Do we hear size or sound? Balls dropped on plates

MASSIMO GRASSI

University of Padua, Padua, Italy and University of Udine, Udine, Italy

The aim of this study is to examine whether it is possible to recover directly the size of an object from the sound of an impact. Specifically, the study is designed to investigate whether listeners can tell the size of a ball from the sound when it is dropped on plates of different diameters (on one, two, or three plates in Experiments 1, 2, and 3, respectively). In this paradigm, most of the sound produced is from the plate rather than the ball. Listeners were told neither how many different balls or plates were used nor the materials of the balls and plates. Although listeners provided reasonable ball size estimates, their judgments were influenced by the size of the plate: Balls were judged to be larger when dropped on larger plates. Moreover, listeners were generally unable to recognize either ball and plate materials or the number of plates used in Experiments 2 and 3. Finally, various acoustic properties of the sounds are shown to be correlated with listeners' judgments.

Imagine the situation: You are in your living room talking to a friend about the size of an object that is in your bedroom. You decide to solve the problem by going into the bedroom to get the object. When you reach the bedroom, your friend in the living room asks you to drop the object on the floor. After listening to the sound of the impact, he says, "It's more or less 50 mm wide!"

Impact sounds are the result of contact between (at least) two objects. The resulting air pressure wave depends on which of the two objects is vibrating. Classically, the two objects assume the names of exciter and resonator. The *exciter* is the object that provides the resonator with the energy necessary for vibration; the *resonator*—or rather, its modes of vibration (Morse, 1981)—characterizes the timbre of the resulting sound. More generally, as the exciter—resonator partition has exceptions, in most impacts we can distinguish between sounding objects (SO) and non- (or less) sounding objects (NSO) according to the object that provides the greater part of the resulting acoustic signal.

In the recent past, different studies have shown that listeners are able to extract from an impact sound information about the physical properties of the objects that produce it. Listeners have been accurate in judging properties of the SO for tasks such as reproducing the length of a rod dropped on the floor¹ (Carello, Anderson, & Kunkler-Peck, 1998) or guessing the ratio dimensions (height and

This work has been supported by the European Commission under contract IST-2000-25287 ("The Sounding Object": www.soundobject.org). I thank Bruno Lucio Giordano, Davide Rocchesso, Giovanni Bruno Vicario, and two anonymous reviewers for many helpful suggestions during preparation of this article. I also thank Roberto Burro for providing the disk-generating software. Correspondence relating to this article may be sent to M. Grassi, Department of Philosophy, University of Udine, via Petracco 8, 33100 Udine, Italy (e-mail: massimo.grassi@uniud.it).

width) and drawing the shape of either bars or plates struck by a mallet (Kunkler-Peck & Turvey, 2000; Lakatos, McAdams, & Caussé, 1997). In other experiments, listeners could also identify properties of the NSO in tasks such as ordering the relative hardness of mallets striking different cooking pans (Freed, 1990) and guessing the gender of walkers from the sound of their footsteps on the hardwood stage of a theatre (Li, Logan, & Pastore, 1991).

Overall, the results of these studies are doubly surprising. The pressure wave reaching the ear is the result of contact between two objects. Nonetheless, listeners are able to segregate from this single pressure wave information concerning the sole object under investigation. Moreover, after listening to the sound, they are also accurate in judging static properties of the physical event that generated it. One explanation for such surprising results comes from direct perception: The acoustic array of everyday sounds contains the necessary information for recovering the static properties of the sound source event (Fowler, 1990, 1991; Gaver, 1993a, 1993b). Consistently, in experiments presenting everyday sounds, listeners directly describe the source rather than providing descriptions of the acoustic characteristics of the sound (Gaver, 1988; Vanderveer, 1979). In addition, subjective estimates are explained better by comparing them directly with the physical properties of the object rather than with the acoustic characteristics of the sound (Carello et al., 1998).

Independently of the approach chosen, there is no doubt that hearing a sound is a necessary prerequisite for being able to say something about the sound source event. Thus, studies focusing on perception of the NSO are of particular interest. In such studies (e.g., Freed, 1990; Li et al., 1991), listeners are asked to judge the properties of an object whose contribution to the resulting acoustic array is marginal. Nonetheless, both Freed (1990) and Li et al.

(1991) have shown that listeners can provide correct estimates about the NSO. However, in both of these studies, listeners received prior information about both the event and its SO. Therefore, it is still unknown whether listeners can provide veridical estimates of the NSO when provided with no prior information about the sound source event and the SO. Furthermore, in such experiments, listeners have to evaluate NSO properties on the basis of a sound that is mostly produced by the SO. Thus, in principle, manipulations of the SO might influence listeners' judgments of the NSO. Freed did find that estimates of the "hardness" of a mallet were not influenced by the different pans, but in his experiment, listeners were informed that "sounds were made by striking *metal* cooking pans [i.e., not a single pan] with different percussion mallets" (emphasis added). Thus, listeners in that experiment knew (1) that the pans were made of metal and (2) that the sounds were produced by various pans (i.e., pans of various shapes and sizes). Consequently, when estimating hardness, they had enough information to be able to ignore timbre changes due to the different pans.

In the present experiments, listeners were asked to judge the size of balls (the NSOs) of different sizes/masses dropped onto one or more plates (the SOs) by drawing computer-generated disks of the same size as the balls that were dropped. In all experiments, listeners received no foregoing information about the sound source event. The vibration of the ball is negligible in comparison with that of the plate because of the geometry of the two objects and whether or not this geometry allows each object to efficiently modulate air pressure. The vibration pattern of a ball is characterized by stress waves. When the ball impacts the plate, it begins to vibrate omnidirectionally, and as a result, the perturbation of the air surrounding the ball is only marginally efficient. In contrast, when the plate is set vibrating by the ball, its vibrational pattern is much more efficient, because it is characterized by bending waves. The geometrical shape of the plate makes it behave like a dipole, so that vibration moves a lot of air along the plate's two flat sides (Morse, 1981). Moreover, the event "ball dropped on a plate" has two major characteristics. Variations in the characteristics of the ball (i.e., its size and/or mass) are mainly responsible for changes in the amplitude of the resulting sound. A variation in one of the plate's characteristics (i.e., its size) is mainly responsible for changes in the frequency content of the sound; both Lakatos et al. (1997) and Kunkler-Peck and Turvey (2000) have indicated that frequency is the acoustic feature that carries information about the shape and size of the SO.

The present study is concerned with the following questions: (1) Is veridical knowledge of the sound source event a necessary prerequisite for a good estimate of the physical properties (size) of the NSO (the ball)? (2) Are listeners' estimates of the NSO sensitive to unrevealed changes in the properties of the SO (the plate)? (3) Does the amplitude of the sound carry information about the size of the NSO? (4) Given the direct and "indirect" ap-

proaches to perception, which will be able to explain the results of the experiments?

EXPERIMENT 1

The goal of this first experiment was to investigate whether listeners could estimate veridically the size of the NSO when provided with no foregoing information about the sound source event—hence, whether a veridical reconstruction of the sound source event is or is not a necessary prerequisite for estimating the size of the NSO.

Method

Subjects. Ten undergraduates of the University of Padua volunteered for the experiment. They all reported having normal hearing.

Apparatus. Seven solid wooden balls (pine) that were 10, 15, 20, 25, 30, 40, and 50 mm in diameter and weighed, respectively, 0.35, 1.1, 2.9, 4.9, 8.6, 22.2, and 44.5 g (mass density, ~647 kg/m³) and a baked clay plate 215 mm in diameter were used for the experiment. Balls were dropped onto the middle of the plate from a height of 150 mm, by means of a plastic pipe placed vertically above the plate and closed at the lower end by a flap that held the ball still. The flap could be opened by a lever and the ball released. The plate was set on a foam block measuring $500 \times 500 \times 40$ mm placed on a table. Therefore, after impact, the plate could vibrate freely without transmitting the vibration to the table. The foam block just slightly damped the vibration of the plate, since its back side was not completely flat but had a circular edge 3 mm high and 125 mm in diameter. The listeners sat facing the apparatus. The entire apparatus and the experimenter were 2 m away from the listener but were hidden from the listener's view by a 1.5×1.5 m frame covered with opaque paper.

Procedure. Before the session started, the experimenter dropped a ball, randomly selected from the ball set, on the plate. Then, the experimenter asked the listener if he or she could tell the shape of the object that had just been dropped. The main experiment then began. On each trial, the experimenter dropped the same ball three times following the listener's request. At each presentation, the ball was allowed to bounce freely until its motion ended. At the same time, during these three successive presentations, the listener had to create and adjust on a computer screen a disk that the listener thought was just as large as the ball that was being dropped. In particular, the listeners were asked to imagine the ball and to produce a disk with a diameter the same size as that of the ball they had imagined. Within the experimental session, each trial (a ball of a single size being dropped on the plate) was repeated five times in randomized order, for a total of 35 trials for each listener. At the beginning of each trial, the computer screen was blank. Custom software allowed disks to be drawn ranging from ~0.5 mm up to ~300 mm in diameter. Disk drawing was controlled by the listener with two pairs of keyboard keys: One pair incremented/decremented the diameter of the disk by ~0.5 mm with each keypress; the other incremented/ decremented the diameter of the disk by ~15 mm with each keypress. At the end of the experiment, and in all subsequent experiments, the listeners were asked questions about what material the impacting object (the ball) was made of and about the surface, shape, and material of the object it was dropped on. These questions were intended to determine what overall knowledge the listener had obtained. Finally, in both this experiment and all subsequent experiments, the listeners received no information about the size of the balls, the height from which the balls were dropped, or the materials of the balls or the plate.

Results and Discussion

To the first question, all listeners replied that a spherical object had been dropped. A first analysis was con-

ducted on the average size (in millimeters) of the disk diameters drawn by the listeners for each ball size. This analysis showed that the relation between subjective and actual diameters could best be described with a power function [linear fit $r^2 = .963$, F(1,5) = 131.11, p < .0001; power fit $r^2 = .985$, F(1,5) = 338.55, p < .0001 (exponent = 1.48)]. The disk diameters drawn by the listeners (log transformed and averaged across listeners for each stimulus) are presented in Figure 1.

Overall, the listeners underestimated the diameter of the ball, and this underestimation decreased as ball size increased (see Figure 1). Furthermore, log—log psychophysical functions of all listeners were positively accelerated, with slopes ranging from 1.26 to 1.71. Interestingly, the listeners' estimates dropped slightly for the largest ball size, for the listeners thought the two largest balls were of similar sizes.

Eight listeners reported that they heard metal balls, 1 reported hearing glass balls, and the last listener reported hearing stone balls. Seven listeners reported that the balls were dropped on a metal plate, and 3 that a pottery plate had been used.

The results show that the listeners were easily able to recognize the sound source event (i.e., a ball dropped on a plate) but failed to recognize its characteristics (i.e., the materials of the objects involved in the sound source event). Nonetheless, the listeners were able to provide reasonable estimates of the diameter of the ball. Overall, ball diameters were underestimated in comparison with their actual diameters, although the absolute error of the underestimation was relatively small (5–7 mm). It may be that this underestimation reflects what listeners thought about the material of the ball. The listeners believed that the balls were made of materials with higher densities than wood, and thus, underestimation of a ball's diameter in comparison with its actual diameter might reflect such an expectation. However, regardless of the perceived material of the ball, any error in material judgment should lead to a *constant* error in evaluating ball diameter (i.e., estimates should be greater or less than the actual size of the ball by a constant factor) rather than a decreasing error, as was found in the present experiment.

In conclusion, when evaluating a sound source event with no foregoing information, listeners can still provide reasonable estimates of properties of the NSO such as its size. Therefore, a veridical reconstruction of the sound source event is not a necessary prerequisite for producing reasonable estimates of the NSO.

EXPERIMENT 2

In the previous experiment, although listeners were not able to reconstruct veridically all of the properties of the sound source event, they were able nonetheless to provide reasonable estimates of NSO size. The goal of the present experiment was to understand whether listeners' estimates of the NSO were sensitive to unrevealed changes in SO properties.

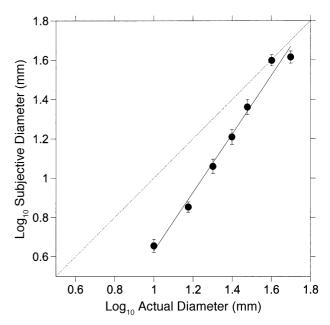


Figure 1. Experiment 1: Log subjective diameter as a function of the log actual diameter of the balls. The solid line represents the linear interpolation of the data. The lighter diagonal dashed line represents perfect match. Vertical bars show ± 1 standard errors of the mean.

For this experiment, listeners performed the same task as in Experiment 1. However, they were not aware that the ball could be dropped onto plates either 215 mm or 185 mm in diameter. In particular, a decrement in plate diameter (but leaving other characteristics, such as the plate's material and shape, unchanged) shifts all of the frequency components of the sound up by a certain ratio.

If listeners are able to extract information concerning the NSO (the ball) from the sound independently of information on the SO (the plate), they should be able to evaluate ball diameter independently of the size of the plate onto which the ball is dropped. Alternatively, if listeners are not able to segregate information about the two objects, we would expect listeners' estimates to be influenced by variations of the SO. In particular, previous research offers one alternative hypothesis. According to Lakatos et al. (1997) and Kunkler-Peck and Turvey (2000), the frequency domain carries information about SO size: Low-frequency sounds are produced by large vibrating objects, and high-frequency sounds are produced by small objects. If listeners do not segregate NSO properties (amplitude domain variations) from SO properties (frequency domain variations), they might add the two types of information and overestimate the ball's diameter when it is dropped onto the large plate.

Method

Subjects. Ten undergraduates of the University of Padua volunteered for the experiment. They all reported having normal hearing. None of the subjects had taken part in the previous experiment.

Apparatus and Procedure. The procedure was the same as in Experiment 1, with the addition of a second baked clay plate, sim-

ilar to the first but 185 mm in diameter. As in the first experiment, before the experimental session listeners heard the sound of a single ball, and were then asked if they recognized the shape of the impacting object from its sound. The rest of the procedure was also the same as in Experiment 1, except that the stimulus set (single combinations of ball and plate) was extended to include 35 new trials in which balls were dropped onto the 185-mm diameter plate. This manipulation resulted in a total of 70 trials per experimental session. Each listener evaluated a random sequence of stimuli, and the duration of the experimental session was approximately 45 min.

Results and Discussion

To the preliminary question, all listeners answered that a spherical object had been dropped. Log transformations of the disk diameters (in millimeters) drawn by listeners were averaged separately for each listener and each stimulus (a single combination of ball and plate). Psychophysical functions for the estimates from the 215- and 185-mm plates were calculated on these log-transformed data. A very close fit for behavioral data was obtained with two linear regressions performed as a function of log transformations of true ball diameters [respectively, $r^2 = .991$, F(1,5) = 567.30, p < .0001 (B = 1.59) for the 215-mm plate; and $r^2 = .989$, F(1,5) = 458.49, p < .0001 (B = 1.47) for the 185-mm plate] (see Figure 2). As in the previous experiment, the listeners' estimates showed a slight drop for the largest ball size.

Analysis of variance was performed on the log subjective diameters. Subjective diameters of balls dropped onto the smaller plate were underestimated as compared with the diameter of the same balls when dropped onto the larger plate $[F(1,9)=77.08,\,p<.0001]$ (see Figure 2). In particular, the different timbres produced by the two plates affected the perceived dimension of all balls in a similar fashion $[F(6,54)=2.14,\,p>.05]$: All balls increased their subjective diameter when dropped onto the larger plate. This pattern of results was similar across all listeners.

The listeners reported that they heard either metal, stone, glass, or wooden (1 listener) balls falling onto one or more plates made either of metal or pottery. Some listeners interpreted the different timbres of the two plates as being due to the ball hitting the same plate at different points. The majority of the listeners reported that within the experiment, either balls or plates could be made of more than one material.

As in the previous experiment, the listeners were easily able to recognize the sound source event (i.e., a ball dropped on a plate) but failed to recognize its characteristics (i.e., the materials of the objects as well as the number of plates). Despite these errors, and also as in the previous experiment, the listeners provided reasonable estimates of ball diameter, with absolute errors within a few millimeters.

The listeners thought that balls were larger when they were dropped on the larger plate. One explanation for this difference lies in their perception that the balls (or the plates) could be made of more than one material of different mass densities. Let us suppose that a listener

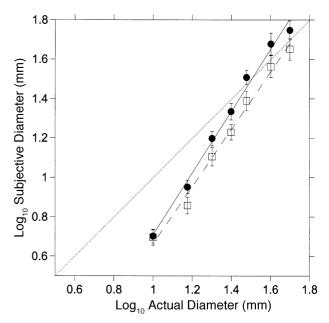


Figure 2. Experiment 2: Log subjective diameter as a function of the log actual diameter of the balls. Filled circles and empty squares represent, respectively, estimates collected when using the 215- and 185-mm plates. The solid and the heavy dashed lines represent, respectively, the linear functions fitting estimates for the 215- and 185-mm plates. The lighter dashed line represents perfect match. Vertical bars show ± 1 standard errors of the mean.

thought that the two timbres heard in the experiment were the results of two sets of balls of different mass densities, rather than of two plates with different diameters. In this case, a lower estimate associated with one timbre (e.g., the presumed higher mass density material) would be consistent with such an expectation. However, such a hypothesis is not backed up by the facts: Only 2 listeners thought that balls (or plates) could be made of two materials. Of the remaining 8 listeners, 4 thought that three materials (for either balls or plates) were used in the experiment, and 4 thought that balls were made of only one material but that only one plate was used, with each ball hitting the same plate at different points. Although the sound source events imagined by these listeners differed, their results were similar, with high estimates when a ball was dropped upon the larger plate and low estimates when it was dropped onto the smaller plate. Furthermore, as in the previous experiment, the error of their estimates was not constant, but large and negative for the small- or medium-sized balls, null or positive for the larger balls. Such a finding is not consistent with any of the listeners' beliefs about the materials of the balls.

These results may be more suitably explained by supposing that listeners coupled the sizes of the ball and the plate. Listeners were unable to segregate the effect of the different balls from that of the different plates, so they added the two types of information together, thus producing higher estimates when a large ball was dropped

upon the larger plate and lower estimates when the same ball was dropped onto the smaller plate.

Definitely, the two different timbres in the experiment led to a certain degree of confusion, as reflected by the data. The data of the present experiment show greater variability than the results of Experiment 1 (compare the standard errors in Figures 1 and 2). In fact, where results in the first experiment were strongly consistent across listeners, all of whom estimated similar perceptual diameters, the listeners' estimates in the second experiment were more variable.

In conclusion, the results of the present experiment show that listeners' estimations of the NSO are influenced by unrevealed changes in the properties of the SO.

EXPERIMENT 3

Experiment 2 showed that listeners' estimates were influenced by unrevealed changes of properties of the plate: The listeners produced higher estimates when a ball was dropped upon the larger plate and lower estimates when the ball was dropped upon the smaller plate. I suggested that this difference might be due to the facts that (1) the listeners were unable to segregate the information concerning the SO (the plate) from the information concerning the NSO (the ball) and (2) the listeners thus coupled the two types of information when producing their estimates. The goal of this third experiment was to test this hypothesis. Furthermore, the experiment was also designed to replicate the results of Experiment 2, as those results had shown a certain variability across listeners.

In the present experiment, the ball could be dropped onto either of the two plates used previously or onto a new 165-mm plate. The further decrement in plate diameter would produce a new and yet higher timbre by shifting all of the frequency components of the sound up by a certain ratio. If listeners coupled the size of the ball with the size of the plate, estimates collected for this smallest plate should be lowest, those for the 185-mm plate should be in the middle, and those for the 215-mm plate should be the highest.

Method

Subjects. Twelve undergraduates of the University of Padua volunteered for the experiment. They all reported having normal hearing. None of the subjects had taken part in the previous experiments.

Apparatus and Procedure. The procedure was the same as in Experiment 2. A third baked clay plate 165 mm in diameter was used in addition to the 215- and 185-mm plates from the second experiment. This third plate had the same characteristics as the plates used in Experiment 2 (material, shape, etc.) except for the diameter. As in Experiments 1 and 2, before the experimental session listeners heard the sound of a single ball dropping, and they were then asked whether they could recognize the shape of the impacting object from its sound. The procedure was similar to the procedure used for the second experiment, except that the sequence of random stimuli for the listeners to evaluate was repeated three times, not five as in the previous experiments. This change resulted in a total of 63 trials (7 balls × 3 plates × 3 repetitions) per experimental session. The duration of the session was approximately 40 min.

Results and Discussion

As far as the preliminary question is concerned, all listeners answered that a spherical object had been dropped. Log transformations of the disk diameters (in millimeters) drawn by the listeners were averaged separately for each listener and stimulus (single combination of ball and plate). Psychophysical functions for the estimates from the 215-, 185-, and 165-mm plates were calculated on these data. Three linear regressions performed on log transformations of the actual diameters of the balls fitted listeners' estimates closely [respectively, $r^2 = .985$, F(1,5) = 338.57, p < .0001 (B = 1.57) for subjective estimates from the 215-mm plate; $r^2 = .994$, F(1,5) =841.12, p < .0001 (B = 1.54) from the 185-mm plate; and $r^2 = .992$, F(1,5) = 677.44, p < .0001 (B = 1.56) from the 165-mm plate] (see Figure 3). As in the previous two experiments, the listeners' estimates showed a slight drop for the largest ball size.

Analysis of variance was performed on the data. As in Experiment 2, subjective estimates depended on the size of the plate: The bigger the plate, the higher the estimate of ball diameter $[F(2,22)=33.95,\,p<.0001]$ (see Figure 3). Furthermore, and also as in Experiment 2, the sounds produced by the three plates affected subjective diameters of all balls in a similar fashion (F<1). This pattern of results could be observed in data from 10 of the listeners. The remaining 2 made some reversals in their estimates for the 185- and 215-mm plates.

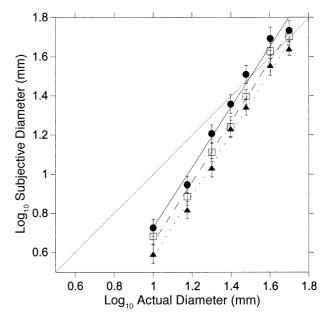


Figure 3. Experiment 3: Log subjective diameter as a function of the log actual diameter of the balls. Filled circles, empty squares, and filled triangles represent, respectively, estimates collected when using the 215-, 185-, and 165-mm plates. The heavier solid, dashed, and dotted lines represent, respectively, the linear functions fitting estimates for the 215-, 185-, and 165-mm plates. The light dashed line represents perfect match. Vertical bars show ± 1 standard errors of the mean.

The listeners reported that they heard either metal, stone, glass, or wooden balls falling on one (the majority of listeners) or more plates (2 listeners) made either of metal or pottery. As in Experiment 2, listeners interpreted the different timbres produced by the three plates as either the result of several materials being used in the balls or the plates, or of the ball being dropped onto different points of a single plate.

Once again in this experiment, listeners were easily able to recognize the sound source event (i.e., a ball dropped on a plate) but failed to recognize its characteristics (i.e., the materials of the objects and the number of plates). Once again, they were also able to provide reasonable diameters for the balls, with absolute errors within a few millimeters. As in Experiment 2, the majority of listeners (10) coupled the size of the ball with the size of the plate, providing the highest estimates for the larger balls dropped onto the larger plates and the lowest estimates for the smaller balls dropped onto the smallest plate.

The results of this experiment are again consistent with the hypothesis that listeners are not able to segregate, within a single sound event, features due to the different plates from those due to the different balls. Thus, an increment in plate diameter adds a constant value to the log-transformed subjective diameter of the ball. In addition, only 1 listener thought that the balls might be made of three materials of different mass densities, and only 1 thought that the plates might be made of three materials. Thus, an explanation of the different estimates recorded for plate diameter becomes difficult in light of the listeners who misinterpreted timbre changes as being due to different materials used in the experiment.

ANALYSIS OF THE PHYSICAL AND ACOUSTICAL EVENT

As reported earlier, regardless of the approach chosen, listeners do need to hear sounds in order to be able to estimate on the basis of them. Therefore, the acoustic signal must convey some information about the physical event that generated the sound. In particular, changes in the physical properties of the NSO and/or the SO must have a counterpart in changes of the properties of the acoustic signal. The aim of the following analysis is to highlight possible acoustic carriers of such information and to compare them with the subjective estimates of size.

The physical interaction between the ball and the plate and its consequences for the resulting acoustic waveform can be described. On one hand, the physical event is characterized by a ball of mass m falling on a plate from height h, subject to gravity acceleration g. When the ball reaches the plate, all of the potential energy (E = mhg) is converted into kinetic energy. After the impact, part of this energy is transformed into acoustic energy, the remaining energy is reconverted into kinetic energy (the ball bounces, the plate vibrates), and so on.

In the present experiments, the height of fall was constant. Therefore, the only energy variation was due to the

different masses of the balls, or, as the listeners believed, to their different mass densities. Ball mass is the result of the product of its volume and the mass density of the material. Moreover, when log transformed, the mass, volume, and diameter of the ball differ from each other only by a constant factor.

Generally speaking, energy is the capacity for doing work, power is the rate of doing work over time, and in the event under analysis, the work is the displacement of air causing sound. Because of all these factors, perception of ball size from its sound may be dependent on the ongoing air pressure wave reaching the ear—its poweras suggested by Grassi (2002). In addition, other acoustic parameters, which may be important for estimating the size, are theoretically predictable. First, the greater the mass impacting the plate, the greater will be the plate's oscillation. Consequently, the amplitude of the acoustic waveform will be greater, and so will the loudness of the resulting sound. Second, the greater the mass of the ball, the longer will be the duration of the oscillation and, consequently, the duration of the resulting waveform. However, although this statement is theoretically true, subsequent impacts may damp the vibration of the plate and thus shorten its oscillation. Third, the greater the ball mass, the longer will be the time of contact between the ball and the plate (Avanzini, 2001). Time of contact alters the frequency components of the sound produced by balls of different masses when they are dropped onto a plate, because a long time of contact damps vibrations whose periods are shorter than the time of contact itself. Consequently, the greater the mass impacting the plate, the greater is the damping of the high-frequency vibrations of the plate, resulting in a weakening of the high-frequency components. Thus, the sound of a light ball impacting a plate is bright, and the sound of a heavy ball impacting the same plate is dull. This difference can be captured by the centroid. The *centroid* is the sum of the frequency components of a sound weighted by their relative amplitudes. This acoustic index is related to the perceptual brightness of the sound (Grey & Gordon, 1978). Finally, substituting a smaller plate, as in Experiments 2 and 3, will affect the frequency of the entire acoustic event; a decrease in plate diameter produces a higher pitch sound, as all frequency components are shifted up by a certain ratio.

An increment in ball mass impacting the plate affects not only the sound on its first impact but also the high-order structure of the event (Warren & Verbrugge, 1984). The high-order structure is defined as the temporal distribution of a signal, and for the present event, as the temporal distribution of a ball's bounces. In fact, for impacts in which h is kept constant and m is increased, bounces will become less spread out in time. The impact velocity is independent of the mass of the ball. Furthermore, the higher the mass m, the longer the time of contact between the ball and the plate. As a consequence, the dissipation of kinetic energy for a long time of contact will be greater than the dissipation for a short contact. Therefore, after the impact, light balls will travel for

longer than heavier balls; thus, the distance between the first and second bounces will be greater for lighter balls, as will the overall number of bounces. After the first bounce, *mutatis mutandis*, the analyses can be repeated identically for all subsequent bounces.

Sounds of the dropped balls were recorded with a Sennheiser MKH 40 P48 microphone onto a portable Tascam DA-P1 DAT recorder with a 44.1 kHz sample rate and 16-bit resolution. About 20 recordings were recorded for each ball/plate combination. Duration of the sound, its amplitude peak, its frequency centroid, its average root mean square (RMS) power, and time between the first and second bounces were calculated from each recording and then averaged for each ball/plate combination (see Figure 4 and the Appendix for details). The modal value for the number of bounces was also recorded. These values were used to compute forward multiple linear regressions on the log-transformed subjective diameters gathered in the experiments.

Analysis of Experiment 1

In the forward multiple linear regression analysis, the following predictors were included: the average RMS power of the sound, its amplitude peak, its centroid, its overall duration, the temporal distance between the first two impacts, and the overall number of bounces. The power of the sound alone was a very good predictor of the performance [$r^2 = .978$, F(1,5) = 217.44, p < .0001].

Analysis of Experiment 2

The average RMS power of the sound could predict well the estimates provided by the listeners for the 215-

mm plate $[r^2 = .989, F(1,5) = 450.56, p < .0001]$. For the 185-mm plate, average duration was the best predictor of the estimates $[r^2 = .977, F(1,5) = 208.36, p < .0001]$, but the power of the sound was the second best predictor $[r^2 = .973, F(1,5) = 179.79, p < .0001]$.

The results of the second experiment showed that the perceptual size of the ball was dependent on the plate on which the ball is dropped. However, the change in plate diameter should, theoretically, affect only the frequency content of the sound. A new forward multiple linear regression was performed in order to understand whether any acoustic predictor could explain the difference between the perceptual diameters of the balls when they were dropped onto the small plate and the large plate. The difference between these log-transformed diameters from the 215-mm versus the 185-mm plate was calculated for each of the seven ball sizes and each of the acoustic indexes. These differences were used, respectively, as a dependent variable and as predictors for the new multiple forward regression. The difference in centroids for the data based on the two plates could provide a hypothesis for the different subjective estimates of size $[r^2 = .594,$ F(1,5) = 9.76, p = .026].

Analysis of Experiment 3

The average RMS power of the signal was the best predictor of performance for the subjective diameters from both the 215-mm plate $[r^2 = .985, F(1,5) = 318.69, p < .0001]$ and the 165-mm plate $[r^2 = .981, F(1,5) = 255.48, p < .0001]$. In contrast, for the 185-mm plate, a model including both the overall duration of the sound and the centroid fitted well with the results recorded

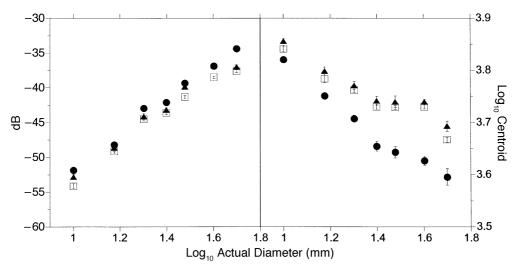


Figure 4. (Left) RMS power of the sound as a function of the log actual diameter of the ball. (Right) Log_{10} of the centroid as a function of the log diameter of the ball. Average RMS power of the sound was calculated over the entire event by means of a 50-msec sliding window. Computation of the centroid was performed on the entire event by averaging centroids computed by means of concatenated temporal windows of 50 msec. Calculation of the centroid included frequencies up to 22050 Hz. In computation of the centroid, all frequency components with an amplitude 30 dB below the highest frequency component were excluded. Filled circles, empty squares, and filled triangles represent powers or centroids for the 215-, 185-, and 165-mm plates, respectively. In both graphs, vertical bars represent ± 1 standard errors of the mean.

and the Log Diameter of the Dan										
	215-mm Plate			185-mm Plate			165-mm Plate			
	r^2	F(1,5)	p	r^2	F(1,5)	p	r^2	F(1,5)	p	
Power	.99***	414.05	.0001	.97***	213.94	.0001	.98***	177.67	.0001	
Peak	.93**	65.75	.0001	.94**	46.04	.0001	.90	83.25	.0001	
Centroid	.97**	144.38	.0001	.93*	57.44	.001	.92*	64.31	.0001	
Duration	.01	<1	_	.92	121.41	.0001	.96**	55.91	.001	
Acceleration	.33	2.45	>.05	.71	2.57	>.05	.34	12.53	.017	
Bounces	.49	4.72	>.05	.32	<1	_	.16	2.37	>.05	

Table 1
Results of the Linear Regressions Between the Acoustic Parameters and the Log Diameter of the Ball

[r^2 = .996, F(2,4) = 529.95, p < .0001]. However, after excluding duration, the power of the sound was shown to be the best predictor [r^2 = .975, F(2,4) = 197.74, p < .0001].

As with the results of the second experiment, an analysis was carried out between the differences in the log subjective diameters measured with the three plates on the one hand and the difference in the acoustical predictors on the other. None of the acoustic predictors could explain the different estimates measured with the three plates.

I then investigated how acoustic predictors vary as a function of the size/mass of the ball. Acoustic parameters were compared one by one with the log diameters of the balls by means of a series of linear regressions. The results are shown in Table 1.

Overall, results show a strong correlation between amplitude domain indexes (power in particular) and size of the ball. Also the centroid correlated highly with the size of the ball. In contrast, overall duration (with the exception of the durations recorded with the 185-mm plate), time between first and second bounce, and overall number of bounces correlated poorly with the size of the ball.

In Figure 4, powers and centroids of the sounds are plotted as a function of the log actual diameter of the ball. The power of the sound increases as the size/mass of the ball increases. On average, the most powerful sound was produced by the 215-mm plate (M = -42 dB), and the least powerful sound by the 185-mm plate (M =-44 dB). The difference in power between the three series of data ranged from a maximum of 3.5 dB to a minimum of 0.9 dB. In the majority of cases, these differences were small (within 1.5 dB). Power seemed to increase monotonically as the size of the ball increased. However, for the largest ball size, both a reversal (with the 165-mm plate) and a power value smaller than expected (with the 185-mm plate) emerged. Overall, the centroid decreased as the size of the ball increased. The sound with the highest centroid (brightest timbre) was the one produced by the 165-mm plate, followed by that produced by the 185mm plate. The centroid of the 215-mm plate was the lowest (dullest timbre).

These analyses suggest that no acoustic predictor is as efficient as the actual size of the balls for fitting listen-

ers' estimates (compare the r^2 s in this section with those for Experiments 1, 2, and 3). Nonetheless, predictors such as the power of the sound (amplitude domain predictor) were almost as effective as the actual size of the balls for predicting listeners' estimates. Overall duration of the sound was also a good predictor of listeners' performance, but only for estimates collected when using the 185-mm plate; the same predictor could not be used as efficiently to scale the estimates for either the 215-mm or the 165-mm plate. Consequently, this result seems purely fortuitous. Furthermore, the different centroids of the sounds produced by the plates suggest a hypothesis for the different subjective estimates associated with the plates (see analysis of Experiment 2).

In summary, variation in the size/mass of balls found a counterpart in variation of the amplitude component of the sound (i.e., its power). In addition, an amplitude domain index (i.e., power) was the best acoustical predictor of listeners' performance.

GENERAL DISCUSSION

The results of the current experiments suggest that it is possible to provide reasonable estimates of NSO properties even with no foregoing information about the sound source event. Furthermore, even if listeners reconstruct a nonveridical sound source event, they can still provide reasonable estimates of the NSO. Moreover, unrevealed changes in the properties of the SO can influence listeners' estimates of the NSO. Comparison between subjective estimates and acoustic parameters showed that sound power was almost as efficient as the actual size of the ball for fitting listeners' estimates. Finally, analysis confirmed that changes in the size/mass of the ball found a counterpart mainly in changes of the amplitude content of the sound, but changes in plate diameter found a counterpart in changes of the frequency content of the sound, leaving the amplitude substantially unaffected.

Overall, both the present results and those of Freed (1990) and Li et al. (1991) have shown that listeners are able to estimate the physical properties of an object (the NSO) whose contribution to the resulting acoustic waveform is marginal. However, results from the present experiments seem even more remarkable than the previous

^{***}Best acoustic predictor. **Second best predictor. *Third best predictor.

results, as listeners provided scaling of the ball size on an absolute basis. Therefore, the listeners' task here was more difficult than guessing the gender of the walker in a two-alternative forced choice task (Li et al., 1991). In addition, listeners' answers were isometric to the stimulus, so that it was possible to compare listeners' estimations directly with the physical parameter under investigation. In contrast, Freed's subjects estimated mallet hardness only on an ordinal scale, so it was impossible to know whether the estimates were scaled similarly to the actual hardnesses of the mallets.

Throughout the experiments, the results showed a steady psychophysical function with an exponent slightly greater than 1.5. In addition, the monotonic increment of such functions fell slightly at the largest ball size. This result reflects nonmonotonic increments and reversals of almost all acoustic parameters with the larger ball size (see Figure 4 and the Appendix). It may be that, with the largest ball, modal vibrations of the plate were saturated. Hence, the largest ball did not lead to an appropriate increment (or decrement) in the acoustical parameters, and the listeners' estimates mirrored the acoustic characteristics of the sound. The sounds themselves, however, did carry information about the changes in the size/mass of the ball, and many acoustic predictors covaried with the size/mass of the ball. In particular, size/mass covaried with power, and power was the most efficient acoustic predictor of listeners' performance. Grassi (2002) demonstrated that manipulation of this acoustic parameter was the most effective at impairing performance in a sizefrom-sound paradigm. In contrast, in studies investigating an SO, authors indicated frequency as the acoustic feature carrying information about the object's shape and size (Kunkler-Peck & Turvey, 2000; Lakatos et al., 1997). Power may be the acoustic clue linking actual and subjective size of the NSO. However, only studies in which acoustic features are directly manipulated (e.g., Grassi, 2002) or environmental sounds are investigated by means of synthetic stimuli (e.g., Klatzky, Pai, & Krotkov, 2000; Lutfi, 2001; Lutfi & Oh, 1997; McAdams, Chaigne, & Roussarie, 2004) can provide final confirmation of this hypothesis.

In Experiments 2 and 3, the listeners coupled the sizes of the balls and the plate and systematically perceived the balls as smaller when dropped onto a small plate. The effect of the plate on subjective estimates was additive: The plate did not affect the exponent of psychophysical functions, but it did affect the intercept of the linear regressions in log—log coordinates. A decrement of the plate diameter resulted mainly in a different timbre of the sound, as all frequency components were shifted up by a certain ratio (see Figure 4). Therefore, a high pitch and bright sound conveyed the impression that the sound was produced by a small ball, and this result was supported by the centroid being the most powerful predictor for the different subjective estimates measured in Experiment 2.

At the beginning of the last century, many researchers studied tonal volume (Boring, 1926; Rich, 1916). Ac-

cording to the tonal volume hypothesis, high-pitch tones are extensively associated with small objects, and low tones are associated with large and voluminous objects (Perrott, Musicant, & Schwethelm, 1980). It is likely that the size of a ball dropped onto a small plate was underestimated when compared with the size of the same ball when dropped onto a larger plate, because the resulting sound was both higher in pitch and brighter. Freed (1990) failed to find an effect of the different pans on hardness estimates of mallets, but the reason for this difference may lie in the information Freed's listeners received before the experiment. Because of the information they had been given, Freed's listeners could interpret the timbre changes due to the different pans and ignore them when estimating hardness.

Can the results of the present experiments be interpreted by a direct perception approach? In the introduction, I mentioned two problems raised by studies using real impact sounds: perceptual segregation of information concerning the two objects involved in the impact, and veridicality of listeners' estimates of the physical properties of the objects. I believe that the solution of both problems is fundamental to answering the question of whether listeners perceive the sound source event directly or not.

In impact sounds, wave pressure reaching the ear is the result of the physical interaction of (at least) two objects. Nonetheless, in both the present and previous experiments (Carello et al., 1998; Freed, 1990; Kunkler-Peck & Turvey, 2000; Lakatos et al., 1997; Li et al., 1991), listeners seemed able to segregate out of this single pressure wave the information concerning the sole object they were asked to evaluate. Did listeners really segregate information concerning the plate from that concerning the ball? How accurate was the reconstruction of the soundproducing event? In the present experiments, listeners easily understood that the sound was the result of a ball being dropped on a circular plate, and their perceptions of both these items can be explained. Balls are clearly distinguishable from other objects because of the regularity of their bounces (Warren & Verbrugge, 1984), and likewise, Kunkler-Peck and Turvey (2000) have demonstrated that we can recognize the shape of an SO (the plate) from its sound. Nonetheless, a certain confusion arose for listeners when they had to recognize the materials of the two objects, especially of the nonvibrating object. Across the experiments, only a few listeners reported hearing wooden balls; all of the rest thought that the balls might be made of metal, stone, glass, or pottery. In addition, in the second and third experiments, only a few listeners understood that sounds were obtained using two (or three) plates. On the contrary, most listeners thought either that the ball was hitting one plate at different points or that the balls might be made of different materials. Nonetheless, although the majority of listeners thought that only one plate was used in the experiment, results showed that their estimates were indeed influenced by the size of the plate. Thus, listeners were unable to segregate from the resulting sound the effects due to the size of the plate and those due to the size of the ball, since their reconstruction of the sound-producing event was only partially veridical.

In the present and in previous experiments (Carello et al., 1998; Freed, 1990; Kunkler-Peck & Turvey, 2000; Lakatos et al., 1997; Li et al., 1991), after listening to the sound, listeners were usually accurate in judging static properties of the physical event that generated it. Furthermore, the best predictors for listeners' performance could be found in the physical properties of the object itself rather than in the acoustic characteristics of the sound (Carello et al., 1998). Did listeners somehow achieve a direct estimate of the size of the ball? In the present experiment, the simplest backward path for recovering ball size included recognition of the shapes of both the NSO and SO and selection of the relevant acoustic feature (e.g., sound power) that conveyed size information about the NSO (the invariant). At this stage, the listener needed to apply the transfer function that links the mass of the ball to the power of the sound. The final step required calculation of the logarithm of ball mass, followed by transformation of the latter into the logarithm of diameter by multiplying its value by a constant. Such an output could then be scaled according to some secondary properties of the event, such as the material of the ball (its mass density), the height from which it was dropped, or the distance of the sound source event. In the present experiments, listeners provided reasonable estimates of the size of the balls, and the balls' actual sizes turned out to be the best single predictor for listeners' estimates. Therefore, at a first glance, results could reflect a direct perception of size. However, the results showed that psychophysical functions were always positively accelerated, and thus demonstrated an evident error in the transfer function of the process. For example, if listeners thought that balls were made out of steel, we would have expected a constant underestimation of the subjective size of the balls when compared with their actual size, as the mass density of steel is higher than that of wood. This was not the case, however; none of the psychophysical functions measured in the experiments showed a constant error in the estimations. In fact, as many acoustic parameters could efficiently scale the subjective size of the ball, it may also be plausible that listeners used the disk-generating software to scale one or more aspects of the auditory sensation rather than evaluating the size of the ball directly.

Moreover, how are the abilities to reconstruct the sound source event and to evaluate the physical properties of the event related to each other? Research has demonstrated that listeners are able to recognize an event among a class of possible events from its sound (Gaver, 1988; Gygi, Kidd, & Watson, 2004; Vanderveer, 1979; Warren & Verbrugge, 1984). Furthermore, as we have seen, listeners are able to estimate the physical properties of an event veridically (Carello et al., 1998; Freed, 1990; Kunkler-Peck & Turvey, 2000; Lakatos et al., 1997; Li et al., 1991). In addition, the results of the present experiments show that a veridical reconstruction of the sound source event is not a

necessary prerequisite for providing almost veridical estimates of the physical properties of an object. Thus, listeners can provide almost veridical estimations of an object even after an erroneous reconstruction of the sound source event. Does this result not suggest that listeners evaluate sounds rather than sources?

In conclusion, we know that human beings are born with limited capabilities, and one way of overcoming this drawback is to maximize existing capabilities. In this sense, I think that the hearing system does not need to provide us with veridical estimates of the size of an object, but rather, with functional estimates of the size of an event we are listening to (e.g., the event "ball dropping on a plate"). An auditory estimate of an event's size may be crucial when objects are occluded from vision. In addition, large and voluminous events may be potentially more dangerous than smaller events. If we are listening to a low-frequency sound, we are likely encountering an event that is generated by a large object. If we are listening to a high-amplitude sound, we are encountering an event that is close to us, or generated by a large object, or both. In any case, a large size estimate of the event (i.e., an estimate coupling SO and NSO sizes) would be advantageous for the perceiver, as he/she might benefit from greater margins of safety if he/she needs to prepare a motor reaction to the event.

REFERENCES

AVANZINI, F. (2001). Computational issues in physically-based sound models. Unpublished doctoral dissertation, University of Padua.

BORING, E. G. (1926). Auditory theory with special reference to intensity, volume, and localization. *American Journal of Psychology*, **37**, 157-188.

CARELLO, C., ANDERSON, K. L., & KUNKLER-PECK, A. J. (1998). Perception of object length by sound. *Psychological Science*, 9, 211-214.
FOWLER, C. A. (1990). Sound-producing sources as objects of percep-

tion: Rate normalization and nonspeech perception. *Journal of the Acoustical Society of America*, **88**, 1236-1249.

FOWLER, C. A. (1991). Auditory perception is not special: We see the world, we feel the world, we hear the world. *Journal of the Acoustical Society of America*, **89**, 2910-2915.

FREED, D. J. (1990). Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events. <u>Journal of the Acousti-</u> cal Society of America, 87, 311-322.

GAVER, W. W. (1988). Everyday listening and auditory icons. Unpublished doctoral dissertation, University of California, San Diego.

GAVER, W. W. (1993a). How do we hear the world? Explorations in ecological acoustics. *Ecological Psychology*, 5, 285-313.

GAVER, W. W. (1993b). What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychology*, 5, 1-29.

GRASSI, M. (2002). Recognising the size of objects from sounds with manipulated acoustical parameters. In J. A. Da Silva, E. H. Matsushima, & P. Ribeiro-Filho (Eds.), Fechner Day 2002: Proceedings of the International Society for Psychophysics (pp. 392-397). Rio de Janeiro: International Society for Psychophysics.

GREY, J. M., & GORDON, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. <u>Journal of the Acoustical Society of America</u>, 63, 1493-1500.

GYGI, B., KIDD, G. R., & WATSON, C. S. (2004). Spectral-temporal factors in the identification of environmental sounds. *Journal of the Acoustical Society of America*, **115**, 1252-1265.

KLATZKY, R. L., PAI, D. K., & KROTKOV, E. P. (2000). Perception of material from contact sounds. <u>Presence: Teleoperators & Virtual Environments</u>, 9, 399-410.

- Kunkler-Peck, A. J., & Turvey, M. T. (2000). Hearing shape. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 279-294.
- LAKATOS, S., MCADAMS, S., & CAUSSÉ, R. (1997). The representation of auditory source characteristics: Simple geometric form. *Perception & Psychophysics*, 59, 1180-1190.
- LI, X., LOGAN, R. J., & PASTORE, R. E. (1991). Perception of acoustic source characteristics: Walking sounds. <u>Journal of the Acoustical So-</u> ciety of America, 90, 3036-3049.
- LUTFI, R. A. (2001). Auditory detection of hollowness. *Journal of the Acoustical Society of America*, **110**, 1010-1019.
- LUTFI, R. A., & OH, E. L. (1997). Auditory discrimination of material changes in a struck-clamped bar. <u>Journal of the Acoustical Society of America</u>, 102, 3647-3656.
- MCADAMS, S., CHAIGNE, A., & ROUSSARIE, V. (2004). The psychomechanics of simulated sound sources: Material properties of impacted bars. *Journal of the Acoustical Society of America*, 115, 1306-1320.
- Morse, P. M. (1981). Vibration and sound (3rd ed.). New York: Acoustical Society of America.

- Perrott, D. R., Musicant, A., & Schwethelm, B. (1980). The expanding-image effect: The concept of tonal volume revisited. *Journal of Auditory Research*, **20**, 43-55.
- RICH, G. J. (1916). A preliminary study on tonal volume. *Journal of Experimental Psychology*, 1, 13-22.
- VANDERVEER, N. J. (1979). Ecological acoustics: Human perception of environmental sounds. *Dissertation Abstracts International*, 40(09), 4543B. (UMI No. 8004002)
- WARREN, W. H., & VERBRUGGE, R. R. (1984). Auditory perception of breaking and bouncing events: A case study in ecological acoustics. *Journal of Experimental Psychology: Human Perception & Perfor*mance, 10, 704-712.

NOTE

1. This example is a typical exception to the exciter–resonator partition: The rod is, at the same time, exciter (it strikes the floor) and resonator (its vibration is the main sound source).

APPENDIX

Average Acoustic Values Extracted From Recordings of
Stimuli Used in Experiments 1, 2, and 3

Sumun Used in Experiments 1, 2, and 3										
Log ₁₀ Diameter of Ball (mm)	Duration (msec)	Peak (dB)	Bounce-to-Bounce Interval (msec)	Number of Bounces						
215-mm Plate										
1.00	765.1	-20.9	213.0	4						
1.18	782.1	-16.3	195.0	3						
1.30	716.1	-9.9	170.8	3						
1.40	707.5	-10.0	167.0	3 3 2						
1.48	671.7	-6.9	137.1	2						
1.60	664.5	-5.5	111.9	2 3						
1.70	903.8	-5.0	195.0	3						
Average	744.4	-10.6	169.9	3						
185-mm Plate										
1.00	825.9	-20.8	204.2	3						
1.18	876.7	-17.1	198.1	3 3 3 3 2 3						
1.30	961.9	-10.5	175.6	3						
1.40	966.9	-9.8	175.3	3						
1.48	1,021.5	-7.0	149.2	3						
1.60	1,123.7	-5.1	146.6	2						
1.70	1,115.2	-6.6	191.8	3						
Average	984.5	-10.9	177.2	3						
165-mm Plate										
1.00	818.1	-20.2	220.9	4						
1.18	844.4	-17.5	198.2	3						
1.30	846.5	-12.6	173.4	3						
1.40	895.8	-10.6	171.4	3 3 2						
1.48	927.1	-8.5	159.0	2						
1.60	979.2	-5.3	153.6	3						
1.70	970.5	-6.7	173.7	3						
Average	897.3	-11.6	178.6	3						

Note—Overall duration of the sound was calculated from onset of the sound to the point at which amplitude of the waveform reached the amplitude level of the background noise. Bounce-to-bounce interval is the time between the first and second bounces.

(Manuscript received January 21, 2003; revision accepted for publication May 18, 2004.)