

Auditory detection of hollowness

Robert A. Lutfi^{a)}

Department of Communicative Disorders and Waisman Center, University of Wisconsin, Madison, Wisconsin 53706

(Received 6 June 2000; revised 13 November 2000; accepted 15 May 2001)

The airborne sounds produced by freely vibrating hollow and solid bars were synthesized according to the equations of bar motion from theoretical acoustics, and were presented to listeners over headphones. In a two-interval, forced-choice task, listeners were asked to distinguish between the hollow and solid bar sounds as bar length was varied at random from one presentation to the next. All other physical properties of the bar were held constant across trials. Listener decision strategies for detecting hollowness in iron, aluminum, and wood bars were determined from regression weights describing the relation between the listener's response and the frequency, intensity, and decay modulus of the individual partials comprising these sounds. The obtained weights were compared to those of a hypothetical listener that bases judgments on the acoustic relations intrinsic to hollowness, as determined from the equations for motion. Results indicate that listeners adopt roughly one of two decision strategies, either basing judgments on the appropriate acoustic relations, or basing judgments predominantly on frequency alone. The decision strategy of some listeners also changed from one type to the other with a change in bar material or upon replication of the same condition. The results are interpreted in terms of the vulnerability of the intrinsic acoustic relations to small perturbations in acoustic parameters, as would be associated with listener internal noise. They demonstrate that basic limits of human sensitivity can have a profound effect on the identification of rudimentary source attributes from sound, even in conditions where acoustic variation is largely dictated by physical variation in the source. © 2001 Acoustical Society of America. [DOI: 10.1121/1.1385903]

PACS numbers: 43.66.Ba, 43.66.Fe [DWG]

I. INTRODUCTION

Everyday experience suggests that the ear is quite good at inferring basic attributes of objects from sound. The “ping” of the wine glass identifies the glass as fine crystal, the “clunk” of the table struck with your knuckle indicates that the table is hollow and made of wood. Such simple examples represent in principle how we use sound to gain information about our surroundings, but even in regard to these simple examples we understand very little of the process. What are the limits of our ability to determine physical attributes of the source in these cases, and what information in the sound is used to make such determinations? These are the questions that motivate the present experiments.

In earlier papers, we described a method for measuring precisely how a listener's judgments regarding the physical attributes of an object are influenced by specific acoustic parameters and their relations (Lutfi, 1995; Lutfi and Oh, 1994, 1995, 1997). The relative influence, in each case, is estimated from correlations of the listener's judgments with small, experimentally introduced perturbations in the values of acoustic parameters from trial to trial. Importantly, the perturbations are introduced in such a way so as not to violate the lawful relations governing the motion of the object, i.e., the sounds, in some sense, remain real. To do this, the sounds are synthesized digitally using a resonant source for which the equations for motion are known. In this study, as in the previous studies, we chose as our resonant source the

struck, clamped bar. This is the preferred choice inasmuch as the equations for motion are relatively simple, yet they apply to a large class of common, freely vibrating objects. Many familiar musical instruments, for example, fall into the category of struck, clamped bars (triangle, tuning fork, wood block, and xylophone).

Our first experiments employed a standard, two-interval, forced-choice procedure to investigate the discrimination of differences in the material composition of bars (Lutfi and Oh, 1997). The goal was to measure the best possible performance under the most favorable listening conditions. Hence, all physical attributes of the bar were held constant except for material, and all listeners were given extensive training in the task before data collection. To provide a standard for evaluating listener performance, the equations for motion were analyzed to yield a maximum-likelihood test for the task, that is, a decision rule that would maximize percent correct. For each material discrimination, the test amounted to a specific weighting of the frequency, amplitude, and decay of individual partials in the sound. Listener decision weights, computed from the response correlations with these parameters, generally followed the maximum-likelihood weights, but were in all cases unduly biased toward frequency. In the most severe cases, the frequency bias resulted in as much as an 80% reduction in performance efficiency. The results suggest that, even under the best circumstances, a listener's ability to discriminate material is far from ideal.

In the present study we investigate yet another basic discrimination, the discrimination of hollow from solid bars.

^{a)}Electronic mail: lutfi@waisman.wisc.edu

The study is a natural outgrowth of the earlier study on material discrimination involving merely different relations among the same acoustic cues. However, in the present study we employ a different procedure for introducing perturbations in these cues. In the earlier procedure acoustic variation was introduced by perturbing bar material, the physical attribute to be discriminated. In the present procedure acoustic variation is introduced by perturbing a physical attribute of the bar unrelated to the discrimination. In the discrimination of solid from hollow bars we perturb bar length. We adopt this new procedure for two reasons. First, as a practical matter, we wished to readdress a specific issue regarding the methodology used in our previous study. In that study, the bars varied continuously along the physical dimension to be discriminated. Hence, the nominal discrimination between silver and iron, for example, was really more akin to choosing between two alloys having different relative concentrations of silver and iron. In the present study we investigate a task where there is no ambiguity regarding the two classes to be discriminated. In this case, the bar is either hollow or solid, and the only ambiguity is regarding the unrelated physical attribute, bar length.

The second reason for adopting this procedure is theoretical and has to do with the larger problem of how listeners determine source attributes from sound. The acoustic information intrinsic to different source attributes is necessarily confounded in the single pressure wave form arriving at the ear. How then might a listener determine from sound that a class of resonant objects is, say, hollow or is made of metal despite differences in the size or shape of individual exemplars or the manner in which they are driven to vibrate? Gibson (1966) has suggested that listeners “pick up” the requisite information identifying such general classes of objects or events in the form of higher order acoustic relations that are invariant within a class. While he did not go beyond a general description of these relations, others have since offered analytic treatments that identify the intrinsic acoustic information necessary to distinguish among certain general classes of resonant objects given few physical constraints (Wildes and Richards, 1988; Kac, 1966; Lutfi, 2000). One such case is the distinction between hollow and solid bars. It can be shown, for example, that specific relations among acoustic parameters that are intrinsic to freely vibrating hollow bars remain constant despite variation in bar size, shape, or material composition (Lutfi, 2000). In this particular case, at least, such relations seem likely candidates for the type of information Gibson had in mind. Will listener judgments of hollowness be based on these relations? The answer might depend on whether they provide a significant detection advantage. An analysis of the intrinsic acoustic relations associated with various bar attributes suggests that these relations can be subtle (Lutfi, 2000). Basic limits in auditory sensitivity might, therefore, exert a greater influence on the listener’s selection of cues than the intrinsic acoustic relations that serve to disambiguate source attributes. We test the possibility in the present experiments. In particular, we compare the decision weights of listeners to those of a hypothetical listener that bases judgments on acoustic relations intrinsic to

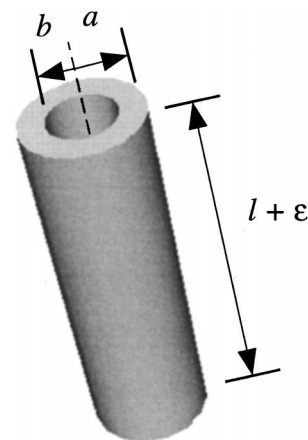


FIG. 1. Representation of bar geometry used in the study. The outer radius a was fixed at 2.0 cm in all conditions. The length l of the bar was either 10 or 25 cm in different conditions and was randomly perturbed for each presentation of a sound, $\epsilon \in N(l, 0.5 \text{ cm})$. The hollow inner radius b was chosen for each listener to yield discrimination performance levels in the range of 70%–90% correct when compared to the solid bar, $b = 0$.

hollowness as determined from the equations for motion of the freely vibrating bar.

II. METHODS

A. Stimuli

The procedure for synthesizing stimuli was identical to that described in Lutfi and Oh (1997) and is briefly reviewed here. The airborne sound of a cylindrical bar rigidly clamped at one end and struck at the other was synthesized over headphones using the theoretical equations describing the motion of the bar. The resultant sound-pressure wave form is a sum of exponentially damped sinusoids whose individual frequencies (ν_n , Hz), amplitudes (C_n , dyn/cm²), and decay moduli (τ_n , s) are uniquely determined by the specific material and geometry of the bar, as well as the manner in which the bar is struck,

$$y = \sum_n C_n e^{-t/\tau_n} \sin(2\pi\nu_n t). \quad (1)$$

Altogether five bar parameters were incorporated in the stimulus synthesis. Related to bar geometry were bar length l , outer radius a , and inner radius b (hollowness), see Fig. 1. Related to bar material were elasticity Q and mass density ρ . The specific values used for the different conditions of the experiment are given in Table I. These values were chosen to represent different resonant sources that might realistically be encountered in everyday listening. The values of b ranged from 0.5 to 1.2 cm and were chosen individually for each listener and condition to yield average performance levels between 70% and 90% correct. Past experience has dictated that performance levels much above or below this range yield unreliable estimates of listener decision weights. The equations relating these physical parameters to the acoustic parameters of Eq. (1) are reviewed by Morse and Ingard (1968, pp. 175–191, 222). For a total force P of an impulse applied to the free end of the bar,

TABLE I. Physical and acoustic parameters associated with different clamped bars used in the study. Entries for acoustic parameters pertain to the first partial. Bar length l was perturbed from one presentation to the next, all other physical properties of the bars were held constant within a block of trials. Values of b are nominal as they were selected for each condition and listener to yield performance levels between 70% and 90% correct. Values of density and elasticity for different materials are taken from Kinsler and Frey (1962). Elasticity values for wood are contained in parentheses as they vary widely depending on factors related to the properties of the wood.

| Bar type | Physical parameters | | | Acoustic parameters | | |
|-----------------|-------------------------|---|---|--------------------------|------------------------|----------------------|
| | Geom, l, a, b (cm) | Density, ρ (g/cm ³) | Elasticity, Q (dyn/cm ² ×10 ¹¹) | Frequency, ν (Hz) | Amp. $ C $ (dB SPL) | Decay, τ (s) |
| Solid iron | 10, 2.0, 0.0 | 7.7 | 10.5 | 2067 | 68.5 | 1.09 |
| Hollow iron | 10, 2.0, 1.0 | 7.7 | 10.5 | 2311 | 70.0 | 0.63 |
| Solid wood | 25, 2.0, 0.0 | 0.72 | (60.0) | 2586 | 69.1 | 0.05 |
| Hollow wood | 25, 2.0, 0.8 | 0.72 | (60.0) | 2785 | 70.0 | 0.04 |
| Solid aluminum | 10, 2.0, 0.0 | 2.7 | 7.1 | 2871 | 68.5 | 0.14 |
| Hollow aluminum | 10, 2.0, 1.0 | 2.7 | 7.1 | 3210 | 70.0 | 0.08 |

$$\nu_n = \frac{\pi}{2l^2} \sqrt{\frac{Q\kappa^2}{\rho}} \beta_n^2, \quad (2a)$$

$$C_n = (-1)^{n-1} \frac{lU}{\pi^2 \beta_n^2} \sqrt{\frac{8\rho}{Q\kappa^2}}, \quad (2b)$$

$$\tau_n = 16 \times 10^8 (\pi\rho/4\kappa^2 \nu_n^3), \quad (2c)$$

where $\beta_1 = 0.597$, $\beta_2 = 1.494$, and $\beta_{n>2} = n - \frac{1}{2}$ determine the frequency ratio of successive partials, $\kappa = \frac{1}{2}\sqrt{a^2 + b^2}$ is the radius of gyration of the bar, and the gain term $U = P/[\rho\pi(a^2 - b^2)]$. The value of P was chosen to fix the sound intensity for hollow bars at 70 dB SPL. Only the first three partials ($n = 1-3$) were synthesized as the higher partials were well beyond the range of audibility. Further details regarding the stimulus synthesis can be found in Lutfi and Oh (1997). Also, examples of these sounds can be heard at <http://abrl.waisman.wisc.edu/>.

As practical matter, a 5 ms cosine-squared ramp was used to truncate signals after 1 s. This kept trials at a reasonable length, while allowing adequate time for the second and third partials of the sounds to decay to inaudibility. Note that in the previous study signals were less than half this duration (400 ms). The longer duration is intended to ensure that the sounds are not too brief to allow for normal identification of bar attributes. All signals were played over a 16 bit, Crystal Audio 4237B DAC at a 40 kHz sampling rate. The output of the DAC was low-pass filtered with rapid rolloff above 16 kHz, asymptoting to a maximum 60 dB of attenuation at 22 kHz. Sounds were delivered to the right ear of listeners over Sennheiser Model 520 II headphones, and were calibrated using a loudness matching procedure. All signals were presented to individual listeners seated in a double-walled, IAC sound-attenuation chamber.

B. Procedure

For each presentation of a signal, a different perturbation in acoustic parameters was imposed without violating the lawful relations governing the values of these parameters. This was achieved by adding a random increment or decrement ϵ on each presentation to bar length, a physical attribute

of the bar unrelated to hollowness.¹ All other physical attributes of the bar were held constant within trial blocks. In each condition, bar length was chosen at random from normal distributions to synthesize a set of 100 wave forms each for the hollow and solid bars. As indicated in Table I, the mean value of bar length l was 25 cm for the wood bars, whereas it was 10 cm for aluminum and iron bars. The longer length for wood bars was necessary to maintain the frequencies of the partials within the audible range. The standard deviation of bar length was fixed for all conditions at 0.5 cm. The exact value chosen is somewhat arbitrary, and is expected to affect the estimates of listener weights only to the extent that it causes performance to fall outside the acceptable range of 70%–90% correct.

Conditions were run in separate blocks of trials with replications across days. A two-interval, forced-choice procedure was used with sounds separated by 500 ms. On each trial one exemplar was chosen at random from the set of solid and hollow bar sounds for presentation. Each sound had equal probability of occurring in the first or second interval of the trial. The listener's task was to indicate by button press which of the two intervals contained the sound corresponding to the hollow bar. Correct feedback was given after each trial. Trials were run in blocks of 50 with short breaks between blocks taken at the discretion of the listener. Listeners practiced conditions for several hundred trials prior to data collection. After practice a total of 400 trials was run for each listener for each condition with at least one 400-trial replication of each condition made on a separate day. The listeners were seven students at the University of Wisconsin–Madison ranging in age from 20 to 38 years. Subject protocol required that listeners be allowed to terminate participation in a study at any time at their discretion. Hence, not all listeners participated in all conditions. Two of these listeners had participated in the earlier study on the discrimination of bar material. All listeners had normal hearing by standard audiometric tests, and all were paid at an hourly rate for their participation. Listeners were instructed that on each trial they would hear the sound of a hollow and a solid bar struck at one end and clamped at other, that the bars were struck with constant force, and that the bars were

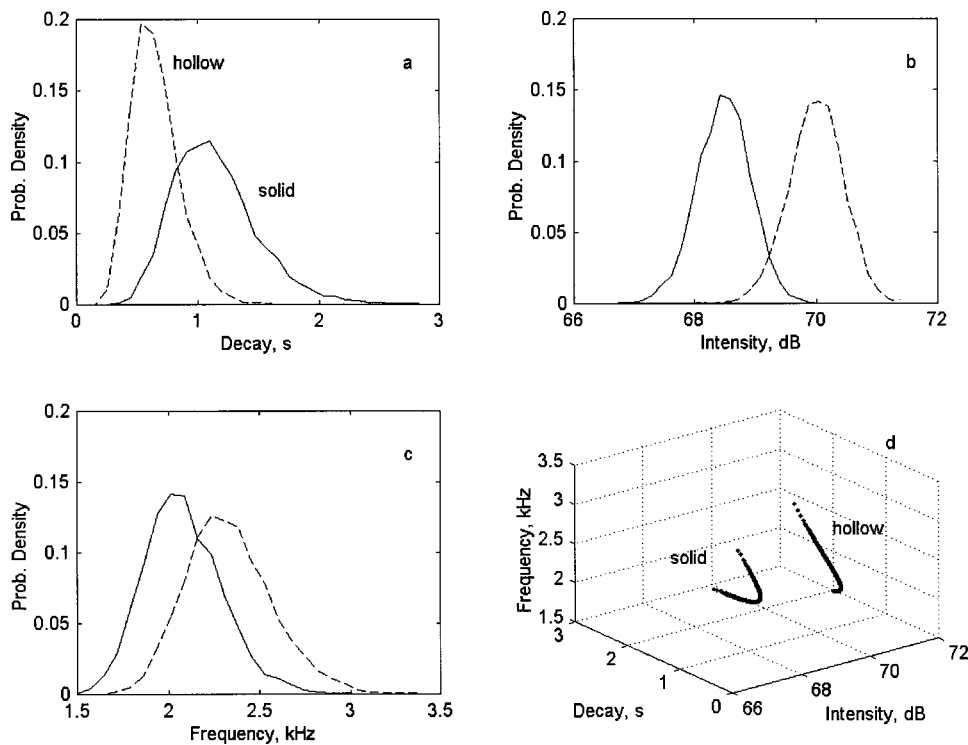


FIG. 2. Distribution of values of acoustic parameters (first partial) for the iron bar. (a)–(c) Values of each parameter for hollow (dashed curves) and solid (continuous curve) bars as bar length varies from one presentation to the next. A judgment of hollowness based on any one acoustic parameter in isolation would be prone to error inasmuch as the distributions overlap. (d) The distribution of parameter values from (a)–(c) are plotted relative to one another so as to reveal the intrinsic relations among the parameters that allow the bar to be identified as hollow or solid without error.

identical in all other respects except for length. They were instructed that the length would vary from one presentation to the next, but that in every case their goal was to identify the sound corresponding to the hollow bar. Before each block of trials the listeners were told the material of the bar, and were shown an example of a real bar having the geometric dimensions corresponding to those used to synthesize the sounds they heard.

C. Analysis of decision weights

Before we can properly analyze listener decision weights it is necessary to identify the relevant information for the detection of hollowness given the particulars of the task. The problem amounts to finding an analytic solution for hollowness using Eqs. (2a)–(2c). The derivation is straightforward. Note first that the information in successive partials is redundant—the frequency ν_n , amplitude C_n , and decay τ_n of the n th partial are merely scaled in each case by the constant factor β_n raised to some power. This means that for the conditions of the present experiments, where bar length l is perturbed, the problem of detecting hollowness reduces analytically to that of solving two equations for two unknowns. The two equations are any two of the equations (2a)–(2c) for any given n . The two unknowns are the inner radius of the bar b ($= 0$ solid, > 0 hollow) and bar length l . The value of b must be solved anew on each presentation given the observed values of ν_n , C_n , and τ_n , but there is always a unique solution on each presentation. Figure 2 demonstrates this result. Figures 2(a)–(c) show the distribution of values for each acoustic parameter (first partial) of the iron bar, both hollow (dashed curves) and solid (continuous curve), as length varies from one presentation to the next. Note that a judgment of hollowness based on any one acoustic parameter in isolation would be prone to error inasmuch

as the distributions overlap. Figure 2(d) shows the same data, but now plotted in a way so as to reveal the relations among parameter values. Here it can be seen that there are unique combinations of ν_n , C_n , and τ_n that allow the bar to be identified as hollow or solid without error. (Note that it is the possibility of error-free performance that distinguishes the present task from that of our earlier study involving the discrimination of bar material.)

Having identified ν_n , C_n , and τ_n for any partial as containing the relevant information for hollowness, we next derive the maximum-likelihood decision rule. There are actually three such rules corresponding to the solutions obtained for each pair of the equations (2a)–(2c). The derivation proceeds identically in each case. Note first that when the outer radius a is constrained to have a single value, as it is in the present experiments, the inner radius b is monotonically related to the radius of gyration κ (or U/κ) as in the case of Eq. (2b). From Eqs. (2a)–(2c), moreover, we see that κ (and U/κ) is proportional to the product of parameter pairs raised to different powers. This means that each optimal decision rule can be expressed in terms of a weighted sum of the log of parameter pairs, where the weights are the corresponding powers. Consider, for example, optimal decisions based on ν_n and τ_n . From Eq. (2c), we see that κ is proportional to $\tau_n^{-1/2} \nu_n^{-3/2}$. A maximum-likelihood decision rule, therefore, is of the form,

$$R = 1 \text{ iff } \text{Dif}(-\frac{1}{2} \log \tau_n - \frac{3}{2} \log \nu_n) < 0 \quad \text{else } R = 2, \quad (3)$$

where R is the response, Dif is the difference between the second and first (i.e., second-first) interval of the forced-choice trial, and $-\frac{1}{2}$ and $-\frac{3}{2}$ are the optimal decision weights for the two acoustic parameters. Optimal decision weights for the other parameter pairs are analyzed in like fashion; for the combination of C_n and ν_n they are, respectively, 1 and

$\frac{1}{2}$, and for the combination of C_n and τ_n they are, respectively, 6 and -1 .

The next step is to determine whether listeners make use of any one of the three optimal decision rules. This can be done using standard multiple regression (Winkler and Hays, 1975). Again, take rule (3) as an example. Since the combination of log-parameter values is linear, responses strictly based on Eq. (3) will correlate perfectly with interval differences between $\log \tau_n$ or $\log \nu_n$ once the relation between τ_n and ν_n is taken into account. That is, the partial correlation of the response with the differences in $\log \tau_n$ and $\log \nu_n$ will be -1 in each case. In fact, we do not expect listener responses to strictly follow Eq. (3) since we have selected conditions (i.e., perturbations in l and values of b) to produce errors in the response. A more realistic regression model is therefore

$$R = 1 \text{ iff } \text{Dif}(w_1 \log \tau_n = w_2 \log \nu_n) + \xi > 0 \text{ else } R = 2, \quad (4)$$

where w_1 and w_2 are the regression weights for interval differences in $\log \tau_n$ and $\log \nu_n$, and ξ is listener error. The values of w_1 and w_2 for each listener can then be computed from the relation of listener's response and the interval differences in $\log \tau_n$ and $\log \nu_n$ on each trial using standard multiple regression analysis.² Now, if listener responses are based on Eq. (3) and the error ξ is unbiased, then the regression weights in Eq. (4) should bear the same relation as the optimal decision weights given in Eq. (3). In particular, w_1 and w_2 should both be negative and have a ratio of one to three. All three of the optimal decision rules can be evaluated in this way, the only difference being differences in the optimal decision weights for each rule.

III. RESULTS

A. Analysis of listener decision weights

Figures 3–5 give the computed regression weights for each listener for the three different bar materials. To facilitate comparisons the regression weights for each decision rule have been normalized so that their unsigned values sum to one. Different listeners are represented by different symbol types with repeated symbols representing replications, some taken several months apart. The end points of the three lines drawn in the top panel of each figure give the decision weights for the three optimal decision rules corresponding to each parameter pair. The end points of the three lines drawn in the bottom panel of each figure give the decision weights of a hypothetical listener that bases decisions on frequency ν_n alone. Listeners appear to fall into two distinct groups, one choosing to base decisions on the optimal combination of ν_n and τ_n (top panel), the other choosing not to apply any of the optimal decision rules, but rather choosing to base decisions largely on frequency alone (bottom panel). Listener membership in either group varies both within and across conditions of bar material. Note that the occasional appearance of the same symbol in both panels shows that some listeners chose to switch from one listening strategy to the other upon replication of the same condition. With the exception of these cases, however, test–retest reliability for individual estimates is good, even for those estimates taken

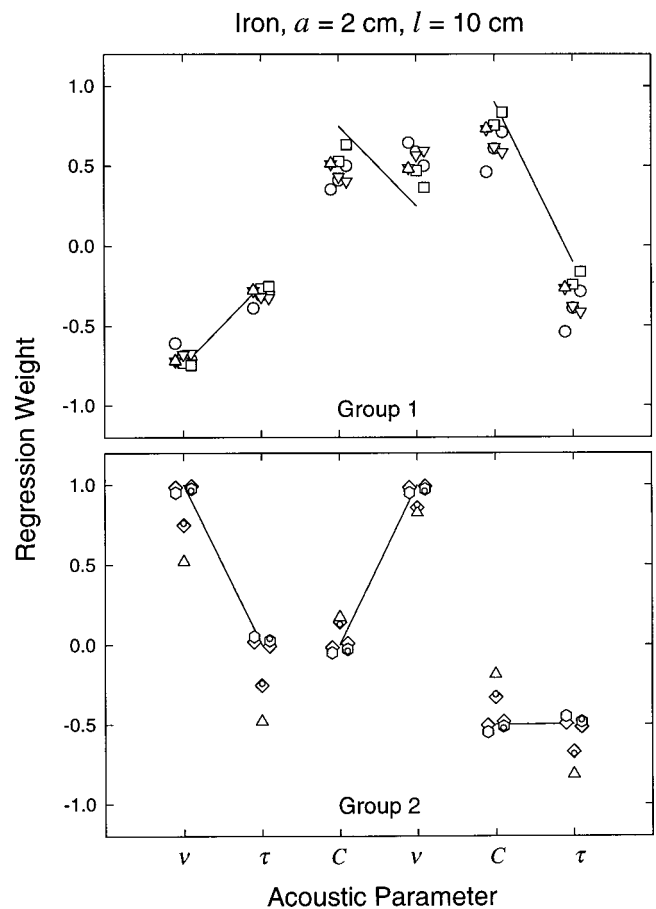


FIG. 3. Regression weights are plotted giving the relation of the listener's response to the pair of acoustic parameters underlying each optimal decision rule. Data are for the iron bar, $a=2$ cm, $l=10$ cm. The raw regression weights for each decision rule have been normalized so that their unsigned values sum to one. Different symbol types represent the regression weights for different listeners, and repeated symbols represent replications of conditions. The end points of the three lines drawn in the top panel give the decision weights for the three optimal decision rules corresponding to each parameter pair. Reading from left to right the parameter pairs are ν_n and τ_n , C_n and ν_n , and τ_n and C_n . The end points of the three lines drawn in the bottom panel give the decision weights of a hypothetical listener that bases decisions on frequency ν_n alone. Listeners' data are plotted in one or the other panel depending on which decision rule most closely agreed with their regression weights. Note, in the bottom panel, that nonzero regression weights for the combination of τ_n and C_n are possible because each parameter covaries with ν_n , which is not included in the regression for τ_n and C_n .

several months apart. Inclusion of the practice trials, moreover, did not change the agreement between estimates. When questioned, listeners in the second group reported without exception that they were choosing as hollow the sound having the higher frequency or pitch. Indeed, Fig. 2(c) shows that better than chance performance can be achieved by basing decisions on frequency alone. Also, for a few listeners in the first group it is unclear as to whether the listening strategy involved an optimal weighting of frequency and decay or of intensity and decay (note for example the squares in Fig. 3). Performance levels in these instances, though within the acceptable range of the experiment, were apparently too high for the regression weights to distinguish among the different optimal decision rules.

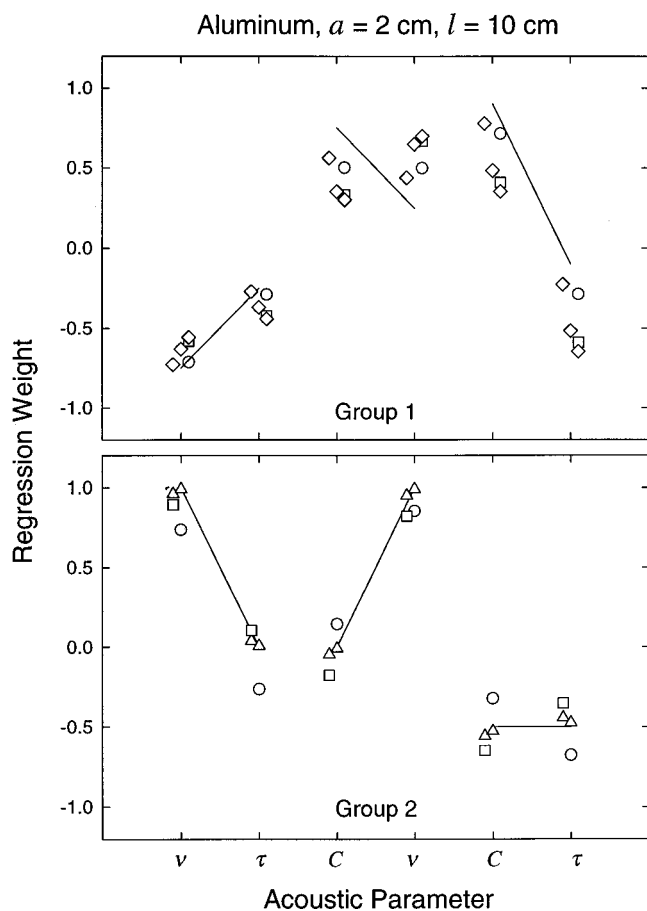


FIG. 4. Same as Fig. 3, except that data are for the aluminum bar, $a = 2$ cm, $l = 10$ cm. Except where a new listener's data are added, individual listeners maintain the same symbol type as in Fig. 3.

B. Effect of limited resolution on discrimination performance

For those listeners who appear to adopt an optimal decision rule it is of interest to consider what factors might have caused their performance to be less than perfect. One obvious factor is limited sensory resolution—that is, the inability on some trials to discriminate small differences in the frequency, amplitude, and decay of the individual partials of these sounds. Note, for example, in Table I, that the difference in mean intensity between the sounds for hollow and solid bars in these conditions is 0.9–1.5 dB. The normal intensity difference limen at these moderate sound levels is at best 0.7 dB (Jesteadt *et al.*, 1977). With the added perturbation in intensity the difference in intensity would have been difficult or impossible to discriminate on some proportion of trials. Table II gives an indication of the relative extent to which limited resolution affected discriminability of the individual acoustic parameters in these experiments. Table II lists, for three listeners, the d' values for the individual parameters expressed relative to those of an ideal observer (an observer with unlimited resolution). The d' values were obtained simply by rerunning conditions with the hollowness parameter b fixed at zero for all but one of the three relevant parameters. If limited resolution were to have no impact on performance in these conditions all values in Table II should equal 1.0. For the discrimination of frequency the

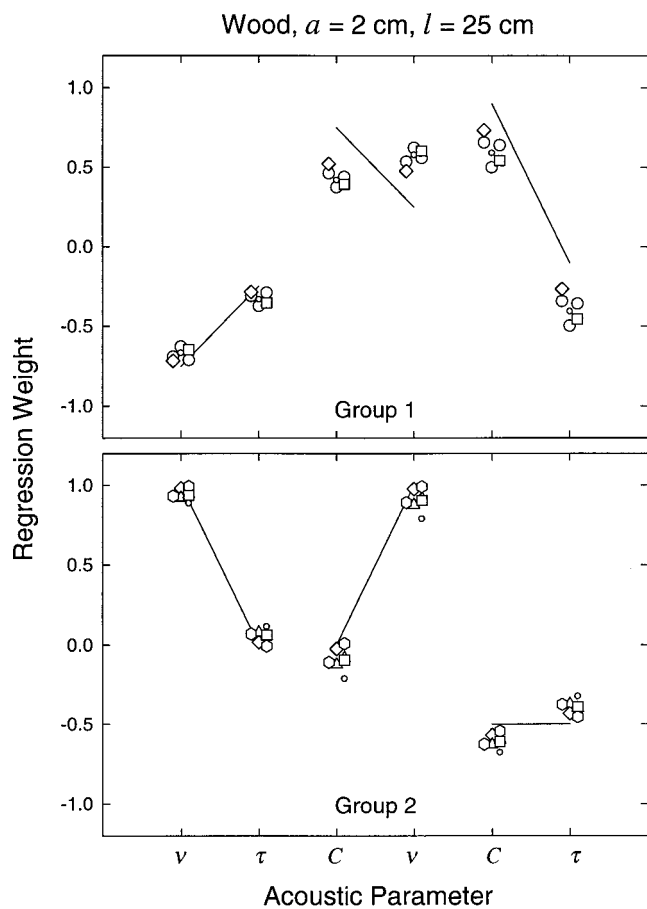


FIG. 5. Same as Fig. 3, except that data are for the wood bar, $a = 2$ cm, $l = 25$ cm. Except where a new listener's data are added, individual listeners maintain the same symbol type as in Fig. 3.

values are, indeed, equal to or very nearly equal to 1.0. However, for the discrimination of intensity, and for the discrimination of decay for the iron bar, the values are significantly less than 1.0. These data suggest, not surprisingly, that limited sensory resolution is likely responsible for less than perfect performance of listeners who appeared to adopt an optimal decision rule. The data do not, however, necessarily imply that limited sensory resolution would significantly affect the selection of decision weights. This is because the analytic solution for hollowness depends on the *relation* among the values of acoustic parameters (cf. Fig. 3), not the value of any single parameter. Consequently, intensity (or decay) by its relation to frequency can convey information

TABLE II. Obtained d' values for the discrimination of individual acoustic parameters. The values are expressed relative to those of an ideal observer with no limit in sensory resolution.

| Material | Listener symbol | d'/d'_{ideal} | | |
|----------|-----------------|------------------------|-----------|-------|
| | | Frequency | Intensity | Decay |
| Iron | Large-circle | 0.92 | 0.36 | 0.31 |
| | Up-triangle | 0.95 | 0.38 | 0.29 |
| | Diamond | 1.00 | 0.38 | 0.42 |
| Wood | Up-triangle | 1.00 | 0.25 | 0.84 |
| | Diamond | 0.98 | 0.30 | 0.85 |

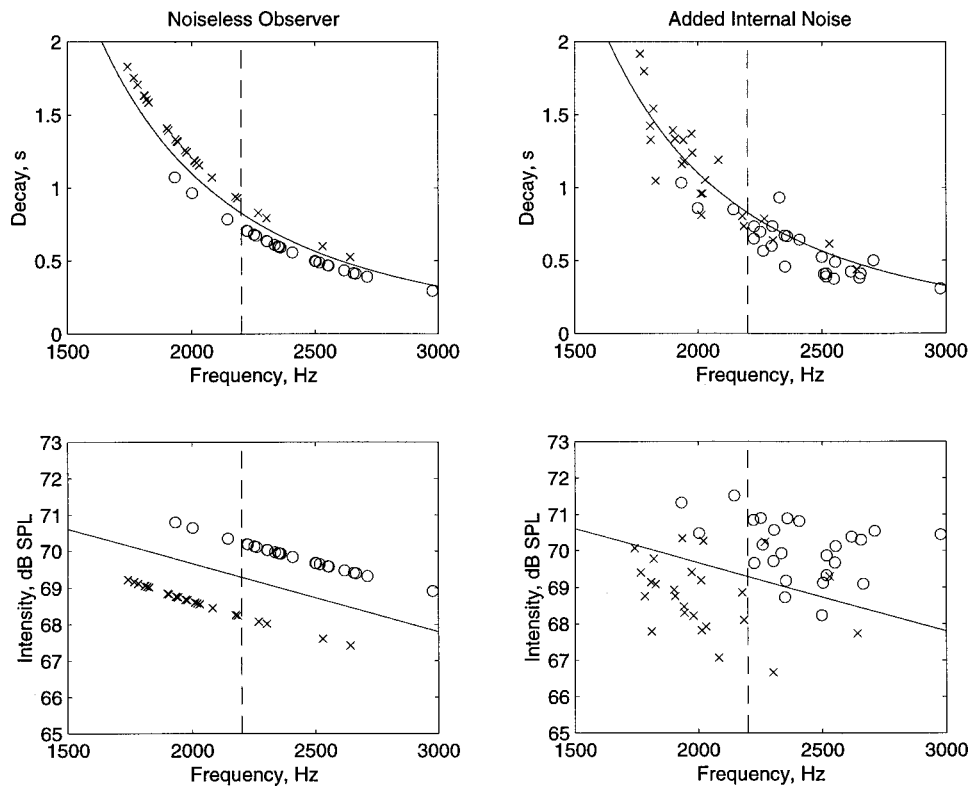


FIG. 6. Simulated effect of listener internal noise for iron bar with hollowness radius $b = 1.0$. Circles and crosses represent, respectively, the parameter values for hollow and solid bars. Right and left panels represent, respectively, parameter values with and without internal noise jitter. The decision border corresponding to the optimal decision rule is given by the continuous curves. The decision border for an observer that bases judgments on frequency alone is given by the dashed curve. See the text for specific details.

regarding hollowness even on trials in which the difference in intensity (or decay) is too small to be detected.

Could limited resolution have affected the listener's selection of decision weights? To evaluate this question we generated expectations for listener performance in these experiments using a standard internal noise model (cf. Durlach *et al.*, 1986). Figure 6 shows the results of independently adding a small amount of jitter to each acoustic parameter so as to simulate the effect of internal noise. For frequency, intensity, and decay the percentage of jitter is, respectively, 0.2, 17, and 25. These values represent the current best empirical estimates of the difference limens for these parameters, for normal-hearing adults (Wier *et al.*, 1977; Jesteadt *et al.*, 1977; Van Heuven and Van Den Broecke, 1979).³ The data shown in Fig. 6 are for the iron bar with hollowness radius $b = 1.0$; similar results, however, were obtained for the aluminum and wood bars. The parameter values with (right) and without (left) jitter are designated separately for the hollow (circles) and solid (crosses) bars. The continuous curve in each panel is the decision border corresponding to the optimal combination of the designated acoustic parameters. The dashed curve is the decision border for an observer that bases judgments on frequency alone. The effect of adding internal noise in each case is to increase the number of crossings of the decision border that would yield an incorrect judgment. Importantly, however, the increase in crossings is much greater for the optimal decision rule than it is for the judgments based on frequency alone. This results in the two decision models producing more similar performance levels than would be anticipated based on the parameter values without internal noise.

Figure 7 shows the expected effect of the internal noise on the psychometric function. Here the dashed and continu-

ous curves represent, respectively, performance based on the optimal weighting of frequency and decay, and performance-based differences in frequency alone. The data of individual listeners (symbols) are plotted for comparison. Note that for the iron bar (upper panel) there is only a small performance advantage for the optimal decision rule over the judgments based on frequency alone. In the range of $b = 1.0$ – 1.3 cm where most of the listeners' data fall the advantage is less than 5%. For the wood bar (lower panel) the optimal decision rule actually yields poorer performance than for the judgments based on frequency alone. This can be understood from the distribution of the values of frequency and decay in Fig. 6 (upper panels), which are for iron, but show a similar pattern for wood. Here the values skirt the decision border of the optimal rule so closely that even a small amount of perturbation can force performance to be near chance.

The analysis of internal noise suggests a likely explanation for the specific pattern of listeners' decision weights. With limited resolution an optimal listening strategy, as would be associated with an analytic solution for hollowness, yields at best a small performance advantage in some conditions over judgments based on frequency alone. Also, since frequency is the most salient single cue, varying over many more just-noticeable differences than either decay or intensity (cf. Wier *et al.*, 1977; Jesteadt *et al.*, 1977; Van Heuven and Van Den Broecke, 1979), there is good reason to expect that listeners would often chose to base decisions on frequency alone in lieu of an optimal combination of acoustic parameters. The fact that performance levels of listeners generally fall within the bounds predicted by the optimal listening strategy and one based on frequency alone lends additional support to this analysis.

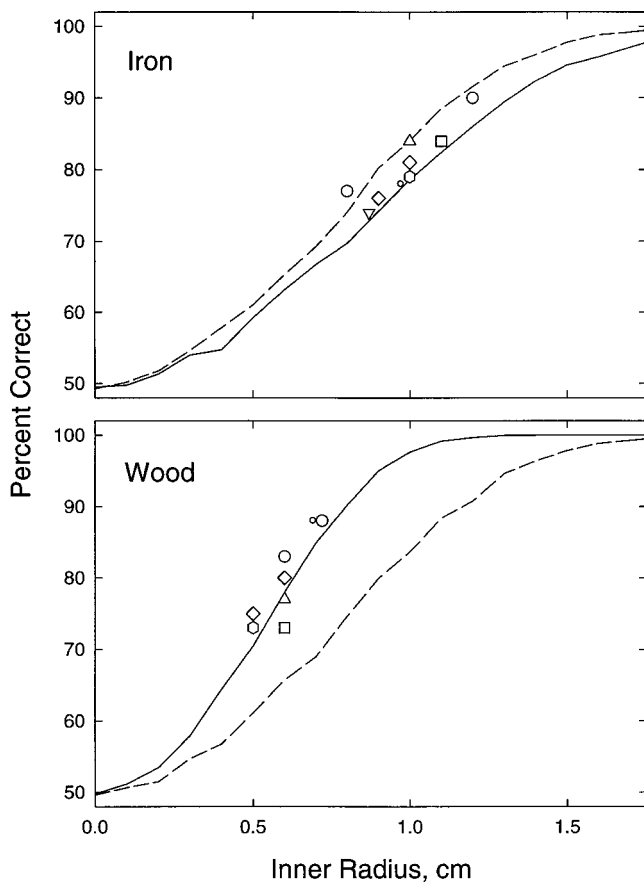


FIG. 7. Psychometric functions for internal noise simulation. Dashed curves are the functions resulting from optimal weighting of frequency and decay [as given by Eq. (3)]. The continuous curve is the function resulting from judgments based on frequency alone. The data of individual listeners are given different symbols consistent with those of Figs. 3–5.

IV. DISCUSSION

The results of these experiments reveal clear individual differences in the decision strategies that listeners use when judging hollowness from the synthesized sound of a struck, clamped bar. Listeners tend to fall into two groups in terms of their cue preference: those giving weight to frequency and decay consistent with an analytic solution for hollowness, and those giving predominant weight to frequency. The results also indicate, for some listeners, a change from one decision strategy to the other with changes in the material composition of the bar or with replication of a condition.

The general pattern of results is consistent with that obtained in the companion to this study involving the discrimination of bar material. There, too, significant individual differences were observed combined with a bias for frequency for many listeners. Apart from the difference in task, that study differed fundamentally from the present study in that the material discrimination was to some degree ambiguous. Unlike the present study, in which the bar was unequivocally hollow or solid, the companion study required, in effect, a decision between two bars of the same material in different relative concentrations. This ambiguity was suspected to have played a role in the outcome of the earlier experiments. However, the similarity in the pattern of results to those of

the present study suggests that the earlier results were not specifically related to this ambiguity.

Several studies in the literature offer similar comparisons to the present study, though they used real or recorded sound sources. Kunkler-Peck and Turvey (2000) challenged listeners to determine from sound the height and width of a rectangular steel plate suspended by a loose support and struck in its center by a hard mallet. While listener judgments consistently underestimated the true values, the ratio of height to width was recovered with reliable accuracy, despite variations in plate material. Kunkler-Peck and Turvey offer both theoretical and acoustical analyses to show that there is sufficient information for the task in the separate modal frequencies associated with height and width. They imply a similar account of their listeners' ability to reliably identify the shape of plates (circle, triangle, or square) of different material. The ability to recover height–width ratios from sound has also been demonstrated for loosely suspended, struck bars made of metal and wood (Lakatos *et al.*, 1997). Acoustical analyses by these authors revealed that for metal bars the frequencies of both torsional and transverse bending modes correlated highly with height–width ratio and with listener judgments, suggesting that these modes may have served as potential cues. For wood bars the role of these modes was less clear as they were often weak or absent. Other physical attributes of objects or events that listeners have been shown to reliably judge from sound include the length of rods dropped on a hard surface (Carello *et al.*, 1998), the hardness of mallets striking metal pans (Freed, 1990), the breaking or bouncing of glass (Warren and Verbrugge, 1984), the position of hands clapping (Repp, 1987), and the gender of walkers (Li *et al.*, 1991). In each of these cases, specific features of the acoustic wave form or spectra have been correlated with the object or event and listener judgments to identify potential cues underlying the listeners' ability to perform the task.

Taken together, the past studies show that there is sufficient information in sound to allow identification of complex source attributes, and moreover, that listeners are capable of such identifications with limited variation in other source attributes. One might ask, therefore, what new information is provided by the present results? The answer lies in the approach to identifying potential acoustic cues. There are two important differences between the past and present studies in this regard. First, it should be noted that a strong correlation of a listener's judgments with a particular cue, as has been demonstrated in past studies, is not a sufficient condition to implicate that cue. Such a result does not rule out the possibility that the listener may use a very different cue that happens to be highly correlated with the one analyzed by the experimenter. Indeed, considering the complexity of the real acoustic events that listeners were asked to judge in past studies, it is almost certain that alternative cues were available in some cases.⁴ The advantage in the present study of using synthesized sounds rather than real sound sources is that *all* relevant information for the specific task is contained in the equations for motion. This allows precise measurement of the independent and relative contribution of each potential source of information to the listener's response,

without lingering questions regarding the possibility of alternative acoustic cues. A second issue pertaining to the measurement of listener cues has to do with the performance levels of listeners in past studies. In many of these studies the focus has been on whether or not the listener is capable of accurately identifying a particular acoustic event, hence, conditions have often been selected where few or no identification errors are made. While high performance levels are to some extent informative regarding the capabilities of listeners in these tasks, they are much less informative regarding the possible cues used by listeners. Consider, for example, the outcome of the present study if inner radius values were selected to yield perfect performance. The correlation of the listener's response with an optimal decision rule would necessarily be one, but, because there are three such rules, it would not be possible to distinguish among them. Indeed, it might not even be possible to conclude that the listener used any one of these rules. If the values of the inner radius were chosen such that individual parameter distributions did not overlap (cf. Fig. 2), listeners could achieve perfect performance by basing decisions on a single acoustic parameter alone. This is why it is essential that listeners make errors when using regression analyses to evaluate listener cues, and why we have chosen the conditions of the present study to yield performance levels between 70% and 90% correct.

Returning to the original motivation for the study, we can evaluate the conjecture that listener judgments would correlate highly with the invariant relations among acoustic parameters that are unique to hollowness. The results appear to provide only partial support. Although many listeners do, indeed, appear to base decisions on such relations, a nearly equal number appear to base decisions on simple differences in frequency. Moreover, there are differences in the way listeners approach the task that depend on the material of the bar, and that change with replication of the same condition. Generally speaking, therefore, the results do not provide clear support for the use of intrinsic cues for hollowness.

One must question how this outcome might have been influenced by the sound synthesis used in this study. Real bars, after all, come in many shapes, sizes, and materials, they are not perfectly symmetric or homogeneous, and they are not typically struck at exactly the same point with exactly the same force. Might such differences prove advantageous for detecting hollowness, and might the synthesized sounds be perceived differently for this reason? We do not rule out the possibility, but we think it unlikely for both analytic and empirical reasons. First, we have specifically chosen conditions of the experiment to maximize the likelihood that listeners would detect the acoustic information intrinsic to hollowness. These choices were based on extensive performance simulations of the type described in Sec. III B, involving a wide range of bar materials and geometries and known listener sensitivity (see Lutfi, 2000). A fixed impulse was used as the driving force because it is generally most diagnostic regarding bar properties. And, all bar properties, except for length, were fixed so that no acoustic variation, other than that resulting from changing length, could serve to confound the change intrinsic to hollowness. We have also

considered the related question as to whether our synthesized sounds can be reliably discriminated from real bar sounds. Preliminary experiments involving direct comparisons by listeners suggest not (Lutfi and Oh, 1997); however, stronger evidence comes from experiments showing that listeners are insensitive to acoustic differences from real bar sounds far greater than those of our synthesized sounds (Lutfi and Oh, 1994).

While the possibility of more advantageous conditions for detection cannot be ruled out, it is instructive to consider why some listeners did not make better use of the information that was available to them in these conditions. Some insight is obtained by considering the effect of limited sensory resolution on judgments. The effect is potentially complex since it must depend on the listener's decision weights as well as the acoustic variation resulting from variation in bar parameters. We have relied, therefore, on an internal noise model to evaluate the potential interaction among these factors. An important outcome of this analysis is that the analytic solution for hollowness is highly vulnerable to the small perturbations in the acoustic parameters that would be associated with listener internal noise. The result is that judgments following the analytic solution yield no large performance advantage over judgments based on differences in frequency alone, indeed, in some cases they yield worst performance. This outcome appears to provide an explanation for both the overall performance levels of our listeners and their specific pattern of decision weights.

That such small perturbations in the acoustic signal, as would be associated with internal noise, could have such profound effects on listener judgments may seem surprising; certainly in the context of the large variation associated with changes in bar length and whether the bar is hollow or solid. Such effects are possible, however, because the intrinsic acoustic relations that would serve to disambiguate hollow from solid bars are subtle and so often do not yield a significant detection advantage over simple differences in individual acoustic parameters (cf. Figs. 6 and 7). The effect is comparable to an analysis of variance where the interaction effects (the intrinsic relations) are small compared to the main effects of variables (the simple differences in individual acoustic parameters). The results underscore the importance of limited sensitivity as a factor in source identification, and raise the question as to whether there are other conditions where limited sensitivity for acoustic relations may similarly play a role. Few studies have specifically addressed the question. Moreover, while analytic treatments have identified the intrinsic acoustic information associated with specific source attributes, they have not generally indicated conditions where, based on known limits of human sensitivity, this information might offer a significant categorical advantage over simple statistical differences in individual acoustic parameters (Jenison, 1997; Kac, 1966; Wildes and Richards, 1988). For the present case, at least, some relevant data are provided by the results of simulations we have used to select stimulus conditions for our experiments (see Lutfi, 2000). These simulations indicate that the results of the present study are not atypical of what would be expected for the discrimination of other rudimentary bar attributes such as

shape, size, or material. In such cases, judgments based on the intrinsic acoustic relations are rarely expected to offer a detection advantage over simple differences in individual acoustic parameters when limited sensitivity is taken into account. For the discrimination of material and hollowness, the results of these simulations are corroborated by data from human listeners (Lutfi and Oh, 1997).

Clearly, many factors can be expected to influence how a listener recovers a particular source attribute from sound, not the least of which are the dynamics of the source, how it is excited, and how it might vary along other physical dimensions. But while the acoustic effect of these factors can be profound, it is often the small changes in the relation among acoustic parameters that serve to disambiguate sources. In such cases, as in the present study, basic limits in human sensitivity can affect not only what a listener hears, but also how they listen.

ACKNOWLEDGMENTS

I would like to thank Dr. Claudia Carello and especially Dr. Daniel J. Freed, and Dr. D. Wesley Grantham for helpful comments on an earlier version of the manuscript. This research was supported by a grant from the NIDCD (No. R01 CD01262-10).

¹We had also experimented with perturbations in parameters related to the material of the bar (density and elasticity). For these parameters, however, the resultant variation among acoustic parameters was too highly correlated to permit reliable estimates of listener decision weights.

²Specifically, the two regression weights in each case are given by the two-element vector $\mathbf{w} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$, where \mathbf{y} is a column vector giving the listener's response on each trial, \mathbf{X} has two columns corresponding to the interval differences for $\log \tau_n$ and $\log v_n$ on each trial, and \mathbf{X}^T denotes the transposition of \mathbf{X} .

³These values may be underestimates as they come from conditions where there is minimal uncertainty regarding the values of stimulus parameters from trial to trial. The possible influence of stimulus uncertainty for conditions most similar to those of the present study is discussed in Lutfi and Oh (1997).

⁴By way of example, the spectral acoustic cues analyzed by Kunkler-Peck and Turvey (2000) almost certainly were correlated with spatial acoustic cues afforded by large suspended plates.

Carello, C., Anderson, K. A., and Kunkler-Peck, A. J. (1998). "Perception of object length by sound," *Psychological Science* **9**, 211–214.

Durlach, N. I., Braida, L. D., and Ito, Y. (1986). "Towards a model for

discrimination of broadband signals," *J. Acoust. Soc. Am.* **80**, 63–72.

Freed, D. J. (1990). "Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events," *J. Acoust. Soc. Am.* **87**, 311–322.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems* (Houghton–Mifflin, Boston).

Jenison, R. L. (1997). "On acoustic information for motion," *Ecological Psychol.* **9**, 131–151.

Jesteadt, W., Wier, C. C., and Green, D. M. (1977). "Intensity discrimination as a function of frequency and sensation level," *J. Acoust. Soc. Am.* **61**, 169–177.

Kac, M. (1966). "Can one hear the shape of a drum?," *Mathematics Monthly* **13**, 21–23.

Kinsler, L. E., and Frey, A. R. (1962). *Fundamentals of Acoustics* (Wiley, New York), pp. 55–78.

Kunkler-Peck, A. J., and Turvey, M. T. (2000). "Hearing shape," *J. Exp. Psychol.* **26**, 279–294.

Lakatos, S., McAdams, S., and Causse, R. (1997). "The representation of auditory source characteristics: Simple geometric form," *Percept. Psychophys.* **59**, 1180–1190.

Li, X., Logan, R. J., and Pastore, R. E. (1991). "Perception of acoustic source characteristics: Walking sounds," *J. Acoust. Soc. Am.* **90**, 3036–3049.

Lutfi, R. A. (1995). "Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks," *J. Acoust. Soc. Am.* **97**, 1333–1334.

Lutfi, R. A. (2000). "Source uncertainty, decision weights, and internal noise as factors in auditory identification of a simple resonant source," *Abstracts of the Association for Research in Otolaryngology* **23**, 171.

Lutfi, R. A., and Oh, E. (1994). "Auditory discrimination based on the physical dynamics of a tuning fork," *J. Acoust. Soc. Am.* **95**, 2967.

Lutfi, R. A., and Oh, E. (1995). "Auditory detection of changes in mass density and elasticity of a tuning fork," *J. Acoust. Soc. Am.* **97**, 3330.

Lutfi, R. A., and Oh, E. (1997). "Auditory discrimination of material changes in a struck-clamped bar," *J. Acoust. Soc. Am.* **102**, 3647–3656.

Morse, P. M., and Ingard, K. U. (1968). *Theoretical Acoustics* (Princeton University Press, Princeton, NJ), pp. 175–191.

Repp, B. H. (1987). "The sound of two hands clapping: An exploratory study," *J. Acoust. Soc. Am.* **81**, 1100–1109.

Van Heuven, V. J., and Van Den Broecke, M. P. R. (1979). "Auditory discrimination of rise and decay times in tone and noise bursts," *J. Acoust. Soc. Am.* **66**, 1308–1315.

Warren, W. H., and Verbrugge, R. R. (1984). "Auditory perception of breaking and bouncing events: A case study in ecological acoustics," *J. Exp. Psychol.* **10**, 704–712.

Wier, C. C., Jesteadt, W., and Green, D. M. (1977). "Frequency discrimination as a function of frequency and sensation level," *J. Acoust. Soc. Am.* **61**, 178–184.

Wildes, R., and Richards, W. (1988). "Recovering material properties from sound," in *Natural Computation*, edited by W. Richards (MIT, Cambridge, MA), pp. 356–363.

Winkler, R. L., and Hays, W. L. (1975). *Statistics: Probability, Inference, and Decision* (Holt, Rinehart, and Winston, New York), pp. 683–695.