Feelings Elicited by Auditory Feedback from a Computationally Augmented Artifact: The Flops

Guillaume Lemaitre, Olivier Houix, Patrick Susini, Yon Visell, and Karmen Franinović

Abstract—This paper reports on emotions felt by users manipulating a computationally and acoustically augmented artifact. Prior studies have highlighted systematic relationships between acoustic features and emotions felt when individuals are passively listening to sounds. However, during interaction with real or computationally augmented artifacts, acoustic feedback results from users' active manipulation of the artifact. In such a setting, both sound and manipulation can contribute to the emotions that are elicited. We report on a set of experimental studies that examined the respective roles of sound and manipulation in eliciting emotions from users. The results show that, while the difficulty of the manipulation task predominated, the acoustical qualities of the sounds also influenced the feelings reported by participants. When the sounds were embedded in an interface, their pleasantness primarily influenced the valence of the users' feelings. However, the results also suggested that pleasant sounds made the task slightly easier, and left the users feeling more in control. The results of these studies provide guidelines for the measurement and design of affective aspects of sound in computationally augmented artifacts and interfaces.

Index Terms—Computationally augmented interface, emotion, auditory displays, acoustical features, manipulation

✦

1 INTRODUCTION

AUDITORY displays are increasingly found in everyday products ranging from coffee makers to electric cars, due to advances in microfabrication and embedded computation. Such an intervention can serve functional purposes, like making an electric car audible to pedestrians, and at the same time improve the emotional appeal of a product, for example, by enhancing the roar and perceived power of the engine through "active automotive sound design" [1]. Thus, auditory displays can also play an important role in determining the emotions generated before, during, and after user interaction, particularly affecting qualities of user experience related to desirability and pleasurability [2]. The study presented here further illustrates the extent to which these two aspects are inseparable.

Emotions generated by an artifact depend on diverse factors, as has been extensively explored in prior literature [3], [4], [5], [6], [7]. In a model proposed by Rafaeli and Vilnai-Yavetz [8], people are asked to appraise an artifact

according to three primary dimensions: the level of instrumentality—how the artifact fits an ascribed function or how *usable* it is, aesthetic qualities, such as visual color or sonic timbre, and symbolism, which relates to how the artifact is interpreted in a certain cultural context. A deeper understanding of the mechanisms through which emotions may be elicited can likewise be expected to aid the design of affective interactions for new computational artifacts with their potentially rich layers of functionality.

Picard notes that "Music is perhaps the most socially accepted form of mood manipulation" [9, p. 234]. There has been considerable prior research on the computational analysis of emotional content in music and on techniques for generating or selecting musical stimuli inducing a chosen mood. By comparison, less prior research has investigated the emotional content in everyday sounds, which is commonly taken to refer to nonmusical, nonspeech sounds, such as those produced by a door closing, an object bouncing or breaking, etc. Such sounds have played essential roles in establishing mood, expectations, and intentions in the dramatic arts, notably in film, radio, and theater [10], [11].

Several empirical studies have highlighted systematic relationships between sound features and the feelings reported by the listeners, when sound is the only potential source of emotion elicitation, and when listeners are required to carefully listen to the sounds. However, the generality of these relationships is questionable. In fact, the sounds of manufactured everyday objects seldom occur alone, and users are rarely in a situation when they pay much attention to them. Rather, they happen in many situations as a consequence of a user interacting with an object. The interaction, itself, can induce emotions, depending, for instance, on how successful the manipulation with respects to the goals of the user is. And it is even likely that the success or failure to

- G. Lemaitre is with the Unità di ricerca Interazione, Facoltà di design e arti, Università Iuav di Venezia, Dorsoduro 2206, 30123 Venezia, Italia. E-mail: GuillaumeJLemaitre@gmail.com.
- O. Houix and P. Susini are with Équipe Perception et Design Sonores, STMS-IRCAM-CNRS, 1 place stravinsky, 75004 Paris, France. E-mail: {houix, susini}@ircam.fr.
- Y. Visell is with the Institut des Systèmes Intelligents et de Robotique (ISIR), Université Pierre et Marie Curie, CNRS UMR 7222, Boite courrier 173, 4 Place Jussieu, 75252 Paris Cedex 05, France. E-mail: yon@zero-th.org.
- K. Franinović is with Zurich University of the Arts, Switzerland.

interact with an object would affect emotions in a more drastic fashion than the acoustical appearance might do. These sources of emotion (sound and manipulation) appear to be at least qualitatively different. One could even wonder whether sounds might affect emotions when subjects are not focusing on the sounds, or if the relationships between sound features and feelings would still hold.

The goal of this paper is to assess these issues and to identify potential stable relationships between sound features and emotions. However, the previous example illustrates that it is necessary to go beyond the simple account that sounds can influence emotions and to consider the influence of sound on emotion in a more general theoretical framework of the processes that elicit emotions in human subjects. We examine such a framework in the following paragraphs.

1.1 The Appraisal Theory of Emotions

Most modern emotion theorists agree on a componential approach to emotions, assuming that an *emotional episode* is a dynamic process consisting of coordinated changes in several *components*: cognitive, neurophysiological, action tendencies, motor expression, subjective feelings [12], [13], [14], [15], [16], [17]. Among these components, *feelings* have a particular status: They serve as a monitoring function, and are consciously accessible. Feelings thus represent the component of an emotion episode that a subject can report.

The appraisal theory advocated by Scherer ("component process model") provides a useful framework to analyze the influence of sound and interaction on a user's emotions (see, for instance, [18] for a recent overview). It proposes that different components that form an emotion episode are the results of different *appraisals*. In this account, emotional episodes are elicited and differentiated on the basis of a person's subjective appraisal of a stimulus or situation on a number of dimensions: novelty, intrisic pleasantness, goal conduciveness, coping potential, and normative significance [19], [20], [21]. These criteria can be readily grouped together into four overlapping sets. The first set involves qualities that are intrinsic to the stimuli: the appraisal of novelty and pleasantness. These criteria are coded at a low level of processing and in an automatic fashion. The second set has motivational bases. These appraisal criteria are related to the needs, goals, and values of the subject, and are highly dependent on the motivational state of the subject. The third set is related to the assessment of the subject's ability to deal with the situation: the appraisal of the subject's power and coping potential. The fourth set deals with the appraisal of the social and normative implications of a situation or stimulus.

When we apply this framework to the sounds of interactive artifacts, we can assume that sounds and interaction are not addressed by the same appraisal criteria: The characteristics of a sound are more easily related to its pleasantness or novelty, whereas manipulating an object can also interfere with the subject's goals, his or her ability to deal with a situation, and potentially has normative and social implications. However, if a sound makes the manipulation of the interface more or less difficult, it may possibly also interfere with the subjects's goals.

1.2 Empirical Assessment of an Emotion Episode

There are two main kinds of approach to the empirical assessment of emotions caused by a stimulus. First, physiological measures (heart rate, skin conductance, facial EMG, startle reflex, etc. [22]) can measure neurophysiological activities, action tendencies, and motor expressions. Second, self-reports can provide insights into the feelings of the subjects. Various methods have been used, from free-form verbalizations to semantic scales. The results of such studies have very often suggested that the feelings observed in or reported by subjects can be accounted for by a low-dimensional framework (see, for instance [23], [24], [25], [26], [22], [27], [13]). Two or three *dimensions* (but sometimes more [28]) of the reported feeling are generally considered: the *valence* dimension, describing negative to positive feelings, the *arousal* dimension, describing the degree of arousal (from calm to excited) felt by the subject, and the *dominance* dimension (also called *power*), describing how dominated or dominant the subjects feel. In the context of the appraisal theory, Scherer proposed an interesting account of how the different types of appraisals of a stimulus or a situation can result in different dimensions of a subject's reported feelings [29]:

- Valence results from the appraisal of intrinsic pleasantness (a feature of the stimulus) and goal conduciveness (the positive evaluation of a stimulus that help in reaching goals or satisfying needs), see also [21].
- Arousal results from the appraisal of the stimulus' novelty and unexpectedness (when action is needed unexpectedly).
- Dominance results from the appraisal of the subject's coping potential.

Therefore, for what concerns the sounds of manipulated objects, the appraisal of the features of sound may mainly have an influence on the valence (appraisal of pleasantness) and arousal (appraisal of novelty) dimensions of the feelings. Possibly, if the sound has a function in the interaction, it also may have an influence on the appraisal of the goal conduciveness (imagine an alarm clock that does not ring loud enough to wake one up).

1.3 Emotions and Auditory Feedback

Sound quality studies provide, indirectly, some insights into relationships between acoustical features of product sounds and emotions ([30]; see also [31] for a review of sound quality evaluation). For example, it has been reported that attractive products are perceived as easier to use [32], [33]. Emotional reactions to the sounds of everyday products have been primarily studied in terms of pleasantness or annoyance [34], or preference [35].

Sensory pleasantness and annoyance are two important concepts for the quality of the sounds of manufactured products. Even though they have generally not been addressed as feelings, they display interesting similarities with the feelings elicited by different appraisal criteria. Sensory pleasantness, a notion that directly comes from the psychoacoustical tradition, is considered as an auditory attribute, and is related to other auditory attributes such as roughness, sharpness, tonality, and loudness [36]. Sensory

pleasantness thus seems very close to Scherer's concept of intrinsic pleasantness for sounds. Annoyance is a concept used to describe the nuisance caused by noises, particularly in the case of exposure to urban noises. Noise annoyance can be related to acoustic variables, but acoustic characteristics do not play the most important role. Psychosociological variables (and particularly how the noise exposure affects the lives of the neighboring inhabitants) are important determinants [34]. In Scherer's terms, annoyance is also therefore linked, through goal-conduciveness, to the valence dimension, to the dominance dimension (coping potential), and to normative and social appraisals.

More recently, Västfjäll et al. have attempted to relate preference, emotions, and acoustical characteristics of product sounds. They have assessed a model proposed by Mehrabian and Russell relating preference for current mood (in this context: "experienced momentary affective state") to valence and arousal [37], and assessed it in relation to users' preferences for exterior and interior car noises [35]. In another study they have found significant correlations between valence and arousal ratings and several psychoacoustical descriptors of aircraft sounds, determining valence to be correlated with loudness and naturalness, and activation with sharpness and tonal content [38]. Naturalness was in this case rated by listeners on two scales: "how artificial and how natural are the sounds?" However, they have also shown that, in the context of sound quality evaluation, individual factors (users' frame of mind, emotional states, expectations) influence affective and preference judgments [39]. Other studies have shown that even the appraisal of simple acoustic characteristics sometimes results in important individual differences [40].

1.4 Emotions and Computational Artifacts

Sounds are also predominantly used in many forms of communication between human users and electronic devices. Assessing the emotions induced by the use of such interfaces takes part of the evaluation of their ergonomy. Among these interfaces, computer displays have been most often evaluated [41], [42], [43] and particularly in the case of affective computing interfaces [44], [45].

Furthermore, because computer interfaces (and, more particularly, computer games) have the potential to induce emotions through different types of appraisals, they can also be used as an experimental technique to elicit emotions in subjects in a laboratory setting, and study emotion processes [46], [47]. The results of these studies confirm that such games can in fact induce strong emotions, particularly because they engage participants in demanding tasks and provide them with rewards or losses [48], [49], [50].

As regards the issues addressed in the present paper, a study reported by van Reekum et al. [51], [52] is particularly interesting. These authors used a computer game to create emotions in adolescents. Successfully completing a level of the game (goal-conducive event) or losing a ship (goal obstructive event) was operationalized as goal conduciveness, and pleasant and unpleasant sounds associated with these events were operationalized as intrinsic pleasantness of the event. Formally, such a game therefore manipulates the two variables that putatively influence emotions in the case of manipulating a sounding object. The results showed that

the manipulation of goal conduciveness and intrinsic pleasantness successfully resulted in physiological changes. It also resulted in variations of the voice characteristics and self-reports were consistent with the hypothesized manipulation of the valence, arousal, and power dimensions of the felt emotions. These results therefore supported Scherer's prediction that the emotions induced in such a framework would result from different types of appraisals (intrinsic pleasantness and goal conduciveness), yet the influence of intrinsic pleasantness was as clearly readable as the influence of goal conduciveness. Such predictions can also be made in our case of a user manipulating a sounding object.

1.5 Goal and Outline of the Studies

Let us now consider the results reported by Västfjäll et al. [38] as a starting point of this study. They found that the feelings reported by listeners required to listen to aircraft sounds exhibited the following relationships: Valence was correlated with the loudness and the naturalness of the sounds (the naturalness was assessed by asking listeners to rate this aspect on a scale in a pilot experiment), and arousal was correlated with the tonality and the sharpness. Loudness and naturalness can fairly be assumed to modulate the appraisal of the sensory pleasantness of the sounds (resulting in variations of the valence of the reported feelings). If the novelty of the sounds is assumed to be the intrinsic characteristic of a sound able to influence the arousal dimension of the feelings, it is however not clear how variations of tonality and sharpness might have influenced the appraisal of the novelty of the sounds.

Just as manipulating a computer interface induces emotions in the users, it is likely that manipulating any kind of tangible artifact (kitchen utensil, vehicle, communication device, etc.), with the purpose of doing something (preparing food, traveling from one place to another, communicating with a remote person), induces emotions as well. In Scherer's framework, manipulating an interface might affect the valence of the feelings (how pleasant is the manipulation, how successfully can one reach the goal that she or he has intended with the interface), arousal (how expected are the required actions), and dominance (how can one cope with this artifact). In the specific case of computational artifacts augmented with sounds, the embedded sounds add their potential influence to the valence and arousal dimensions of the feelings. But is their influence still noticeable, compared to the potentially stronger influence of the manipulation, especially if one considers that the users were not specifically required to evaluate the sounds?

To address this issue, an interactive computational artifact was designed, called the *Flops*. It is a glass embedded with an orientation sensor allowing it to control the generation of virtual impact sounds when tilted. It implements the metaphor of a glass full of virtual items that may be poured out of it.

Two questions are here addressed: 1) Can the relationships between sound parameters (naturalness, sharpness, and tonality) and emotional reaction, like those found in [38], be generalized to another type of sounds: impacts? 2) Do sounds still influence users' feelings while they are performing tasks with a device instead of passively listening to

Fig. 1. A video showing a user using the Flops glass.

Fig. 2. Model used for the interaction.

sounds? We were, as in [38], interested in relationships common to different listeners and not in differences in individual preferences. Two experiments were conducted to address these questions: In Study 1, we observed the relationships between the feelings reported by listeners passively listening to a set of sounds associated with the Flops interface. Three sounds were selected from these results, inducing positive, neutral, and negative feelings. These sounds were used in Study 2, in which users had to interact with the Flops to perform several tasks that were more or less difficult, and we observed the contribution of the sounds and the difficulty of the manipulation to the feelings reported by the users.

2 DESIGN OF THE FLOPS GLASS

The Flops glass is a computationally augmented artifact in the shape of a plastic container. In addition to its experimental role presented in this paper, the Flops glass has been developed and exhibited as an interactive object in artistic and design contexts [53]. For the purposes of the experiments, the Flops glass was designed to contain virtual objects that can be poured out of the glass when it is tilted, generating impact sounds as they hit a virtual surface under the glass. The impact sounds were designed to test the results reported in [38] (see below). The interaction model used to generate sounds from the user's gesture is based on the metaphor of balls poured out of the Flops glass.

2.1 Physical Design

The physical interface of the Flops glass is shown in Fig. 1. Its shell was modeled in 3D software and was 3D-printed in ABS plastic. The interface contains an accelerometer (Analog Devices model ADXL 320) that is used to sense gestures performed with the glass. The sensor data are captured with a microcontroller board (Arduino BT) and transmitted wirelessly via Bluetooth to a remote computer (see Fig. 6 in the Appendix, which can be found in the Computer Society Digital Library at http://doi.ieeecomputersociety.org/ 10.1109/T-AFFC.2012.1). The sensor data processing and sound playback are managed in software (Cycling'74 Max/ MSP 5.0.6) running on the computer.

2.2 Interaction Design

The interaction model, transforming tilt angle to a flow of objects falling, is based on the model of spherical items rolling without slidin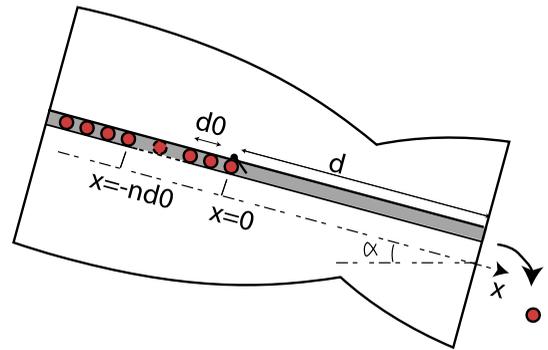g on a tilted rod, and falling when reaching the extremity of the rod (see Fig. 2). A reservoir of virtual items is situated at the inside of the Flops glass (at a distance d from the mouth of the Flops glass), where virtual items are stored (regularly separated by a distance d_0). When the Flops glass is tilted with an angle $\alpha(t)$, the position of the nth item $x_n(t)$ is determined by numerically integrating the equation

$$\ddot{x}_n(t) = \frac{2}{3} g \sin \alpha(t) - b \dot{x}_n(t), \qquad (1)$$

where g is the gravitational acceleration and b is the coefficient of fluid friction.

2.3 Sound Selection

The model described above is used to drive the generation of impact sounds: When the Flops glass is tilted, impacts are generated each time $x_n(t) = d$. Each cue triggers the playback of an impact sound. Note that this model assumed that the receptacle into which the objects fell was immediately below the lip of the glass. The motion of items sliding in the glass did not produce sound.

A set of sounds was selected to vary the features studied in [38]. Sharpness and tonality are classical psychoacoustical features for which computational descriptors are available [36]. The naturalness of a sound is a "feature" of the sound which has received little attention in the past. We have simply manipulated this aspect by using recordings of natural sounds on the one hand, and sounds synthesized with the specific purpose of sounding nonnatural on the other hand.

Thirty-two impact sounds were created. Sixteen sounds ("natural sounds") were high-quality recordings of real-impact sounds, from collisions of everyday objects to musical percussions. These recordings were extracted from data-bases of sounds used at IRCAM in various musical creations. Sixteen ("synthetic sounds") were synthesized by several different algorithms based on additive and subtractive sound synthesis. These latter were specifically created with the purpose of sounding synthetic (beeps and noises), although we did not formally test if listeners could differentiate the two categories. The creation and the selection of these sounds was made by homogeneously sampling across two psychoacoustical descriptors: spectral centroid and tonality index. These descriptors were computed using the IrcamDescriptor toolbox [54]. Spectral centroid is a descriptor that many studies have found to be correlated with the sensation of sharpness [55]. For a signal
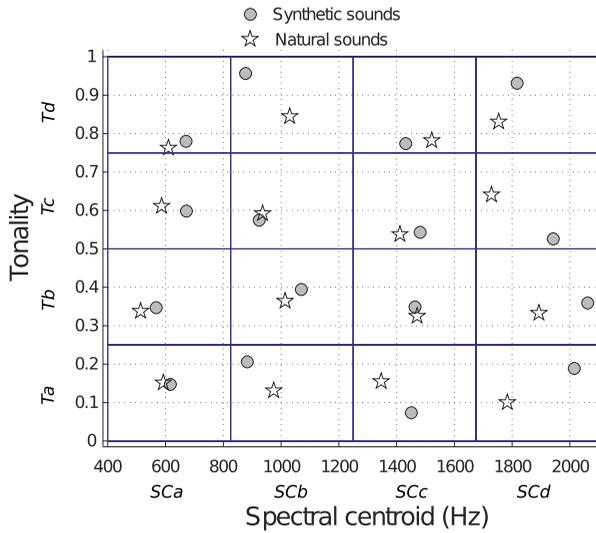
Fig. 3. The 32 sounds used in Study 1 homogeneously sample the spectral centroid/tonality space.

$x(t)$ with a Fourier transform $X(f)$, it is computed as the barycenter of the frequency bins:

$$SC = \frac{\sum X(f).f}{\sum X(f)}. \qquad (2)$$

The tonality index T is computed from the spectral flatness measure SFM, a measure of the noisiness of the sound spectrum:

$$T = min\left(\frac{log(SFM)}{-6}, 1\right), SFM = \frac{\left(\prod_1^N X(k)\right)^{\frac{1}{N}}}{\frac{1}{N}\sum_1^N X(k)}. \qquad (3)$$

The sounds were selected so as to correspond to one of the four nominal values of spectral centroid and tonality defined in Fig. 3.

The spectral centroid of the sounds varied from 410 to 1,890 Hz, and the tonality of the sounds varied from 0.07 to 0.96 (the index of tonality could theoretically vary from 0 to 1, with 0 corresponding to a white noise and 1 to a pure tone). All the sounds were created to have a rather short attack time (from 44 to 90 ms). All samples had the same duration and last approximatively 350 ms (see Table 4 in the Appendix, available in the online supplemental material).

3 STUDY 1: EMOTIONS AND SOUNDS ONLY

The experimental study reported in this section aimed at testing the assumption that the parameters used to create the Flops glass sounds (spectral centroid, tonality, naturalness) influenced the feelings reported by participants listening to these sounds. The main difference with [38] was that instead of aircraft noises, we used bursts of impacts sounds and that the sounds were accompanied by a video.

In this experiment, participants were not only listening to the sounds, but were watching videos of users manipulating the Flops glass. This was done for to facilitate a comparison with Study 2, in which participants were actually manipulating the Flops. This procedure allowed us to insure that we had neutralized a potential influence of

the visual aspect of the Flops glass (the images displayed in the videos were all the same).

To report their feelings, participants were required to use three scales of valence, arousal, and dominance. Although Västfjäll et al. [38] used only valence and arousal scales, we also used the dominance dimension to allow the comparisons with Study 2.

3.1 Method

3.1.1 Participants

Twenty-five participants (14 women and 11 men) volunteered as listeners and were paid for their participation. They were aged from 19 to 45 years old (median: 28 years old). They were selected from their age and on the basis of the Spielberger trait anxiety inventory [56] because age and anxiety have been shown to influence affective reactions [26]. Only participants with a score below 39 (low trait anxiety) were selected.

3.1.2 Stimuli

Thirty-two videos were generated, corresponding to the 32 sounds described above, with the same images of a user pouring the glass. They showed a user manipulating the Flops glass (with only the hand visible, see Fig. 1). All the videos were the same, except for the soundtrack. Each video was 8 s long. All the soundtracks had been equalized for loudness in a preliminary study (because we were not interested in this parameter).

The levels of the sounds varied from 52 to 79.9 dB(A) (median 71.5 dB(A)). The video showed a user tilting the Flops glass three times: first, slowly dropping three items out of the glass, and then tilting the Flops glass more quickly two times to increase the rate of items dropped out of the Flops glass (see Fig. 7 in the Appendix, available in the online supplemental material). A total of 28 items were dropped in each video.

3.1.3 Apparatus

The stimuli were amplified over a pair of Yamaha MSP5 loudspeakers. Participants were seated in a double-walled IAC sound-isolation booth. The experiment was run using the PsiExp v3.4 experimentation environment including stimulus control, data recording, and graphical user interface [57]. The sounds and videos were played with Cycling'74 Max/MSP/Jitter version 5.0.6. The scales were presented on an Elo Touch Screen. This allowed the participants to interact with the interfaces simply by touching the screen.

3.1.4 Procedure

The participants were first presented with a text explaining the procedure and explaining the meaning of the three scales of valence, arousal, and dominance. This text was adapted from [26]. It emphasized that the scales were to be used to report what the participant felt and not how he or she would evaluate the stimulus or the situation. Then, they were presented with a selection of eight pictures selected from the IAPS set of images [58] and required to report their feelings. Theses eight images were selected so as to cover a broad range of arousal, valence, and dominance norms, but did not make use of any erotic or violent content to keep the range of emotions comparable to what we

TABLE 1
Study 1. Correlation between the Three Scales

Scale	valence	arousal	dominance
valence	1.00	-.78**	.92**
arousal		1.00	-.76**
dominance			1.00

**p<.01/3.

expected to result from the experimental manipulations. The goal of this step was to familiarize each subject with the use and meaning of the three scales. Then, they were presented with all the videos, played one after the other, to get an overview of the range of sounds used in the study. Finally, they watched each video again, and had to report their feelings for each video.

Both for IAPS images and the videos, participants had to indicate their feelings by selecting an item on each of three 9-point scales (valence, arousal, dominance), using the Self-Assessment Manikins (SAM, [26]). SAM is a nonverbal pictorial assessment technique that directly measures the valence, arousal, and dominance associated with a person's affective reaction. During the experiment, the participants could consult a text reminding them of the meaning of each scale.

3.2   Analysis

For the eight images, the judgments averaged across the subjects were significantly correlated with the norms provided by [58] (valence: $r(6) = 0.96$, $p < 01$; arousal: $r(6) = 0.91$, $p < 0.01$; dominance, $r(6) = 0.85$, $p < 0.01$), indicating that the participants understood correctly the meaning of the three scales.

For the videos, the design had three within-subject factors (naturalness of the sounds, with two nominal values: synthetic and natural; spectral centroid, with four ordered values: SCa, SCb, SCc, SCd; and tonality, with four ordered values: Ta, Tb, Tc, Td, see Fig. 3). The dependent variables were the judgments on the valence, arousal, and dominance scales.

For the 32 sounds, the standard deviation of the judgments made by the participants varied from 1.27 to 2.2 for the valence scale, from 1.38 to 2.13 for the arousal scale, and from 1.36 to 1.91 for the dominance scale, which is consistent with the data gathered for the IAPS image sets [58]. This indicates that the feelings elicited by the sounds tended to be rather consistent across the participants. This was confirmed by computing the two-way random effect interclass correlations (ICC) between the 25 subjects across the 32 sounds. For the valence scale, ICC = 0.87 ($F(31, 744) = 7.8$, $p < 0.01/3$), for the arousal scale, ICC = 0.80 ($F(31, 744) = 5.0$, $p < 0.01/3$), and for the dominance scale ICC = .73 ($F(31, 744) = 3.7$, $p < 0.01/3$). In the following, the judgments will therefore be averaged across participants.

The judgments on three scales were significantly correlated: Valence and dominance were strongly related, and valence and arousal a little less (see Table 1). Therefore, there were systematic patterns in the judgments: The sounds causing positively valued feelings (a high value on the valence scale) tended to cause calm and dominant feelings, whereas the sounds causing feelings that were negatively

valued (a low value on the valence scale) tended to systematically cause feelings judged as excited and dominated. This was confirmed by the informal postexperimental interviews with the participants: Many participants had indeed spontaneously indicated that "shrill" sounds tended to irritate them (i.e., positively valued, excited, and dominated judgments), whereas "soft and low" sounds tended to be felt as more relaxing (negatively, calm, and in-control). To analyze the contributions of the independent factors (naturalness, spectral centroid, tonality) to the three dependent scales (valence, arousal, dominance), the data were submitted to three repeated-measure analyses of variance (ANOVA), with the degree of freedom corrected by the Geisser-Greenhouse procedure to account for possible violations of the sphericity assumption. In the following, all the probabilities are reported after this correction. The α values were corrected by the Bonferroni procedure to account for multiple comparisons. For a significance level of 0.05 and 0.01 for the family of three tests, the adjusted levels were therefore $\alpha = 0.05/3$ (0.0167) and $\alpha = 0.01/3$ (0.0033) for each individual test [59]. Because this procedures turned out to be very conservative, we have, however, considered the probability values that reached significativity without ($p < 0.05^+$, $p < 0.01^{++}$) and with adjustment ($p < 0.0167^*$, $p < 0.0033^{**}$). In the ANOVA tables, we have also reported the *experimental* η^2 values in addition to the traditional η^2 values. These values indicate the percentages of the variance only due to the experimental factors.

Table 5 in the Appendix, available in the online supplemental material, reports the results of the ANOVA for the valence scale. Naturalness and spectral centroid both had a significant effect on the judgments of valence ($F(1, 24) = 35.9$, $Exp. \eta^2 = 32.6\%$ and $F(3, 72) = 24.3$, $Exp. \eta^2 = 37.2\%$, $p < .01/3$). These factors did not interact significantly ($F(3, 72) = 0.9$, $p = 0.415$). The interaction between naturalness and tonality, as well as the three-way interaction were also significant ($F(3, 72) = 3.9$, $p < .05/3$, and $F(9, 216) = 3.8$, $p < 0.01/3$), but contributed to a lesser extent to the variance of the data ($Exp. \eta^2 = 5.7\%$ and 13.4%, respectively). Note that the interaction between spectral centroid and tonality could also be considered significant with a more permissive criterion ($F(9, 216) = 2.2$, $p < 0.05$, $Exp. \eta^2 = 6.7\%$).

The upper panel of Fig. 4 represents the influence of the spectral centroid and naturalness of the sounds on the judgments of valence. Overall, the judgments on the valence scale varied from 2.08 to 5.8 (the range of the judgments is therefore 3.1 on a scale of 9) with an average of 4.39, indicating that the participants have mainly used the center of the scale for all the videos. The judgments were rather concentrated, and skipped toward the low-valence part of the scale. The lowest values the spectral centroid led to the highest values of the valence judgments (i.e., more positively valued feelings). Natural sounds also led to larger valence judgments than the synthetic sounds.

Table 6 in the Appendix, available in the online supplemental material, reports the results of the ANOVA for the arousal scale. The spectral centroid of the sounds had the only significant principal effect on the arousal judgments ($F(3, 72) = 11.9$, $p < 0.01/3$, $Exp. \eta^2 = 42.4\%$.
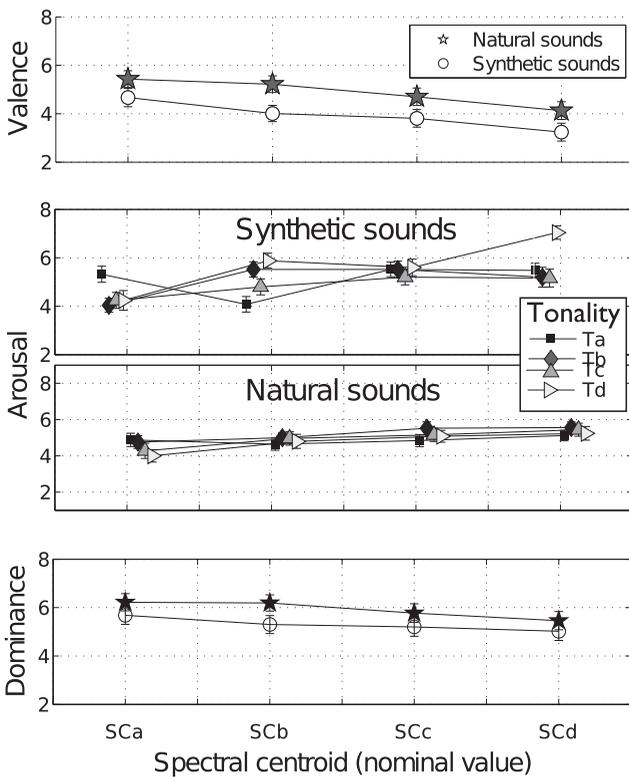
Fig. 4. Influence of the spectral centroid and of the naturalness on the sounds of the judgments of dominance. Vertical bars represent the 95 percent confidence interval.

The interactions between spectral centroid and tonality, and between naturalness and tonality were also significant $(F(9, 216) = 5.2, \eta^2 = 24.6\%$ and $F(3, 72) = 6.4, \eta^2 = 11.9\%$, $p < 0.01/3)$. Note that the three-way interaction was also almost significant $(F(9, 216) = 2.36$, $p < 0.05$, $Exp. \eta^2 = 11.8\%$. The middle panels of Fig. 4 represent the influence of these three factors on the arousal scale. The judgments of arousal were concentrated on the middle of the scale. Across all 32 sounds, the judgments of arousal varied from 4 to 7.04 (the range of the judgments was therefore 3.0 on a scale of 9), with an average of 5.06.

Overall, the arousal judgments seemed to be slightly but systematically influenced by the spectral centroids of the sounds (the higher the spectral centroid, the more aroused felt the participants). However, comparing the two middle panels of Fig. 4 shows that the influence of spectral centroid depended on the naturalness the sounds. Whereas for the natural sounds, the arousal judgments slightly increased with the value of the spectral centroid, independently of the tonality value (upper panel), the influence of spectral centroid on the arousal judgments seems to have been cluttered by the unsystematic effect of the tonality for the synthetic sounds (lower panel). This might explain why the interaction between these two factors was significant, and the nonsignificativity of the principal effect of the tonality.

Table 7 in the Appendix, available in the online supplemental material, reports the results of the ANOVA for the dominance scale. Naturalness and spectral centroid both had a significant effect on the dominance judgments $(F(1, 24) = 26.4, Exp. \eta^2 = 31.9\%$, and $F(3, 72) = 6.9$, $Exp. \eta^2 = 24.5\%$, $p < 0.01/3)$, as well as the three-way interaction between the

three factors $(F(9, 216) = 3.1$, $Exp. \eta^2 = 23.4\%$, $p < 0.05/3)$. With a less stringent criterion, the interaction between spectral centroid and tonality could also be deemed significant, but it contributed to only a small proportion of the variance $(F(9, 216) = 0.9$, $p < 0.05$, $Exp. \eta^2 = 6.5\%)$. The lower panel of Fig. 4 shows that the dominance judgments were influenced by naturalness and spectral centroid in the same way that the valence judgments were: The participants felt more dominant when hearing natural sounds with a low spectral centroid, as compared to synthetic sounds with a higher spectral centroid. However, the range of the judgments was smaller for the dominance judgments than for the valence judgments: The judgments of dominance varied from 3.96 to 6.6 (the range of the judgments was therefore 2.2 on a scale of 9) for all 32 sounds, with an average value of 5.60.

3.3 Discussion

Participants' feelings in response to the 32 sounds were distributed over a small portion of the valence-arousal-dominance space, and were centered around the neutral positions of each scale. This is not really surprising because only the sounds varied during the experiment. It could therefore not be expected that a set of videos displaying a user dropping virtual items out of a glass would cause emotional reactions comparable with those caused by sounds or images with a strong semantic content (e.g., violent images, etc.).

Only spectral centroid and naturalness directly influenced the judgments on the three scales. These observations were further confirmed when studying the correlations between the judgments on the scales and several other acoustic descriptors. We tested not only the spectral centroid measure and tonality index (which were used to create and select the sounds), but also all the descriptors contained in the IrcamDescriptor toolbox [54]. Overall, the judgments on the three scales were correlated with the many variants of the spectral centroid, the best correlations being obtained by the "perceptual" spectral centroid (this descriptor is computed as the centroid computed using the specific loudness of the Bark scale).

No scale was significantly influenced by tonality alone, even though it was assumed to be correlated with the arousal judgments. Nevertheless, tonality interacted significantly with naturalness and spectral centroid to influence the three scales. In particular, the interaction between tonality and naturalness, and tonality and spectral centroid contributed to the variance of the arousal judgments with an amount comparable to the influence of spectral centroid only. A possible explanation is that tonality is probably a relevant parameter for long and continuous sounds such as aircraft noise, but not for short impact sounds, particularly those without any kind of resonance. The significant interactions might indicate that the manipulation of the "tonality" parameter had slightly modulated the influence of other acoustic factors. However, this parameter was not relevant enough to have a main effect of its own.

The contributions of specific acoustic features to specific dimensions of the reported feelings are interesting to interpret. Both spectral centroid and naturalness influenced the valence and dominance scale, but only the spectral centroid influenced the judgments of arousal. Following Scherer's appraisal theory, the valence judgments reflected

the subjects's appraisal of the sound pleasantness. Both the spectral centroid and the naturalness therefore contributed to the pleasantness of the sounds. The arousal judgments reflected the subjects' appraisal of the novelty of the sounds. The novelty of the sounds here has to be considered as resulting from the unexpectedness of the sounds in the sequence because the participants had previously heard all the sounds. The results therefore suggest that only the spectral centroid influenced this appraisal.

The smaller but significant variations of the dominance judgments are intriguing. If we assume that the dominance judgments reflected the listeners' appraisal of their coping potential, these judgments indicate that certain sounds made the listeners feel more in control than some others. One interpretation is that of an experimental artifact. If the dominance scale was completely irrelevant to the subjects, they might have systematically mapped their dominance to the other scales. The correlations between the three scales is in line with this idea, but systematic patterns of variations in the arousal-valence space (i.e., "boomerang-shaped") have also been reported for other acoustic stimuli [22]. Furthermore, the postexperimental interviews suggested that the subjects found the dominance scale relevant. Another explanation is that the subjects were imagining themselves manipulating the Flops, and anticipated the interface to be more or less difficult depending in the different sounds. Such a possibility is tested in Study 2.

4 STUDY 2: SOUNDS, MANIPULATION, AND EMOTIONS

Study 1 demonstrated that the manipulations of sounds (spectral centroid and naturalness) resulted in relatively homogeneous appraisals of intrinsic pleasantness (artificial sounds with a high spectral centroid resulting in negative appraisal, and conversely). However, the influence of sound in such an interaction might be very different when they are actually manipulating the Flops from when passively listening to the sounds. Particularly if this interface is used to reach a goal, the manipulation is likely to also induce emotional reactions by influencing the user's appraisal of her or his goal conduciveness and coping potential. Whereas Study 1 used a "passive listening" paradigm similar to those found in the literature, the goal of the second experimental study was therefore to study the influence of the sounds on the reported feelings when the users were manipulating the Flops glass. We decided to focus on the pleasantness of the sounds. Three sounds were selected from the previous results: a sound eliciting positively valued feelings (pleasant), a sound eliciting negatively valued feelings (unpleasant), and a sound eliciting feelings with a neutral valence. Note that because of the correlation between the scales in Study 1, the sounds also elicited dominant, dominated, and neutral feelings. A game with the Flops was designed. The difficulty of the game was varied in four steps, providing another source of influence on the emotions. These four difficulty levels and three sounds resulted in 12 combinations that made it possible to compare the influence of the intrinsic pleasantness of the different sounds, and of the different levels of difficulty of the manipulation on the feelings

reported by the users (12 combinations proved to be the maximum in a reasonable amount of time). It addressed the following questions: What are the respective influences of sound pleasantness and difficulty of manipulation on the reported feelings?

4.1 The Flops Game

The Flops game was designed so as to engage the users in a task requiring them to learn how to manipulate the Flops. Furthermore, the game was specifically designed to induce emotions by providing rewards when the users won, engaging them in difficult and long tasks, requiring them to concentrate, to hurry, etc. The difficulty of the Flops game was varied so as to influence the subjects' appraisal of the goal conduciveness of the game and of their coping potential.

4.1.1 Rules

The objective of the game consisted of pouring exactly 10 virtual balls in less than 20 seconds, four times (the virtual balls were signaled by their sound). The game ended when the user had succeeded in pouring 10 balls four times. The score was the number of trials needed by the user to succeed (the best score was therefore 4—the first four trials successful—and increased when the performance decreased). The user had a maximum of 20 seconds to pour the balls; otherwise the trial was lost. The game kept on running until the users performed four successful trials. However, to prevent never-ending sessions, the game also ended when the user reached 25 trials. The worst possible score was therefore 25: 25 trials required for four successes or less.

This game therefore required the user to learn how to dynamically control the tilting of the Flops glass to pour balls at a rate that was not too quick to be able to count the falling balls and not too slow, so as to succeed in less than 20 seconds. We previously demonstrated that users can learn how to control an interface based on its sound feedback [60]. Depending on how quickly they learned to control the interface, the game could last a few trials or last long without successful trials. Note that, unlike [51] and [52], we hypothesize that varying the difficulty of the task not only manipulates the goal conduciveness (by helping or preventing users to reach the assigned goal). It also manipulates the user's appraisal of their coping potential by making the device more or less easily controllable. Because we observe the reported feelings after each trial, we can only observe the combination of these different appraisals.

4.1.2 Different Difficulty Levels

The parameters of the interaction model allowed the experimenter to set up the Flops game so as to make it more or less difficult. Here, two parameters were manipulated to provide different difficulty levels. These parameters were the length d between the reservoir of balls and the mouth of the Flops, and the distance d_0 between two balls (see Fig. 2). Increasing d increased the time required for the first ball to fall out of the Flops. Small values of d made the Flops easy to use because the balls fell out of the glass immediately when the user tilted the Flops. Larger values of d made the game more difficult because one had to wait for the balls to roll

along the rod before hearing a sound. Small values of d_0 made the game more difficult because the balls fell with a very rapid pace that might prevent a user from properly count the balls falling out [61].

To select increasing difficulty levels, a preliminary experiment was conducted with four participants. It consisted of letting participants play the Flops game at different difficulty levels, and asking them to rate the difficulty of the task. The sound used in this preliminary experiment was a percussion sample (N25) selected from the results of the previous experiment to produce medium values of valence (5.08), arousal (4.88), and dominance (6.12). Ten presets of the interaction model were created by varying d and d_0 (arbitrary units), in order to set different difficulty levels (increasing d and decreasing d_0 to increase the difficulty level). Care was taken to choose presets that let the game feasible.

These presets are reported in Table 3 in the Appendix, available in the online supplemental material, with preset 1 the easiest difficulty level and preset 10 the most difficult. The participants played 10 Flops games, each with a different interaction preset. From these results, we selected the presets 1, 4, 6, and 7, taking care to select presets for which the four participants agreed on the difficulty (for instance, preset 9 was not selected because two participants found it easy and two difficult). In the remainder of this paper, these presets will be referred to as four difficulty levels (Da to Dd).

The different sounds were a priori not supposed to objectively influence task difficulty: The different sounds resulted in sequences of impacts that were equivalently easy to count. However, the results of Study 1 might suggest that the subjects could feel differently in control of the interface, depending on the different sounds.

4.2 Experiment

4.2.1 Method

Participants. Twenty-nine participants (16 women and 13 men) volunteered as listeners and were paid for their participation. They were aged from 19 to 44 years old (median: 30.5 years old). They were selected on the basis of the Spielberger trait anxiety inventory [56].

Sounds. Three sounds were selected from the results of Study 1: N36, inducing the most positive feelings, and S165, inducing the most negative feelings, and N25 already used in the preliminary experiment (see Table 4 in the Appendix, available in the online supplemental material) and inducing a rather neutral feeling (5 is exactly the center of the valence scale). Note, however, that the judgments on the three scales obtained for this sound were closer to N36 than S165. These three sounds are referred to as, respectively, pleasant, neutral, and unpleasant.

Procedure. The preliminary step explaining the judgment scales was similar to that in Study 1. After this step, the experimenter demonstrated the Flops game, using the easiest difficulty level. Then, the participant had to play the Flops game at the easiest difficulty level and at the highest difficulty level, with the possibility of talking to the experimenter (the participants were aware that these games had the easiest and most difficult levels). This step was intended to give an overview of the range of difficulty levels to the participant

and to ensure that the participant manipulated the Flops and the experimental interface correctly.

Then, the participants performed 12 Flops games (three sounds times four difficulty levels, the order of which was randomized for each participant). After each game, they had to report their feelings by selecting an item on each of the three 9-point scales, using the Self-Assessment Manikins [26]. During the experiment, the participants could consult a text reminding them of the meaning of each scale. They also had to judge the difficulty of each game on a horizontal slider.

4.2.2 Analysis

The design had two within-subject factors (the pleasantness of the sounds, with three levels, and the difficulty level of the interaction, with four ordered levels), and five dependent variables: score, judged difficulty, valence, arousal, and dominance. The computation of the two-way random effect ICC between the 29 subjects and across the 12 games indicated that the subjects' answers were reliable (difficulty: ICC = 0.96, $F(11, 308) = 23.0$, $p < .000$; score: ICC = 0.97, $F(11, 308) = 33.6, p < 0.000$; valence: ICC = 0.94, $F(11, 308) = 15.8$, $p < 0.000$; arousal: ICC = 0.88, $F(11, 308) = 8.7$, $p < 0.000$; dominance: ICC = 0.95, $F(11, 308) = 21.4$, $p < 0.000$).

To analyze the influence of the within-subject factors on the five dependent variables, the data were submitted to five analyses of variance. In the following, all the probabilities are reported after the Geisser-Greenhouse correction. The α values were corrected by the Bonferroni procedure to account for multiple comparisons. For a significance level of 0.05 and 0.01 for the family of three tests, the adjusted levels were therefore $\alpha = 0.05/5$ (0.01) and $\alpha = 0.01/5$ (0.002) for each individual test [59]. Because these criteria turned out to be very conservative, we have, however, considered the probability values that reach significativity without and with adjustment: $p < 0.05^+$, $p < 0.01^{++}$, and $p < 0.002^*$.

Score and judged difficulty. The scores and the judgments of difficulty, averaged across all participants, were correlated: $r, (10) = 0.93 p < 0.01$, indicating that the difficulty judged by the participants was directly related to their performance. Table 8 in the Appendix, available in the online supplemental material, reports the results of the analysis of variance for the score. This analysis showed that the scores were only influenced by the difficulty levels of the game, $(F(3, 84) = 76.3, p < 0.01/5, Exp. \eta^2 = 97.6\%$, and not by the interaction between the two factors. Note, however, that the effect of the pleasantness was almost significant $(F(2, 56) = 4.7, p < 0.05)$, indicating that the pleasantness of the sounds might have very slightly influenced the scores ($Exp. \eta^2 = 1.8\%$).

Table 9 in the Appendix, available in the online supplemental material, reports the results of the analysis of variance for the difficulty of the games judged by the participants. The results found for the score were qualitatively the same as for the difficulty judgments: only the difficulty level influenced the difficulty judged by the participants ($F(3, 84) = 66.1$, $p < 0.01/5$, $Exp. \eta^2 = 94.4\%$). Together, these results indicate that almost only the variations of the parameters of the interaction model have objectively (the score) and subjectively (the judged difficulty) influenced the difficulty of the
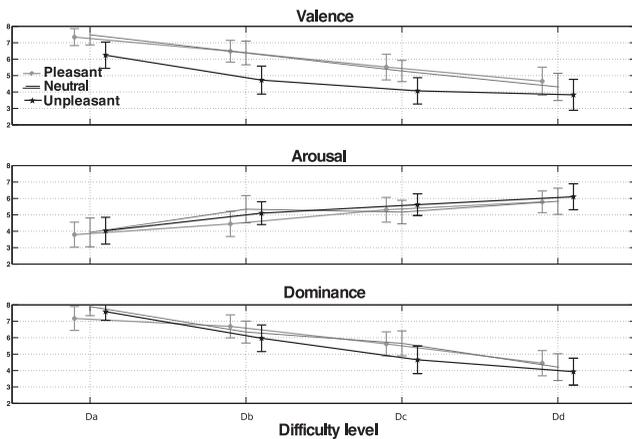
Fig. 5. Emotions during the Flops game. Vertical bars represent the 95 percent confidence intervals.

task. However, the influence of the pleasantness was close to be significant (F(2, 56) = 4.40, p < 0.05), even if the effect was small ($Exp.\ \eta^2 = 2.6\%$). The two effects did not interact (F(6, 168) = 1.3, p = 0.27).

Valence, arousal, and dominance. The average judgments of valence varied from 3.83 to 7.48 (the range was therefore 3.65 on a scale of 9), with a median of 5.40. The average judgments of arousal varied from 3.79 to 6.10 (the range is therefore 2.31), with a median of 5.24. The average judgments of dominance varied from 3.93 to 7.90 (the range was therefore 3.97), with a median of 5.81. The range of variation of the judgments along the three scales were greater than in Study 1.

The averaged judgments on the three scales were correlated (valence versus arousal, $r(10) = -0.89$, p < 0.01, valence versus dominance, $r(10) = 0.92$, p < 0.01, arousal versus dominance, $r(10) = -0.95$, p < 0.01), as in Study 1.

However, the analyses of variance for the three scales report interesting results. The results of these analyses are reported in Tables 10, 11, and 12 in the Appendix, available in the online supplemental material. The effect of the difficulty level was significant for the three scales: The more difficult the game was, the more the participants reported negative feelings (F(3, 84) = 123.9, p < 0.01/5), felt aroused (F(3, 84) = 22.4, p < 0.01/5), and dominated ($F(3, 84) = 55.8$, p < 0.01/5); the easier the game, the more the participants reported positive feelings, felt calm, and in control. For the three scales, the difficulty level accounted for the most of the variance in the data ($Exp.\ \eta^2 = 74.3\%$, 91.1%, and 93.4%).

The pleasantness levels influenced only the valence of the feelings reported by the participants (F(2, 56) = 15.78, p < 0.01/5, $Exp.\ \eta^2 = 23.1\%$), and not the arousal (F(2, 56) = 2.5, p = 0.09). The effect of pleasantness on the dominance scale was almost significant (F(2, 56) = 3.6, $p < 0.05$), but the size of the effect was very small ($Exp.\ \eta^2 = 3.0\%$). This is illustrated by Fig. 5. This figure shows that the unpleasant sounds led to more negative feelings than the pleasant or neutral sounds, but did not affect the arousal. The pleasant and neutral sounds resulted in very slightly more dominant feelings. The difference of valence between the feelings induced by the pleasant and unpleasant sounds was 1.3 on average. The difference is less than the difference that was

induced by the two sounds in Study 1 (passive listening): 3.7. Note that the feelings induced by the pleasant and neutral sounds were confounded, which is not surprising as they had induced similar valence in Study 1.

4.3 Discussion

Table 2 compares the results of the two studies. The first notable difference is that the range of judgments was larger in Study 2 than in Study 1. This is in particular true for the dominance judgments, for which the judgments extended in Study 2 over a range almost twice as large as in Study 1. This can be attributed to the participants' being in control of the interface in Study 2.

The influence of the difficulty of the task is therefore consistent with the prediction that the dominance judgments account for the appraisal of the users's goal conduciveness (reflected by the valence dimension) and coping potential (reflected by the dominance dimension). That the successes or failures of the game also influenced the arousal dimension has been shown in the case of video games [50]: In the case of the Flops game, the most difficult settings forced the participants into long, demanding, and unsuccessful trials. It is not surprising that they felt particularly aroused in these cases.

The influence of the sounds on the valence judgments remained important in Study 2, though smaller than that of the difficulty levels. On the contrary, the influence of the sounds on the arousal judgments completely disappeared in Study 2: Only the difficulty of the task influenced this scale.

It is interesting to note that the sounds had a slight but significant (with a forbearing criterion) effect in Study 2 on both the scores and the difficulty judged by the participants. The different sounds used in Study 2 also had a tiny effect on the dominance judgments. This suggests that the different sounds also modified the ability of the users to successfully interact with the interface: Unpleasant sounds made the interface slightly more difficult to use than pleasant sounds. This in turn might also explain the significant influence of the sound parameter in Study 1 on the dominance scale: The subjects probably anticipated this effect.

These results can be summarized by saying that the *aesthetics* of the sounds mainly influenced the valence of the emotions, whereas the *functionality* of the device also influenced the arousal and the dominance dimensions of the feelings of the users (this result is in line with the position of Scherer [62]). However, the aesthetics of sounds also had a minor influence on the objective and subjective difficulty of the task, and thus contributed to the dominance judgments.

TABLE 2
Comparison between Studies 1 and 2

Scale		Valence		Arousal		Dominance	
Study		1	2	1	2	1	2
Range		3.1	3.6	3.0	3.8	2.2	4.0
N		33%**		3%		32%**	
SC	P	37%**	23%**	42%**	4%	24%**	3%+
T		3%		4.4%		4%	
-	DL	-	74%**	-	91%**	-	93%**

5 GENERAL DISCUSSION

The purpose of the study reported in this paper was to examine how the sounds of computationally augmented artifacts affect users' emotions. The idea developed here is that it is important to consider that the sounds of objects never occur alone (a subject has to interact with them), and that subjects are rather seldom paying close attention to them. We therefore raised the issue of the generality of the sound/emotions relationships found in the literature for sounds occurring as the consequence of a user interacting with a manufactured object or a tangible interface. A radical hypothesis might be that, when occurring in situations overwhelmed by many potentially stronger sources of emotion elicitation, the sounds, if barely noticed, may not cause any emotions.

Because the emotions induced by the aesthetics of a sonic feedback and by the functionality of an interactive object appear very unlikely to be qualitatively equivalent, it was necessary to consider different potential mechanisms of emotion elicitation. Scherer's appraisal theory therefore provided us with a particularly well-adapted theoretical framework, clearly distinguishing those emotions resulting from the appraisal of the intrinsic pleasantness or novelty of a sound, and those resulting from the appraisal of the goal conduciveness and coping potential of the user manipulating the object.

We first manipulated the features of a set of impact sounds to test the relationships between these features and the subjects' feelings, and compared them to previously reported relationships [38]. Impact sounds are essentially different from the continuous aircraft sounds used in [38] because of their very short duration. This practically makes properties such as pitch and tonality irrelevant, particularly for the impacts with no resonance. However, the properties common between impact and aircraft sounds influenced the reported feelings. Naturalness and sharpness clearly influenced the valence dimension: Participants reported more positively valued feelings for natural than for synthetic sounds, and negatively valued feelings for the sharper sounds. Sharpness also influenced the arousal dimension. More surprisingly, naturalness and sharpness also influenced the dominance dimension. Even if the possibility of an experimental artifact cannot be completely ruled out, an interesting interpretation is that the subjects anticipated pleasant sounds to make the manipulation easier than unpleasant sounds.

In Study 2, we used a paradigm inspired by van Reekum et al. [51] and [52]. Instead of passively listening to the sounds, the participants were engaged in game using the Flops interface. The results showed that, when embedded in a game in which the successes and failures also create emotions in the participants, the sounds mainly influenced the valence dimension, whereas the difficulty of the game also influences the dominance and arousal dimensions. In the account of Scherer [62] and Grandjean and Scherer [21] of the emotional processes, the intrinsic pleasantness of the stimuli may be appraised early, on an unconscious, automatic, and possibly schematic level of processing, and only influence the valence of the reported feelings, whereas the ability to achieve the task required by the game (goal conduciveness) tends to be evaluated later in the sequence,

on a conscious, controlled, and propositional level, and influences the valence and the dominance dimension of the reported feelings. The ability to deal with the outcomes of the situation (coping potential) also influences the dominance dimension. Therefore, our results are very consistent with this account. The difficulty of the game, influencing goal conduciveness and participants' coping potential, had a clear effect on the valence and dominance dimensions of the reported feelings. The pleasantness of the sounds had a robust effect on the valence dimension, though smaller. It is important to note that the effect of the sound pleasantness and of the task difficulty on the valence dimension did not interact, which strengthens the idea of different appraisals coming into play. The influence of the task difficulty on the arousal dimension is less easy to picture because we did not introduce a specific variable to affect this dimension. However, it has been suggested that the unexpectedness of the action required to deal with a situation would affect the arousal [29]. In our case, more difficult settings resulted in longer unsuccessful sessions and puzzled the participants, who had to figure out how to control a very unstable interface. Together with the more intense physical engagement required, the unexpectedness of the required control might therefore explain the differences of reported arousal.

Besides these large effects, it is interesting to note that the pleasantness of the sounds had a small but persistent influence on the objective and subjective difficulty of the task, and on the dominance judgments in both Studies 1 and 2. This effect was not expected because we thought the different sounds could not change the difficulty of the task. However, theoreticians of product design have argued that "what is beautiful is usable" [33]. Furthermore, we have recently reported that users tend to find that an interface is more pleasant and more usable when natural sounds are used as auditory feedbacks, compared to arbitrary sounds [63].

From an applied perspective, an interesting conclusion is that the sharpness, measured by the spectral centroid, and the naturalness of the sounds have a great influence on the emotions felt by the users. This influence appears to be robust to the change of context and perception introduced by the manipulation. These features therefore appear to be strong predictors of the users' emotions to sounds, independently of whether he or she is solely passively listening to the sounds or generating the sounds when interacting with an object: high spectral centroids (and/or synthetic sounds) induce negative emotions, and low spectral centroids (and/or natural sounds) induce positive emotions. The influence of sharpness was reported for passive listening of aircraft sounds [38], and is also consistent with the results reported by Kumar et al. [64], who showed a relationship between unpleasantness of a set of sounds in the frequency region between 2,550 and 5,000 Hz. This is also consistent with the sounds chosen by van Reekum et al. [51], [52]. And it is, of course, consistent with the definition of Zwicker's sensory pleasantness [36]. Furthermore, the influence of sharpness on valence judgments has been reported for vocal expression [65], [66], musical expression [67], and synthesized musical sounds [68]. Sharpness is, in fact, a very important dimension of the timbre of several musical instruments and everyday sounds [55]. Naturalness also seems to be an important

determinant in the appraisal of the sounds of everyday objects [63], [69]. However, it can be considered to be a high-level auditory attribute that cannot be indexed by a few psychoacoustical descriptors.

If sounds only influence the pleasantness of an interface, this aspect is, nonetheless, very important for the aesthetics of computationally augmented artifacts, whereas, though the sharpness (and other low-level sound features) can be easily measured and manipulated, the question of how to handle naturalness remains open.

ACKNOWLEDGMENTS

REFERENCES

[1] R. Schirmacher, R. Lippold, F. Steinbach, and F. Walter, "Praktische Aspekte Beim Einsatz Von ANC-Systemen in PKW," 2007.

[2] D. Norman, "Emotion and Design: Attractive Things Work Better," *Interactions,* vol. 9, no. 4, pp. 36-42, 2002.

[3] D. Norman, *The Design of Everyday Things.* Basic Books, 2002.

[4] D. Norman, *Emotional Design: Why We Love (or Hate) Everyday Things.* Basic Civitas Books, 2004.

[5] P. Desmet, "A Multilayered Model of Product Emotions," *The Design J.,* vol. 6, no. 2, pp. 4-13, 2003.

[6] F. Spillers, "Emotion as a Cognitive Artifact and the Design Implications for Products that Are Perceived as Pleasurable," *Proc. Fourth Conf. Design and Emotion '04,* 2004.

[7] N. Tractinsky and D. Zmiri, "Exploring Attributes of Skins as Potential Antecedents of Emotion in HCI," *Aesthetic Computing,* pp. 405-422, MIT Press, 2006.

[8] A. Rafaeli and I. Vilnai-Yavetz, "Instrumentality, Aesthetics and Symbolism of Physical Artifacts as Triggers of Emotion," *Theoretical Issues in Ergonomics Science,* vol. 5, no. 1, pp. 91-112, 2004.

[9] R.W. Picard, *Affective Computing.* The MIT Press, 2000.

[10] M. Chion, C. Gorbman, and W. Murch, *Audio-Vision: Sound on Screen.* Columbia Univ. Press, 1994.

[11] D. Blumstein, R. Davitian, and P. Kaye, "Do Film Soundtracks Contain Nonlinear Analogues to Influence Emotion?" *Biology Letters,* vol. 6, pp. 751-754, 2010.

[12] A. Ortony and T.J. Turner, "What's Basic about Basic Emotions?" *Psychological Rev.,* vol. 97, no. 3, pp. 315-331, 1990.

[13] J.A. Russell and L.F. Barrett, "Core Affect, Prototypical Emotional Episodes, and Other Things Called Emotion: Dissecting the Elephant," *J. Personality and Social Psychology,* vol. 76, no. 5, pp. 805-819, 1999.

[14] J.A. Russell, "Core Affect and the Psychological Construction of Emotion," *Psychological Rev.,* vol. 110, no. 1, pp. 145-172, 2003.

[15] J.A. Russell, J.-A. Bachorowski, and J.-M. Fernández-Dols, "Facial and Vocal Expressions of Emotions," *Ann. Rev. of Psychology,* vol. 54, pp. 329-349, 2003.

[16] K.R. Scherer, "What Are Emotions? And How Can They Be Measured?" *Social Science Information,* vol. 44, no. 4, pp. 695-729, 2005.

[17] P.N. Juslin and D. Västfjäll, "Emotional Responses to Music: The Need to Consider Underlying Mechanisms," *Behavioral and Brain Sciences,* vol. 31, pp. 559-621, 2008.

[18] K.R. Scherer, "The Dynamic Architecture of Emotion: Evidence for the Component Process Model," *Cognition and Emotion,* vol. 23, no. 7, pp. 1307-1351, 2009.

[19] K.R. Scherer, "Appraisal Theory," *Handbook of Cognition and Emotions,* chapter 30, T. Dalgleish and M. Power, eds., pp. 637-663, John Wiley and Sons, 1999.

[20] P.C. Ellsworth and K.R. Scherer, "Appraisal Processes in Emotion," *Handbook of Affective Sciences,* chapter 29, R.J. Davidson, K.R. Scherer, and H.H. Goldsmith, eds., pp. 572-595, Oxford Univ. Press, 2003.

[21] D. Grandjean and K.R. Scherer, "Unpacking the Cognitive Architecture of Emotion Processes," *Emotion,* vol. 8, no. 3, pp. 341-351, 2008.

[22] M.M. Bradley and P.J. Lang, "Affective Reactions to Acoustic Stimuli," *Psychophysiology,* vol. 37, pp. 204-215, 2000.

[23] H. Schlosberg, "Three Dimensions of Emotion," *The Psychological Rev.,* vol. 61, no. 2, pp. 81-88, 1954.

[24] C.E. Osgood, "Dimensionality of the Semantic Space for Communication via Facial Expressions," *Scandinavian J. Experimental Psychology,* vol. 7, pp. 1-30, 1966.

[25] A. Mehrabian, "A Semantic Space for Nonverbal Behavior," *J. Consulting and Clinical Psychology,* vol. 35, no. 2, pp. 248-257, 1970.

[26] M.M. Bradley and P.J. Lang, "Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential," *J. Behavior Therapy and Experimental Psychiatry,* vol. 25, pp. 49-59, 1994.

[27] J.A. Russell, "A Circumplex Model of Affect," *J. Personality and Social Psychology,* vol. 39, no. 6, pp. 1161-1178, 1980.

[28] J.R.J. Fontaine, K.R. Scherer, E.B. Roesch, and P.C. Ellsworth, "The World of Emotions Is Not Two-Dimensional," *Psychological Science,* vol. 18, no. 12, pp. 1050-1057, 2007.

[29] K.R. Scherer, E.S. Dan, and A. Flykt, "What Determines a Feeling's Position in Affective Space? A Case for Appraisal," *Cognition and Emotion,* vol. 20, no. 1, pp. 92-113, 2006.

[30] D. Västfjäll and M. Kleiner, "Emotion in Product Sound Design," *Proc. les Journées du Design Sonore,* 2002.

[31] G. Lemaitre, P. Susini, S. Winsberg, B. Letinturier, and S. McAdams, "The Sound Quality of Car Horns: A Psychoacoustical Study of Timbre," *Acta Acustica United with Acustica,* vol. 93, no. 3, pp. 457-468, 2007.

[32] D.A. Norman, *Emotional Design: Why We Love (or Hate) Everyday Things.* Basic Books, 2004.

[33] N. Tractinsky, A. Katz, and D. Ikar, "What Is Beautiful Is Usable," *Interacting with Computers,* vol. 13, pp. 127-145, 2000.

[34] R. Guski, I. Felscher-Suhr, and R. Schuemer, "The Concept of Noise Annoyance: How International Experts See It," *J. Sound and Vibration,* vol. 223, no. 4, pp. 513-527, 1999.

[35] D. Västfjäll, M.-A. Gulbol, M. Kleiner, and T. Gärling, "Affective Evaluations of and Reactions to Exterior and Interior Vehicle Auditory Quality," *J. Sound and Vibration,* vol. 255, no. 3, pp. 501-518, 2002.

[36] E. Zwicker and H. Fastl, *Psychoacoustics Facts and Models.* Springer Verlag, 1990.

[37] D. Västfjäll, T. Gärling, and M. Kleiner, "Does It Make You Happy Feeling This Way? A Core Affect Account of Preference for Current Mood," *J. Happiness Studies,* vol. 2, pp. 337-354, 2001.

[38] D. Västfjäll, M. Kleiner, and T. Gärling, "Affective Reactions to Interior Aircraft Sounds," *Acta Acustica United with Acustica,* vol. 89, pp. 693-701, 2003.

[39] D. Västfjäll, "Contextual Influences of Sound Quality Evaluation," *Acta Acustica United with Acustica,* vol. 90, pp. 1029-1036, 2004.

[40] P. Susini, S. McAdams, S. Winsberg, I. Perry, S. Vieillard, and X. Rodet, "Characterizing the Sound Quality of Air-Conditioning Noise," *Applied Acoustics,* vol. 65, no. 8, pp. 763-790, 2004.

[41] D. Fallman and J. Waterworth, "Dealing with User Experience and Affective Evaluation in HCI Design: A Repertory Grid Approach," *Proc. Computer-Human Interface Workshop Evaluating Affective Interfaces,* Apr. 2005.

[42] B. Cahour, P. Salembier, C. Brassac, J.L. Bouraoui, P. Pachoud, P. Vermersch, and M. Zouinar, "Methodologies for Evaluating the Affective Experience of a Mediated Interaction," *Proc. Computer-Human Interface Workshop Evaluating Affective Interfaces,* 2005.

[43] C. Swindells, K.E. MacLean, K.S. Booth, and M. Meitner, "A Case-Study of Affect Measurement Tools for Physical User Interface Design," *Proc. Conf. Graphic Interface,* 2006.

[44] C. Wiberg, "Affective Computing versus Usability," *Proc. Computer-Human Interface Workshop Evaluating Affective Interfaces,* 2005.

[45] R.W. Picard and S.B. Daily, "Evaluating Affective Interactions: Alternatives to Asking What Users Feel," *Proc. Computer-Human Interface Workshop Evaluating Affective Interfaces,* 2005.

[46] S. Kaiser, T. Werle, and P. Edwards, "Multimodal Emotion Measurement in an Interactive Computer-Game: A Pilot-Study," *Proc. Eighth Conf. Int'l Soc. for Research on Emotions,* N.H. Frijda, ed., 1994.

[47] A. Kappas, "BRIEF REPORT Don't Wait for the Monsters to Get You: A Video Game Task to Manipulate Appraisals in Real Time," *Cognition and Emotion,* vol. 13, no. 1, pp. 119-124, 1999.

[48] T. Johnstone, C.M. van Reekum, T. Bänziger, K. Hird, K. Kirsner, and K.R. Scherer, "The Effects of Difficulty and Gain versus Loss on Vocal Physiology and Acoustics," *Psychophysiology,* vol. 44, pp. 827-837, 2007.

[49] N. Ravaja, T. Saari, J. Laarni, K. Kallinen, M. Salminen, J. Holopainen, and A. Järvinen, "The Psychophysiology of Video Gaming: Phasic Emotional Response to Game Events," *Proc. Digitial Game Research Assoc. Conf.: Changing Views—World in Play,* 2005.

[50] N. Ravaja, M. Turpeinen, T. Saari, S. Puttonen, and L. Keltikangas-Järvinen, "The Psychophysiology of James Bond: Phasic Emotional Responses to Violent Video Games," *Emotions,* vol. 8, no. 1, pp. 114-120, 2008.

[51] C. van Reekum, T. Johnstone, R. Banse, A. Etter, T. Wehrle, and K.R. Scherer, "Psychophysiological Responses to Appraisal Dimensions in a Computer Game," *Cognition and Emotion,* vol. 18, no. 5, pp. 663-688, 2004.

[52] T. Johnstone, C.M. van Reekum, K. Hird, K. Kirsner, and K.R. Scherer, "Affective Speech Elicited with a Computer Game," *Emotion,* vol. 5, no. 4, pp. 513-518, 2005.

[53] M.-H. Lemaire, "Blurred and Playful Intersections: Karmen Franinović's flo)(ps," *Wi: J. Mobile Media,* vol. 10, http://wi.hexagram.ca/?p=51, 2009.

[54] G. Peeters, "A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project," Institut de Recherche et de Coordination Acoustique Musique (IRCAM), Paris, France, Cuidado Projet Report, 2004.

[55] N. Misdariis, A. Minard, P. Susini, G. Lemaitre, S. McAdams, and E. Parizet, "Environmental Sound Perception: Meta-Description and Modeling Based on Independent Primary Studies," *EURASIP J. Speech, Audio, and Music Processing,* vol. 2010, article 6, 2010.

[56] R.B. Cattell and I.H. Scheier, *The Meaning and Measurement of Neuroticism and Anxiety.* Ronald, 1961.

[57] B.K. Smith, "PsiExp: An Environment for Psychoacoustic Experimentation Using the IRCAM Musical Workstation," *Proc. Soc. for Music Perception and Cognition Conf.,* 1995.

[58] P.J. Lang, M.M. Bradley, and B.N. Cuthbert, "International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual," technical report A-7, Univ. of Florida, 2008.

[59] H. Abdi, "Bonferroni and Šidák Corrections for Multiple Comparisons," *Encyclopedia of Measurement and Statistics,* N. Salkind, ed., pp. 103-107, 2007.

[60] G. Lemaitre, O. Houix, Y. Visell, K. Franinović, N. Misdariis, and P. Susini, "Toward the Design and Evaluation of Continuous Sound in Tangible Interfaces: The Spinotron," *Int'l J. Human-Computer Studies,* special issue on sonic interaction design, vol. 67, pp. 976-993, 2009.

[61] P.G. Cheatham and C.T. White, "Temporal Numerosity: III. Auditory Perception of Number," *J. Experimental Psychology,* vol. 47, no. 6, pp. 425-428, 1954.

[62] K.R. Scherer, "Which Emotions Can Be Induced by Music? What Are the Underlying: Mechanisms? And How Can We Measure Them?" *J. New Music Research,* vol. 33, no. 3, pp. 239-251, 2004.

[63] P. Susini, N. Misdariis, G. Lemaitre, and O. Houix, "Naturalness Influences the Perceived Usability and Pleasantness of an Interface's Sonic Feedback," *J. Multimodal User Interfaces,* submitted for publication, 2011.

[64] S. Kumar, H.M. Forster, P. Bailey, and T.D. Griffiths, "Mapping Unpleasantness of Sounds to Their Auditory Representation," *J. Acoustical Soc. of Am.,* vol. 124, no. 6, pp. 3810-3817, 2008.

[65] K.R. Scherer, R. Banse, H.G. Wallbott, and T. Goldbeck, "Vocal Cues in Emotion Encoding and Decoding," *Motivation and Emotion,* vol. 15, pp. 123-148, 1991.

[66] K.R. Scherer, T. Johnstone, and G. Klasmeyer, "Vocal Expression of Emotion," *Handbook of Affective Sciences,* chapter 23, R.J. Davidson, K.R. Scherer, and H.H. Goldsmith, eds., pp. 433-456, Oxford Univ. Press, 2003.

[67] P. Juslin and R. Timmers, "Expression and Communication of Emotion in Music Performance," *Handbook of Music and Emotion: Theory, Research, and Applications,* P. Juslin and J. Sloboda, eds., pp. 453-489, Oxford Univ. Press, 2010.

[68] K.R. Scherer and J.S. Oshinsky, "Cue Utilization in Emotion Attribution from Auditory Stimuli," *Motivation and Emotion,* vol. 1, no. 4, pp. 331-346, 1977.

[69] P. Susini, N. Misdariis, O. Houix, and G. Lemaitre, "Does a 'Natural' Feedback Affect Perceived Usability and Emotion in the Context of Use of an ATM?" *Proc. Sound and Music Computing Conf.,* 2009.

Guillaume Lemaitre received the PhD degree in acoustics from the Université du Maine, Le Mans, France, in 2004. He is a postdoctoral researcher at the Università Iuav di Venezia, Italy. His research interests concern the perception of everyday sounds, and how listeners can identify and interpret the cause of the sounds. He is in particular interested in the interactions between sound perception and action execution. The results of this research apply to the development of evaluation methods for the design of sounds of everyday objects and computer interfaces. He has worked with the Auditory Lab at Carnegie Mellon University, Pittsburgh, Pennsylvania, the French National Institute for Computer Science (INRIA), and Ircam, where he was involved in the European projects and actions CLOSED, SID, and MINET.

Olivier Houix received the PhD degree in acoustics in 2003 from the Université du Maine, Le Mans, France. His research interests concern the perception of environmental sounds and the gesture-sound relationship applied to sound design. He teaches audio engineering and psychoacoustics. He is a member of the French Acoustical Society. He has been involved in the Psychology Department of the University of Paris 8 and collaborates regularly with Ircam, where he was involved in national and European projects such as CLOSED.

Patrick Susini received the PhD degree in acoustics in 1999. He has been the head of the Sound Perception and Design team at Ircam since 2004. His research activities include auditory perception and cognition of everyday sounds, sound quality, loudness of nonstationary sounds, and sonic interaction design. He organized the First and the Second International Sound Design Symposiums in Paris. He teaches psychoacoustics and environmental acoustics. He is an active member of the French Acoustical Society. He has coordinated the European project CLOSED.

Yon Visell received the PhD degree in electrical and computer engineering from McGill University, Canada. He is a postdoctoral researcher at the Institut des Systèmes Intelligents et de Robotique, Université Pierre et Marie Curie. His research concerns the engineering of haptic and auditory interfaces, multisensory contributions to the perception action. He has served as a Canadian national delegate to the ESF COST Action on Sonic Interaction Design, and was a local coordinator for the European project CLOSED at the University of Applied Arts, Zurich. He conducted research on automatic speech recognition at Loquendo, and on sonar signal processing at ARL, Austin, Texas. He was employed at Ableton AG, Berlin, Germany, where he led audio signal processing development for award-winning computer music software. In 2004, he founded the art-technology organization Zero-Th, Pula, Croatia.

Karmen Franinović received the Laurea degree (summa cum laude) from the Istituto Universitario di Architettura di Venezia, the master's degree from the Interaction Design Institute Ivrea, and the PhD degree in 2011 from the School of Computing, Communications, and Electronics at the University of Plymouth. She teaches and conducts research in the Interaction Design program at Zurich University of the Arts. She focuses on sonic and embodied interaction, and on the ways these can affect phenomenological experience and social dynamics. She has been leading research projects at the European, Swiss and local levels. She is the Swiss delegate for the ESF COST Action on Sonic Interaction Design and the Action on Soundscape of European Cities and Landscapes. Previously, she worked as an architect on large scale public projects at AltenArchitekten, Berlin, and Studio ArchA, Torino. Since 2004, she has been the director of the Zero-Th organization whose works have been exhibited internationally.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.