (DR), and to evaluate intersubject variability and intrasubject test-retest reliability for the respective measures of loudness discomfort. To establish the normal auditory DR, audiometric thresholds and loudness discomfort levels (LDLs) were measured from a group of 59 normal-hearing adults without sound tolerance problems. The resulting estimates of the LDL and DR were on the order of 100 dB HL and 95 dB, respectively. A subset (n = 18) of this larger group participated in further studies in which loudness growth functions and the upper limit of the auditory DR were measured by categorical scaling judgments. The findings revealed no significant differences between the test methods for absolute (LDL) and relative (categorical scaling) judgements of loudness discomfort, intersubject variability, or intrasubject test-retest reliability, and suggest that the simple LDL estimate of loudness discomfort is an efficient and valid clinical measure for characterizing the "threshold of discomfort."

Key Words: Auditory dynamic range, categorical scaling of loudness, Contour Test of Loudness, loudness discomfort level, test-retest reliability

Abbreviations: DR = dynamic range; LDL = loudness discomfort level

Sumario

El propósito de esta serie de experimentos fue el establecer valores normativos de referencia para juicios objetivos o subjetivos de molestia ante la intensidad subjetiva (loudness) y para el rango dinámico auditivo (DR), y evaluar la variabilidad intersujeto, así como la confiabilidad del test-retest intrasujeto, para las medidas respectivas de la molestia en la intensidad subjetiva. Para establecer el DR auditivo normal, se midieron los umbrales audiométricos y los niveles de molestia en la intensidad subjetiva (LDL) en un grupo de 59 adultos normoyentes sin problemas de tolerancia al sonido. Los estimados resultantes para el LDL y el DR estuvieron en el orden de 100 dB y 95 dB, respectivamente. Un sub-grupo (n = 18) de este grupo más grande participó en estudios posteriores donde las funciones del crecimiento de la intensidad subjetiva y el límite superior del DR auditivo fueron medidos por una escala de juicios de categorización. Los hallazgos no revelaron diferencias significativas entre los métodos de evaluación absolutos (LDL) y los juicios relativos (escala de categorización) para la molestia en la intensidad subjetiva, la variabilidad intersujeto, o para la confiabilidad test-retest intrasujeto, y sugieren que el simple

---

*University of Maryland, Baltimore

LaGuinn Sherlock, University of Maryland Frenkil Building, 16 S. Eutaw St., Suite 400, Baltimore, MD 21201; Phone: 410-328-5947; Fax: 410-328-1886; E-mail: gsherlock@smail.umaryland.edu

estimado de un LDL es una medida clínica eficiente y válida para caracterizar el "umbral de molestia".

Palabras Clave: Rango dinámico auditivo, escala de categorización de la intensidad subjetiva, Prueba de Contorno de la Intensidad Subjetiva, nivel de molestia en la intensidad subjetiva, confiabilidad test-retest

Abreviaturas: DR = rango dinámico; LDL = nivel de molestia en la intensidad subjetiva

Judgements of loudness discomfort are characterized by appreciable intersubject response variability (Stephens and Anderson, 1971; Geller and Margolis, 1984; Bentler and Pavlovic, 1989; Elberling and Nielsen, 1993). This fact complicates the production of normative loudness data for clinical and rehabilitative use (Hawkins, 1980; Pascoe, 1988; Valente et al, 1997; Elberling, 1999; Brand and Hohmann, 2001). Normative data, however, are useful and necessary for establishing abnormalities of loudness perception. Loudness data depend on a number of factors (Skinner, 1988), including instruction set (Beattie et al, 1979; Geller and Margolis, 1984; Cox et al, 1997), stimulus properties (e.g., bandwidth, frequency, and duration) (Bentler and Pavlovic, 1989; Ricketts and Bentler, 1996), and psychophysical procedure (Elberling and Nielsen, 1993; Jenstad et al, 1997; Rasmussen et al, 1998). The most common psychophysical procedures for assessing loudness perception are categorical scaling and magnitude estimation. The two procedures differ primarily in the nature of the response used to define the loudness continuum. Arbitrary numbers are assigned to a given sound intensity in magnitude estimation, whereas adjectives (or numerical correlates), corresponding to a given response category (e.g., soft, comfortable, loud, etc.), are used in categorical scaling. Both procedures yield suprathreshold perceptual judgements to sound intensities along a continuum extending from soft to (uncomfortably) loud (Yost and Nielsen, 1985).

Typically, because of time and cost constraints, clinicians do not measure or sample judgements across the full continuum for loudness. Instead, the clinician usually estimates the endpoints of the loudness continuum. These endpoints correspond generally to absolute judgements for very soft sounds at the threshold of audibility, presumably the lower limit of the loudness continuum, and to absolute judgements of loudness discomfort for intense sounds, representing the functional upper limit of the dynamic range for sound intensity. The latter estimate will be referred to in this study as the loudness discomfort level (LDL). The absolute difference between the lower and upper limits of this continuum for sound intensity provide an estimate of the auditory dynamic range (DR).

Because the auditory DR for sound intensity is both historically and clinically significant for audiology, it was surprising to us that an informal review of the modern literature revealed only a couple of studies in which estimates of the normal DR for sound intensity have been explicitly documented (Geller and Margolis, 1984; Elberling and Nielsen, 1993). Moreover, it was surprising that in these recent studies the estimates were based on small sample sizes, typically on the order of ten subjects. Perhaps the limited interest in this index stems from the fact that the normal auditory DR for sound intensity is commonly believed to be well known and established. Indeed, one often hears audiologists and hearing scientists report that the auditory DR is on the order of 120 dB. It is unclear what the source is for this "nominal" value, but it may derive from classic figures of the auditory DR in textbooks such as Davis and Silverman's *Hearing and Deafness*. For example, in Figure 2.5, Davis (1970) presents composite data derived from various sources (now more than 60 years old) to illustrate an idealized auditory DR with various upper limit thresholds for "discomfort" (120 dB SPL), "tickle" (130 dB SPL), and "pain" (140 dB SPL). Alternatively, perhaps the assumed value of 120 dB for the auditory DR is more simplistic and can be ascribed to the limits of most commercially available audiometers, which are on the order of 120 dB HL for the midrange audiometric frequencies. Whatever the source of the "nominal" value for the normative auditory DR, this value needs to be updated and validated for a large sample of normal-hearing persons from whom estimates are measured for both ends of the response continuum.

We would expect absolute and relative estimates of loudness discomfort to be generally similar, but we also might reasonably expect systematic differences in these estimates in the same individuals because of inherent differences in the nature of the measurements (Skinner, 1988; Elberling and Nielsen, 1993; Jenstad et al, 1997; Rasmussen et al, 1998). To the extent that differences in the estimates of loudness discomfort exist for absolute and relative judgements, we also may expect related differences in estimates of the listener's auditory DR, which is based on judgements of loudness discomfort. However, we could find only two studies in the literature that directly compared absolute and relative judgements of loudness discomfort in the same group of listeners. Geller and Margolis (1984) measured and compared judgements of loudness for the upper end of the auditory DR using an ascending LDL method and a magnitude estimation protocol. They reported surprisingly good agreement (within ~2 dB) between their LDL data and magnitude estimation judgements of loudness discomfort. Elberling and Nielsen (1993) compared absolute judgements of LDLs (measured using an ascending method of limits) with relative judgements of loudness discomfort measured by magnitude estimation and categorical scaling. They reported for their sample of ten normal-hearing listeners that the presentation levels measured for the LDL were lower (by ~4–7 dB) than the discomfort levels reported for either of the two loudness scaling methods.

Categorical scaling of loudness is complicated by response variability as assessed in estimates of intrasubject, test-retest reliability. Estimates of reliability across the response categories of the loudness continuum differ somewhat within a given study (Robinson and Gatehouse, 1996; Cox et al, 1997), but there appears to be no consistent pattern of reliability differences across categories when one examines data from past loudness-scaling reports. For instance, the results from Cox et al (1997) revealed that reliability differences were greater at the upper end of the categorical scale than at the lower end. In contrast, Robinson and Gatehouse (1996) reported greater reliability differences at the lower end of the categorical scale than at the upper end. Rasmussen et al (1998) and Dirks and Kamm (1976) reported greater reliability differences for the middle category of the loudness scale, with less variation at the lower and/or upper ends of the scale.

The purpose of our study was four-fold: (1) to generate normative values and ranges of intersubject variation for the auditory DR; (2) to compare measures of absolute (i.e., LDL) and relative (i.e., categorical scaling) loudness discomfort; (3) to evaluate and compare intersubject response variability and intrasubject, test-retest reliability for each procedure; and (4) to compare our loudness data with other related loudness data reported in the literature, especially with those from Cox et al (1997) whose categorical-scaling protocol we used here. We also sought to address a general clinical issue, namely, the loose practice of treating the LDL as an absolute threshold of discomfort analogous to the audiometric threshold of audibility. For a number of reasons, including susceptibility to extraneous conditions of measurement and associated sizable response variation, many investigators do not assume the LDL response to be an absolute threshold estimate. Nonetheless, the LDL is effectively an absolute judgement that is often used by clinicians as a de facto threshold for sound tolerance when deriving information for limiting hearing aid output. In fact, one can find LDL terminology used in the audiology literature interchangeably and synonymously with "threshold of discomfort" (Davis, 1970;

Skinner, 1988; Bentler and Pavlovic, 1989; Henry et al, 2002).

EXPERIMENT 1—NORMATIVE AUDIBILITY, LDL, AND DR ESTIMATES

Our purpose in this first experiment was to generate normative data to characterize audiometric thresholds, LDLs, and the auditory DR among adults with normal auditory function. Audiometric thresholds and LDLs were measured from a pool of subjects who were being screened to assess eligibility for a separate research project. The resulting data were used here to generate the normative estimates for audiometric thresholds, LDLs, and DR.

METHODS

Subjects

Audiometric thresholds and LDLs were measured from 59 subjects who self-reported normal-hearing sensitivity and denied sound tolerance problems. The participant age range was 19 to 40 years (mean age = 28 years). There were 21 males and 38 females in our sample.

Procedures

Subjects were seated for evaluation in a double-walled audiometric test suite (IAC, series 1400ATT). The pure-tone stimuli were produced by a clinical audiometer (Grason-Stadler, model GSI-10, calibrated quarterly according to standard procedures [ANSI S3.6-1996]) and presented to the subject via headphones (Telephonics, model TDH-50P in MX41/AR cushions). Audiometric thresholds were measured for each octave frequency between 250 and 8000 Hz, using the modified Hughson-Westlake procedure (Carhart and Jerger, 1959). LDLs were measured in each ear at 500, 1000, 2000, and 4000 Hz. Stimulus duration was approximately 1000 msec, with an interstimulus interval of approximately 400 msec. Subjects were instructed to press the handheld response button when the signal level became uncomfortable (see specific instructions in the Appendix). The starting presentation level was varied between 60 and 70 dB HL for each ascending run, using a 5 dB step size. Two ascending runs were completed per frequency, and the higher presentation level was recorded and used for data analysis. We recorded a value 5 dB greater than the limiting value (e.g., 120 dB HL) for the LDL response when the discomfort level exceeded the limits of the audiometer. The procedure used to evaluate LDLs closely followed that used for patient assessment by the University of Maryland Tinnitus and Hyperacusis Center (Gold et al, 1999).

RESULTS AND DISCUSSION

Average audiometric threshold and LDL data are shown in Table 1. Average audiometric thresholds are on the order of 7–9 dB across frequency. These threshold data are characterized by standard deviations of 5.16 to 6.53 dB, reflecting clinically acceptable intersubject response variability. On average, the LDL estimates slightly exceeded 100 dB

Table 1. Audiometric Thresholds, LDLs, DRs, and Associated Standard Deviations as a Function of Frequency for 59 Normal-Hearing Listeners

Frequency	500 Hz	1000 Hz	2000 Hz	4000 Hz
Audiometric Threshold (dB HL)	8.66	9.01	7.11	8.23
SD (dB)	5.17	5.16	6.53	5.83
Average LDL (dB HL)	102.20	103.86	101.65	100.85
SD (dB)	11.81	10.67	11.95	13.58
DR (dB)	93.69	95.00	94.66	92.75
SD (dB)	13.13	11.88	13.88	14.92

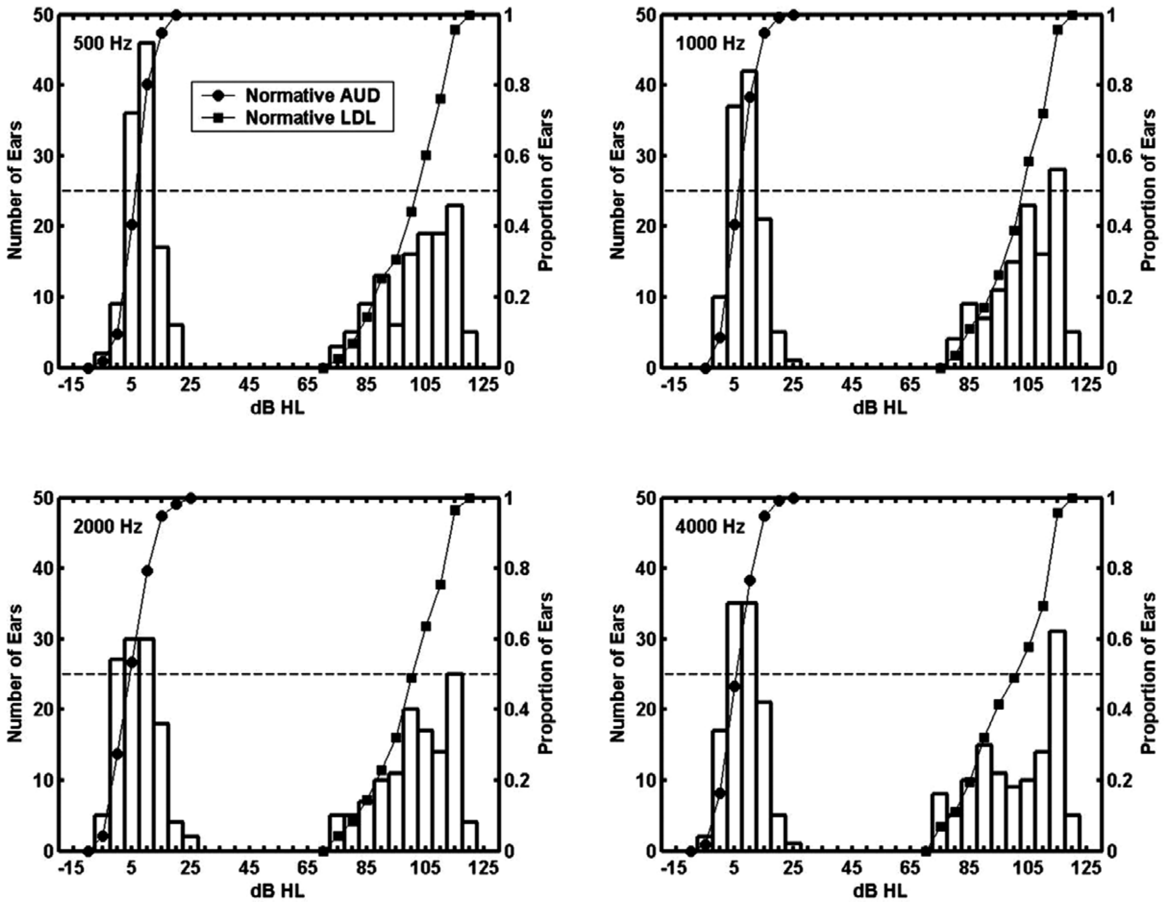

Figure 1. Frequency distribution histograms for the test frequencies 500, 1000, 2000, and 4000 Hz. Each bar represents the number of ears yielding audiometric thresholds (left-side bars) and LDLs (right-side bars) within 5 dB-interval bins for each ear of 59 normal-hearing adults. Superimposed on the respective histograms are the cumulative distribution functions representing the corresponding proportions of ears in the total sample for the audiometric (filled circles) and LDL (filled squares) data.

HL across frequency. The associated standard deviations, 10.67 to 13.58 dB, are about twofold larger than those for the corresponding audiometric thresholds and represent the expected larger intersubject response variability for the LDL.

To evaluate the ranges and distributions of the audiometric thresholds among our normative listener sample, we constructed the frequency distribution histograms shown in Figure 1. Displayed for each test frequency condition are the numbers of ears in our sample that yielded audiometric thresholds within various 5 dB-interval bins across the measurement range (in dB HL). Superimposed on each histogram is a cumulative distribution function representing the corresponding proportions of ears in our total sample summed across this range of threshold levels. For purposes of comparison, the associated histograms and cumulative distribution functions also are presented for the ranges and distributions of the LDL values for each test frequency condition in Figure 1. The audiometric thresholds span comparable ranges of about 30 dB across each test frequency, typically extending from -5 to 25 dB HL. The corresponding ranges of the LDL estimates extend from about 75–120 dB HL, representing intersubject variation over 45 dB for a given frequency condition. Accordingly, the cumulative distribution function for the audiometric thresholds is steeper than the corresponding LDL function for each test frequency condition. Thus, these normative patterns reveal less intersubject response variability at the lower limit of the DR (audiometric threshold) compared to the variation in responses across subjects represented by the LDL at the functional upper limit of the DR. There is little or no effect of test frequency on the ranges and distributions of the data for either the audiometric thresholds or the LDL data.

The absolute difference between the LDL estimate and the audiometric threshold for

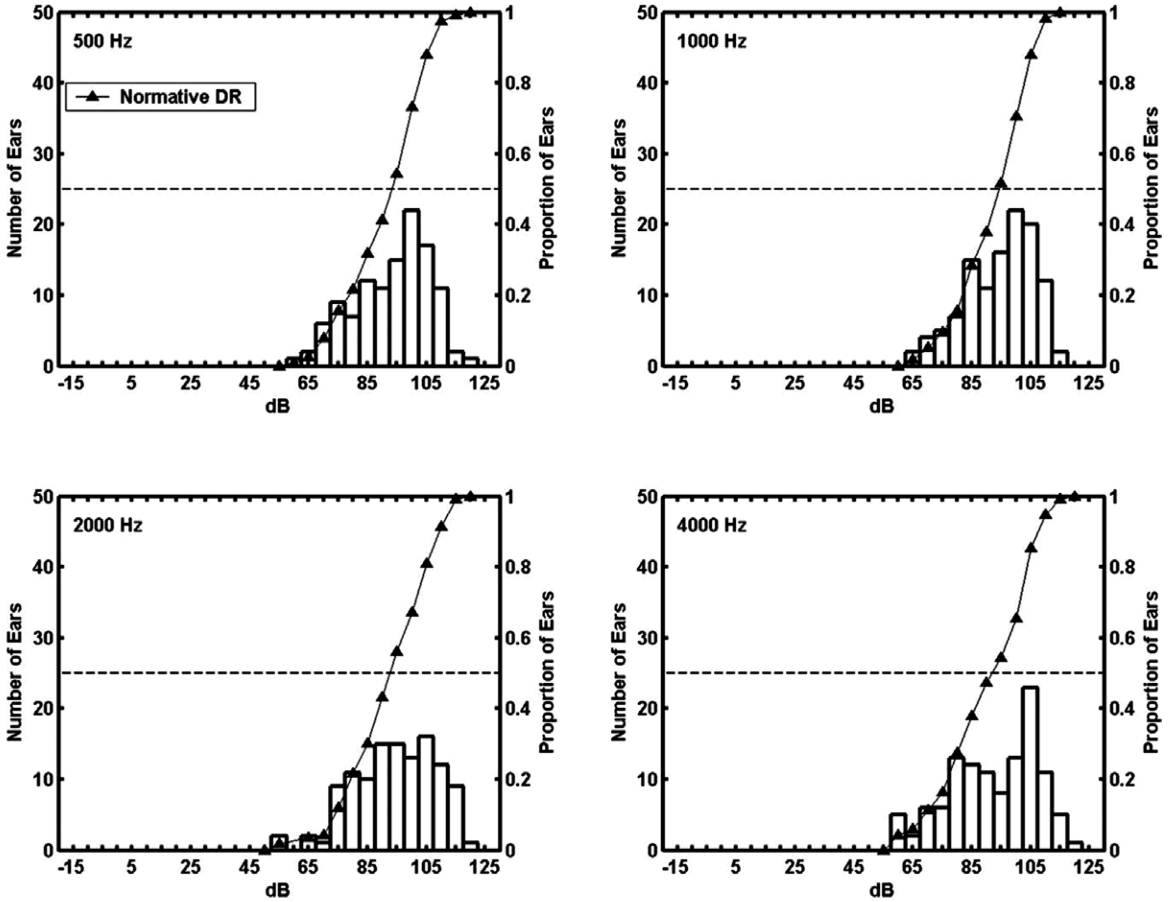

Figure 2. Frequency distribution histograms and cumulative distribution functions for the auditory DR estimates for the test frequencies 500, 1000, 2000, and 4000 Hz.

each frequency condition in Table 1 provides a measure of the corresponding DR for our normal sample of 59 subjects. These DR values are presented in the lower portion of Table 1, along with the associated standard deviation estimates. The resulting DR values are relatively invariant across frequency, with the smallest value (92.75 dB) measured for 4000 Hz and the largest value (95.00 dB) measured for 1000 Hz. Thus, the average DR for our normative sample routinely exceeds 90 dB, with a standard deviation on the order of 12 to 15 dB.

The corresponding frequency distribution histograms and cumulative distribution functions for the DR estimates are shown in separate panels of Figure 2 for each frequency condition. There is little, if any, meaningful difference in the ranges and distributions of the DR values across frequency. However, the remarkable observation in reviewing these data is that the intersubject variability across the DR estimates can be as great as 60 dB (i.e., range from ~60 to 120 dB) for a given frequency condition. This is a surprisingly large range of normal intersubject response variability for the auditory DR, which, to our knowledge, has not been appreciated or recognized heretofore.

EXPERIMENT 2—COMPARISON OF METHODS

Our second objective was to compare different methods for measuring the upper limit of the auditory DR. It is common clinical practice to measure the LDL rather than the full loudness growth function, to characterize the upper end of the auditory DR (Martin et al, 1998; Medwetsky et al, 1999). At issue in this experiment was whether absolute and relative judgements of loudness discomfort from the same listeners would yield significant differences in the estimate of the upper limit of the auditory DR.

METHODS

Subjects

A subgroup of 18 normal-hearing subjects from Experiment 1 participated in Experiment 2. This set of subjects was comprised of 13 females and 5 males, ages 22 to 40 years (mean age of 28 years). Hearing thresholds for all participants were between 0 and 25 dB HL in the frequency region from 500–6000 Hz.

Procedures

Loudness Discomfort Level

The absolute estimates of loudness discomfort, the LDLs, were measured in a single-wall test booth (IAC, series 400) with bilateral ER-3A insert earphones (calibrated in a 2-cc coupler according to ANSI standards). The LDLs were otherwise measured using the same procedures described in Experiment 1. LDL measurement was completed before the Contour Test of Loudness.

Contour Test of Loudness

Categorical scaling judgements of loudness were measured following the protocol described by Cox et al (1997) for the Contour Test of Loudness. These relative measurements were performed on the same day that the LDLs were measured for each listener, again, in a single-wall audiometric test booth using ER-3A insert earphones. The warble-tone stimuli (FM ± 5% of center frequency) were produced by a clinical audiometer (Grason-Stadler, model GSI-10). Loudness judgements were measured independently for 500 and 2000 Hz using an ascending-level method. The standard Contour instructions were provided in writing and were read aloud to each subject (see Appendix). The initial starting level was 20 dB HL for all subjects. The presentation level was incremented in 5 dB steps, as recommended for subjects with normal-hearing sensitivity (Cox et al, 1997). Four 200 msec pulses of the warble tone were presented at each stimulus level. The interstimulus interval was typically on the order of 1000 msec but varied somewhat based upon each subject's response time. The subject responded to the perceived loudness of the tones after presentation of the series of tones at a given level. The initial presentation level of 20 dB HL routinely yielded a response of "very soft" (category 1). Thereafter, stimulus level was increased systematically and a categorical judgement of loudness was obtained for each presentation level until the subject reported a response of "uncomfortably loud" (category 7), which terminated the trial sequence. Three consecutive ascending trial sequences were conducted at each frequency for each subject. The median value for each rating category was determined from the three trial sequences and used for final analysis. In Experiment 2, the right ear was always tested first, but the frequency order was randomized.

RESULTS AND DISCUSSION

Average LDLs for Experiment 2 are shown in Table 2. Two of the subjects had missing LDL data for Experiment 2, so the values are reported here for 16 subjects. Values ranged from a low of 83 dB HL at 4000 Hz to a high of 91 dB HL at 500 Hz, with standard deviations on the order of 7.50 to 9.07 dB.

Table 2. LDLs in dB HL and Corresponding Standard Deviations (in dB) for the Frequencies Tested in Experiment 2

Frequency	Experiment 2 (N = 16)
500 Hz	91.41 (8.91)
1000 Hz	91.25 (9.07)
2000 Hz	87.97 (7.50)
4000 Hz	83.44 (8.93)

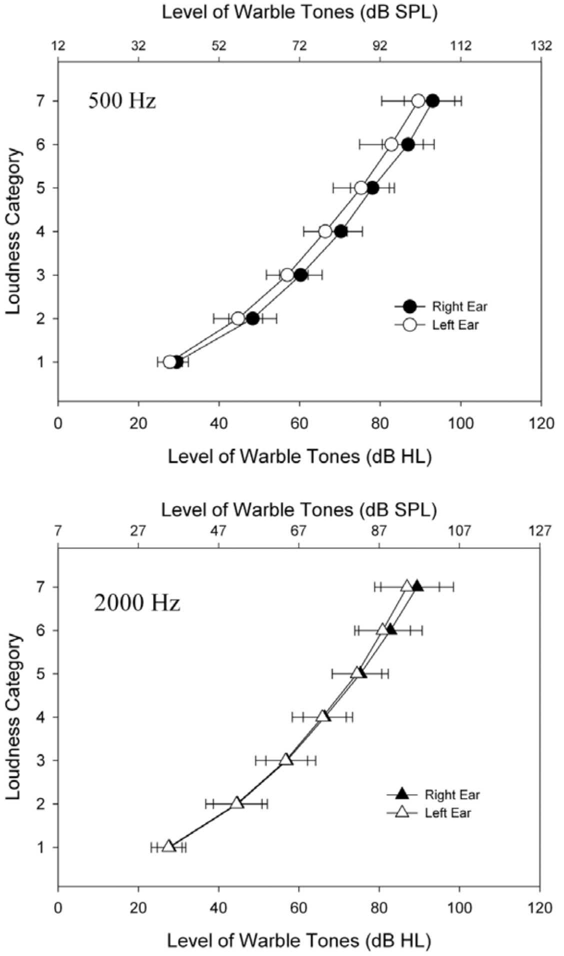

Figure 3. Average presentation levels assigned by normal-hearing adults (N = 18) as a function of loudness category, from "very soft" (category #1) to "uncomfortably loud" (category #7) for 500 Hz (top panel) and 2000 Hz (bottom panel). Presentation level is represented in dB HL on the lower axis and in dB SPL on the upper axis. Left-ear values are represented by open symbols and right-ear values are coded by closed symbols.

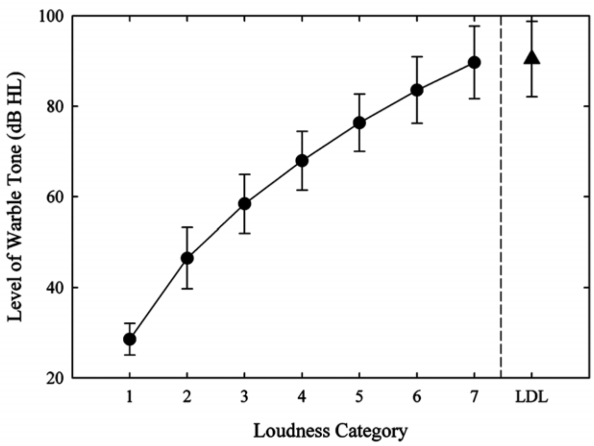

Figure 4. Average presentation levels for each loudness category measured with the Contour Test are shown for the group of normal-hearing adults (N = 18) from Figure 3. Their loudness growth data for the right and left ears and for the two frequencies, 500 and 2000 Hz, are combined here. Right/left ear and 500/2000 Hz LDL data were combined for the same listeners and are shown here on the right side of the panel alongside the value for Contour loudness category "7."

The Contour data for the right and left ears of all 18 subjects are shown in Figure 3 for 500 Hz (top panel) and for 2000 Hz (bottom panel). The categorical loudness judgements, ranging from "very soft" (category 1) to "uncomfortably loud" (category 7), are displayed as a function of presentation level in dB HL along the bottom axis and in dB SPL along the top axis of each panel. The loudness growth functions are generally comparable for the two frequencies. The average presentation level of 28 dB HL for the category of "very soft" is the same for 500 and 2000 Hz, while the average presentation level corresponding to a rating of "uncomfortably loud" is 92 dB HL for 500 Hz and 88 dB for 2000 Hz. The range of standard deviations is 2.91 to 4.29 dB for a rating of "very soft" and is 7.10 to 9.06 for a rating of "uncomfortably loud." This finding reflects about a two-fold increase in response variability from the lower to the upper limit of the auditory dynamic range.

The two measures of loudness discomfort, the absolute estimate of the LDL and the relative estimate of "uncomfortable loudness" from the Contour Test of Loudness, were compared in a three-factor analysis of variance. The results revealed no significant main effects or interactions (F = 1.35, df = 7, p = .231) for ear, test method, or frequency. Consequently, data from the right and left ears and for 500 and 2000 Hz were combined for each test method and plotted together in Figure 4. The loudness response category for the Contour Test is plotted as a function of presentation level (in dB HL). The upper limit of the auditory DR is virtually identical for the LDL and the categorical judgement of uncomfortable loudness, both in terms of the mean values (90.40 vs. 89.56 dB HL) and intersubject response variability (8.34 vs. 8.02 dB). This finding indicates that, on average, estimates of the upper limit of the auditory DR are comparable for absolute and relative judgements of loudness discomfort in the same listeners.

An apparent discrepancy is evident in the mean LDL values between Experiment 1, based upon our large subject pool (n = 59), and Experiment 2, for our subgroup (n = 16) selected from the larger subject pool. This discrepancy can be seen by comparing the corresponding results in Tables 1 and 2. The primary reason for this discrepancy is that the subjects in Experiment 2 were purposely recruited because of their normal-but-lower LDLs, which were necessary to minimize ceiling effects when assessing treatment outcomes among these participants in a separate and unrelated study (Formby et al, 2002, 2003a, 2003b). This discrepancy was irrelevant in Experiment 2. That is, our purpose in Experiment 2 was to compare absolute and relative judgements of loudness discomfort for assessing the upper limit of the auditory DR within the same set of subjects on the same measurement day. These are the values that we have reported here. A possible secondary reason that may explain part of the discrepancy is that earphones were used in Experiment 1, whereas insert receivers were used in Experiment 2 to present the sound stimuli. This latter explanation seems unlikely because Valente et al (1997) showed only small differences of 2–3 dB in LDL estimates measured with the two different transducers. Consistent with the findings of Valente et al (1997), a comparison of the average LDL values for our 16 subjects who participated in both Experiment 1 (headphones) and Experiment 2 (insert phones) revealed differences of 1–3 dB. Thus, a transducer effect was probably not a significant contributing factor to the differences in LDLs between our large group and our smaller subgroup of listeners.

EXPERIMENT 3—TEST-RETEST RELIABILITY

Our third objective was to evaluate and compare the intrasubject test-retest reliability of the two test methods used in Experiment 2 for assessing loudness discomfort.

Subjects

The same 18 subjects who were tested in Experiment 2 participated in Experiment 3.

Procedures

The procedures for measuring the LDLs and the categorical scaling judgements of loudness were the same as those described in Experiment 2. The average time interval between test sessions was 10 days (SD = 7.07).

RESULTS AND DISCUSSION

The mean test and retest LDL and Contour loudness data are presented in Tables 3 and 4, respectively, for each ear as a function of test frequency. Also shown are the average

Table 3. LDL Test vs. Retest Levels, Test-Retest Differences, and Corresponding Standard Deviations for the Right and Left Ears as a Function of Frequency for 16 Normal-Hearing Subjects

		Frequency			
Test Session	Ear	500 Hz	1000 Hz	2000 Hz	4000 Hz
Test 1 dB HL (SD)	R	92.19	91.56	89.69	82.81
		(9.12)	(9.44)	(8.06)	(9.66)
	L	90.63	90.94	86.25	84.06
		(8.92)	(8.98)	(6.71)	(8.41)
Test 2 dB HL (SD)	R	95.00	94.17	91.94	87.50
		(8.22)	(8.62)	(6.89)	(6.00)
	L	93.53	94.12	90.88	86.76
		(6.56)	(7.12)	(7.75)	(7.06)
Test-Retest Difference dB (SD)	R	2.19	1.56	1.56	4.06
		(9.12)	(7.24)	(7.00)	(8.21)
	L	2.33	2.33	4.67	2.33
		(6.23)	(6.23)	(6.94)	(5.63)

Table 4. Contour Test vs. Retest Levels, Test-Retest Level Differences, and Corresponding Standard Deviations (SD in dB) as a Function of Loudness Response Category for the Right and Left Ears for 500 and 2000 Hz for 18 Normal-Hearing Subjects

500 Hz		Loudness Response Category						
Session	Ear	1	2	3	4	5	6	7
Test 1 dB HL (SD)	R	29.44	48.33	60.28	70.28	78.06	86.94	93.06
		(2.91)	(5.94)	(5.28)	(5.28)	(5.46)	(6.45)	(7.10)
	L	27.78	44.72	56.94	66.39	75.28	82.78	89.44
		(3.08)	(6.06)	(5.18)	(5.37)	(6.96)	(7.90)	(9.06)
Test 2 dB HL (SD)	R	28.06	44.44	56.11	65.83	75.00	82.78	90.28
		(3.04)	(6.39)	(6.76)	(6.47)	(6.86)	(6.91)	(7.57)
	L	27.50	43.33	54.72	64.17	73.33	81.94	87.78
		(2.57)	(5.69)	(6.96)	(7.72)	(7.28)	(9.10)	(9.58)
Test-Retest Difference dB (SD)	R	-1.39	-3.89	-4.17	-4.44	-3.06	-4.17	-2.78
		(2.87)	(5.02)	(6.00)	(5.91)	(5.98)	(4.93)	(6.00)
	L	-0.28	-1.39	-2.22	-2.22	-1.94	-0.83	-1.67
		(2.70)	(6.14)	(6.00)	(7.12)	(7.10)	(3.09)	(8.22)
2000 Hz		**Loudness Response Category**						
Session	Ear	1	2	3	4	5	6	7
Test 1 dB HL (SD)	R	29.44	48.33	59.72	69.17	77.50	83.61	89.17
		(3.38)	(6.64)	(7.37)	(6.91)	(6.47)	(7.24)	(7.12)
	L	27.50	44.44	56.67	65.83	74.44	80.83	86.94
		(4.29)	(7.65)	(7.48)	(7.52)	(6.16)	(6.91)	(8.07)
Test 2 dB HL (SD)	R	27.78	43.89	56.39	65.00	74.72	82.22	88.33
		(3.08)	(6.76)	(7.03)	(6.64)	(6.96)	(8.08)	(8.40)
	L	28.33	44.17	55.83	65.00	73.06	80.00	86.39
		(3.83)	(6.91)	(8.09)	(7.86)	(7.51)	(8.40)	(9.67)
Test-Retest Difference dB (SD)	R	-1.67	-4.44	-3.33	-4.17	-2.78	-1.39	-0.83
		(4.20)	(7.05)	(7.67)	(7.52)	(6.00)	(6.82)	(7.12)
	L	0.83	-0.28	-0.83	-0.83	-1.39	-0.83	-0.56
		(3.09)	(4.99)	(6.00)	(6.47)	(5.89)	(6.00)	(6.84)

test-retest differences for each measurement procedure. Test-retest differences are comparable for the two measurement procedures and are routinely less than 5 dB for all frequency conditions.

A three-factor analysis of variance (ANOVA) was completed on the LDL data (ear, frequency, and session) and revealed a significant main effect only for frequency (F = 11.94, df = 4, p < .01), with no interactions. The LDLs measured for 4000 Hz were about 4 dB lower than those measured for 500, 1000, and 2000 Hz. A four-factor ANOVA of the Contour test data (ear, frequency, loudness category, and measurement session) revealed a significant main effect for ear (F = 24.64, df = 1, p < .01), loudness category (F = 1487.63, df = 6, p < .01), and measurement session (F = 23.26, df = 1, p < .01), with no interactions. The average value for the right ear was 2 dB higher than the left ear. As expected, mean values across rating categories were significantly different. The mean value in retest session 2 was 2.03 dB lower than that measured in initial test session 1. However, these small, statistically significant differences are not clinically meaningful in view of the standard practice among audiologists of using 5 dB measurement step sizes.

GENERAL DISCUSSION

The auditory DR estimated from the differences between the LDLs and audiometric thresholds for our full sample of 59 subjects was on the order of 95 dB and was relatively invariant across frequency from 500–4000 Hz. This normative estimate is slightly smaller than Elberling and Nielsen (1993) reported for a smaller set of 10 normal-hearing subjects. Their respective mean values for 500 and 2000 Hz were 101.95 (SD = 11.37 dB) and 99.25 (SD = 13.90 dB) dB HL.

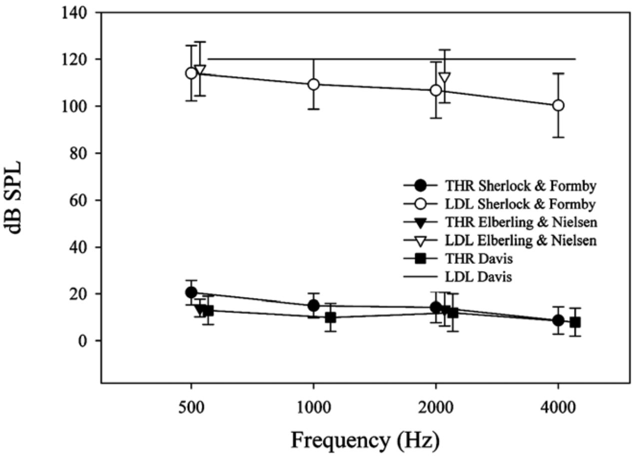

Figure 5. Audiometric threshold and LDL (sound tolerance) estimates as a function of frequency depict various estimates of the auditory DR. The closed symbols represent audibility thresholds from the current study, from Elberling and Nielsen (1993), and from Davis (1970). All data have been converted into dB SPL using correction factors for the corresponding transducer. The open symbols represent estimates of the upper limit of sound tolerance (i.e., LDLs in the current study and from Elberling and Nielsen, 1993). The straight line at 120 dB represents the approximate median LDLs reported by Davis (1970), which were measured at CID by Silverman (1947).

Our corresponding mean values for 500 and 2000 Hz were 93.69 (SD = 13.13) and 94.66 (SD = 13.88) dB, respectively.

Our DR estimates, however, are substantially less than those that one calculates from the differences in the thresholds of discomfort and audibility shown by Davis (1970) in *Hearing and Deafness*, figure 2.5. Consider our data in relation to the comparable data shown by Davis, which we have replotted in Figure 5. Davis assumed invariant discomfort thresholds across frequency at 120 dB SPL (based upon approximate median LDL values from Silverman [1947]). Davis also presented median audibility thresholds reported by the National Physical Laboratory in Britain for 99 otologically normal men, 18 to 25 years of age. These latter values, shown here in Figure 5, are comparable to the average audiometric thresholds measured for our sample. The resulting auditory DR values derived from Davis are on the order of 108–112 dB over the range from 500–4000 Hz. His values, therefore, actually are somewhat less than the popularly cited value of 120 dB and are about 15 dB greater than our corresponding empirical estimates, which we measured from a single sample of listeners.

We are not aware of previous reports describing such large intersubject response variation as that shown in Figure 2 for the normal auditory DR. This sizable variation reflects a relatively small amount of intersubject response variability associated with estimates of the audiometric threshold and an appreciably larger contribution from the intersubject response variability associated with LDL judgements at the upper limit of the DR. The respective contributions from the ranges of the audiometric thresholds and the LDL estimates in Figure 1 to the larger intersubject variability in the auditory DR values in Figure 2 are quantified for each test frequency condition in Table 5. Shown in

Table 5. Normative Percentiles for Audiometric Threshold, LDLs, and DR Values Based on Responses from a Group of 59 Normal-Hearing Adults

Audiometric Threshold (dB HL)	5%	25%	50%	75%	95%
500 Hz	0	5	10	10	20
1000 Hz	0	5	10	10	20
2000 Hz	0	0	5	10	20
4000 Hz	0	5	10	10	20
LDL (dB HL)					
500 Hz	80	95	105	110	120
1000 Hz	85	95	105	115	115
2000 Hz	80	95	105	110	115
4000 Hz	75	90	105	115	115
DR (dB)					
500 Hz	70	85	95	105	110
1000 Hz	70	85	95	105	110
2000 Hz	75	85	95	105	115
4000 Hz	65	80	95	105	115

Table 5 are the audiometric thresholds and LDL values representing the 5th, 25th, 50th, 75th, and 95th percentiles that we estimated from the cumulative distribution functions for the corresponding test frequency conditions in Figure 1. Also shown are the corresponding percentiles for the auditory DR values that we estimated from the cumulative distribution functions in Table 2. Whereas the 5th to 95th percentile range for the audiometric thresholds consistently spans 20 dB, the comparable range of LDL estimates is on the order of 35 to 40 dB. In turn, the resulting 5th to 95th percentile ranges for the auditory DR values span 40 to 50 dB across our test frequency conditions.

A related and remarkable finding in this research is that a few subjects in our larger sample of 59 normal volunteers had LDLs as low as 70 dB HL. This result means that for these subjects their LDLs were 30 dB below the group mean LDL. One might reasonably expect some of these individuals would be bothered by moderately loud sounds and would suffer hyperacusis-like symptoms. However, none of the subjects in our sample reported inordinate sound tolerance problems. This fact is perplexing, but it is not surprising in light of growing evidence that there may be an underlying psychological component for many individuals who report inordinate distress to sounds that normally are judged to be comfortable by most persons (see Nelting et al, 2002).

Results for our two methods for measuring loudness discomfort judgements (whether by a simple absolute estimate of the LDL or a relative judgement from categorical scaling) reflect statistically comparable measures of loudness discomfort, despite differences in instruction set, stimulus properties, and psychophysical procedure. In fact, the methods are virtually identical in terms of the estimates of average loudness discomfort (Contour category “7” vs. LDL), intersubject response variability, and intrasubject test-retest reliability. This finding supports the validity of using simple LDL measures rather than categorical scaling to estimate the upper limit of sound tolerance when clinical time is limited.

A comparison of our Contour results with the findings of Cox et al (1997) reveals only slight differences. The normal-hearing subjects who participated in Experiments 2 and 3 assigned the softer ratings to slightly more intense stimuli and the louder ratings to slightly less intense stimuli than were reported by Cox et al. In other words, the DR for the listeners in this study was slightly smaller than that reported by Cox et al. The difference for the lower end of the rating scale probably can be attributed to differences in the starting levels used by us and by Cox et al for the Contour test. The starting presentation level used by Cox et al was 5 dB SL, whereas the listeners in this study were started at 20 dB HL to maintain consistency across test sessions. Examination of the audiometric threshold data for our listener group reveals that the 20 dB HL starting level typically corresponded to 5–10 dB SL for most of our sample of listeners. The difference for the upper end of the rating scale in this study and in Cox et al is likely related to our very specific subject selection criterion. As discussed earlier, our smaller sample of subjects reported normal sound tolerance and produced LDLs that were toward the low end of the normal range. These selection criteria minimized ceiling effects, allowing us to observe a wider range of treatment-related changes in categorical loudness judgements (Formby et al, 2002, 2003a, 2003b).

An important point of agreement between our categorical loudness judgements and those of Cox et al (1997) is the common trend shown in Table 6 for change in intersubject variability across response categories. Cox et al reported a general trend for variability to increase systematically as the response changed from the “very soft” to the “loud” and “uncomfortably loud” categories with increasing sound intensity. We see a similar trend in our data, with lowest response variability (SD = 3.50 dB) measured for the “very soft” category and the highest response variability (SD = 8.02 dB) measured for the “uncomfortably loud” category. However, in our data the intersubject variability is otherwise fairly constant across the intermediate categories (6.34–7.34 dB). These trends are consistent with Skinner’s observation that, in many studies, listeners make more reliable responses at threshold (SD = 2–4 dB) than at the LDL (SD = 4–6 dB). This is a potentially significant observation, and we will revisit this issue later. However, before concluding this discussion of intersubject variability, we note that other studies of categorical scaling do not report a consistent trend in the pattern of variability

Table 6. Intersubject Variability Standard Deviations in dB for Categorical Loudness Scaling Measured in This Study and in Related Studies

Loudness Category	Sherlock and Formby	Cox et al (1997)	Rasmussen et al (1998)
Very soft	3.50	5.75	7.60
Soft	6.74	8.50	12.15
Comfortable, but slightly soft	6.49	10.05	
Comfortable	6.49	10.45	8.20
Comfortable, but slightly loud	6.34	11.70	
Loud, but OK	7.34	12.20	5.95
Uncomfortably loud	8.02	13.50	2.10

across response categories. The data from one such study are highlighted, along with our variability data and those of Cox et al, in Table 6. Specifically, Rasmussen et al (1998) show less intersubject variability at the extremes of the response continuum and maximum variability for "soft" and "comfortable" categories (which is consistent with Dirks and Kamm [1976] observations for greater variability at MCL). Thus, the published data are in disagreement.

Previous studies have suggested that the level of a preceding stimulus (i.e., the "anchor effect" of a starting stimulus level) influences the ultimate judgement of loudness (Beattie et al, 1997; Jenstad et al, 1997; Keidser et al, 1999). Our results do not appear to support this idea. The average Contour "7" rating was routinely within 5 dB of the LDL rating by the same subject. This consistent agreement occurred despite the fact that for the Contour test the subjects were judging the stimuli beginning from a "very soft" level, systematically providing categorical judgements as sound intensity was raised to an "uncomfortably loud" level. In contrast, for the LDL measurements, subjects were judging the loudness discomfort of stimuli

Table 7. Survey of LDLs in dB SPL and Associated Standard Deviation (SD) Estimates Measured as a Function of Frequency in This Study and in Other Studies for the Transducer and Method of Measurement Indicated

Study	Transducer	Method	No. of Subjects	500 Hz (dB SPL)	1000 Hz (dB SPL)	2000 Hz (dB SPL)	4000 Hz (dB SPL)
Sherlock and Formby	TDH-50P	Ascending	59	114.10 (11.81)	109.39 (10.67)	106.85 (11.95)	100.35 (13.58)
Elberling and Nielsen (1993)	ER-3A	Ascending	10	119.35 (11.49)		111.45 (11.32)	
Bentler and Pavlovic (1989)	ER3/50	Ascending method of limits	15	*105.02 (9.98)	104.22 (10.09)	**104.24 (9.53)	95.68 (11.66)
Geller and Margolis (1984)	TDH-49	Single interval yes/no	10			97.90 (11.7)	
Dirks and Kamm (1976)	TDH-49	Single interval yes/no	2	102.60		99.20	
Morgan et al (1974)	TDH-49	Method of constant stimuli	6	113.10 (5.1)	108.20 (3.4)	108.30 (4.8)	108.50 (7.0)
Stephens and Anderson (1971)	TDH-39	Ascending method of limits	16			93.40 (11.4)	

*570 Hz **2150 Hz

Table 8. Intrasubject Test-Retest Standard Deviations in dB for Each Loudness Category (data combined across frequency) Measured in This Study and in Related Studies

Loudness Category	Sherlock and Formby	Cox et al (1997)	Rasmussen et al (1998)	Robinson and Gatehouse (1996)
Very soft	1.53	2.1	4.4	
Soft	4.17	3.1	8.3	7.3
Comfortable, but slightly soft	3.75	4.0		
Comfortable	4.31	4.4	7.3	5.6
Comfortable, but slightly loud	2.92	4.7		
Loud, but OK	2.78	6.2	4.6	3.5
Uncomfortably loud	1.81	5.9	3.8	

only at and near the upper end of the DR, using some form of absolute internal criterion.

Normal intersubject response variability for the LDL is appreciable in this study and routinely approached or exceeded a standard deviation of 10 dB for our sample of 59 subjects. This finding is consistent with the variability noted in other studies in Table 7 (Stephens and Anderson, 1971; Morgan et al, 1974; Dirks and Kamm, 1976; Geller and Margolis, 1984; Bentler and Pavlovic, 1989; Elberling and Nielsen, 1993). The results from Experiment 2, for our smaller, more homogenous subject pool, reflected somewhat less intersubject variability. These values are characterized by standard deviations on the order of 7–9 dB for both the LDL and category "7" of the Contour Test of Loudness (i.e., "uncomfortably loud").

The results from Experiment 3, in which test-retest reliability was examined, revealed average intrasubject test-retest differences that were usually less than 5 dB for both the absolute and relative loudness measurement procedures. This finding, which suggests comparable test-retest reliability for the two measurement procedures, is consistent with previous reports that intrasubject reliability differences typically are less than 10 dB for both absolute and relative judgements of loudness discomfort (Allen et al, 1990; Ricketts and Bentler, 1996; Robinson and Gatehouse, 1996; Cox et al, 1997; Rasmussen et al, 1998).

Examination of the standard deviations associated with the intrasubject test-retest differences in Table 8 reveals a trend of smaller response variability at the "very soft" end of the categorical scale (SD = 1.53 dB) and at the "uncomfortably loud" end of the scale (SD = 1.81), with greater variability in the middle of the loudness continuum for the "comfortable" levels (SD = 4.31). This trend is consistent with that reported by Rasmussen et al (1998), whose data are shown for comparison in Table 8. By contrast, Cox et al (1997) found increasing intrasubject variability from the "very soft" category to the "uncomfortably loud" category. Her standard deviation values also are shown in Table 8 along with those from Robinson and Gatehouse (1996), whose data reverse the variability trend described by Cox et al. Robinson and Gatehouse's values reflect greater variability for the "soft" response category and systematically less variability for the "comfortable" and "loud" categories. Thus, as we noted for intersubject response variability, results across studies for intrasubject test-retest reliability are inconsistent across loudness response categories. However, in view of the small differences between the absolute and relative judgements of loudness discomfort and the small differences between the associated estimates of intersubject response variability and intrasubject test-retest reliability, we conclude that the simple measure of the LDL can be used as a valid and efficient tool for estimating the upper limit of the auditory DR.

Finally, in concluding this report, we return to our introductory theme. Namely, we revisit the idea that intersubject response variability is appreciable for judgements of loudness discomfort and, therefore, this undesirable property may limit its clinical applications. Implicit in this concept is some reference or standard to which response variability for loudness judgements is being compared. Presumably, the "gold standard" clinical reference for intersubject response variability is the known small response variability for the audiometric threshold. Certainly intersubject response variability in this study is greater for measurements of loudness discomfort, at the upper end of the auditory DR, than for the audiometric threshold, at the lower end of the DR. However, performing a loudness judgement per se probably cannot explain the greater response variability associated with this task. We know this because categorical scaling judgements for "very soft" tones in this study yielded standard deviations across listeners that were of similar magnitude to (if not smaller than) standard deviations measured for audiometric thresholds from the same listeners. This observation, along with the fact that absolute and relative judgements of loudness discomfort are virtually identical in this study, supports the idea that a "threshold of discomfort" may be just as valid clinically as is the audiometric threshold of audibility. However, we would be ill-advised to push this argument beyond reasonable limits because there is ample evidence to the contrary, including some evidence of our own on adaptive recalibration of loudness (Formby et al, 2002, 2003a, 2003b). In the psychological literature, loudness is considered a response proclivity (Watson, 1973), meaning a natural inclination or tendency for judging sound quality but not an absolute threshold judgement per se. We also know that numerous extraneous factors influence and confound loudness judgements (see Skinner, 1988). Nonetheless, our findings do not appear to discourage the idea that, under carefully controlled conditions, the LDL represents a "threshold of discomfort" having clinical validity and significance.

To sum up, normative auditory DR estimates provide a reference for audiologists to use for diagnostic and rehabilitative purposes. Clinicians can use these estimates to identify patients who may be unusually sensitive to moderate and loud sounds when LDLs for a given patient fall appreciably below the normal range. Accordingly, these patients most likely will report discomfort from amplified sound and will have a higher likelihood of rejecting amplification. These patients likely will require intervention above and beyond a traditional hearing aid fitting. Patients identified with reduced sound tolerance may benefit from a gradual increase in hearing aid gain, or alternative intervention methods such as sound therapy to modify sound sensitivity (Formby et al, 2003). Our results also indicate that a simple measurement of the LDL is as reliable and accurate as categorical scaling of loudness for estimating the upper limit of the auditory DR and for identifying "normal" sound tolerance.

Acknowledgment. This research was supported by NIH awards R21DC04514 and R01DC04678 from the National Institute on Deafness and Other Communication Disorders. We gratefully acknowledge Michael Keaser for providing statistical assistance and aid in preparation of Figures 1 and 2. Monica Hawley also provided statistical assistance and helped in preparing Figures 3 to 5.

REFERENCES

Allen JB, Hall JL, Jeng PS. (1990) Loudness growth in 1/2-octave bands (LGOB)—a procedure for the assessment of loudness. *J Acoust Soc Am* 88:745–753.

Beattie RC, Edgerton BJ, Gager DW. (1979) Effects of speech materials on the loudness discomfort level. *J Speech Hear Disord* 44:435–458.

Beattie RC, Huynh RC, Ngo VN, Jones RL. (1997) IHAFF Loudness Contour Test: reliability and effects of approach mode in normal-hearing subjects. *J Am Acad Audiol* 8:243–256.

Bentler RA, Pavlovic CV. (1989) Comparison of discomfort levels obtained with pure tones and multitone complexes. *J Acoust Soc Am* 86:126–132.

Brand T, Hohmann V. (2001) Effect of hearing loss, centre frequency, and bandwidth on the shape of loudness functions in categorical loudness scaling. *Audiology* 40:92–103.

Carhart R, Jerger JF. (1959) Preferred method for clinical determination of pure tone thresholds. *J Speech Hear Disord* 24:330–345.

Cox RM, Alexander GC, Taylor IM, Gray GA. (1997) The Contour Test of Loudness Perception. *Ear Hear* 18:388–400.

Davis H. (1970) Acoustics and psychoacoustics. In: Davis H, Silverman SR, eds. *Hearing and Deafness*. New York: Holt, Rinehart and Winston, 9–46.

Dirks DD, Kamm C. (1976) Psychometric functions for loudness discomfort and most comfortable loudness levels. *J Speech Hear Res* 19:613–627.

Elberling C. (1999) Loudness scaling revisited. *J Am Acad Audiol* 10:248–260.

Elberling C, Nielsen C. (1993) The dynamics of speech and the auditory DR in sensorineural hearing impairment. In: Beilin J, Jensen GR, eds. *Recent Developments in Hearing Instrument Technology.* Proceedings of the 15th Danavox Symposium. Copenhagen, Denmark: Stougaard Jensen, 99–134.

Formby C, Sherlock LP, Gold SL. (2002) Adaptive recalibration of chronic auditory gain: interim findings. In:

Patuzzi R, ed. *Proceedings of the Seventh International Tinnitus Seminar*. Perth, Australia: University of Western Australia, 165–169.

Formby C, Sherlock LP, Gold SL. (2003a) Adaptive plasticity of loudness induced by chronic attenuation and enhancement of the acoustic background. *J Acoust Soc Am* 114:55–58.

Formby C, Sherlock LP, Gold SL. (2003b) Adaptive recalibration of chronic auditory gain. Abstracts of the 26th Midwinter Meeting of the Association for Research in Otolaryngology. Mt. Royal, NJ: Association for Research in Otolaryngology.

Geller D, Margolis RH. (1984) Magnitude estimation of loudness I: application to hearing aid selection. *J Speech Hear Res* 27:20–27.

Gold SL, Frederick EA, Formby C. (1999) Shifts in DR for hyperacusis patients receiving Tinnitus Retraining Therapy (TRT). In: Hazell J, ed. *Proceedings of the VIth International Tinnitus Seminar*. Cambridge, UK: Tinnitus and Hyperacusis Centre, London, 297–301.

Hawkins DB. (1980) Loudness discomfort levels: a clinical procedure for hearing aid evaluations. *J Speech Hear Disord* 45:3–15.

Henry JA, Jastreboff MM, Jastreboff PJ, Schechter MA, Fausti SA. (2002) Assessment of patients for treatment with Tinnitus Retraining Therapy. *J Am Acad Audiol* 13:523–544.

Jenstad LM, Cornelisse LE, Seewald RC. (1997) Effects of test procedure on individual loudness functions. *Ear Hear* 18:401–408.

Keidser G, Seymour J, Dillon H, Grant F, Byrne D. (1999) An efficient, adaptive method of measuring loudness growth functions. *Scand Audiol* 28:3–14.

Martin FN, Champlin CA, Chambers JA. (1998) Seventh survey of audiometric practices in the United States. *J Am Acad Audiol* 9:95–104.

Medwetsky L, Sanderson D, Young D. (1999) A national survey of audiology clinical practices, part 1. *Hear Rev* 6:24–32.

Morgan DE, Wilson RH, Dirks DD. (1974) Loudness discomfort level: selected methods and stimuli. *J Acoust Soc Am* 56:577–581.

Nelting M, Rienhoff NK, Hesse G, Lamparter U. (2002) The assessment of subjective distress related to hyperacusis with a self-rating questionnaire on hypersensitivity to sound. *Laryngorhinootologie* 81:327–334.

Pascoe DP. (1988) Clinical measurements of the auditory DR and their relation to formulas for hearing aid gain. In: Jensen JH, ed. *Hearing Aid Fitting: Theoretical and Practical Views*. 13th Danavox Symposium. Copenhagen, Denmark: Stougaard Jensen, 129–151.

Rasmussen AN, Olsen SO, Borgkvist BV, Nielsen LH. (1998) Long-term test-retest reliability of category loudness scaling in normal-hearing subjects using pure tone stimuli. *Scand Audiol* 27:161–167.

Ricketts TA, Bentler RA. (1996) The effect of test signal type and bandwidth on the categorical scaling of loudness. *J Acoust Soc Am* 99:2281–2287.

Robinson K, Gatehouse S. (1996) Test-retest reliability of loudness scaling. *Ear Hear* 17:120–123.

Silverman SR. (1947) Tolerance for pure tones and speech in normal and defective hearing. *Ann Otol Rhinol Laryngol* 56:658–677.

Skinner MW. (1988) Determining an individual's auditory area. In: *Remediation of Communication Disorders*. Englewood Cliffs, NJ: Prentice Hall, 118–148.

Stephens SDG, Anderson CMB. (1971) Experimental studies on the uncomfortable loudness level. *J Speech Hear Res* 14:262–270.

Valente M, Potts LG, Valente M. (1997) Differences and inter-subject variability of loudness discomfort levels measured in sound pressure level and hearing level for TDH-50P and ER-3A earphones. *J Am Acad Audiol* 8:59–67.

Watson CS. (1973) Psychophysics. In: Wolman BB, ed. *Handbook of General Psychology*. Englewood Cliffs, NJ: Prentice Hall, 275–306.

Yost WA, Nielsen DW. (1985) *Fundamentals of Hearing*. New York: Holt, Rinehart and Winston.

APPENDIX

Instructions for LDL:

I will be presenting tones that get louder and louder. I want you to press the button when the sound is uncomfortably loud.

Instructions for the Contour Test of Loudness:

The purpose of this test is to find your judgements of the loudness of different sounds. You will hear sounds that increase and decrease in volume. You must make a judgement about how loud these sounds are. Pretend you are listening to the radio at that volume. How loud would it be? After each sound, tell me which of these categories best describes the loudness. Keep in mind that an uncomfortably loud sound is louder than you would ever choose on your radio no matter what mood you are in.