

# The psychomechanics of simulated sound sources: Material properties of impacted thin plates<sup>a)</sup>

Stephen McAdams<sup>b)</sup>

Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Schulich School of Music, McGill University, 555 Sherbrooke Street West, Montreal, Quebec H3A 1E3, Canada

Vincent Roussarie

Direction de la Recherche et de l'Innovation Automobile, PSA Peugeot-Citroën, Route de Gizy, F-78943 Vélizy-Villacoublay, France

Antoine Chaigne

École Nationale Supérieure de Techniques Avancées (ENSTA Paris Tech), Unité de Mécanique, chemin de la Humière, F-91761 Palaiseau, France

Bruno L. Giordano

Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Schulich School of Music, McGill University, 555 Sherbrooke Street West, Montreal, Quebec H3A 1E3, Canada

(Received 13 October 2009; revised 23 June 2010; accepted 23 June 2010)

Sounds convey information about the materials composing an object. Stimuli were synthesized using a computer model of impacted plates that varied their material properties: viscoelastic and thermoelastic damping and wave velocity (related to elasticity and mass density). The range of damping properties represented a continuum between materials with predominant viscoelastic and thermoelastic damping (glass and aluminum, respectively). The perceptual structure of the sounds was inferred from multidimensional scaling of dissimilarity judgments and from their categorization as glass or aluminum. Dissimilarity ratings revealed dimensions that were closely related to mechanical properties: a wave-velocity-related dimension associated with pitch and a damping-related dimension associated with timbre and duration. When asked to categorize sounds, however, listeners ignored the cues related to wave velocity and focused on cues related to damping. In both dissimilarity-rating and identification experiments, the results were independent of the material of the mallet striking the plate (rubber or wood). Listeners thus appear to select acoustical information that is reliable for a given perceptual task. Because the frequency changes responsible for detecting changes in wave velocity can also be due to changes in geometry, they are not as reliable for material identification as are damping cues.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3466867]

PACS number(s): 43.75.Kk, 43.66.Jh, 43.66.Lj [DD]

Pages: 1401–1413

## I. INTRODUCTION

In order to categorize, identify, and interact with sound sources in the environment, we need to perceive their geometric and material properties and the interactions among them (McAdams, 1993). Previous studies have used a variety of approaches such as classification, recognition or identification of the object (Avanzini and Rocchesso, 2001; Cabe and Pittenger, 2000; Houix *et al.*, 1999; Krotkov *et al.*, 1996; Warren and Verbrugge, 1984), discrimination among sounds (Lakatos, *et al.*, 1997; Lutfi, 2001; Lutfi and Oh, 1997), scaling sounds according to prespecified properties such as perceived hardness, length, mass, and speed (Freed, 1990; Carello *et al.*, 1998; Kunkler-Peck and Turvey, 2000) or es-

timating the relative similarity of sound pairs (Klatzky *et al.*, 2000; McAdams *et al.*, 2004). The last technique is of particular interest as it lends itself to an exploratory data-analysis approach that imposes few *a priori* hypotheses on how the mechanics or acoustics of the sound source influences the perceptual structure. An examination of differences among experimental tasks could lead us to better understand what kinds of information are relevant for specific context-dependent perceptual tasks in the everyday world. To this end, the current study employed both dissimilarity ratings (involving a more general perceptual comparison) and material categorization (involving a greater focus on source-related acoustical properties).

We used physical models that simulate the vibratory behavior of objects, because one can control precisely the mechanical parameters of the synthesis model, such as those related to damping, all the while obtaining the complex relations among analytic signal parameters that specify the mechanical nature of the object. Phenomena related to energy loss are present in all physical vibrations. In any physical system that is excited, the energy introduced into the system

<sup>a)</sup> Preliminary results leading to the present study were first reported at the 5th French Acoustics Congress, Lausanne (McAdams, 2000a) and the 139th Meeting of the Acoustical Society of America, Atlanta, GA (McAdams, 2000b). Initial work for this study was performed at STMS-Ircam-CNRS and ENST in partial fulfillment of the requirements for the Ph.D. dissertation of V.R. at the Université du Maine, Le Mans.

<sup>b)</sup> Author to whom correspondence should be addressed. Electronic mail: smc@music.mcgill.ca

will dissipate over time in the form of the generation of heat internally and sound radiation externally. Energy loss processes also affect the sound and are perceived by listeners as a diminishing of the sound level and by certain changes in timbre, due, for example, to the relatively faster rates of decay of higher partials compared to those of lower partials. A realistic physical model must thus produce changes in the sound that resemble those in natural physical phenomena. The physical justification and mathematical formulation of these phenomena can be very complex at times. It thus behooves us to start with simple general models.

We have shown in a previous study on impacted bars (McAdams *et al.*, 2004) that a simple damping model including only terms for viscoelastic loss and fluid damping gives rise to a single perceptual dimension that is salient for listeners. It is therefore natural to turn now to plates, which require a more complex model with a greater number of damping parameters, in order to determine whether this mechanical phenomenon is relevant for another class of vibrating objects. We will describe the influence of the various dissipative phenomena in order to better understand their respective roles in the perceptual characterization of materials. After having presented the damping models and their physical meaning for each material, we will describe the hybrid model used to synthesize the experimental stimuli. In the first perceptual experiment, we characterize listeners' sensitivity to material-related mechanical phenomena inherent to the physical system of an impacted plate. Listeners are shown to be directly sensitive to various aspects of energy dissipation in the materials. In a second experiment, we observe how these physical phenomena affect listeners' categorization of the sounds according to their materials.

Perceptual studies of material properties of struck objects have demonstrated a strong dependence of perceptual judgments on acoustical features related to the dissipation of energy within the vibrating object. These acoustical features vary widely in complexity, ranging from the relatively simple measure of the duration of the sound (Giordano and McAdams, 2006; Giordano *et al.*, 2010), through more sophisticated measures of the temporal decay of the overall sound level (Freed, 1990; Giordano and McAdams, 2006; Giordano *et al.*, 2010), to measures of the decay time for single spectral components of the impact sound (Lutfi, 2001; Lutfi and Liu, 2007) and global measures of the damping spectrum (damping as a function of the frequency of the constituent partials; Klatzky *et al.*, 2000; Avanzini and Rocchesso, 2001; Giordano and McAdams, 2006; Giordano *et al.*, 2010). It is still unclear which of these variables are perceptually relevant and which, instead, appear as strongly associated with the behavioral outcome simply because they are correlated with variables that are indeed the object of perception. Here, we propose a model for the extraction of the acoustical correlates of damping, test its ability to explain the behavioral outcome for two different perceptual tasks, and test the correlation of its parameters with several of the measures previously associated with material perception.

Another issue addressed in this study concerns the nature of differences among listeners in the relative contributions of various acoustical properties to a given perceptual

task, and how these contributions vary with the perceptual task to be completed. Lutfi and Liu (2007) have documented participant-specific listening styles across different source-perception tasks (size and material, mallet hardness, and presence/absence of external damping). In particular, they observed that, across variations in both the sound set and the judged property of the sound source, interindividual variability explained an important proportion of the variability in the perceptual weights for the different properties of the investigated impact sounds. This result was interpreted as supporting the presence of listener-specific strategies in the weighting of acoustical information, independent of task and sound set.

Lutfi and Liu (2007) also found that the disagreement among observers with respect to the weighting strategies increased when the task was less constrained (size and material or mallet hardness vs presence/absence of external damping). Giordano *et al.* (2010) proposed that acoustical features for which learning abilities are at their best and for which sensory noise is at its minimum are more relevant in determining the behavioral response. Further, inspired by the veridicality assumption of the ecological approach to perception, they proposed that acoustical features are weighted perceptually in proportion to the extent to which they accurately specify the source property being judged. In the present study, we measure the degree to which participants use the same weighting strategy when evaluating two different sound sets (plates having different material properties struck with wood or rubber mallets) with two different perceptual tasks (dissimilarity rating and material categorization). Because Lutfi and Liu (2007) varied both task and sound set across experimental conditions, it is unclear which of these two methodological sources is responsible for introducing variation in the perceptual weighting strategies. In contrast to their findings, the results of the present study suggest that individual listening strategies remain relatively constant across variations in the sound set, but are strongly affected by variations in the task.

The remainder of this article is organized as follows: in Sec. II we present the physical model used to synthesize the impact sounds investigated in the perceptual experiments; in Sec. III we introduce an analysis system for the extraction of the acoustical features of impact sounds; in Secs. IV and V we present dissimilarity-rating and material-identification experiments, respectively, conducted on the same two sets of impact sounds; and in Sec. VI we present an analysis of the acoustical correlates for data from Experiments 1 and 2 and analyze the extent to which participants weight acoustical information consistently across variations in sound set and task.

## II. SYNTHESIS OF IMPACTED PLATE SOUNDS BY PHYSICAL MODELING

The sound signals used in this study result from the simulation of a sphere-plate impact. In the following, we will describe the physical models of the plate, of the sphere-plate contact governed by Hertz's law, and of the acoustic radiation. A full description of the model is available in Lambourg *et al.* (2001). Here, we present only aspects of the simulation

TABLE I. Parameters used for the plate model.

Symbol	Definition (units)
$l_x, l_y, h$	Length, width and thickness of the plate (m)
$D_i$	Rigidity constants (MPa)
$\rho$	Plate mass density (kg/m <sup>3</sup> )
$\rho_a$	Air density (kg/m <sup>3</sup> )
$c_0$	Air velocity (m/s)
$R_{vm}, s_m$	Parameters for viscoelastic loss model
$R_1, c_1$	Parameters for thermoelastic loss model
$a_n, b_m$	Parameters for model of acoustic radiation loss
$\omega_c$	Critical frequency (rad/s)
$c_i$	Wave velocity ( $E_i/\rho$ ) <sup>1/2</sup> (m/s)
$m$	Mallet mass (kg)
$V$	Mallet speed (m/s)
$K$	Stiffness constant (N m <sup>-3/2</sup> )
$\alpha$	Damping factor (s <sup>-1</sup> )
$P_\alpha$	Slope of the $\alpha(f)$ curve (rad/s/Hz)

relevant to the current study, focusing on damping and wave velocity (resulting from a combined variation of elasticity and mass density). The notations used to describe the physical parameters in the equations are listed in Table I.

## A. The plate model

### 1. Equations of the general model

The physical system of the model is composed of a thin plate with free edges and a sphere that collides with the plate. The model is based on the theory of flexural vibrations for thin plates (Kirchhoff-Love assumptions; Leissa, 1969, pp. 331–340). The action of the impactor is represented by an external force in the equation of motion. In the general case, the elasticity of the plate is represented by four rigidity constants  $D_i$  (in MPa; see Lambourg *et al.*, 2001). This reduces to only two constants  $D_1$  and  $D_4$  in the case of isotropic materials such as glass and aluminum. In our case, the rigidity constants include internal and radiation losses, which we will now describe.

### 2. Modeling the dissipative phenomena

We considered three major dissipative mechanisms that are representative of a large class of materials: viscoelasticity, thermoelasticity, and radiation. The simulations are made in the time-domain. In this case, the losses are represented by derivatives of forces and displacement versus time (Chaigne and Lambourg, 2001). In the Laplace domain, a time derivative is represented by a multiplication by the Laplace variable  $s$  (see, for example, Schiff, 1999, pp. 2–22), and we will use this notation throughout the text. The Laplace transforms  $\tilde{D}_i$  of the rigidity constants (also called complex rigidities in what follows) are then written

$$\tilde{D}_i = D_i(1 + \tilde{d}_{iv}(s) + \tilde{d}_{it}(s) + \tilde{d}_{ir}(s)) \quad (1)$$

The term  $\tilde{d}_{iv}$  describes viscoelastic losses resulting from the fact that there is a delay between the instant of application of the stress and the resulting strain. This is the main source of loss in glass below a critical frequency  $f_c$  described

below. We then use for the expression of  $\tilde{d}_{iv}$  a particular class of the differential representation of Wiechert's two-cell model (Chaigne and Lambourg, 2001, a spring in parallel with two Maxwell models, each composed of a spring and a damper in series). For this type of model,  $\tilde{d}_{iv}$  is written

$$\tilde{d}_{iv}(s) = \sum_{m=1}^2 \frac{sR_{vm}}{s + s_m} \quad (2)$$

where  $R_{vm}$  and  $s_m$  are the real intrinsic parameters of the material.

The term  $\tilde{d}_{it}$  represents the thermoelastic losses due to the coupling between the flexural waves and the thermal waves in the thickness of the plate,  $h$ . This is the main source of loss for aluminum at low frequencies ( $f < f_c$ ). The law of thermoelastic behavior, as well as the heat transfer equation, allow us to establish the expression of  $\tilde{d}_{it}$  in an approximate model of thin isotropic plates

$$\tilde{d}_{it}(s) = \frac{R_1 s}{s + c_1/h^2} \quad (3)$$

where  $R_1$  and  $c_1$  are the thermoelastic coefficients (Chaigne and Lambourg, 2001). Note in this expression that the thermal losses decrease with the square of the plate's thickness,  $h^2$ .

The term  $\tilde{d}_{ir}$  represents the losses due to an acoustical coupling between the plate and the external fluid (air). We presume that the influence of the fluid charge on a plate of finite dimensions can be modeled asymptotically for higher order modes by performing a Padé development of the rigidity (Chaigne and Lambourg, 2001). Radiation losses are dominant above a critical frequency  $f_c$  written

$$f_c = \frac{c_0^2}{2\pi} \sqrt{\frac{\rho}{h^2 D_1}} \quad (4)$$

In the range of frequencies lower than  $f_c$ , the other losses usually dominate. With the formalism of viscoelasticity (in the form of polynomials of the Laplace variable  $s$ ), the radiation loss model is written:

$$\tilde{d}_{ir}(s) = \frac{2\rho_a c_0}{\omega_c \rho h} \frac{\sum_{m=1}^3 b_m \left(\frac{s}{\omega_c}\right)^m}{\sum_{n=0}^3 a_n \left(\frac{s}{\omega_c}\right)^n} \quad \text{with} \quad \omega_c = 2\pi f_c. \quad (5)$$

### 3. The specific cases of glass and aluminum

The relevance of the different damping mechanisms varies across materials. We have chosen to study two isotropic materials, glass and aluminum, specifically because the main losses of each are of different origins. We will present the glass model, in which the viscoelastic losses are preponderant for  $f < f_c$ , and the aluminum model, in which the main losses are thermoelastic below  $f_c$ . For both materials, losses are mainly due to radiation above that frequency. For these materials (as for any isotropic material) the set of rigidity constants reduces to  $D_1$  and  $D_4$ . The thermoelastic and radia-

tion parameters are directly derived from the physical models of thermal and sound-structure coupling. The viscoelastic parameters (for glass) are derived from experiments in which the damping factors for all frequency components are determined, i.e., the damping spectrum, and then the model values are determined by fitting a curve to these empirical values.

In the case of glass, the damping model includes three terms: a constant fluid damping term, a viscoelastic damping term, and a radiation damping term. The damping model for aluminum includes three terms as well, replacing the viscoelastic damping term with a thermoelastic damping term.

## B. Computation of radiated pressure

The computation of the radiated pressure is based on the resolution of Rayleigh's integral. This method allows the calculation of the pressure at an arbitrary point in space by considering only the normal velocity profile at the source (cf. Lambourg *et al.*, 2001).

## C. Hertz's law and the sphere-plate interaction

Doutaut *et al.* (1998) performed finite-difference simulations of sphere-bar collisions in order to simulate mallet percussion instrument sounds. The same method is used here to simulate the impact of spheres on plates. The contact is governed by Hertz's law for elastic bodies (see, for example, Johnson, 2003, pp. 84–106) where the interaction force  $F$  is a nonlinear function of the compression  $\delta$ . This compression is equal to the difference between sphere and plate displacement at the contact point. Thus, we write:

$$F = K\delta^{3/2} \quad (6)$$

where  $K$  (in  $\text{N m}^{-3/2}$ ) is the impact stiffness that depends on the rigidity of both sphere and plate. This constant is crucial because it determines the width and amplitude of the contact force and thus the number of modes excited in the plate.

## D. Parameters of the synthesis model

The plate model allows us to control separately the contributions of the different losses, which have been shown to be independent (Lambourg *et al.*, 2001). In order to evaluate the global and separate influence, we have simulated the hybrid materials by interpolating the two viscoelastic and thermoelastic models that are characteristic of glass and aluminum, respectively. The independence of the two damping models, on the one hand, and the characterization of the materials glass and aluminum with these two models, on the other hand, allows us to study two crucial aspects of this type of vibrating object with the same set of stimuli: the perception of damping and the identification of materials on the basis of this factor.

### 1. Control of the damping parameters

We have extended the synthesis model presented in Lambourg *et al.* (2001) with a complex rigidity [Eq. (1)] that takes into account simultaneously the terms  $\tilde{d}_{it}$  and  $\tilde{d}_{iv}$  by interpolation. We replace  $\tilde{d}_{it}(s) + \tilde{d}_{iv}(s)$  in Eq. (1) by  $\tilde{d}(s)$ , and  $\tilde{d}$  is defined in Table II as a linear interpolation between the

TABLE II. Interpolation and extrapolation functions used to construct the hybrid materials.

Values of $H$	Complex rigidity model	Simulated material
$H < 0$	$\tilde{d} = (1-H)\tilde{d}_{iv}$	Ultra-glass
$H = 0$	$\tilde{d} = \tilde{d}_{iv}$	Glass
$0 < H < 1$	$\tilde{d} = H\tilde{d}_{it} + (1-H)\tilde{d}_{iv}$	Hybrid
$H = 1$	$\tilde{d} = \tilde{d}_{it}$	Aluminum
$H > 1$	$\tilde{d} = H\tilde{d}_{it}$	Ultra-aluminum

two constituent terms. Table II gives the complete function of the hybrid damping model in which the parameter  $H$  is the interpolation factor. This approach also allows us to explore conditions beyond the normal contributions for each material, thus simulating “ultra-glass” ( $H < 0$ ) and “ultra-aluminum” ( $H > 1$ ) by extrapolating beyond the continuum between them. The values of the constants  $R_i$  and  $s_i$  of the interpolated loss models are those taken from the synthesis program and are estimated from physical measures on real glass and aluminum plates (see Lambourg *et al.*, 2001). Their values are given in Table III. The term  $\tilde{d}_{ir}$  is held constant in the interpolation process because it is common to both materials.

To reduce the number of mechanical parameters characteristic of the respective contributions of viscoelastic and thermoelastic losses, we have computed two new parameters derived from the interpolation function ( $HP_\alpha, HR_1$ ):

- The curve representing damping ( $\alpha$ ) as a function of frequency is nearly linear in the lower frequencies for viscoelastic losses. It can thus be approximated by its slope ( $P_\alpha$ ). The contribution of the viscoelastic losses in the interpolation function is approximated by the product of  $H$  and the slope.
- Similarly, the thermoelastic loss model can be characterized by the parameter  $R_1$  defined in Eq. (3). As a mechanical parameter characteristic of thermoelastic loss, we compute in our stimuli  $HR_1$ , a weighting of  $R_1$  by the interpolation factor  $H$ .

Thermoelastic decay and viscoelastic decay behave differently and are perceived differently, as is clear from the perceptual distinction of sounds 6 (aluminum,  $H=1$ ) and 11 (glass,  $H=0$ ) in the stimulus set we studied<sup>1</sup>. Further, they have been formalized, and the physical synthesis model has been developed, in such a way as to allow interpolation be-

TABLE III. Parameters used in the loss models [taken from Eqs. (1)–(5)].

Loss model	Parameter values	
Thermoelastic	$R_1 = 0.0545$	$s_1 = 31\,000 \text{ rad/s}$
Viscoelastic	$R_{v1} = 1.625 \times 10^{-3}$	$s_1 = 5180 \text{ rad/s}$
	$R_{v2} = 1.962 \times 10^{-3}$	$s_2 = 55\,100 \text{ rad/s}$
Radiation	$a_0 = 1.1669$	$b_1 = 0.0620$
	$a_1 = 1.6574$	$b_2 = 0.5950$
	$a_2 = 1.5528$	$b_3 = 1.0272$
	$a_3 = 1.0$	$\rho_a = 1.2 \text{ kg/m}^3$
	$c_0 = 344 \text{ m/s}$	

TABLE IV. Elastic parameters in the models for glass and aluminum plates. The constant term for fluid damping,  $R_f$ , is identical for the two materials.

Parameter	Glass	Aluminum
$D_1$ (MPa)	6700	6500
$D_4$ (MPa)	10 270	8600
$E$ (MPa)	6692	6473
$\rho$ (kg/m <sup>3</sup> )	2550	2660
$c$ (km/s)	1.62	1.56
$R_f$ (/s)	0.4	0.4

tween them. Aluminum and glass are two archetypal materials for these two dissipation phenomena. Therefore, we feel that it is worth enhancing these phenomena artificially with the sound synthesis model by creating “virtual” materials with exaggerated thermoelasticity or viscoelasticity, as well as chimæric combinations of them, in continuity with the real world and in order to evaluate our perception with regard to these mechanical phenomena. This is a classic approach to perception that has been used in speech research, interpolating between acoustical characteristics of different classes of consonants, for example (cf. Liberman and Studdert-Kennedy, 1978).

## 2. Control of other material parameters

The rigidity constants  $D_1$  and  $D_4$  were estimated by a classic approximation based on the measured frequency of the lowest vibrational modes of real plates (Lambourg *et al.*, 2001; McIntyre and Woodhouse, 1988). Starting from these values, the sounds used in the perceptual experiments were synthesized varying the value of the rigidity constants around those estimated in the real plates, so as to obtain hybrid materials in terms of both wave velocity and damping. The plate sounds are much more rich spectrally and more strongly inharmonic than the bar sounds studied in McAdams *et al.* (2004). The other parameters that were manipulated in these stimuli included the elastic modulus  $E$  and the mass density  $\rho$  of the plate. Wave velocity ( $c = \sqrt{E/\rho}$ ) affects  $D_1$  and  $D_4$  and was varied linearly. The decision to vary wave velocity linearly is not based on the perceived pitch of the sounds, but only on a linear distribution of the first modes. The values, as for damping, are distributed around the corresponding values for glass and aluminum materials reported in Table IV.

## 3. Choice of mallet

For the simulated plate sounds, we also wanted to explore the perceptual effect of the mallet material. The hybrid sounds were thus synthesized in two groups, one being struck by a wooden mallet and the other by a rubber mallet. The reason for this choice is that the softer rubber mallet has considerable attenuation of the higher frequencies compared to the wood mallet. The constants introduced in the two excitation models are reported in Table V along with the geometric parameters of the simulated plates. Note that the two excitation models are distinguished by the value of the impact stiffness coefficient,  $K$ . The tasks performed by the listeners were thus systematically repeated in two separate

TABLE V. Parameters of the simulated mallets and plate.

Wood mallet	
$m$	$3 \times 10^{-3}$ kg
$K$	$42.0 \times 10^7$ N m <sup>-3/2</sup>
$V$	1.0 m/s
Rubber mallet	
$m$	$3 \times 10^{-3}$ kg
$K$	$1.2 \times 10^7$ N m <sup>-3/2</sup>
$V$	1.0 m/s
Plate geometry	
Width $l_y$	22 cm
Length $l_x$	23 cm
Thickness $h$	2 cm

blocks of trials, once with each mallet, in order to estimate the effect of mallet hardness on the perception of the plates without having listeners focus on this parameter.

## III. SPECTROTEMPORAL ANALYSIS OF IMPACT SOUNDS

We characterized the sound stimuli in terms of acoustical features by combining and extending previous methods (McAdams *et al.*, 2004; Giordano and McAdams, 2006). We extracted a total of 17 sound features based on the analysis of three different representations: the amplitude envelope, the damping spectrum, and a simulation of the signal transformations taking place in the peripheral auditory system.

The amplitude envelope  $E(t)$  was defined as:

$$E(t) = |x(t) + iH[x(t)]| \quad (7)$$

where  $H$  is the Hilbert transform of the signal  $x(t)$  (Hartmann, 1997, pp. 421–422). Hearing-range amplitude fluctuations were attenuated by forward-reverse low-pass filtering  $E(t)$  using a zero-phase-distortion Butterworth filter (cutoff = 16 Hz).  $E(t)$  was converted to dB relative to the maximum digital amplitude of 1. We extracted four descriptors of the amplitude envelope (Fig. 1):  $\alpha_1$  is the slope of the least-squares line fitted to a 30-ms envelope portion starting from 2.5 ms after the peak level;  $\alpha_2$  is the slope of the least-squares line fitted to the last 1 s of the envelope;  $ED_3$  dB and  $ED_{10}$  dB measure the amount of time the signal exceeds a given level threshold of  $-3$  and  $-10$  dB with respect to the envelope maximum, respectively ( $ED$ =equivalent duration).

The damping spectrum was computed by modeling the signal as a sum of exponentially damped sinusoids with given frequency, phase, starting amplitude, and damping factor  $\alpha$ , the inverse of the time needed for amplitude to decay to  $1/e$  of its initial value. We analyzed the first 150 ms of the signal using the ESPRIT method (Laroche, 1993, N sinusoidal components=1000). We discarded components with a starting amplitude lower than  $-65$  dB from the maximum digital amplitude of 1, with frequency outside the 40–16000 Hz range, and with  $\alpha$  greater than  $200$  s<sup>-1</sup>. Figure 2 shows the damping spectrum for the signals generated by striking the same plate with two different mallets. Except for a lower number of sinusoidal components extracted from the signals created with the rubber mallet, the shapes of the damping

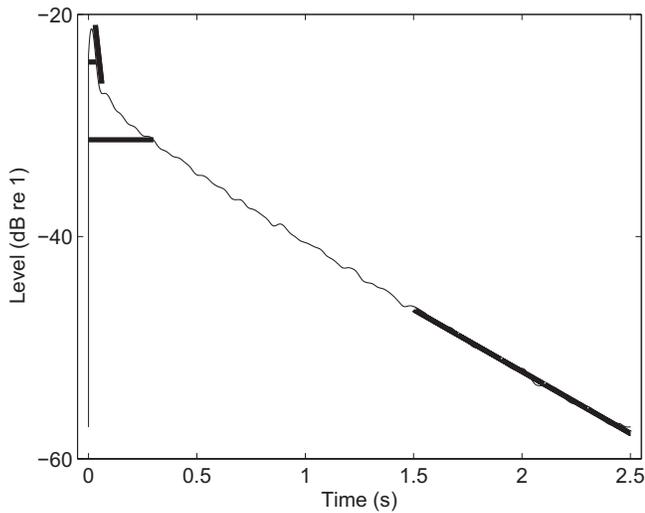


FIG. 1. Amplitude envelope for plate number 6 (aluminum reference) struck with a wood mallet. The slanted solid lines show the least-squares function used to extract measures of decay:  $\alpha_1$  (upper left) and  $\alpha_2$  (lower right). The horizontal solid lines show the equivalent durations  $ED_{3 \text{ dB}}$  (above) and  $ED_{10 \text{ dB}}$  (below). For display purposes, the level at time 0 has been set to the end value.

spectra were largely unaffected by variations in the mallet. For this reason, further modeling used ESPRIT data collected for the rubber and wood mallets striking the same plate. The resulting damping spectrum was thus modeled as

$$\alpha(f) = \frac{p_a f^3}{f^3 - p_b f + \frac{2}{3} p_b f_c} \quad (8)$$

where  $f$  is the frequency of the sinusoidal component and  $p_a$ ,  $p_b$  and  $f_c$  are model parameters estimated with a nonlinear least-squares fitting technique. Finally, the damping-spectrum descriptors were  $p_a$  in  $s^{-1}$ ,  $p_b$  in  $s^{-2}$ ,  $f_c$  in Hz, and  $f_1$  (the frequency of the first sinusoidal component extracted from the ESPRIT representation) in Hz. Note that  $p_a$  mea-

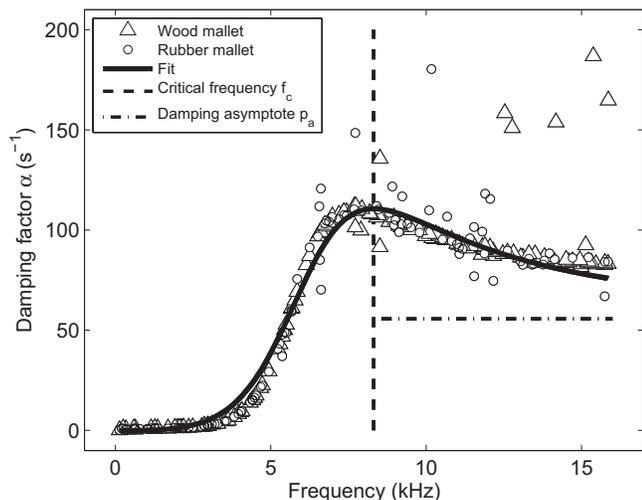


FIG. 2. Damping spectrum for plate number 6 (aluminum reference) struck with a wood mallet. The continuous line shows the modeled damping spectrum. The dashed vertical line represents the critical frequency,  $f_c$ , of the plate estimated from the model. The horizontal dashed-dotted line indicates the estimated damping asymptote,  $p_a$ .

asures  $\alpha$  as frequency asymptotically approaches infinity, and  $f_c$  corresponds to the critical frequency of the plate (in Hz) as estimated from our model of the ESPRIT data. Further, within the sample of investigated stimuli,  $p_b$  has the strongest influence on the extent to which damping factors decrease progressively at frequencies removed from  $f_c$ : faster decreases in damping are observed with increasing values of  $p_b$ ; for example, glass would have a higher value of  $p_b$  than would aluminum.

A final set of descriptors was extracted from a simulation of the signal processing taking place in the peripheral auditory system (McAdams *et al.*, 2004). The simulation produced the time-varying power at the output of a set of cochlear filters (sampling rate: 44.1 kHz). The center frequencies of the cochlear filters (range: 30–16 000 Hz) were uniformly spaced on a frequency scale derived from measures of masked detection thresholds in normal-hearing listeners: the ERB-rate scale (Moore and Glasberg, 1983) reflects the spacing of auditory filters along the basilar membrane. A first descriptor extracted from this representation characterizes the damping spectrum. Wildes and Richards (1988) proposed  $\tan \phi = \alpha / (\pi f)$ , which is essentially the slope of the damping spectrum, as a unitless measure of the damping of vibration in struck solids. In our case, a similar measure  $\tan \phi_{\text{aud}}$  is computed on the output of the cochlear filters rather than on the damping spectrum itself in order to take into account basic properties of peripheral auditory processing (see Giordano and McAdams, 2006; Giordano *et al.*, 2010, for discussion of a strong association between this descriptor and human perception of material properties and for computational details). Intuitively, the higher  $\tan \phi_{\text{aud}}$ , the faster the relative energy decay in cochlear filters with higher center frequencies. More specifically,  $\tan \phi_{\text{aud}}$  is given by

$$\tan \phi_{\text{aud}} = \frac{\sum_i \frac{\alpha_i}{\pi F_i} p_i}{\sum_i p_i} \quad (9)$$

where  $\alpha_i$  is the damping factor for the power output at the  $i$ th cochlear filter, as estimated with linear regression methods,  $F_i$  is the center frequency of the  $i$ th cochlear filter in Hertz, and  $p_i$  is the total power at the output of the  $i$ th cochlear filter (see Fig. 3).

The temporal resolution of the output at each cochlear filter was subsequently decreased to 100 Hz (10 ms  $\approx$  auditory window of temporal integration in humans, Plack and Moore, 1990) and raised to the power of 0.25 to yield an approximate measure of the specific loudness within each cochlear filter (Hartmann, 1997, p. 66). Finally, the time-varying loudness and spectral center of gravity (SCG) were defined as the sum of the specific loudnesses and as the specific-loudness-weighted average frequency on the ERB-rate scale, respectively (see Fig. 4). SCG captures the auditory attribute of brightness (Grey and Gordon, 1978; McAdams *et al.*, 1995). The measurement unit for loudness was termed pseudo-sones (p.s.), because it was calculated without taking into account the actual presentation levels. Four loudness descriptors were derived: the attack value,  $Lou_{\text{att}}$ , the loudness of the first 10 ms of the signal; the mean value

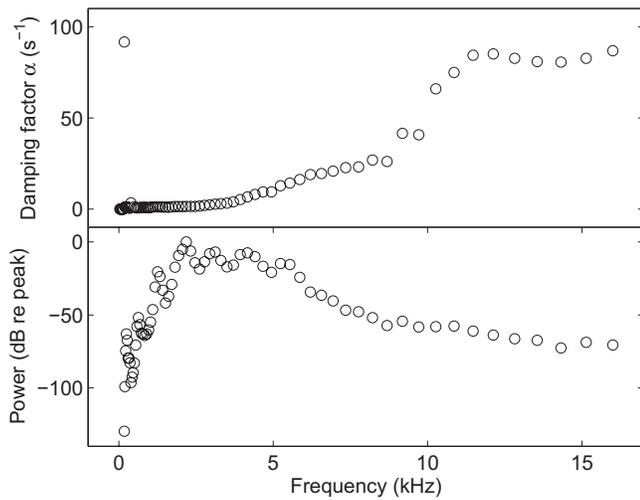


FIG. 3. Damping factors (upper panel) and total power (lower panel) at the output of the cochlear filters extracted from a simulation of the processing taking place in the peripheral auditory system as a function of the filter center frequencies for plate number 6 struck with a wood mallet.

over the whole duration,  $Lou_{mea}$ , and the slope of the initial and final portions of the temporal function of loudness,  $Lou_{sl1}$  and  $Lou_{sl2}$ , respectively. Three SCG-related descriptors were extracted: the attack value  $SCG_{att}$ , the SCG of the first 10 ms of the signal; the mean value over the whole duration,  $SCG_{mea}$ , and the slope of the initial portion of the temporal function of SCG,  $SCG_{slo}$ . The slope measures were extracted by means of linear regression over a portion of the temporal function (see Giordano and McAdams, 2006, for details). Finally, the effective duration of the signal,  $Dur$ , was defined as the temporal extent over which the loudness exceeded a fixed threshold of 0.2 pseudo-sones.

#### IV. EXPERIMENT 1: DISSIMILARITY RATINGS

##### A. Methods

##### 1. Subjects

Twenty listeners participated in the experiment (9 females, 11 males; aged 20–32 years;  $M=24$  years). They

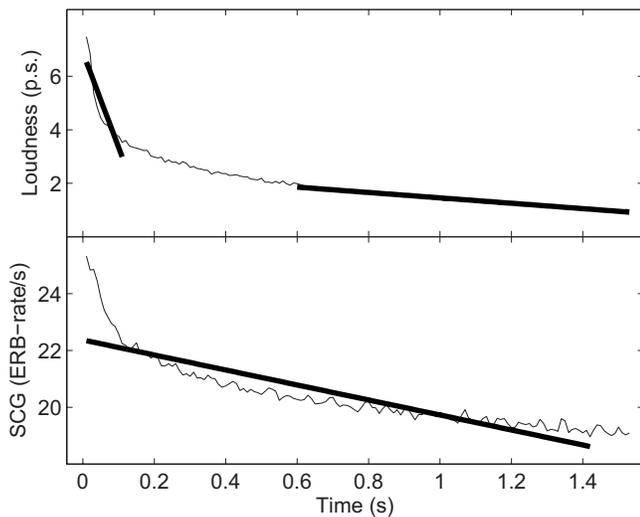


FIG. 4. Time-varying loudness and spectral center of gravity (SCG) for plate 6 struck with a wood mallet. The solid black lines show the least-squares regression used to derive the slope measures ( $Lou_{sl1}$ ,  $Lou_{sl2}$  in the upper panel,  $SCG_{slo}$  in the lower panel).

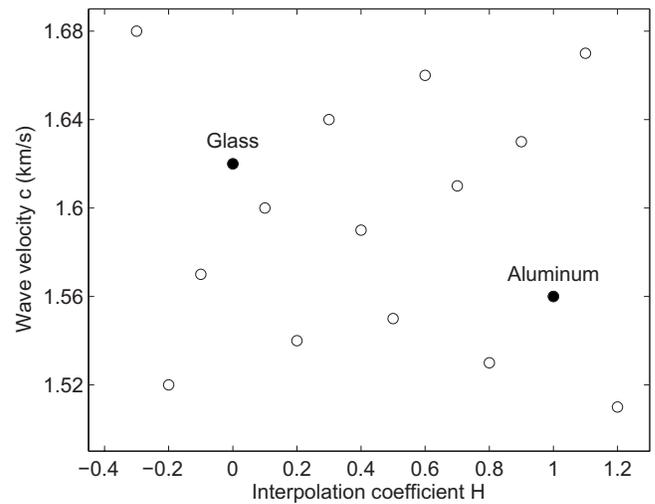


FIG. 5. Representation of the stimulus space in terms of the mechanical parameters  $H$  (interpolation coefficient between viscoelastic and thermoelastic damping) and  $c$  (wave velocity). The reference values for aluminum and glass are shown.

were paid for their participation. All listeners reported having normal hearing. None of them was a professional musician.

##### 2. Apparatus

Listeners were seated in a Soluna S1 double-walled sound booth. Sounds were generated on a NeXT workstation equipped with an ISPW card and running the Max sound-synthesis environment (Lindemann *et al.*, 1991). The PsiExp experimental programming environment (Smith, 1995) was used to present stimuli and collect responses. Sounds were converted with a ProPort digital-to-analog converter before being amplified with a Canford stereo amplifier and presented diotically (same signal to both ears) over AKG-1000 headphones at an average level of 70 dB SPL as measured with a Bruel & Kjaer 2209 sound level meter (A-weighting, fast response).

##### 3. Stimuli

Two sets of 16 sounds were synthesized simulating a plate being struck by either a wooden or rubber mallet (hereafter referred to as wood and rubber sets).<sup>1</sup> The stimuli varied in two properties of the plate: its wave velocity and its damping characteristics [interpolation between viscoelastic (glass) and thermoelastic (aluminum) conditions] (see Table II). The stimulus set was chosen so that across the 16 sounds, 16 values of  $c$  and  $H$  were included, and the stimuli were homogeneously distributed in a space represented by these two dimensions (Fig. 5). The same set of  $H$  and  $c$  values was used for the wood and rubber sound sets. Stimuli were sampled at 48 kHz with 16-bit resolution.

##### 4. Procedure

Listeners rated the dissimilarity of pairs of sounds by moving a slider along an on-screen scale marked “very similar” and “very dissimilar” at the two extremes. Listeners were allowed to replay the pair of sounds as many times as needed before entering their final response. They were in-

structured to use the full scale over the set of sound pairs. The wood-mallet and rubber-mallet sound sets were separately evaluated in two experimental sessions with set order being counterbalanced across participants. Within each of the experimental sessions, listeners evaluated all 120 possible different-sound pairs and 16 same-sound pairs. The order of presentation of all pairs and of the sounds within a pair were randomized. Each session took about 30 min to complete.

## B. Results

For each of the participants, and for both sound sets, the average same-sound dissimilarity was significantly lower than the average different-sound dissimilarity [ $M=0.014$  and  $0.505$ ,  $SD=0.044$  and  $0.282$ , respectively, unpaired  $t(134) \leq -10.24$ ,  $p < 0.001$ , equal variance not assumed]. As such, we conclude that participants correctly understood the use of the rating scale.

Subsequent analyses focused on the different-sound pairs. Our aim was to determine the perceptual representation of the experimental sounds and, in particular, how the parameters of the sound-synthesis model are mapped onto perception. We therefore fitted a distance model to the perceptual dissimilarities using the multidimensional scaling (MDS) algorithm CLASCAL (Winsberg and De Soete, 1993; McAdams et al., 1995). CLASCAL models dissimilarities as distances within a Euclidean space that is common to all of the sounds, plus distances along dimensions specific to each of the sounds (called “specificities”). Each dimension of the Euclidean space and the set of specificities are weighted differently by separate latent classes of listeners, which are also estimated by the algorithm. The weights are considered to estimate the relative perceptual salience of each dimension and the whole set of specificities for a given class of listeners with similar data patterns. The distance model used in CLASCAL and the steps taken to select the best model are described in McAdams et al. (1995, 2004).

For the wood-mallet set, CLASCAL selected a model with three dimensions without specificities and two latent classes of participants. For the rubber-mallet set, the model included two dimensions without specificities and two latent classes of participants. Both models accounted well for the observed data ( $R^2$  between model distances and average dissimilarities  $\geq 0.85$ ). Despite the fact that the two MDS models were estimated from dissimilarities generated with different sound stimuli, the two solutions were similar in structure (correlations between the first and second dimensions of the two models,  $r(14) \geq 0.97$ ,  $p < 0.001$ ; correlation between rubber and wood model distances,  $r(118) = 0.94$ ,  $p < 0.001$ ).

We investigated the mechanical nature of the MDS dimensions, regressing the parameters of the physical model onto the coordinates of the sound stimuli along each of the dimensions of the common Euclidean spaces ( $df=14$  in all cases). In particular, for each of the MDS dimensions we fitted two univariate rank regression models (Iman and Conover, 1979) with either  $H$  or  $c$  as predictors. For both the wood and rubber sets, Dim1 was well predicted by the damping-related parameter  $H$  ( $R^2 > 0.99$ ,  $p < 0.001$ ) and was independent of the wave-velocity parameter  $c$  ( $|R^2|$

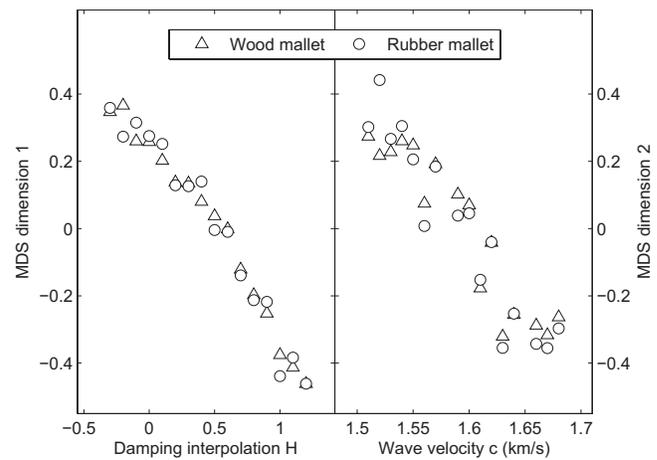


FIG. 6. Association between the dimensions of the MDS models and the parameters of the physical model:  $H$  and Dimension 1 on the left,  $c$  and Dimension 2 on the right.

$< 0.04$ ,  $p \geq 0.54$ ), whereas Dim2 was independent of  $H$  ( $|R^2| < 0.10$ ,  $p > 0.41$ ), but was well accounted for by  $c$  ( $R^2 > 0.89$ ,  $p < 0.001$ ). Dim3 of the wood-mallet MDS model was independent of both  $H$  and  $c$  ( $R^2 \leq 0.07$ ,  $p \geq 0.43$ ). We are unable to explain the nature of this third dimension. Figure 6 shows Dimensions 1 and 2 for both the wood and rubber sets as a function of the sound synthesis parameters  $H$  and  $c$ , respectively.

Finally, listeners were assigned to latent classes by a Bayesian *post hoc* analysis (Winsberg and De Soete, 1993). For the wood set, the two latent classes contained twelve and eight participants, respectively. For the rubber set, the first and second latent classes contained eleven and nine participants, respectively. The distribution of participants within latent classes remained fairly constant across the two sound sets: only three of the 20 participants were assigned to different latent classes in the wood and rubber sound sets. The agreement between the partitioning of listeners in latent classes across sound sets is higher than would be expected by chance (adjusted Rand index = 0.46; Hubert and Arabie, 1985).

Within the MDS framework, the range of variation of the stimulus coordinates along the different dimensions estimates the relative weight accorded to each dimension in rating between-stimulus dissimilarities for each latent class of listeners (the greater the range, the more salient a dimension is). Independently of the particular stimulus set (wood vs. rubber), participants in Class 1 focused primarily on Dim1 (ranges: wood = 0.96; rubber = 0.73) and attributed a secondary relevance to the other dimension(s) (ranges: wood, Dim2 = 0.31, Dim3 = 0.13; rubber, Dim2 = 0.40). Also independently of the particular stimulus set, participants assigned to Class 2 focused primarily on Dim2 (ranges: wood = 0.88; rubber = 1.20) and attributed a secondary relevance to Dim1 (ranges: wood = 0.69; rubber = 0.91), although Dim3 of the wood set had a similar weight to Dim2 (0.85).

## C. Summary

Stimuli generated with parameters related to two mechanical properties of the simulated vibrating objects gave

rise to two primary perceptual dimensions. One dimension corresponded to variations in wave velocity ( $c$ ), most likely associated with the perceived pitch derived from the modal frequencies of the plate that are determined in part by this mechanical property. The other dimension corresponded to variations in damping properties, according to an interpolation function ( $H$ ) between viscoelastic and thermoelastic damping, most likely associated with timbre and duration properties. An analysis of the acoustical nature of these dimensions will be addressed in Sec. VI. Latent-class analysis revealed two classes of listeners with opposite patterns of weights on the two dimensions, one class being more sensitive to the wave-velocity-related perceptual dimension and the other more sensitive to the damping-related dimension. There was a marked similarity in perceptual ratings between the wood and rubber sound sets, suggesting that the perception of the plate properties was independent of those of the mallet.

## V. EXPERIMENT 2: MATERIAL IDENTIFICATION

In the second experiment, we focused on a different perceptual task applied to the same sound sets: the identification (or categorization) of the material of the struck object. We were particularly interested in how the results from both experiments would compare across the two tasks. Lutfi and Liu (2007) have argued for participant-specific fingerprints in the weighting of acoustical information, independent of the task requested of the listener. Accordingly, one would expect participants who estimated dissimilarities with a particular focus on the wave-velocity-related acoustical properties to still focus on these properties when identifying the material of the sounding object. Similarly, listeners who focused more on damping-related properties in the dissimilarity task should also preferentially base their identifications on these properties.

### A. Method

At the end of Experiment 1, the same listeners participated in Experiment 2. They were presented the same sounds as in Experiment 1, one at a time, and were asked to identify the material of the struck object as being made of glass or aluminum by clicking on the appropriate button on the screen. They were further instructed to carry out the task ignoring the properties of the mallet. This instruction is justified because we are primarily interested in the perception of the material properties of the plates and how they generalize across different mallets, rather than in the perception of the mallet properties per se.

Prior to beginning the main part of the experiment, the participants were presented with a demonstration of real aluminum and glass plates that were struck with wood and rubber mallets while suspended by nylon threads. The plates had the same geometrical and material characteristics as reference stimuli 6 and 11, respectively, in the synthesized stimulus set. They were struck several times by the experimenter inside the sound booth. The presentation of the real plate sounds was meant to familiarize the listeners with the sound of free-vibrating plates of both materials.

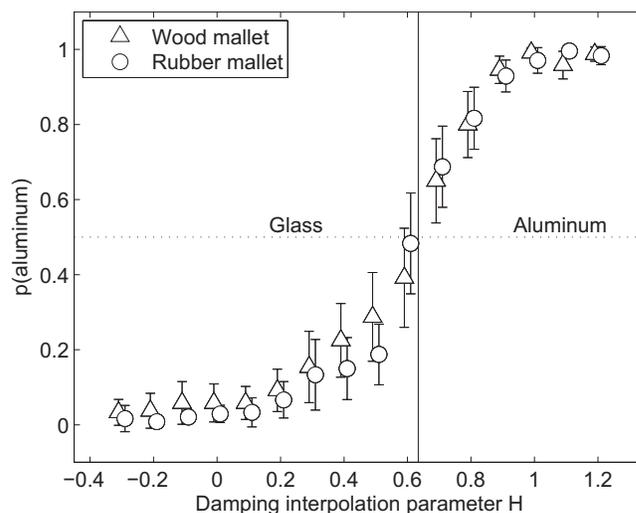


FIG. 7. Population probability of the “aluminum” response as a function of the damping interpolation parameter  $H$ . Error bars bracket 95% confidence intervals for the “aluminum” response distribution across participants. The vertical line represents the value of  $H$  at which “aluminum” and “glass” responses are equiprobable, corresponding to the average of the estimate obtained from the rank-regression modeling of participant-specific data, with  $H$  as predictor. Data for the wood and rubber sets have been displaced slightly along the abscissa for ease of visualization.

Synthesized sounds simulating strikes with either the wood or rubber mallet were presented in separate blocks. The order of the blocks was counterbalanced across subjects. During each block, each of the 16 sounds was presented 12 times in random order for a total of 192 trials. Each block took about 15 min to complete.

## B. Results and discussion

The influence of the parameters of the physical model on the auditory identification of the plate material was assessed by means of rank regression models ( $df=14$  in all cases). The advantage of rank regression is that it is insensitive to the form of the relation between dependent and independent variables (which is often sigmoidal in identification functions such as those we are examining, as can be seen in Fig. 7). We used univariate rank regression to test the extent to which the sound synthesis parameters accounted for the proportion of “aluminum” responses. The analysis was performed independently for each of the participants with data from the wood and rubber sound sets. For all of the listeners and for both sound sets, identification responses were independent of the wave velocity parameter  $c$  ( $R^2 \leq 0.05, p \geq 0.17$ ). However, the damping interpolation parameter  $H$  strongly influenced identification ( $R^2 \geq 0.54, p \leq 0.004$ ) and accounted well for the observed proportions of “aluminum” responses (across-participant average  $R^2=0.92$  and  $0.93$ ,  $SD=0.10$  and  $0.05$  for the wood and rubber sets, respectively). Figure 7 shows the proportion of “aluminum” responses averaged across participants for both the wood and rubber sets as a function of the damping interpolation factor,  $H$ . Note once again that the results are unaffected by the mallet material.

In the dissimilarity rating task of Experiment 1, listeners’ ratings revealed a sensitivity to both damping and wave-velocity variations, even though the relative weights of these

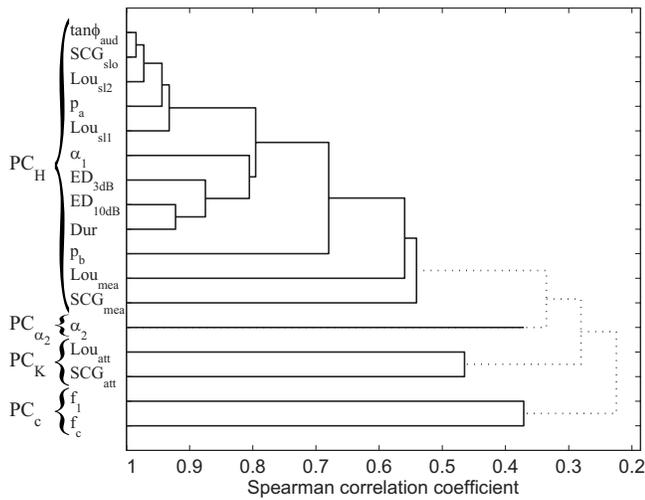


FIG. 8. Hierarchical clustering analysis of the absolute values of rank correlation coefficients among acoustical descriptors (average linkage). Descriptors that join further to the left are more highly correlated. Solid lines connect strongly correlated descriptors that were reduced to their principal component (four clusters).

sources of information varied across classes of listeners. In this experiment, to the contrary, listeners overwhelmingly used damping information and ignored wave-velocity information in performing the material identification task, independently of the latent class to which they belonged in Experiment 1. In the next section, we quantify the acoustical information used by each of the participants in Experiments 1 and 2 to carry out the perceptual tasks. These measures are further analyzed to assess the extent to which variations in the task and sound set influence the weighting strategies used by a given participant.

## VI. STRATEGIES IN THE WEIGHTING OF ACOUSTICAL INFORMATION

We quantified the perceptual weight given by each of the participants to the different acoustical features<sup>2</sup> in each sound set for each task using a multivariate rank regression framework. For data from Experiment 1, the dependent variable was the dissimilarity rating for the different-sound pairs, and the independent variables were the absolute values of the differences in the acoustical features for the paired sounds. For data from Experiment 2, the dependent variable was the probability of choosing the “aluminum” response, and the independent variables were the acoustical features.

Several acoustical descriptors were strongly correlated with each other, within both the wood and rubber sets. Figure 8 presents a hierarchical clustering analysis of the correlations between pairs of descriptors across the wood and rubber sound sets. Strong correlations among predictors are problematic for regression modeling, because they cause inaccuracies in the estimates of the regression coefficients. It also becomes impossible to decide which among a set of highly correlated descriptors is the “right” one. We therefore took an agnostic approach to this problem by reducing groups of correlated descriptors to a single variable, following the methodology presented in Giordano *et al.* (2010). First, descriptors were correlated using rank correlation in order to avoid issues arising from the nonlinear relation be-

tween the correlated variables. Rank correlation is an ordinal data analysis technique that linearizes nonlinear monotonic relations between variables. Second, hierarchical clustering analysis was used to determine which groups of descriptors were highly correlated. Correlated descriptors were then reduced to a single variable by means of a Principal Component Analysis (PCA) carried out on the ranks of the acoustical variables. For each of the clusters, we then applied PCA on the rank-transformed values of the descriptors and used the first principal component (PC) as the final reducing variable for all the descriptors within a cluster. The choice of which descriptors had to be reduced to the same PC was based on an agglomerative hierarchical cluster analysis (average linkage) carried out on a matrix of distance measures defined as one minus the absolute value of the Spearman rank correlation coefficient ( $|\rho_S|$ ) between each pair of acoustical descriptors. To avoid confusion, we will use the notation  $\rho_S$  in this article to distinguish this statistic from the symbol for mass density. Starting from the condition in which each of the descriptors was in a separate cluster, the number of clusters was decreased by one at each succeeding step. For each of the steps, descriptors belonging to the same cluster were reduced to a single PC, and the correlations among the different reduced variables were computed. The procedure was terminated when the reduced variables were uncorrelated with each other, i.e., when  $|\rho_S| \leq 0.6$ . This analysis revealed four predictor variables that were independent of one another with respect to our stimulus sets. Each one captured the shared aspects of a group of highly intercorrelated acoustical descriptors.

The raw acoustical descriptors used to predict the identification data set and the absolute paired differences used to predict the dissimilarity-rating data set were independently reduced to their PCs for both the wood and rubber sets. However, the data reduction procedure was constrained so that it yielded the same number of PCs for each of the data sets and the same partitioning of acoustical descriptors into PCs (labeled  $PC_H$ ,  $PC_{\alpha_2}$ ,  $PC_c$ , and  $PC_K$  in Table VI). This constraint was necessary to ensure that the acoustical models for perceptual data were comparable across the data sets and sound sets. Indeed, it would have been problematic otherwise to compare the behavioral weight of a PC that reduced the acoustical variables A and B in one data set with that of a PC that reduced the acoustical variables A and C in another data set. To this purpose, the agglomerative hierarchical clustering solution that guided the data-reduction process was computed for the absolute between-descriptors correlation averaged across data sets. Note, however, that throughout the data-reduction process, the PCs and the correlations between PCs were computed independently for each of the data sets.

The data-reduction procedure yielded four independent reducing PCs for each of the four data sets (two mallets in two tasks). The maximum and average  $|\rho_S|$  values between PCs across the four data sets are 0.54 and 0.23, respectively. Overall, the PCs accounted well for the original acoustical descriptors (grand average  $|\rho_S|$  between acoustical descriptors and respective PCs=0.86, SD=0.16; see Table VI).<sup>3</sup> A first PC included 12 parameters related to the effects of energy decay over the course of the sound stimulus on the

TABLE VI. Acoustical predictors that were included in the derivation of the four main principal component (PC) predictors and their individual Spearman rank correlations with the resulting PC predictor.

Predictor	PC	$\rho_S$ (wood)	$\rho_S$ (rubber)
$\alpha_1$	$PC_H$	-0.92	-0.96
$ED_3$ dB	$PC_H$	-0.93	-0.99
$ED_{10}$ dB	$PC_H$	-0.95	-0.99
$p_a$	$PC_H$	0.96	0.99
$p_b$	$PC_H$	-0.88	-0.90
$\tan \phi_{aud}$	$PC_H$	0.99	1.00
$Lou_{mea}$	$PC_H$	0.93	-0.46
$Lou_{s11}$	$PC_H$	-0.96	-0.99
$Lou_{s12}$	$PC_H$	-0.99	-0.99
$SCG_{mea}$	$PC_H$	0.56	-0.86
$SCG_{slo}$	$PC_H$	-0.99	-1.00
$Dur$	$PC_H$	-0.99	-1.00
$\alpha_2$	$PC_{\alpha_2}$	1.00	1.00
$f_1$	$PC_c$	0.89	0.89
$f_c$	$PC_c$	-0.83	-0.90
$Lou_{att}$	$PC_K$	-0.81	-0.89
$SCG_{att}$	$PC_K$	-0.84	-0.91

perception of duration, loudness and timbre. This PC was the one that most strongly correlated with the damping interpolation parameter  $H$  for all data sets ( $|\rho_S| \geq 0.97$ ,  $p < 0.001$ ;  $|\rho_S| \leq 0.53$  for the other PCs). It will thus be referred to as  $PC_H$ . A second PC included two frequency-related parameters,  $f_1$  and  $f_c$ . The primary pitch of the plate sound is determined by  $f_1$ . Because for all data sets this PC was the one that most strongly correlated with the wave-velocity parameter  $c$  ( $\rho_S \geq 0.71$ ,  $p < 0.001$ ;  $|\rho_S| \leq 0.25$  for the other PCs), it will be referred to as  $PC_c$ . A third PC included only one acoustical descriptor,  $\alpha_2$ , modeling the rate of decay in the last portion of the amplitude envelope and will be referred to as  $PC_{\alpha_2}$ . The last PC included two descriptors related to the attack of the signal,  $SCG_{att}$  and  $Lou_{att}$ . These very same attack parameters are strongly associated with the impact-stiffness parameter  $K$  in this study and are considered accurate acoustical specifiers of the properties of the mallet/plate interaction and, to a lesser degree, of the mallet material, respectively (Giordano *et al.*, 2010). This last PC will therefore be referred to as  $PC_K$ .

We then quantified the extent to which each of the participants focused on each of the PCs when carrying out either the dissimilarity ratings or the material identification tasks with both the wood and rubber sound sets. We used multivariate rank regression and predicted behavioral data from each of the four experimental conditions using all of the acoustical PCs as predictors. We adopted the partial  $R^2$  ( $R_p^2$ ) for each of the PCs as a measure of the perceptual weight.  $R_p^2$  for a variable A measures the ratio of the gain in the proportion of explained variance when A is included in the set of all predictors to the variance left unexplained when A is not in the regression model.  $R_p^2$  equals one when the full model explains 100% of the variance in the dependent variable and can assume a negative value if the quality of the prediction increases when A is taken out of the model (negative  $R_p^2$  values may be observed for predictors that are independent of the modeled outcome).  $R_p^2$  values were not computed from

the prediction of the ranks, but from the prediction of the untransformed data based on the rank model (cf. Iman and Conover, 1979).<sup>4</sup>

Consistently with the statistical analyses presented in Sec. IV, the majority of the participants in the dissimilarity-rating experiment showed significant effects of  $PC_H$ , but were also influenced by  $PC_c$ . Independently of the particular sound set, around half of the participants rated dissimilarities focusing on  $PC_H$  (55% for the wood set; 50% for the rubber set), whereas the other half focused on  $PC_c$  (45% for the wood set; 50% for the rubber set). Finally, a very low percentage of participants showed significant effects of either  $PC_K$  or  $PC_{\alpha_2}$ . The totality of the participants in Experiment 2 identified the material of the plate focusing on  $PC_H$  for both the wood and rubber sets, whereas only a few individuals appeared to take into account  $PC_c$  or  $PC_K$  (5 and 10% of the participants across the rubber and wood sets, respectively). It is interesting to note that around 50% of the participants in the material identification experiment showed significant but secondary effects of  $PC_{\alpha_2}$ ,  $\alpha_2$  being moderately correlated [ $\rho_S(14) = -0.50$ ,  $p < 0.05$ ] with the damping interpolation parameter  $H$ .

Finally, we analyzed the extent to which the sound set and the experimental task affected the perceptual weighting of acoustical information. In particular, we constructed a  $2 \times 2 \times 4$  repeated-measures analysis of variance (ANOVA) model with the  $R_p^2$  measures of perceptual weight as dependent variable and with task (dissimilarity ratings, identification), sound set (wood, rubber) and acoustical PC ( $PC_H$ ,  $PC_c$ ,  $PC_{\alpha_2}$ ,  $PC_K$ ) as within-subjects factors. Violations of sphericity were corrected with the Greenhouse-Geisser  $\epsilon$  where appropriate. We argued that if weighting strategies depended on task or sound set, the ANOVA model would show a significant task  $\times$  PC or sound set  $\times$  PC interaction, respectively. The three-way interaction between the within-subject factors was not significant [ $F(3, 57) < 1$ ]. Neither the effect of sound set [ $F(1, 19) < 1$ ] nor the sound set  $\times$  PC interaction [ $F(3, 57) = 2.19$ ,  $\epsilon = 0.850$ ,  $p = 0.11$ , respectively] were significant. This result suggests that the weights did not vary as a function of the effects of the two mallets on the sounds. The higher weights for identification than for dissimilarity ratings suggest that the rank regression model accounts better for identification than for dissimilarity-rating data [ $F(1, 19) = 48.15$ ,  $p < 0.001$ ]. The main effect of PC was statistically significant [ $F(3, 57) = 85.01$ ,  $\epsilon = 0.707$ ,  $p < 0.001$ ], but most importantly, the task  $\times$  PC interaction was highly significant [ $F(3, 57) = 43.50$ ,  $\epsilon = 0.730$ ,  $p < 0.001$ ]; although participants rated dissimilarities focusing on acoustical parameters represented by  $PC_H$  and  $PC_c$ , they identified the materials focusing strongly on  $PC_H$  and secondarily on  $PC_{\alpha_2}$  (Fig. 9). This pattern holds for both sound sets. Overall, participants kept a constant weighting strategy across sound sets, but modified it significantly across tasks.

Changes in wave velocity create changes in the modal frequencies present in the resulting sound. These result in pitch changes for a listener. Changes in plate size would also cause changes in modal frequencies, which would in turn result in pitch changes for a listener. If the two were played

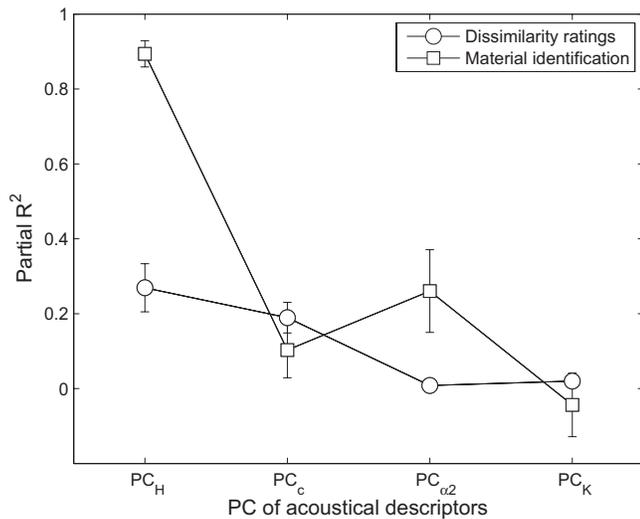


FIG. 9. Influence of task on the perceptual weighting ( $R_p^2$ ) of the principal components (PC) of acoustical descriptors. Error bars bracket 95% confidence intervals around the mean.

against one another so that important frequencies were held constant, logically the listener would hear no pitch changes, and would thus be insensitive to changes in elasticity that were compensated for by changes in geometry, suggesting (in theory anyway) that there are an infinite number of combinations of certain elasticity and geometrical properties that would give indistinguishable sounds. Pitch cues are thus not reliable sources of change in material, because they could also be due to changes in geometry, including size.

## VII. CONCLUSIONS

This study examined the perceptual role of variations in two sound-source properties related to their material composition (damping and wave velocity) in two different perceptual tasks (dissimilarity ratings and material identification). Two primary perceptual dimensions were recovered from dissimilarity ratings in the set of stimuli generated with two parameters related to the mechanical properties of the simulated vibrating objects. One dimension corresponded to variations in wave velocity ( $c$ ), which create differences in the frequency content of the signal that are heard as pitch differences. The other dimension corresponded to variations in damping properties according to an interpolation function between viscoelastic and thermoelastic damping. This variation creates changes that are related to timbre and duration perception and includes amplitude-envelope characteristics and differences in the frequency-damping curves, which affect the rate of change of the spectral centroid. A latent-class analysis revealed two classes of listeners with opposite patterns of weighting on the two primary dimensions, one class being more sensitive to the wave-velocity-related perceptual dimension and the other to the damping-related dimension.

We used a recently developed technique for relating acoustical descriptors to the perceptual dimensions (Giordano *et al.*, 2010). A set of highly intercorrelated candidate acoustical descriptors was reduced to a single principal component, which, in turn, was considered to capture the common behavior of the set, without obliging us to select arbitrarily a single descriptor (or combination of descriptors)

from the set (e.g., is damping better characterized by the rational function estimating the damping spectrum parameters or by  $\tan \phi_{aud}$ ?). This agnostic approach allowed us to interpret the perceptual dimensions as potentially capturing a complex behavior in physical terms that is itself the result of several related properties. The advantage is that many such parameters operate in concert in the physical world. The disadvantage at this stage in the study of psychomechanical relations is that we cannot estimate the relative contributions of the individual descriptors or the nature of their potential interactions. What emerges clearly, however, from this analysis is that the descriptors related to wave velocity and damping are not of the same family and are independent of one another.

There was a marked similarity in perceptual ratings between the wood and rubber sound sets. This result suggests that in spite of the low-pass filtering effect of the rubber mallet compared to the wood mallet, an acoustical fingerprint carrying information related to the material properties (damping and wave velocity) remained invariant through variations in other mechanical sources of the sound event (mallet hardness).

In the material identification experiment, we again found an independence in the perceptual data from the material of the mallets used to strike the plates, suggesting that this task can also be performed independently of that source of mechanical variation, at least when the stimuli are blocked according to mallet material type. We also found that listeners' perceptual weights on the potential acoustical sources of information used to perform the task were very different compared to those in the dissimilarity-rating task. Listeners essentially ignored the pitch information resulting from differences in wave velocity and focused exclusively on perceptual attributes resulting from differences in damping properties.

The material identification results appear to be at odds with those reported by Klatzky *et al.* (2000) and Giordano *et al.* (2010). In the current study, we found virtually no effect of the frequency of the signal on identification responses, whereas participants in the above-mentioned studies clearly took this feature into account. As Giordano and McAdams (2006) have observed, results from a source-identification study depend not only on the criteria for perception that a participant brings to the experimental context from everyday life experience, but also on the acoustical properties of the stimuli to be judged, and in particular on their range of variation within the experimental context, which determines their perceptual relevance.

Taken together, the results from these two experiments demonstrate that perceptual evaluation is robust. In other words, a given perceptual judgment can be carried out more or less consistently on a given set of vibrating objects, independently of the material of the exciter that sets the objects into vibration. To the contrary, different perceptual evaluations (dissimilarity estimation and identification) carried out on the same sound set yield different results in terms of the acoustical information that is used to perform the task. Listeners seem to select acoustical information that is relevant to and reliable for a given perceptual task. Although both

wave-velocity and damping properties are relevant to dissimilarity ratings, only the damping properties are reliable for material identification. This preponderance of damping in identification may be because the timbral and durational properties that result from damping are only affected by changes in material, whereas pitch-related properties that accompany changes in various material properties such as density, elasticity or their combination in wave velocity, can also be created by changes in geometry. Damping properties are thus more reliable for material categorization.

## ACKNOWLEDGMENTS

This work was supported by a grant from the ACI Sciences Cognitives of the French Ministry of Research to S.Mc. and A.C., and grants from the Natural Sciences and Engineering Research Council of Canada and the Canada Research Chair program to S.Mc.

<sup>1</sup>See supplementary material at <http://dx.doi.org/10.1121/1.3466867> Document No. E-JASMAN-128-022009 for audio files in .wav format for all the sound stimuli. For more information, see <http://www.aip.org/pubservs/epaps/html>.

<sup>2</sup>See supplementary material at <http://dx.doi.org/10.1121/1.3466867> Document No. E-JASMAN-128-022009 for Table S1. For more information, see <http://www.aip.org/pubservs/epaps/html>.

<sup>3</sup>See supplementary material at <http://dx.doi.org/10.1121/1.3466867> Document No. E-JASMAN-128-022009 for Table S2. For more information, see <http://www.aip.org/pubservs/epaps/html>.

<sup>4</sup>See supplementary material at <http://dx.doi.org/10.1121/1.3466867> Document No. E-JASMAN-128-022009 for Table S3. For more information, see <http://www.aip.org/pubservs/epaps/html>.

Avanzini, F., and Rocchesso, D. (2001). "Controlling material properties in physical models of sounding objects," in Proceedings of the International Computer Music Conference, La Habana, pp. 91–94.

Cabe, P. A., and Pittenger, J. B. (2000). "Human sensitivity to acoustic information from vessel filling," *J. Exp. Psychol. Hum. Percept. Perform.* **26**, 313–324.

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

Chaigne, A., and Lambourg, C. (2001). "Time-domain simulation of damped impacted plates: I. Theory and experiments," *J. Acoust. Soc. Am.* **109**, 1422–1432.

Doutaut, V., Matignon, D., and Chaigne, A. (1998). "Numerical simulations of xylophones. II. Time-domain modeling of the resonator and of the radiated sound pressure," *J. Acoust. Soc. Am.* **104**, 1633–1647.

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

Giordano, B. L., and McAdams, S. (2006). "Material identification of real impact sounds: Effects of size variation in steel, glass, wood and plexi-glass plates," *J. Acoust. Soc. Am.* **119**, 1171–1181.

Giordano, B. L., Rocchesso, D., and McAdams, S. (2010). "Integration of acoustical information in the perception of impacted sound sources: The role of information accuracy and exploitability," *J. Exp. Psychol. Hum. Percept. Perform.* **36**, 462–476.

Grey, J. M., and Gordon, J. W. (1978). "Perceptual effects of spectral modifications on musical timbres," *J. Acoust. Soc. Am.* **63**, 1493–1500.

Hartmann, W. M. (1997). *Signals, Sound and Sensation* (American Institute of Physics, Melville, NY), pp. 66 and 421–422.

Houix, O., McAdams, S., and Caussé, R. (1999). "Auditory categorization of sound sources," in *Studies in Perception and Action V*, edited by M. A. Grealy and J. A. Thomson (Lawrence Erlbaum Associates, Inc., Mahwah, NJ), pp. 47–51.

Hubert, L. J., and Arabie, P. (1985). "Comparing partitions," *J. Classif.* **2**, 193–218.

Iman, R., and Conover, W. (1979). "The use of the rank transform in regression," *Technometrics* **21**, 499–509.

Johnson, K. (2003). *Contact Mechanics* (Cambridge University Press, Cambridge), pp. 84–106.

Klatzky, R. L., Pai, D. K., and Krotkov, E. P. (2000). "Perception of material from contact sounds," *Presence: Teleop. Virt. Environ.* **9**, 399–410.

Krotkov, E., Klatzky, R., and Zumel, N. (1996). "Robotic perception of material: Experiments with shape-invariant acoustic measures of material type," in *Experimental Robotics IV: Lecture Notes in Control and Information Sciences*, edited by O. Khatib and K. Salisbury (Springer-Verlag, New York, NY), Vol. **223**, pp. 204–211.

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

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

Lambourg, C., Chaigne, A., and Matignon, D. (2001). "Time-domain simulation of damped impacted plates. II. Numerical models and results," *J. Acoust. Soc. Am.* **109**, 1433–1447.

Laroche, J. (1993). "The use of the matrix pencil method for the spectrum analysis of musical signals," *J. Acoust. Soc. Am.* **94**, 1958–1965.

Leissa, A. (1969). *Vibrations of Plates, SP-160* (NASA, Washington, D.C.), pp. 331–340.

Lieberman, A. M., and Studdert-Kennedy, M. (1978). "Phonetic perception," in *Handbook of Sensory Physiology*, edited by R. Held, H. Leibowitz, and H. L. Teuber (Springer-Verlag, New York, NY), Vol. **8**, pp. 143–178.

Lindemann, E., Dechelle, F., Smith, B., and Starkier, M. (1991). "The architecture of the IRCAM musical workstation," *Comput. Music J.* **15**, 41–49.

Lutfi, R. A. (2001). "Auditory detection of hollowness," *J. Acoust. Soc. Am.* **110**, 1010–1019.

Lutfi, R. A., and Liu, C. J. (2007). "Individual differences in source identification from synthesized impact sounds," *J. Acoust. Soc. Am.* **122**, 1017–1028.

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

McAdams, S. (1993). "Recognition of sound sources and events," in *Thinking in Sound: The Cognitive Psychology of Human Audition*, edited by S. McAdams and E. Bigand (Oxford University Press, Oxford), pp. 146–198.

McAdams, S. (2000a). "La psychomécanique des sources sonores simples (The psychomechanics of simple sound sources)," in Actes du 5ème Congrès Français d'Acoustique (Proceedings of the 5th French Acoustics Congress), Société Française d'Acoustique, Paris, France.

McAdams, S. (2000b). "The psychomechanics of real and simulated sound sources," *J. Acoust. Soc. Am.* **107**, 2792.

McAdams, S., Chaigne, A., and Roussarie, V. (2004). "The psychomechanics of simulated sound sources: Material properties of impacted bars," *J. Acoust. Soc. Am.* **115**, 1306–1320.

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., and Krimphoff, J. (1995). "Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes," *Psychol. Res.* **58**, 177–192.

McIntyre, M. E., and Woodhouse, J. (1988). "On measuring the elastic and damping constants of orthotropic sheet materials," *Acta Metall.* **36**, 1397–1416.

Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750–753.

Plack, C. J., and Moore, B. C. J. (1990). "Temporal window shape as a function of frequency and level," *J. Acoust. Soc. Am.* **87**, 2178–2187.

Schiff, J. L. (1999). *The Laplace Transform: Theory and Applications* (Springer-Verlag, New York, NY), pp. 2–22.

See supplementary material at <http://dx.doi.org/10.1121/1.3466867> Document No. E-JASMAN-128-022009 for audio files in .wav format for all the sound stimuli, Table S1, Table S2 and Table S3. For more information, see <http://www.aip.org/pubservs/epaps/html>.

Smith, B. (1995). "PsiExp: An environment for psychoacoustic experimentation using the IRCAM musical workstation," in Proceedings of the Society for Music Perception and Cognition Conference, Society for Music Perception and Cognition, Berkeley, CA, pp. 83–84.

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

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

Winsberg, S., and De Soete, G. (1993). "A latent class approach to fitting the weighted Euclidean model, CLASCAL," *Psychometrika* **58**, 315–330.