Identification of walked-upon materials in auditory, kinesthetic, haptic, and audio-haptic conditions

Bruno L. Giordano
Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Schulich School of Music, McGill University, 555 Sherbrooke Street West, Montréal, Québec H3A 1E3, Canada

Yon Visell
Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Centre for Intelligent Machines (CIM), McGill University, 3840 University Street, Montréal, Québec H3A 2A7, Canada

Hsin-Yun Yao
Tactile Labs, Saint-Bruno, Québec, Canada

Vincent Hayward
Institut des Systèmes Intelligents et de Robotique, UPMC Université Paris 06, UMR 7222, 4 place Jussieu, 75005 Paris, France

Jeremy R. Cooperstock
Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Centre for Intelligent Machines (CIM), McGill University, 3840 University Street, Montréal, Québec H3A 2A7, Canada

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

(Received 8 October 2010; revised 29 February 2012; accepted 29 February 2012)

Locomotion generates multisensory information about walked-upon objects. How perceptual systems use such information to get to know the environment remains unexplored. The ability to identify solid (e.g., marble) and aggregate (e.g., gravel) walked-upon materials was investigated in auditory, haptic or audio-haptic conditions, and in a kinesthetic condition where tactile information was perturbed with a vibromechanical noise. Overall, identification performance was better than chance in all experimental conditions and for both solids and the better identified aggregates. Despite large mechanical differences between the response of solids and aggregates to locomotion, for both material categories discrimination was at its worst in the auditory and kinesthetic conditions and at its best in the haptic and audio-haptic conditions. An analysis of the dominance of sensory information in the audio-haptic context supported a focus on the most accurate modality, haptics, but only for the identification of solid materials. When identifying aggregates, response biases appeared to produce a focus on the least accurate modality—kinesthesia. When walking on loose materials such as gravel, individuals do not perceive surfaces by focusing on the most accurate modality, but by focusing on the modality that would most promptly signal postural instabilities. © 2012 Acoustical Society of America. [http://dx.doi.org/10.1121/1.3699205]

PACS number(s): 43.66.Lj, 43.66.Wv, 43.66.Jh, 43.66.Ba [CJP]

I. INTRODUCTION

Perceiving objects in the environment is one of the fundamental functions of sensory systems. Various studies reveal that humans are capable of perceiving object properties even in the absence of visual information when passively hearing sound-generating objects (Lutfi, 2007), when manipulating them in a purely haptic or active touch context, and in an audio-haptic context (Lederman and Klatzky, 2004). The study of the auditory and haptic perception of object properties has focused on manually generated events (e.g., hitting and scraping) and has left largely unexplored a highly ecological action that generates rich multisensory information about the environment: locomotion. It is thus notable that although footstep sounds are among the non-speech sounds with the highest ecological frequency (Ballas, 1993, Appendix B), they have been investigated only in a handful of studies (e.g., Li et al., 1991; Pastore et al., 2008). This study aims to assess the extent to which non-visual information generated during locomotion contributes to our knowledge of objects in the environment. More specifically, we aimed to (i) measure the ability to identify and discriminate walked-upon materials in various non-visual conditions (auditory, kinesthetic, haptic, and audio-haptic), (ii) ascertain what sensory modality dominates identifications in audio-haptic conditions, and (iii) ascertain the extent to which sensory dominance is determined by response biases (e.g., tendency to answer “A” more frequently...
than “B” in an experiment in which A and B are equally frequent stimuli). From a methodological standpoint, the measurement of kinesthetic perception in this study relied on a novel technique for disrupting tactile material-related information based on a vibromechanical masker. The masking approach is a viable alternative to previously adopted invasive methods based on hypothermia or pharmacology (Srinivasan and LaMotte, 1995; Perry et al., 2001).

Motor activity generates a great deal of multisensory information about objects in the environment. For instance, manipulation produces mechanical responses in objects that can stimulate the auditory system and the tactile and kinesthetic systems, which constitute the sense of active touch (haptics). The hand thus serves a dual function: sensing and manipulation (Jones and Lederman, 2006). The stepping foot, like the manipulative hand, is a crucial link in the sensorimotor locomotion loop and is characterized by a dual function: sensing and support/traction. Stable and efficient locomotion is known to require the gathering of information about the ground serving as support (Ferris et al., 1998), and a large amount of sensory information is available for this purpose. Walking on a surface usually produces strong mechanical responses and, depending on the sole and ground materials, clearly audible acoustic emissions: the forces of interaction have a magnitude that is at least equal to the weight of a person. For example, when a shoe steps onto a solid material (e.g., marble), it produces transient interaction forces generated by impacts between the shoe sole and the floor surface. Previous studies of auditory and audio-haptic objects suggest that the spectrotemporal nature of these impacts contains information about the hardness of the ground material (DiFranco et al., 1997; Giordano et al., 2010). The impulsive nature of the solid-ground signals can be contrasted with the temporally extended nature of those generated when stepping onto loose aggregate ground materials (e.g., sand): impacts may be less defined and can be accompanied by high-frequency textural components produced through processes such as the fracture of brittle structures and the inelastic displacement of load-bearing ground components (Galbraith and Barton, 1970; Ekimov and Sabatier, 2006, 2008; Visell, 2011). Research on the audio-haptic perception of the size of granules inside a container (Pittenger et al., 1997) suggests that walkers can perceive the size of the aggregate-ground elements based on the properties of the acoustical and vibromechanical signals.

A relatively small number of studies has assessed the perception of non-visual walking events and, in particular, of walked-upon objects. Focusing on audition, only two studies have assessed the perception of walker characteristics such as gender and posture (Li et al., 1991; Pastore et al., 2008), whereas several studies have assessed the perception of various properties of mechanically excited objects such as their size, hardness, material class, and manner of excitation (e.g., Warren and Verbrugge, 1984; Freed, 1990; Grassi, 2005; Giordano and McAdams, 2006; Lufti, 2007; Giordano et al., 2010). This asymmetry is perhaps less evident within the haptics literature: although the perception of the properties of manually explored objects has been frequently investigated (e.g., weight, geometry, texture, compliance, Klatzky et al., 1985; Klatzky and Lederman, 2002; Ballesteros and Heller, 2008), several studies have assessed the perception of surface attributes that directly impact locomotion and balance (e.g., slipperiness and slant, compliance, discriminability of tactile ground surface indicators, Kinsella-Shaw et al., 1992; Ferris et al., 1998; Donelan et al., 2002; Pan et al., 2005; Kobayashi et al., 2008; Rosburg, 2008; Courtney and Chow, 2000). However, none of these studies has focused on the perception of highly heterogeneous ground surfaces such as those encountered on a daily basis.

An overarching issue in the study of object perception is how information from different sensory modalities is combined in a multisensory context: Which modality has the strongest influence? Does the availability of information from multiple modalities improve perceptual accuracy? Is multisensory perception dominated by the most accurate modality? To what extent do response biases modulate modality dominance? Our current knowledge of multisensory walking events is lacking in these respects because previous studies focused on those attributes of the environment that subserve navigation and disregarded information about object properties (e.g., Perry et al., 2001; Souman et al., 2009; see Visell et al., 2011 for a notable exception to this trend). In contrast, several previous studies investigated the perception of objects in non-walking audio-haptic interactive conditions. Overall, these studies revealed that although auditory information can modulate the haptic perception of object properties (e.g., crispness of potato chips, roughness of tactile textures, stiffness of tapped objects, wetness and roughness of rubber hands (Zampini and Spence, 2004; Spence and Shanker, 2010; Kim et al., 2007; Suzuki et al., 2008; DiFranco et al., 1997; Avanzini and Crosato, 2006; Reyes-Lecuona and Cañadas-Quesada, 2009; Jousmäki and Hari, 1998; Guest et al., 2002), haptic information allows more accurate perception than audition (Lederman, 1979; Jansson, 1993; Pittenger et al., 1997; Pittenger and Mincy, 1999; Heller, 1982), and dominates auditory information in audio-haptic contexts (Lederman, 1979; Lederman et al., 2002; Altinsoy, 2008; Giordano et al., 2010a).

Overall, the literature on audio-haptic object perception, and in particular on manually explored audio-haptic objects, is not sufficient to outline a set of clear expectations concerning how audition, haptics and, within the haptic modality, touch and kinesthesia are combined to yield a multisensory percept of walked-upon materials. First, physiological differences between the somatosensory systems in the hand and foot (e.g., Weinstein, 1968; Kekoni et al., 1989; Kennedy and Inglis, 2002; Well et al., 2003), and the direct vs mediated contact between the skin and the object with manual exploration and walking with shoes, respectively, can result in differences in the processing of haptic information. Second, the mechanical responses of surfaces to walking are much stronger than those resulting from manual touch (it is difficult to walk stealthily on most surfaces, whereas silent manipulation of objects is relatively effortless) and are likely to produce sounds that are more intense than those generated during manual exploration (e.g., Lederman, 1979). Importantly, the reliance on auditory information in audio-haptic conditions is
modulated by the intensity of touch-generated sounds (Lederman et al., 2002; Lederman and Klatzky, 2004). Finally, to our knowledge none of the previous studies on the audio-haptic perception of objects has disentangled the contribution of tactile and kinesthetic information.

This study assessed the identification of walked-upon materials (solids and aggregates) in various non-visual sensory contexts: auditory (passive listening to walking sounds), kinesthetic (walking with masking of sound information and vibromechanical perturbation of touch information); haptic (walking with masking of sound), and audio-haptic (walking). Analyses assessed: (i) identification performance, (ii) identification-based material discriminability (referred to as discriminability in the following) independent of response biases, and (iii) the similarity of identification confusions and patterns of material discriminability from different conditions. This last group of analyses ascertained patterns of sensory dominance in the audio-haptic context and determined whether dominance is accounted for by modality-specific discrimination abilities or, instead, is produced by response biases. Analysis (ii) and part of analysis (iii) relied on measures of bias-independent material discriminability as derived from General Recognition Theory models of identification confusions (Ashby and Townsend, 1986).

II. EXPERIMENT

A. Methods

1. Participants

Five males participated in the experiment [average age = 25 yr, standard deviation (SD) = 4 yr; average weight = 74.9 kg, SD = 4.7; average height = 173.8 cm, SD = 1.8; average shoe size (US) = 10, SD = 1.4]. None of them reported motor or haptic deficits; all had normal hearing (ISO, 2004; Martin and Champlin, 2000).

2. Apparatus

The kinesthetic, haptic and audio-haptic conditions took place inside a mildly reverberant room with a wooden floor. The apparatus for these conditions is illustrated in Fig. 1. Curtains separated a response area from a ground-presentation area. Ground materials were installed on top of two 240 × 120 cm plywood panels, each supported with 16 heavy-duty rollers. The panels were displaced on top of a vibration-dampening layer of soft rubber. Wood strips divided each panel into four 60 × 120 cm areas, each containing one ground sample (ground-to-floor height ≈10 cm). Two 12 cm-high platforms were placed on the opposite sides of the rubber sheet, one in the response area and one in the presentation area. They were covered with vinyl tiles and were laid on top of a vibration-dampening layer of soft rubber. Ropes hanging from wood poles were secured to the sides of the platforms, and delimited the walking perimeter in the presentation area. On each trial, a ground material was placed between the two platforms. Blindfolded participants walked on the ground sample while holding the ropes with their hands. Participants wore appropriately sized rubber-sole leather shoes of the same model. A computer was used to collect the identification responses.

In the kinesthetic condition, walking participants were presented with a vibrotactile, pseudo-random noise signal meant to perturb tactile information. This was produced by an array of four tactile actuators for each shoe (two for the heel, two for the toe), each placed in the middle of a carbon-fiber tube installed in the top cork layers of a 25-mm outer sole. These coil-type actuators were similar in design to those described by Yao and Hayward (2010). The orientation of the magnets inside the coil maximized the magnitude of displacements in the lateral direction with respect to the plantar surface of the foot. The outer sole was firmly strapped to the shoes worn by participants. The vibrotactile actuators were driven by a lightweight, 20 W, battery-powered amplifier, connected to a portable media player, all placed inside a backpack. Participants used an external switch to turn the amplifier on and off as needed. In the audio-haptic condition, participants used the same outer sole as in the kinesthetic condition, but without vibrotactile actuators. This equalized eventual effects of the outer sole on the gait dynamics across conditions. Walking sounds generated during this condition were recorded with a Beringer (Willich, Germany) ECM8000 microphone, feeding into a Motu (Cambridge, MA) 828 MKII digital audio interface. The microphone was located 8 cm above the center of the edge of a given ground sample, pointing down at an angle of 22°. In the haptic condition, participants used the same unactuated outer sole as in the audio-haptic condition. Walking participants were presented with a continuous 94 dB SPL (sound pressure level) white noise (sampling frequency = 48 kHz, 16-bit resolution) over Sennheiser (Wedemark, Germany) wireless RS146 headphones. Participants wore ear plugs to prevent hearing damage (noise reduction rating = 29 dB; ANSI, 1974).

For the auditory condition, sound stimuli were stored on the hard disk of a Macintosh G5 computer equipped with an M-Audio Audiophile 192 S/PDIF interface for digital-to-analog conversion. Audio signals were amplified with a Grace Design m904 monitor system and presented binaurally through Sennheiser HD280 headphones. Participants sat inside an Industrial Acoustics Company (IAC, Winchester, UK) double-walled soundproof booth. Signal peak level at the headphones ranged from 35 to 80 dB SPL, as measured with a Bruel & Kjær (Nørum, Denmark) Type 2205 sound-level meter coupled with a Bruel & Kjær Type 4153 artificial ear. The level of reproduction corresponded approximately to that of the live sounds, as measured with a Bruel & Kjær Type 2205 sound-level meter located approximately at the height of the head of the walker.
3. Stimuli

Walking-ground stimuli included four aggregate materials [very small gravel (8–10 mm), medium gravel (13 mm), and large gravel (16–20 mm)], and four solid materials [a different vinyl than the one used to cover the response- and presentation-area platforms, plywood, ceramic, and marble (from softest to hardest)]. The ground materials were selected based on the results of a preliminary identification experiment carried out with the same participants so as to increase the likelihood of identification confusions within the aggregate and solid categories while still focusing on materials encountered during everyday walking. Comparatively high confusion rates were required by the general recognition theory analysis of identifications.

In the auditory condition, participants were presented with recordings of their own walking sounds from the audio-haptic trials, including one preliminary audio-haptic session. For each of the materials, at least three recordings were selected after low-level signals and foot-dragging sounds were discarded. Each sound stimulus included two initial steps on the walking ground, followed by the steps taken on the platforms while the participant turned back toward the ground material, and two final steps on the walking ground. The tactile masker presented during the kinesthetic condition was a pseudo-random noise (sampling frequency 2.0 kHz) synthesized to maximize the spectral level of the lateral acceleration (see Fig. 2, for details).

The vibromechanical masking stimulation to the feet amounts to very small oscillatory movements, much smaller than 1 mm, that could not have had any direct mechanical consequences on posture and gate. In the kinesthetic condition, exposure to the tactile masker was limited to when the participant walked on the ground stimuli and was not sufficient to numb the feet. The masking signal was highly attenuated at frequencies below 100 Hz, and thus primarily masked stimuli transduced by subclasses of cutaneous tactile receptors exhibiting a higher-frequency, phasic response (fast-adapting receptors), which are most sensitive to transient or vibrating stimuli. The masker would have had little effect on the slow-adapting receptors that preferentially respond to slowly varying stimulation, notably the quasi-static foot-ground forces that are most important for regulation of balance and locomotion. Thus, the procedure minimized the likelihood of adaptation effects of tactile receptors, and subsequent effects on posture and gait. In sum, the consequences of the masking stimuli were predominantly sensory.

4. Procedure

On each trial, participants were presented with one of the eight grounds and were asked to identify the material by clicking with the mouse on one of eight on-screen buttons labeled “vinyl,” “wood,” “ceramic,” “marble,” “very small gravel,” “small gravel,” “medium gravel,” or “large gravel.” Response time was not limited, and feedback on identification performance was never given. Each participant identified the walking grounds in each of the kinesthetic, auditory, haptic, and audio-haptic conditions.

At the beginning of each haptic and audio-haptic trial, participants stood blindfolded on the response-area platform. At the signal of the experimenter, they opened the curtains, took two steps on the walking ground, turned around while on the presentation-area platform, took two more steps on the walking ground, reached the response-area platform, closed the curtains, and removed the blindfolds. Participants were instructed to walk with a normal pace and to avoid dragging their feet on the ground. The walking ground for the next trial was put in position while participants were in the response area. In order to mask the sounds generated while preparing the next trial, as long as participants remained in the response area they were presented with a 94 or 70 dB SPL white noise in the haptic and audio-haptic conditions, respectively (participants in the audio-haptic condition did not wear ear plugs). The procedure for the kinesthetic condition was the same as for the haptic condition, with the difference that participants started the playback of the vibrotactile noise immediately before taking the first step on the ground material and terminated it immediately after returning to the response-area platform. On each trial of the auditory condition, participants heard one recording of the target material chosen at random from among the available samples. They could listen to the sound stimulus only once. In the kinesthetic, haptic, and audio-haptic conditions, they were shown labeled samples of the walking grounds arranged on a table located in an unreachable position of the response area. Participants were not allowed to touch them. Participants in the auditory condition were shown a picture of the same labeled ground samples.

Each individual took part in twelve experimental sessions on different days, divided into three blocks of four sessions each (one session per condition). The order of presentation of the experimental conditions was randomized within each block of four sessions. On each experimental session, the eight walking grounds were presented in random order on each of eight blocks of trials, for a total of 64 trials. Each session lasted approximately 90 min. All participants identified the same walking grounds as in the current experiment during a short preliminary audio-haptic session, meant to collect recordings for the auditory conditions, and in one.

![FIG. 2. Vertical and lateral acceleration spectra at the center of the toe for the vibrotactile noise (thick black line), and for the wood and small gravel grounds (thin gray and black lines, respectively; sampling frequency = 2 kHz; fast Fourier transform window size = 256 samples). Measurement equipment: PCB (Depew, NY) 352C42 accelerometer connected to a National Instruments (Austin, TX) USB-6218 board for analog-to-digital conversion.](image-url)
additional preliminary experiment (conditions: haptic, auditory, and audio-haptic) meant to select the stimuli for the current experiment, and to allow participants to adapt to the experimental procedure and apparatus (e.g., walking blind with shoes slightly higher than usual). The set of stimuli in this preliminary experiment included half of the walking grounds investigated in the current experiment. No feedback on performance was given.

B. Results

Table I reports the matrices of identification confusions in the population of participants in each of the experimental conditions. Solid materials were rarely confused with aggregates and vice versa (grand average proportion of confusion errors across experimental conditions = 0.015, range across participants = 0.00–0.03). The patterns of identification confusions also reveal an influence of the hardness of solid materials and of the size of the gravel. For each participant and for each experimental condition, the confusion probabilities were rank-correlated with the absolute difference in the ranked hardness or gravel size between each pair of either solid or aggregate materials (e.g., difference between very small gravel and large gravel = 4). For each condition and material category (solid vs aggregates), a paired t-test was used to assess whether the average correlation ρ in the population of participants was significantly different than zero. The t-tests were carried out on the Fisher-z transformed correlations (Fisher, 1915). For both material categories, and in all experimental conditions, confusions were less likely between materials of highly diverse hardness or gravel size (grand-average ρ = −0.540, SD = 0.434). In particular, the association between identification confusions and material properties was significant for aggregates in all experimental conditions [ρ ≤ −0.647, t(4) ≤ −4.16, p ≤ 0.014], whereas for solids a significant association emerged only in the haptic condition [ρ = −0.555, t(4) = −3.37, p = 0.028; ρ ≥ −0.301, t(4) ≥ −2.59, p ≥ 0.061 for the other experimental conditions]. Subsequent analyses considered the matrices of identification confusions within the categories of solid and aggregate materials and assessed (1) identification performance, (2) bias-independent discrimination abilities, and (3) the similarity of the patterns of identification confusions and discriminability among different experimental conditions.

1. Identification performance

Analyses tested for better-than-chance identification performance and assessed the effects of experimental condition and material category (solids vs aggregates) on performance (Fig. 3: average condition- and category-specific performance). Considering chance performance to be 12.5% correct, performance was better than chance in all experimental conditions and for both material categories [single-sample t(4) ≥ 4.24, p ≤ 0.007]. Because of the almost perfect ability to distinguish between solid and aggregate materials, a more appropriate measure of chance-level performance would be to focus on the ability to identify materials within the solid and aggregate categories (chance performance = 25% correct). Performance levels higher than 25% correct were observed for all experimental conditions and material categories [single-sample t(4) ≥ 2.18, p ≤ 0.048], with the exception of the performance for solids in the kinesthetic condition [t(4) = 0.91, p = 0.792].

The effect of experimental condition and material category on identification performance was assessed within a linear mixed-effects model (LMM, Verbeke and Molenberghs, 2000; West et al., 2006) fitted using the SAS® PROC MIXED routine (Littell et al., 2006). In general terms, LMMs can potentially include both fixed and random effects measuring the average effect and its variability within a unit of interest (e.g., population), respectively. With LMMs, both the fixed and random

<table>
<thead>
<tr>
<th>Experimental Condition</th>
<th>Kinesthetic</th>
<th>Auditory</th>
<th>Haptic</th>
<th>Audio-haptic</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>27 36 25 28</td>
<td>1 3 32 23 36</td>
<td>29</td>
<td>49 37 14 20</td>
</tr>
<tr>
<td>CE</td>
<td>24 23 30 38</td>
<td>1 1 20 54 25 19</td>
<td>1 1 27 41 25 26</td>
<td>1</td>
</tr>
<tr>
<td>WO</td>
<td>26 19 28 40</td>
<td>2 4 1 35 21 43 20</td>
<td>1</td>
<td>24 19 42 34</td>
</tr>
<tr>
<td>VI</td>
<td>29 23 32 33</td>
<td>2 1 28 25 29 38</td>
<td>1</td>
<td>15 19 34 51 1</td>
</tr>
<tr>
<td>SG</td>
<td>1 2 1 2 1 62 27 20 5 47 46 20 7 1 78 29 9 3</td>
<td>60 39 16 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG</td>
<td>1 1 1 1 1 19 44 36 18</td>
<td>1 17 32 40 30</td>
<td>1 27 44 34 14</td>
<td>1</td>
</tr>
<tr>
<td>LG</td>
<td>2 2 1 2 8 36 69 2</td>
<td>2 5 16 36 61</td>
<td>1 3 33 83</td>
<td>3 7 33 77</td>
</tr>
</tbody>
</table>

FIG. 3. Average identification performance for each material category and experimental condition. Error bar: ±1 SEM.
effects and the structure of the model covariances are the object of model selection procedures (West et al., 2006, pp. 39–41). The LMMs presented in this study have a simplified structure: they include both significant and nonsignificant fixed effects, have a random effect for the intercept, modeling the variation of the average of the participant-specific data across individuals, and adopt a compound-symmetry structure for the individual-specific covariance matrix of the residuals. Conclusions warranted by more complex models, resulting from model-selection procedures, were equivalent to those warranted by the simplified models presented here. All the models presented here satisfied the normality assumption for participant-specific residuals (Shapiro-Wilk $W \geq 0.95$, $p \geq 0.281$).

Identification performance was thus analyzed within an LMM that included the fixed effects of experimental condition, material category, condition $\times$ category, and ground material (category), i.e., material nested within category (the factor ground material is nested within the factor material category because, for instance, “material” is “marble” only when “category” is “solid”). The material (category) effect was significant [$F(6, 142) = 14.02$, $p < 0.001$], indicating at least one significant difference in the identification performance for two materials in the same category. The condition $\times$ category effect was not significant [$F(3, 142) = 2.21$, $p = 0.090$]. The effect of category was significant [$F(1, 142) = 17.04$, $p < 0.001$], indicating better identification performance for aggregates than for solids. Finally, the effect of experimental condition was significant [$F(3, 142) = 5.52$, $p = 0.001$], with significantly worse performance in the kinesthetic condition than in the haptic and audio-visual conditions. The condition $\times$ experimental condition $\times$ category effect was not significant [$F(3, 142) = 2.21$, $p = 0.090$]. The effect of category was significant [$F(1, 142) = 17.04$, $p < 0.001$], indicating better identification performance for aggregates than for solids. Finally, the effect of experimental condition was significant [$F(3, 142) = 5.52$, $p = 0.001$], with significantly worse performance in the kinesthetic condition than in the haptic and audio-visual conditions [$F(1, 142) = 3.32$, Bonferroni-adjusted (Bonf.-adj.) $p \leq 0.001$; abs($r(142)$) $\leq 1.83$, Bonf.-adj. $p \geq 0.069$ for the other post hoc pairwise contrasts between experimental conditions].

2. Discriminability

Measures of the bias-independent ability to discriminate the ground materials were derived from general recognition theory (GRT, Ashby and Townsend, 1986; Ashby and Perrin, 1988; Ashby, 1992) models of identification data. GRT is a multidimensional extension of signal detection theory (SDT, Green and Swets, 1966) capable of analyzing confusion matrices from multi-category identification experiments characterized by strong response biases. It models the perceptual effects $f_A(\cdot)$ for the stimulus $S_A$ as a multivariate normal distribution with given mean and covariance matrix. In an identification experiment, participants are assumed to partition the space of sensory effects into independent regions using decision boundaries that are often modeled as a linear function of the space dimensions. Perceptual effects that fall within the same region will receive the same identification response $R$. Assuming a two-dimensional sensory space with dimensions $x$ and $y$, the probability $p(R_B|S_A)$ that $S_A$ will receive the response $R_B$ corresponds to

$$p(R_B|S_A) = \int_{r_B} f_A(x, y) \, dx \, dy,$$

where $r_B$ is the region of the space associated with response $R_B$. GRT emphasizes the notion of between-stimuli proximity rather than that of between-stimuli distance. The bias-independent proximity between two stimuli is termed similarity, $s$, and is defined as

$$s(S_A, S_B) = \int_{f_A < f_B} f_A(x, y) \, dx \, dy, \quad s \in [0, 1],$$

where $f_A < f_B$ denotes the region of the space where the sensory effects for $S_A$ are less likely than those for $S_B$. The GRT measures $s$ of sensory similarity are thus analogous to $d'$ in SDT. Within the current study, these measures can be interpreted as a response-bias-free measure of the ability to discriminate ground materials based on identification responses, where $s = 0$ and $1$ for perfectly accurate and perfectly inaccurate discrimination, respectively.

One GRT model was estimated for each of the participants in each of the experimental conditions. The almost perfect solid-aggregate discrimination resulted in around half of the 64 cells in each confusion matrix having a value of zero. For this reason, a maximum of 32 degrees of freedom were available for model fitting. The number of parameters in each GRT model was thus reduced to 26 by assuming (i) a two-dimensional space (two location parameters for the mean of each stimulus distribution), (ii) the same covariance matrix for materials within the same category (three parameters for each of the two covariance matrices), (iii) two fixed orthogonal decision boundaries with one of them separating the two material categories (no free parameter), and (iv) two free decision boundaries, one for the solids and one for the aggregates (two parameters for each boundary). Model fitting was based on the method described in Ennis and Ashby (2003), combined with an iterative routine for the minimization of the sum of the squared model-prediction errors (SSE). The starting configuration used by the SSE minimization routine was the same for all confusion matrices: the free decision boundaries were parallel to the solid-aggregate boundary; the mean of the stimulus distributions was located at the intersection of the free boundary with the fixed boundary being orthogonal to the solid-aggregate boundary; all the covariance matrices were an identity matrix. Thus constrained and fitted, the GRT models accounted well for the data in all the confusion matrices ($SSE = 0.004–0.090$; $RSQ = 0.910–0.996$). Figure 4 shows the GRT models for one representative participant in each of the experimental conditions. Figure 5 shows the average between-stimuli GRT similarity for solid and aggregate materials in each of the conditions.

GRT similarities were analyzed within an LMM with the fixed effects of experimental condition, material category, condition $\times$ category, and materials-pair (category). Only GRT similarities for materials within the same category were considered. The materials-pair (category) effect was significant [$F(10, 218) = 8.20$, $p < 0.001$], indicating that at least two of the similarities within the same category were different. The condition $\times$ category effect was not significant [$F(3, 218) = 0.65$, $p = 0.585$]. The effect of category was significant [$F(1, 218) = 179.67$, $p < 0.001$], indicating a better ability to discriminate among aggregates than among solids. Finally, the effect of experimental condition was
significant \( F(3, 218) = 15.55, p < 0.001 \). In particular, all post hoc pairwise contrasts between average condition-specific similarities were significant \[ \text{abs}(t(218)) > 3.45, \text{Bonf.-adj.} \ p < 0.001 \], with the exception of no significant difference between the auditory and kinesthetic conditions, and between the haptic and audio-haptic conditions \[ \text{abs}(t(218)) = 1.24, \text{Bonf.-adj.} \ p > 0.228 \].

### 3. Comparison of experimental conditions

A last analysis compared the patterns of identification confusions or the bias-independent GRT measures of discriminability from the different experimental conditions in order to (i) assess significant between-condition similarities, (ii) analyze the variation of the between-condition similarity across material categories and data types (identifications vs discriminabilities), (iii) assess significant patterns of sensory dominance in the audio-haptic condition, and (iv) establish discriminabilities, as based on the assessment of which among the non-audio-haptic conditions yielded identifications or discriminabilities that most strongly resembled audio-haptic data.

The similarity between condition-specific data was quantified in terms of the concordance correlation \( \rho_c \) (Lin, 1989), a special case of the Pearson correlation that measures the departure from the equality line \[ \rho_c(A, B) = 1 \text{ and } -1 \] if \( A = B \) and \( A = -B \), respectively. Focusing either on the identification confusions or on the discriminabilities, one \( \rho_c \) coefficient was computed to compare the data for the same participant from each pair of experimental conditions (6 pairs) and for each of the solid and aggregate categories, for a total of 24 \( \rho_c \) coefficients for each participant. The comparison of confusion matrices did not consider the correct responses in order to minimize the effects of identification performance and to help interpret the results in terms of response biases. The comparison of discriminabilities did not consider the self-similarities (e.g., similarity between marble and itself) which, by definition, equal 1. All \( \rho_c \) measures were Fisher-Z transformed (Fisher, 1915) prior to conducting any statistical analysis and modeling. Figure 6 displays the two-dimensional metric multidimensional scaling (MDS) model of each of four \( \rho_c \) matrices of condition similarity averaged across participants: identifications-solids, discriminabilities-solids, identifications-aggregates, and discriminabilities-aggregates. The MDS models were fit to a distance between experimental conditions defined as \( 1 - \rho_c \) \( (R^2 \geq 0.95 \text{ across the four } \rho_c \text{ matrices}) \).

Significant between-condition similarities were assessed based on a one-tailed one-sample \( t \)-test that ascertained whether the across-participant average of each \( \rho_c \) coefficient was significantly higher than zero. For solid materials, no significant similarity emerged between the discriminabilities for the different experimental conditions \[ t(4) < 1.67, p > 0.086 \], and significant similarities between identification confusions emerged only between the haptic and kinesthetic
and between the haptic and audio-haptic conditions \([t(4) \geq 3.43, p \leq 0.013; t(4) \leq 1.53, p \geq 0.100\) for the other comparisons]. For aggregate materials, and for both the discriminabilities and identifications, all between-condition similarities were significant \([t(4) \geq 2.23, p \leq 0.044]\), with the exception of the similarity between the discriminabilities for the auditory condition, on the one hand, and the haptic and audio-haptic conditions, on the other \([t(4) \leq 1.94, p \geq 0.062]\).

The \(\rho_c\) measures of between-condition similarity were analyzed within an LMM that included the fixed effects of material category, data type (identifications vs discriminabilities), pair of experimental conditions (e.g., auditory-haptic), and all the possible interactions between these factors. The three-way interaction was significant, indicating that whereas for solids identifications in the kinesthetic and haptic conditions were more similar to each other than the discriminabilities for these conditions \([F(1, 92) = 7.34, \text{Bonf.-adj.}\ p = 0.049; F(1, 92) \leq 5.33, \text{Bonf.-adj.}\ p \geq 0.139\) for the other post hoc contrasts], aggregate identifications in the kinesthetic and multisensory conditions were more similar to each other than the discriminabilities for these conditions \([F(1, 92) = 14.71, \text{Bonf.-adj.}\ p = 0.001; F(1, 92) \leq 4.92, \text{Bonf.-adj.}\ p \geq 0.174\) for the other post hoc contrasts]. Both the main effect of pair and the data-type \(\times\) pair interaction were significant \([F(5, 92) \geq 2.74, p \leq 0.024]\), indicating that not all the similarities between experimental conditions were equal, and that they varied depending on whether the experimental conditions were compared based on identifications or on discriminations. The main effects of material category and data type, as well as the interaction between these factors were significant \([F(1, 92) \geq 6.27, p \leq 0.014]\) (see Fig. 7). For both identifications and discriminations, data from different experimental conditions were more similar for aggregates than for solids \([F(1, 92) \geq 29.20, \text{Bonf.-adj.}\ p < 0.001]\). Whereas aggregates identifications from different conditions were more similar to each other than were discriminations \([F(1, 92) = 19.68, \text{Bonf.-adj.}\ p < 0.001]\), for solid materials the similarity of the identifications from different conditions was not significantly different than what was observed for the discriminations \([F(1, 92) = 0.80, \text{Bonf.-adj.}\ p = 1.000]\).

A final analysis determined which among the kinesthetic, auditory, and haptic conditions yielded data that most strongly resembled those for the audio-haptic condition. For each material category and for both identifications and discriminations, paired-sample \(t\)-tests were used to ascertain significant pairwise differences among the \(\rho_c\) coefficients between the audio-haptic condition, on the one hand, and all the other conditions, on the other. For aggregate materials, identifications in the kinesthetic condition resembled identifications in the audio-haptic condition more closely than those in any other condition \([t(4) \geq 4.68, \text{Bonf.-adj.}\ p \leq 0.028; t(4) = -2.33, \text{Bonf.-adj.}\ p = 0.240\) for the auditory vs haptic comparison]. For solid materials, no condition yielded data that most strongly resembled those for the audio-haptic condition \([\text{abs}(t(4)) \leq 2.56, \text{Bonf.-adj.}\ p \geq 0.188]\). The same analysis was carried out by focusing on the GRT discriminabilities rather than on the raw identifications. For both solids and aggregates, none of the non audio-haptic conditions yielded GRT discriminabilities that resembled those in the audio-haptic context more strongly than those in any other non audio-haptic context \([\text{abs}(t(4)) \leq 2.01, \text{Bonf.-adj.}\ p \geq 0.343]\).

**C. Discussion**

In all experimental conditions solid materials were seldom confused with aggregate materials. This result suggests either a large between-category diversity in sensory information (e.g., stepping on solids produces not more than two vibromechanical impacts; stepping on aggregates produces a sequence of temporally distributed impacts), or the fact that the perceptual system is well-equipped for differentiating between solids and aggregates because of the high relevance of this distinction for the control of locomotion. After the nearly perfect differentiation between solid and aggregate materials was accounted for, better-than-chance identifications emerged for all the experimental conditions except for the solid materials in the kinesthetic condition. Intact touch information, which was perturbed in the kinesthetic condition, was thus strictly necessary for identifying solids but not for identifying aggregate materials. Similar effects of experimental condition emerged for both solids and aggregates, with significantly worse performance in the kinesthetic condition. It is unclear whether identification-performance results are the product of the discriminability of sensory information or of response biases (e.g., both better discrimination abilities and weaker response biases could produce superior identification performance for aggregates).

Bias-independent measures of material discriminability were estimated within the GRT framework. Consistent with a discrimination-based interpretation of the superior aggregate-identification performance, aggregates were better discriminated than solids in all experimental conditions. This result suggests the presence of more sensory information concerning the identity of aggregate materials than is the case with solids (e.g., with aggregates, information can be extracted from a larger number of impacts produced by the interaction of the heel and toe with multiple objects, and also from their temporal patterning, whereas with solids, information must be extracted from a lower number of impacts whose temporal patterning is largely independent of the ground material). Experimental condition affected similarly the discriminability of materials in both categories. This result is somewhat surprising.
because for mechanically different phenomena such as walking on solids and aggregates (see the Introduction), discrimination abilities are not necessarily maximized or minimized by the same modalities. We have no clear explanation for this result. Importantly, a significant increase in discriminability was observed from the auditory and kinesthetic conditions to the haptic and audio-haptic conditions. Superior discrimination in the haptic and audio-haptic conditions would be expected because the auditory and kinesthetic conditions were both unimodal and thus likely poorer in sensory information. Interestingly, the same level of discriminability was observed for the haptic and audio-haptic contexts, i.e., adding auditory information to a haptic context produced weak gains in perceptual processing.

Overall, the analysis of discrimination is consistent with previous observations of a higher accuracy of haptic perception than auditory perception during the manual exploration of objects (Lederman, 1979; Jansson, 1993; Pittenger and Minck, 1999). Contrary to what was observed for the measures of discriminability, the analysis of identification performance failed to support this hypothesis of higher accuracy. Response biases are likely the reason for this because they influence identification performance but not discriminability. Finally, the fact that the perturbation of tactile information impaired discrimination down to the level observed for the auditory condition suggests that the superiority of haptic compared to auditory perceptual processing is, at least in part, due to the multisensory nature of the haptic modality.

A subsequent analysis assessed the similarity of the patterns of (bias-independent) discriminability and (bias-dependent) identification confusions from different experimental conditions. The comparison of results for discriminabilities and identifications made it possible to infer the effects of the response biases on the similarity of perceptual processes in the different conditions (e.g., response biases are likely the reasons for why conditions A and B yield different discriminabilities but similar identifications). The discriminations in different experimental conditions resembled each other more strongly for aggregate than for solid materials. This result is consistent with a greater cross-modal redundancy of sensory information for aggregate materials (information from different modalities is perfectly redundant if it affords the exact same pattern of discriminabilities). This result is evocative of the superior discriminability and identifiability of aggregate materials because redundant information across integral sensory dimensions (stimulus dimensions that cannot be processed independently of each other) leads to superior perceptual performance (Ashby and Townsend, 1986; Klatzky et al., 1989). It is, however, important to recall that explanations for the superior processing of aggregate materials are based on the cross-modal redundancy of information are of limited value because they do not explain why aggregates are also better identified than solids in the unimodal conditions. Further, this account would require a demonstration of cross-modal integrality of ground-identity information based on different experimental methods than those adopted here. Finally, identifications from different conditions were more similar to each other than were the discriminations only for aggregates but not for solids. This result suggests a strong cross-modal consistency of response biases for the identification of aggregate materials, which results in a greater similarity of identification data from different experimental conditions than is to be expected on the basis of discrimination alone.

The dominance of modality-specific information in the audio-haptic context was also assessed by comparing data from different conditions. For solid materials, only haptic identifications were significantly similar to audio-haptic identifications. Given the superior discriminability of materials in the haptic condition, this result thus appears consistent with the modality-appropriateness hypothesis according to which the dominant information in a multisensory context is the one that maximizes perceptual performance when presented alone (Welch and Warren, 1980). Note, however, that the support for the modality-appropriateness hypothesis is only partial. Indeed, pairwise comparisons failed to reveal a non audio-haptic condition that yielded identification or discriminabilities that resembled those in the audio-haptic condition more strongly than those in any other non audio-haptic condition.

Different results emerged for the aggregates. Most likely because of response biases that were strongly consistent across all conditions, audio-haptic identifications were significantly similar to identifications in all of the other conditions (note that the statistical framework assessed significant between-condition similarities, and not significant between-condition differences). Importantly, subsequent pairwise comparisons revealed that audio-haptic identifications more strongly resembled identifications in the kinesthetic condition than those in any other condition. Notably, the audio-haptic dominance of kinesthetic information was not explained by the similarity of discrimination patterns: the audio-haptic discriminabilities were equally similar to the discriminabilities in all the other conditions. Overall, the results for aggregate materials are in disagreement with what is predicted by the modality-appropriateness hypothesis: first, because the dominant information in the audio-haptic context affords the least accurate identifications and discriminations; second, because the dominance of the kinesthetic information appears to be a product of biases alone, i.e., is not explained by the patterns of material discriminability related to the quality of the sensory information. These results thus support the hypothesis that modality dominance is also influenced by modality-weighting biases developed in everyday multisensory contexts (Lederman and Klatzky, 2004). It is indeed plausible that everyday walking on aggregate materials promotes a focus on kinesthetic information because this modality is highly critical to the control of locomotion: when walking on materials such as gravel, it is perhaps safest to focus on the sensory channels that best signal unstable locomotion (e.g., vestibular system, joint capsule receptors), rather than on those modalities that afford a finer discrimination of the gravel grain.

A few final considerations are necessary about the interpretation of the results for the auditory and kinesthetic conditions. The auditory condition was the only passive one (participants were presented with pre-recorded stimuli). It might be speculated that walking participants allocate attentional resources more efficiently at foot impact time because they actively control the locomotion program, and might thus be capable of better processing interactively produced
sounds than recorded sounds. Further, the sound recordings were taken at around ground level, not at ear level, and did not contain binaural information or information about the filtering effects of the body and head. For these reasons, the estimates of auditory discrimination and identification performance might be at best an underestimate of what would be observed in a hypothetical active auditory condition in which normal walking with selective (and silent) suppression of kinesthetic and tactile information is possible. Note, however, that these interpretational caveats do not apply to the non-significant difference in perceptual performance between the haptic and audio-haptic conditions. Focusing on the kinesthetic condition, the exact origin of the observed effects of the vibromechanical masker is unknown. A working hypothesis was adopted according to which the vibromechanical noise selectively impaired the ability to process tactile information about ground identity. Further experimentation is necessary to establish the exact extent of the masking of tactile information, to assess the eventual effect of non-sensory factors (e.g., attentional resources) and to ascertain the potential effects of the vibromechanical masker on the processing of information from non-tactile modalities (e.g., audition).

III. CONCLUSIONS

Locomotion generates a great deal of multisensory information about walked-upon objects. To date, little or no information is available concerning either our ability to use this information to get to know the environment or the role of different non-visual sensory modalities in the shaping of such a knowledge. This study measured the identification of walking-ground materials in auditory, kinesthetic, haptic, and audio-haptic conditions. A novel method was developed to selectively perturb tactile ground-identity information while leaving the kinesthetic information intact.

In line with previous studies of manually explored objects, haptic perception was in general more accurate than auditory perception. Haptic superiority was potentially the product of the multisensory nature of this modality because the perturbation of tactile information in the kinesthetic condition impaired perceptual performance down to the level observed for the unimodal auditory condition. In line with the hypothesis of a greater multisensory focus on the most accurate information (Welch and Warren, 1980), the analysis of sensory dominance for the identification of solid grounds gave partial evidence in support of a dominance of the most accurate haptic information. In contrast with the same hypothesis, the audio-haptic identification of aggregate materials appeared to focus on the least accurate kinesthetic information. The dominance of kinesthetic information appeared to be a product of biases, and was not explained by discrimination processes. A bias toward focusing on kinesthetic information could be consistent with an optimal decision-making strategy (Kording and Wolpert, 2006) in which individuals attempt to minimize a measure of the expected loss as a result of a potential fall.

ACKNOWLEDGMENTS

The authors wish to thank F. Gregory Ashby for feedback on GRT fitting procedures, and Richard E. Pastore and one anonymous reviewer for useful comments on previous versions of this paper. This work was supported by the Canada Research Chair in Music Perception and Cognition, a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN 312774-05), and a Special Research Opportunity Grant from NSERC to S.M., by a grant to J.R.C. from the Québec Ministère du Développement Économique, de l’Innovation et de l’Exportation (MDEIE) associated with the European Community 7th Framework Program Project Natural Interactive Walking (NIW, No. 222107), and by a grant to V.H. from NSERC (RGPIN 6670-07).


Giordano et al.: Non-visual identification of walking grounds 4011

[Proceedings of the 10th French Acoustics Congress, Lyon (2010b).]


