A SOUND-BASED PROTOCOL TO STUDY THE EMOTIONS ELICITED BY PRODUCT APPEARANCE

Weihua Lu, Jean-François Petiot
LUNAM, Ecole Centrale de Nantes, Institut de Recherche en Communications et Cybernétique de Nantes (UMR CNRS 6597), France.

ABSTRACT
Emotions influence how a customer interacts with the product. To be able to instill emotional value in product design, the understanding of user emotions and the measure of emotions are interesting challenges. Several measuring methods use visual stimuli as assessment scales. Until recently, hearing was an ill-explored part for emotion measurement. This paper describes a new protocol based on sounds for eliciting user emotions. The method uses a set of sounds and association tests, made by a panel of participants. The same objects, cars pictures, were evaluated by two user-tests based on both this new protocol and a classical protocol, the Semantic Differential. We describe in the paper the main stages of the new method, and we compare the results with the Semantic Differential Method using Principal Component Analysis and Generalized Procrustes Analysis. The new protocol seems to be a useful means to collect the intuitive emotion of users.

Keywords: emotional design, non-verbal assessment, Semantic Differential Method, Generalized Procrustes Analysis, Multidimensional Scaling

1 INTRODUCTION
Implicit in many business strategies is that consumers come to their decisions with an objective and economic-based reasoning. In our modern world, this thinking is rapidly unraveling. As markets become flooded with products, consumers are overwhelmed with choices. More and more businesses are realizing that to gain a competitive edge, they must win the customers' hearts as well as their minds. The success of a product in the marketplace is not only determined by its aesthetic appeal, the pleasure it creates, but also by the satisfaction it brings to the user. In an effort to maximize their market potential, companies have to satisfy the user's inner emotions to provide an increased likelihood of product purchase. Obviously, design for emotion plays a more and more important role in product design [1].
Design to evoke and control emotions becomes much more effective than the design focusing on usability. This is due to the affinity the user feels for an object that attracts him/her, due to the formation of an emotional connection with the object. This design style is called Emotional Design [2], which is both the title of a book by Donald Norman and the concept it represents. The main issue is that emotions have a crucial role in the human ability to understand the world, and how people learn new things. Norman states three levels of emotions, labeled "visceral" , "behavioral" , "reflective" , and asserts that "a successful design has to excel at all levels" .
Emotions are defined as an acute and intentional state, which involves a relationship between the person and the stimulus. A difficulty of affective concepts such as emotions is that they are probably as intangible as they are appealing. How to understand and measure user emotions have attracted the attention of many researchers [3]. There are different approaches to reach this goal. Many researches on emotion measurement are based on verbal methods or on visual stimulation. A typical verbal method, the Semantic Differential [4], consists in listing attributes of the product to analyze, and to carry out user-tests in which the subject must assess the product according to these attributes. The Self-Assessment Manikin [5] is a non-verbal pictorial assessment technique that directly measures the pleasure, arousal, and dominance associated with a person's affective reaction to a wide variety of stimuli [6]. In the same category, the Product Emotion Measure *(PrEmo)* [7] is based on visual pictures. These methods are based on the assessment of visual stimulation, and few are concerned with other sensory modalities. Most people agree that auditory sensations, sounds and music, can arouse profound and deep emotional reaction [8].

This research focuses on a new protocol using auditory stimuli to elicit user emotions, the Auditory Parameters Method. The objectives of this paper are to describe this new protocol based on sounds, and to compare it with the classical Semantic Differential Method. In this paper, we describe the two user-tests we carried out, applied to the perception of cars, given by their pictures. In section 2, we present the description of materials and methods, and we describe the application which concerns the appearance of cars. Section 3 presents an analysis of the results. Conclusions and perspectives are drawn in section 4.

2 MATERIALS AND METHODS

2.1 Description of the application

The products chosen to describe our work are cars. Such a familiar product is suitable for user-tests because they are very common and most of us have user experience with it. This example has to be considered as a pilot study to illustrate the method.

To build the product space, we searched several different images of cars on the Internet (more than 100). Next, we selected 10 cars with outstanding characteristics (from our point of view) both in form and color. In order to provide a sufficiently informative picture to the subjects, we chose images with a shot angle showing three profiles of the car (the face, the side and the roof). In an attempt to focus the subjects' attention only on the car itself, we removed the background of the picture with an image processing software. The same 10 car samples (shown in Figure 1) were used for two assessment tests, the Semantic Differential Method and the Auditory Parameters Method.

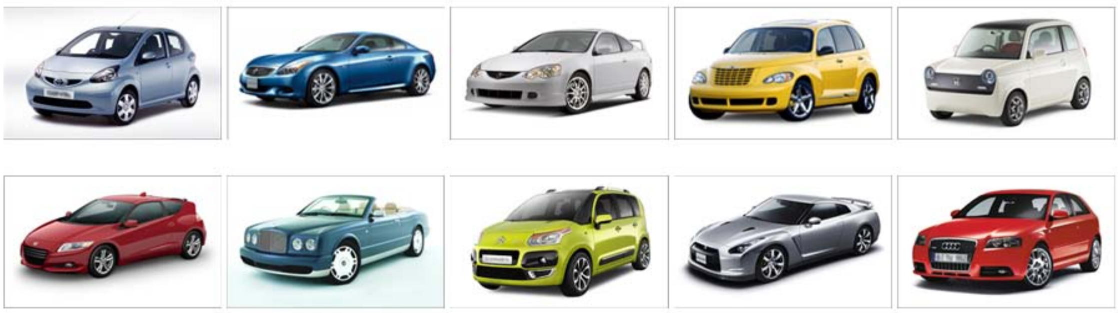


Figure 1. Pictures of the 10 cars used for the tests

2.2 The Semantic Differential Method (SD)

2.2.1 Description of the method

Semantic Differential Method (SD) [4] consists of defining a list of semantic attributes, and carrying out user-tests in which the user must assess the product on measurement scales. The attributes are often defined by pairs of antonymous adjectives, which lie at either end of a seven point qualitative scale. A semantic space, Euclidean and multidimensional, is then postulated. Factor Analysis and Principal Components Analysis may be used to reduce the dimensions of the space and to find the underlying dimensions. They are used for the analysis of families of products or for the detailed analysis of a particular product.

2.2.2 Selection of semantic attributes

The choice of the semantic attributes has been made by analyzing previous papers in Kansei Engineering [9], on the perception of cars [10]. Ten French adjectives were finally selected (shown in Table 1), and their translation in English is given for information.

2.2.3 Semantic Differential Test

The Semantic Differential Test (SD Test) was carried out according to the conventional SD method. The interface is illustrated in Figure 2 (assessment of the product according to an attribute on an unstructured scale). The subjects were asked to rate each car on the semantic scales from “not at all” to “very much” on their opposing ends. The subjects were supposed to express intuitively their assessments according to the attributes. The matrix V_{SD} of the average value of the assessment was computed ($m*p$ matrix, m refers to the number of cars of the product space, p is the number of semantic attributes).

Table 1. List of the semantic attributes

No.	Adjectives	No.	Adjectives
1	Mignonne (Cute)	6	Moderne (Modern)
2	Sportive (Sporty)	7	Robuste (Robust)
3	Classique (Classic)	8	Spacieuse (Spacious)
4	Formelle (Formal)	9	Jeune (Young)
5	Puissante (Powerful)	10	Luxueuse (Luxurious)

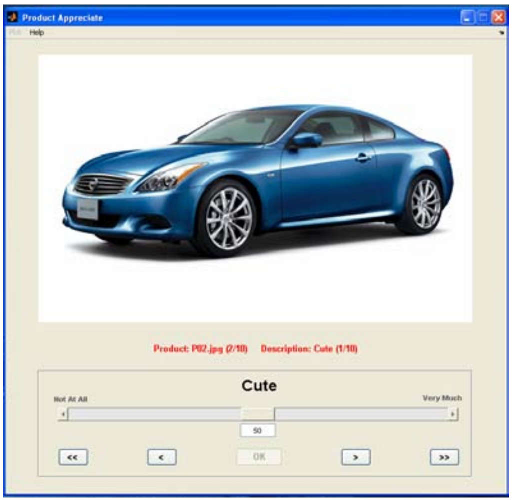


Figure 2. Interface of the SD Test

2.3 The Auditory Parameter Method (AP)

2.3.1 Description

The Auditory Parameter Method (AP) is the new protocol that we propose for products assessments. It uses sounds to elicit user emotions and to develop an emotion measurement. This method is inspired by the Kansei Parameter Method (KP) [11], developed by Prof. Kashiwazaki in Tokyo Denki University. The KP Method is a non-verbal technique for evaluating a set of stimuli (e.g. objects or perfumes), in order to obtain evaluations of these stimuli according to a set of variables. The principle of the method is based on an association test: given a stimulus, the subject is asked to select, among a set of proposed figures, the "most representative figure of the stimulus". The selections have to be in accordance with the emotion evoked by the stimuli. Originally, the KP method involves the vision sense. We adapt this method to the hearing and to the use of sounds to represent the stimuli. Compared to vision, hearing possesses particular features that we must take into account (for example sounds are embedded in time, not pictures). In order to make sure that the auditory stimulations are not over the subjects' cognitive load, we limit the number of proposed sounds to three.

Given a set of sounds (cf. § 2.3.2 for their description), and a set of objects of the product space, the experiment is based on the following stages:

1. Presentation to the subject of one object of the product space (picture),
2. Random selection by the algorithm of 3 sounds among the set of sounds,
3. After hearing the 3 sounds, the subject is asked to select the most "representative" sound of the object.

After this, the experimental protocol makes several iterations of stage 2 and 3, and proposes to the evaluation all the objects of the product space. After running this test with *s* subjects, we counted the number of times a given sound has been associated to a given product. This is described by the matrix *F*:

$$F=\begin{pmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & \ddots & \vdots \\ f_{m1} & \cdots & f_{mn} \end{pmatrix}$$

F: $m*n$ matrix, m refers to the number of objects of the product space, n is the number of sounds

f_{ij}: frequency ratio of the association of sound j to object i (number of time the sound is selected/number of time the sound is presented).

The sounds are described by a set of variables (cf. § 2.3.3 for the chosen variables). The values of the variables are given by the matrix *P*:

$$P=\begin{pmatrix} p_{11} & \cdots & p_{1k} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nk} \end{pmatrix}$$

*P: n*k matrix, n refers to the number of sounds, and k is the number of variables*

Finally, the objects of the product space are described by the matrix V_{AP} ($m*k$ matrix, product of matrix *F* and matrix *P*)

$$V_{AP}=F*P \tag{1}$$

The key point of the method is that if the associations of the sounds to the objects are consensual enough, and if the variables chosen to describe the sounds make sense for the emotions, then the matrix V_{AP} can be a relevant description of the objects. The objective of this paper is to study on an example the relevance of the results provided by this method.

2.3.2 Selection of sounds

A few works use sounds for the study of emotions [8], [12], [13]. After pilot tests and a study of these works, we took an interest in a sound bank, the International Affective Digitized Sounds (IDAS-2) [13], developed by the NIMH Center for the Study of Emotion and Attention. This bank consists of a set of normative standardized, emotionally-evocative, internationally-accessible sound stimuli for experimental investigations of emotion and attention. The collection includes 167 different sounds, which are recorded in the everyday life, with contents across a wide range of semantic categories and physical properties. One of the main interests of this bank is that all the sounds are assessed by a panel of participants according to 3 representative dimensions of emotions: pleasure, arousal, and dominance. The first dimension "pleasure" (ranging from pleasant to unpleasant) represents the degree of satisfaction of the subject (Happy vs. Unhappy). The second dimension "arousal" (ranging from calm to excited) represents the degree of stimulation of the subject (Excited vs. Calm). The last dimension "dominance" represents the degree of control of the subject (Dominated vs. In-control).

To select a subset of sounds for our user-test, we made an Agglomerative Hierarchical Clustering (AHC) of the sounds, described by the 3 dimensions of emotion (pleasure, arousal, dominance). The collection of IDAS-2 was divided into 25 categories by the AHC. We chose finally 24 sounds from different categories, depending on whether they were suitable to represent the design of cars. Meanwhile, we wanted to make sure that the selections cover as much as possible the different dimensions of emotion.

2.3.3 Description of the sounds (Matrix P)

The 24 sounds selected from IDAS-2 were first described by the 3 emotional dimensions (pleasure, arousal, dominance). Pilot tests showed that considering only these dimensions did not provide an enough discriminant description of the objects (matrix V_{AP}).

For this reason, we decided to consider a second category of variables to describe the sounds: perceptual variables. To obtain these perpetual dimensions, we made a Multidimensional Scaling (MDS) study, which uses dissimilarity assessments to create a geometrical representation of the perceptual space related to the family of objects. This method, developed initially for psychometric analysis [14], is a process whereby a distance matrix among a set of stimuli is translated into a representation of these stimuli inside of a perceptual space. Taking all the possible pairs of stimuli (here pairs of sounds) into account, each subject evaluates their degree of similarity on a quantitative scale. The main advantage of this method is that the tests are based on instinctive dissimilarity assessments, which do not impose any criteria or predefined semantic scale. This method provides a space for the perception of these 24 sounds.

A pairwise comparison task of these 24 sounds was conducted with 55 participants, in order to obtain an average dissimilarity matrix of the 24 sounds. A metric-MDS algorithm was used to process the dissimilarity matrix, and to define 2 perceptual dimensions. In order to interpret the two perceptual dimensions, the 55 participants were asked to describe the 24 sounds with terms (free verbalization task). For each sound, the most occurring terms were selected, after a merging of the synonyms. A semantic interpretation of each perceptual dimension was finally made, according to the position of the sounds on dimensions, and the most relevant terms. The first dimension was labeled as “Relax” (represents the degree of comfort of the subjects) and the second as “Uneasy” (represents the degree of acceptance of the subjects).

Finally, the 24 sounds are described in the matrix P by 5 variables (as shown in Table 2): 3 emotional dimensions (Pleasure, Arousal, and Dominance) and 2 perceptual dimensions (Relax and Uneasy).

Table 2. List of dimensions of the 24 sounds

Dimensions		**Interpretation**
Emotional Dimensions	Pleasure	Happy vs. Unhappy
	Arousal	Excited vs. Clam
	Dominance	Dominated vs. In-control
Perceptual Dimensions	Relax	Relaxing vs. Disturbing
	Uneasy	Aggressive vs. Peaceful; Annoying vs. Joyful

2.3.4 Auditory Parameter Test

The Auditory Parameter Test (AP Test), based on the new AP Method, was conducted with a panel of 20 participants. Subjects were asked to assess the 10 cars (associate sounds to cars) with a friendly interface, shown in Figure 3. The image of the car was located at the upper-left, and the choice has to be made at the lower-right.

Figure 3. Interface for the Auditory Parameter Test

For each car, a 4-stage selection process was designed:

1. Random selection of 3 sounds by the algorithm, among the bank of 24 sounds. The subject had to select the most representative one,
2. Random selection of 3 sounds by the algorithm (different of those of stage 1). The subject had to select the most representative one,
3. Random selection of 3 sounds by the algorithm (different of those of stage 1 and stage 2). The subject had to select the most representative one,
4. Presentation of the sounds chosen at stage 1-2-3. The subject had to select the most representative one.

The subjects had to listen to the sounds patiently and carefully, and to select one of them according to their feeling about the car.

2.4 Subjects and procedure

20 subjects (males), students of Ecole Centrale de Nantes, participated to the tests. All the subjects did both the Semantic Differential Test and the Auditory Parameter Test. In order to balance the tests' order, they were divided into 2 equal groups of 10 participants (10 started with the SD test, 10 started with the AP test). The order of presentation of the cars was the same for all participants (we assume that the order-effect is negligible). In the end, all the subjects were required to fill in a questionnaire in order to collect some information about their opinion about the tests (duration, difficulties, interface, improvements…). The questionnaire was administered in writing.

3 RESULTS AND DISCUSSIONS

After a verification of the validity of the data for each subject, the two matrices V_{SD} and V_{AP}, representative of the 10 cars, were computed. The cars are described by two matrices:

- V_{SD} ($m*p$ matrix): average rating of the *m* cars by the SD method (p=10 semantic attributes)
- V_{AP} ($m*k$ matrix): rating of the *m* cars, obtained by the AP method (equation 1) (k=5 dimensions)

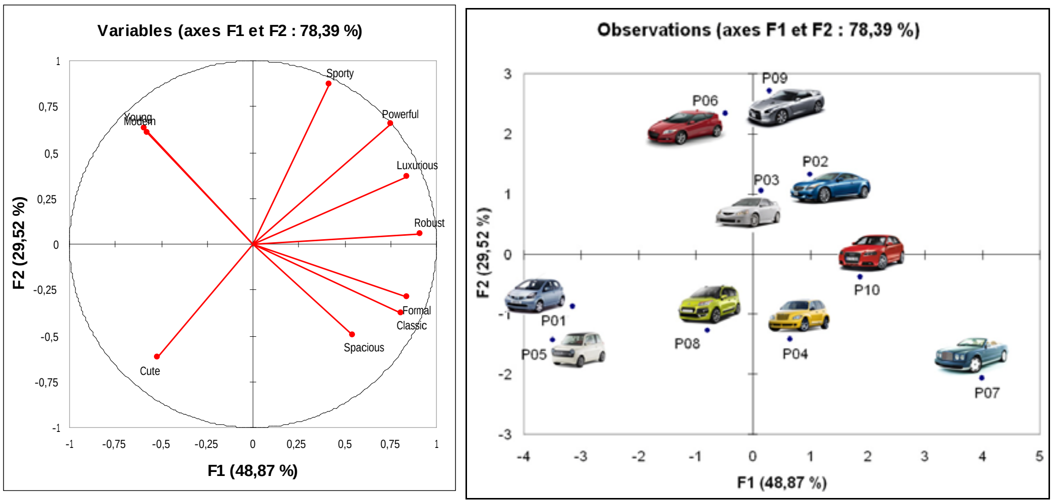
3.1 Results of SD Test

The V_{SD} matrix was analyzed using a standardized Principal Components Analysis (PCA). The variability represented by the two first factors is 78.39%: only two factors are considered to explain differences between the cars. The plane of the variables is given in Figure 4 (a). Factor 1 is mainly created by the variables "Robust", "Luxurious", "Formal" and "Classic". Factor 2 is mainly created by "Sporty", "Powerful", opposed to "Cute". The position of the 10 cars is given in Figure 4 (b). The cars P01 and P05 (particularly cute), are opposed to P07 (formal and classic). The cars P06 and P09 are particularly sporty. These results are finally not surprising: the two dimensions make sense and the assessments of the subjects are in agreement with the overall image of the cars, set up by the makers, advertising or by specialist publications.

3.2 Results of AP Test

The matrix V_{AP} is shown in Table 3. The 10 cars were represented by 5 variables of both emotional and perceptual dimensions.

The matrix V_{AP} was analyzed using standardized PCA. The two first factors represent 77.12% of variability: again, only two factors are considered to represent the differences between the cars. As shown in Figure 5 (a), the variables "Dominance", "Uneasy", and also "Pleasure" and "Relax", contribute to the first factor. The second factor is mainly created by the variable "Arousal". The position of the 10 cars is given in Figure 5 (b). Car P05 is very particular because its location is far from the others (its contribution to factor 1 is 55%, and 23% to factor 2). The car P06, a sporty car with a red color, has an important score on the "Pleasure" variable, and so evokes the feeling "happy". The car P03 evokes particular excitement, while the cars P07 and P01 look like extremely in-control. The cars P01 and P10 elicit typical relaxation. The five dimensions make sense for representing the emotions elicited by the 10 cars.

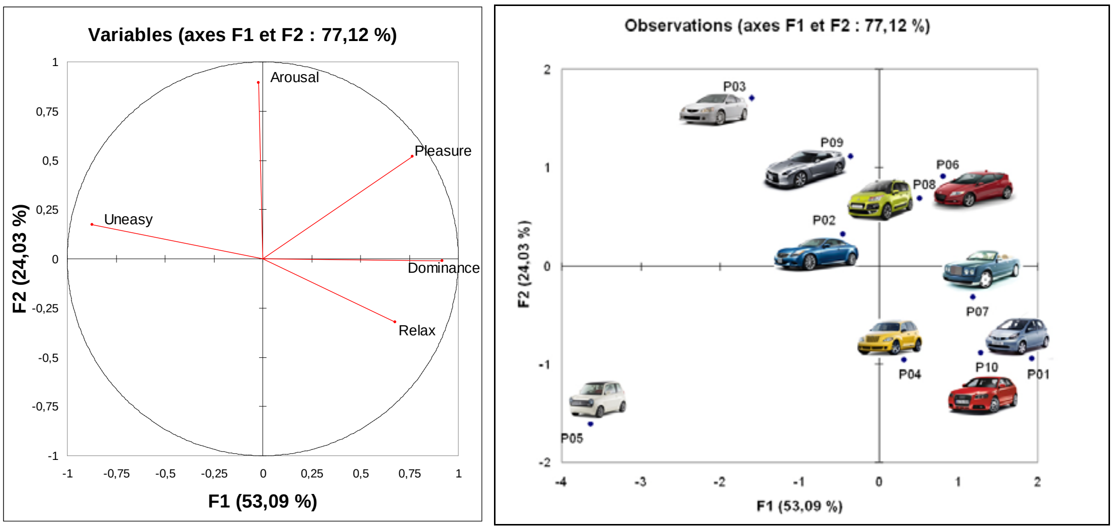


(a) (b)

Figure 4. PCA of the V_{SD} matrix. (a) Plane of the variables; (b) Position of the cars

Table 3. Matrix V_{AP}, evaluation of the cars with the AP Test

Cars	Emotional Dimensions			Perceptual Dimensions	
	Pleasure	Arousal	Dominance	Relax	Uneasy
P01	6.15	5.16	5.43	0.14	-0.31
P02	5.99	5.63	5.18	0.26	0.30
P03	5.80	6.17	5.07	-0.21	0.22
P04	5.76	5.50	5.20	0.80	-0.09
P05	5.28	5.15	5.00	-0.44	0.63
P06	6.19	5.78	5.28	0.29	0.03
P07	5.99	5.58	5.25	0.59	-0.33
P08	6.04	5.78	5.26	0.14	-0.08
P09	6.25	5.64	5.24	-0.16	0.48
P10	5.94	5.51	5.24	1.23	-0.14



(a) (b)

Figure 5. PCA of the V_{AP} matrix. (a) Plane of the variables; (b) Position of the cars

The PCA of the matrices V_{SD} and V_{AP} are interesting to understand the underlying structure of the data, for each of the two tests. But the comparison of the results is difficult because the projections made by PCA are different for these two cases. PCA is interesting to see the differences and similarities between the products for each test, but it is not able to assess the degree of agreement between the two tests.

3.3 Agreement between the two tests with GPA

3.3.1 Background on Generalized Procrustes Analysis (GPA)

Generalized Procrustes Analysis (GPA) is a multivariate technique commonly used in sensory evaluation to analyze free-choice profiling data, to study the consensus among experts in classic sensory analysis, to assess scale use, attribute interpretation, panel performance, monitoring… It also allows one to compare the proximity between the terms that are used by different experts to describe products. The GPA method was first described by Gower in 1975 [15]. GPA considers the description of the products according to a set of configurations, each configuration being described by a matrix which represents the assessment of a given expert. For each configuration, the variables which describe the products are not necessarily the same. The number of variables can also be different. GPA is a method for producing a consensus configuration from the set of different individual data matrices, and to represent the consensus via PCA (Principal Component Analysis). The principle of GPA is to apply transformations (translation, isotropic scaling, rotation and reflection) to the configurations, so as to minimize a goodness of fit criterion (the distance between the transformed configuration and the consensus configuration). GPA only allows ‘rigid-body’ transformations to the datasets and respects the relative distances between products. The individual and consensus configurations are typically submitted to PCA and projected onto a lower dimensional space. This space provides a vantage point to compare individual data and to visualize the consensus.
The degree of consensus is assessed by studying the variance of the datasets. The total variance V_T can be partitioned as follows (equation 2):

$$V_T = V_C + V_W + V_R \qquad (2)$$

where V_C denotes the variance of the consensus, V_W the within-product variance in the projection space and V_R the residual variance. By dividing by V_T, and sharing the within variance V_W among the n products, the equation becomes (equation 3):

$$1 \quad 100\% = R_c + \sum_{j=1}^{n} r_{jW} + R_R \qquad (3)$$

R_c corresponds to the consensus ratio: a large R_c indicates a good consensus.

r_{jW} indicates the within variance of product *j*. A small r_{jW} indicates a good consensus for this particular product *j*.

3.3.2 Agreement between the two tests

The consensus between the results of AP Test and SD Test was analyzed using GPA. The consensus ratio R_c between the two tests (see equation 3) is 20.3%. It signifies that after transformation of the data (rotation, translation, scaling), 20% of the total variance is represented by the consensus configuration. We can conclude that the general agreement between the two tests is weak. The statistical test (permutation test) on R_c indicates that the consensus is significant (confidence = 96%), even if the variance of the consensus is low. Figure 6 represents the values of r_{jW}, the within variance of product *j* (equation 3). The cars P02, P03 and P06 are the most consensual (the within variance is low). The cars P01 and P10 have the highest residuals (the disagreement between the two methods on these two cars is important). Figure 7 presents the PCA of the consensus configuration (the position of the consensus are in red points – the initial configurations SD and AP are projected as additional individuals).

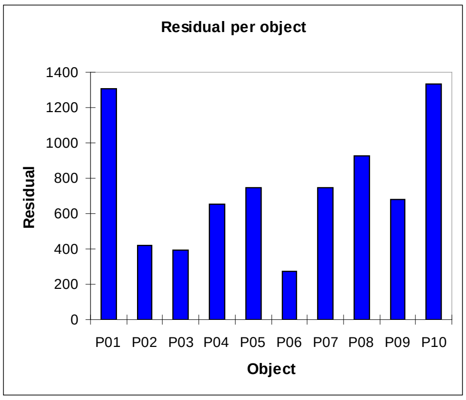


Figure 6. Within variance r_{jW} for each product

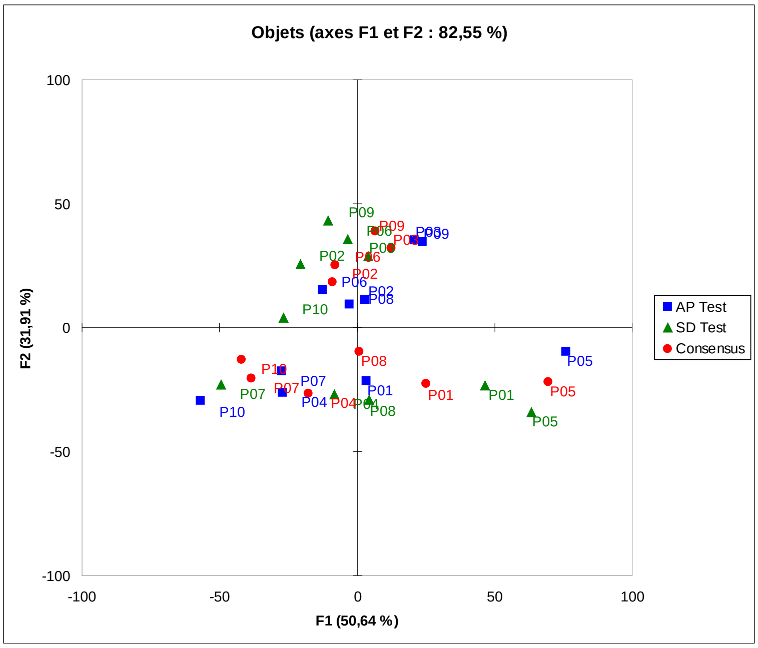


Figure 7. PCA of the consensus configuration (position of the products)

3.4 Discussion

3.4.1 General comparison between two user-tests

Concerning the opinion of the subjects about the two tests, from the questionnaires, both tasks of the tests were considered as clear and easy to fulfill. The interfaces generated with Matlab were regarded as friendly and helpful for carrying out the task, especially the interface of the AP test which was judged as simple and funny to use. The duration of the SD test was considered as acceptable, but subjects felt the duration of the AP test as a little bit long. The semantic attributes proposed in the SD Test were considered as well chosen to represent the design of cars, even if some subjects were fuzzy with some terms whose meanings were similar in some aspect. A majority of subjects agreed that the AP Test was more intuitive and more adapted to measure the emotions elicited by the product design.

3.4.2 Comparison of the information in input and output for the two user-tests

It is interesting to compare the amount of information in input and output for the two user-tests, showed in Table 4. For SD Test, the number of independent evaluations is $m*p*s$ (m refers to the number of cars of the product space, p is the number of semantic attributes, and s is the number of subjects). The number of output information of SD Test is given in the matrix V_{SD} (an $m*p$ matrix), and the ratio between the input and the output for SD Test is "s". For AP Test, the number of independent evaluations is $4*m*s$ (for each car, each subject needs to make 4 choices, while there is m cars and s subjects). The number of output information of AP Test is given in matrix V_{AP} (an $m*k$

matrix, k is the number of variables of the sounds), which is calculated by the product of matrix F and matrix P (equation 1). The ratio between the input and output for AP Test is "$4s/k$". In our AP Test, k=5, so the ratio for AP test equals to $0.8s$, which is almost the same as the ratio of SD Test.

Table 4. Comparison of input and output information for the two user-tests

User-tests	Input	Output	Ratio
SD Test	$m*p*s$	$m*p$	s
AP Test	$4*m*s$	$m*k$	$4s/k$ (k=5)

3.4.3 Significance of AP Test

Even if the positioning of the cars with the AP Test seems to be consistent, we wanted to know if the results of the AP Test are statistically significant. For this, we defined a statistical test, by comparing the results of the AP Test to randomly generated results. The statistical test for the calculation of the significance of the AP Test is described in APPENDIX.

We noticed that, for each car, the test is significant at the 5% level (p-value = 5%) for at least two sounds. It signifies that the results obtained are generally far from results that would have been obtained randomly.

3.4.4 Agreement between AP Test and SD Test

The results show that the agreement between the two tests is weak (only 20.3% of common variance for the two tests). Two hypotheses can be proposed to explain this weak consensus; (1): the consensus is weak because the AP method is not developed enough, and further improvements are needed. (2): the consensus is weak because the two methods do not measure the same concepts.

Concerning (1), our results prove that the method measures stable dimensions. The statistical test shows that the choice of the sounds is not made randomly (some of them are over represented for certain cars). The agreement between subjects is sufficiently good to provide significant results. But further developments are needed, in particular for the selection of the sounds, for the setting-up of the parameters of the AP Test, and for the interpretation of the MDS dimensions.

Concerning (2), the SD Test is made on semantic attributes [16], which are not a direct representation of emotion. This may produce a discrepancy between the results of the two methods. The information we got from AP Test may be concentrated on the feeling and the experience of the cars, which is more closed to subjects' mind, while the information we got form SD Test may be focused on the cognition of the cars, which is more representative of the products themselves. Further studies are needed to confirm these assumptions.

4 CONCLUSIONS

In this paper, we described a new experimental protocol for the measurement of emotions elicited by products appearance. The method, called Auditory Parameter Method, uses auditory stimuli and association tests. The method was used to assess the emotions elicited by a set of 10 cars. First results show that the protocol has a good feedback from the participants and that it makes sense to ask to associate sounds to pictures.

The results of the Auditory Parameter Method were next compared to those of the Semantic Differential Method. The similarities between the results were evaluated using GPA. The results show that the consensus between the two methods is weak. Two hypotheses must be studied to explain this weak consensus. First, the tuning of the Auditory Parameter Method is perhaps not finished. Several parameters of the method must be studied: number of sounds proposed, number of associations asked to the subjects, variables for the description of the sounds. Second, the consensus is weak because the two methods measure two different things (SD measures product semantics – AP measures emotions). To verify this assumption, we envisage measuring emotions with an experimental protocol given in the state of the art, for example [7].

Nevertheless, we think that the AP Method can be a language-free alternative to the Semantic Differential Method, and that these first results are encouraging for developing a new non-verbal experimental protocol.

ACKNOWLEDGEMENTS

The authors would like to thank the NIHM Center for Emotion and Attention (CSEA) at the University of Florida for providing the sound bank IADS-2 for our research. The authors also thank the contributions of the students of Ecole Centrale de Nantes who took part in the two tests.

REFERENCES

[1] McDonagh D., Bruseberg A., Haslam C. *Visual product evaluation: exploring users' emotional relationships with products.* Applied Ergonomics 33 (2002) 231–240.

[2] Norman D., *Emotional Design: why we love or hate everyday things*. Basic Books, New York, 2004.

[3] Plutchik R., *Emotion: Theory, research, and experience*: Vol. 1. Theories of emotion, 1, New York: Academic Press, 1980

[4] Osgood C.E., Suci G.J., Tannenbaum P.H. *The measurement of meaning*, Urbana, USA: University of Illinois Press, 1957.

[5] Lang P. J. *Behavioral treatment and bio-behavioral assessment: Computer applications* . In J. B. Sidowski, J. H. Johnson, E. Awilliams (Eds.), Technology in mental health care delivery systems. Norwood, NJ: Ablex. 1980, pp. 119–137.

[6] Bradley M.M., Lang P.J. *Measuring emotions: the self-assessment manikin and the semantic differential*. Int. J. Behav. Theory & Psychiat. Vol. 25, No. I. pp. 49-59, 1994.

[7] Desmet, P.M.A., Hekkert, P., Jacobs, J.J. (2000). *When a car makes you smile: Development and application of an instrument to measure product emotions*. In: S.J. Hoch and R.J. Meyer (Eds.), Advances in Consumer Research, 27, 111-117.

[8] Bradley, M. M., & Lang, P. J. (2000a). *Affective reactions to acoustic stimuli*. Psychophysiology, 37, 204-215.

[9] Nagamachi M. *Kansei engineering: a new ergonomic consumer-oriented technology for product development*. International Journal of Industrial Ergonomics. 15 (1995), pp. 3-11.

[10] Norgren A.K. *Exploring automotive shape with Kansei design – A systematic approach to building design support systems with shape sensibility*, Doctoral Dissertation, Keio University, 2007.

[11] Otsuka S., Inoue H., Kashiwazaki N., Nowura M., Kubota M., Motoyama T. *Simultaneous evaluation of fragrance and pictures using Kansei Parameter Method*. Proceedings of KEER 2010, March 2010, Paris, France.

[12] Belin P., Fillion-Bilodeau S., Gosselin F. *The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing*. Behavior Research Methods 2008, 40 (2), 531-539 doi: 10.3758/BRM.40.2.53.

[13] Bradley M.M., Lang P.J. *The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective Rating of Sounds and Instruction Manual*, Technical report B-3. University of Florida, Gainesville, F1, 2007.

[14] Shepard R.N., Romney K., Nerlove S.B. *Multidimensional scaling: Theory and applications in the behavioral sciences*. New York: Seminar Press, Vol. I: Theory, 1972.

[15] Gower J.C. *Generalized Procrustes Analysis*. Psychometrika 50, 1975, pp. 33-51.

[16] Schütte S.T.W., Eklund J., Axelsson J.R.C., Nagamachi M. *Concepts, methods and tools in Kansei engineering*, Ergonomics Science, May–June 2004, vol. 5, no. 3, 214–231

APPENDIX: statistical test for the significance of AP Test

In an attempt to verify that AP Test can gather informative and significant information, we defined a statistical test. The random selection process can be seen as the repetition (*s* times, with *s* independent subjects) of a random test with 3 possibilities:

$p1$: probability that a sound is not selected, for a given subject

$p2$: probability that a sound is selected once, for a given subject

$p3$: probability that a sound is selected twice, for a given subject

The test is based on the calculation of the probability to have, by random choices, the number of chosen sounds equal to a given value *y*. Let *Y* be a random variable and y the number of time one sound is selected, for the AP Test.
The computation of the probability $P(Y = y)$ is given by (multinomial law):

$$P(Y = y) = \sum_{x_2 + x_3 = y} \frac{s!}{x_1!x_2!x_3!} p_1^{x_1} . p_2^{x_2} . p_3^{x_3} \tag{4}$$

with:

x1: number of time that the sound is not selected

x2: number of time that the sound is selected once

x3: number of time that the sound is selected twice

x1+*x2*+*x3* = *s*

The results of the AP Test are considered as significant if this probability is lower than 5%. Using equation (4), we have $P(Y = y)$ <5% for y>7. The matrix *F* of the AP Test is given below. The significant results are grayed in this table.

	son1	son2	son3	son4	son5	son6	son7	son8	son9	son10	son11	son12	son13	son14	son15	son16	son17	son18	son19	son20	son21	son22	son23	son24
P1	2	8	2	6	1	2	3	4	2	3	3	0	1	3	7	1	0	11	3	0	2	2	4	10
P2	2	3	1	0	3	0	1	12	0	1	8	3	2	3	2	7	2	1	3	6	4	9	4	3
P3	4	2	0	2	7	2	3	10	3	3	4	5	1	1	0	6	3	0	2	2	1	7	9	3
P4	0	3	2	4	2	3	6	0	4	4	1	7	3	3	2	3	2	2	7	2	5	3	2	10
P5	3	8	7	4	0	13	2	0	2	6	5	1	3	3	9	0	6	1	2	0	0	1	2	2
P6	0	1	1	2	7	1	1	10	3	3	2	2	1	0	1	6	2	2	8	3	5	7	5	7
P7	0	2	3	2	2	0	1	8	4	1	2	8	3	2	2	4	2	4	5	7	3	6	5	4
P8	2	4	3	4	8	3	1	2	2	6	6	1	3	2	4	2	1	0	0	0	3	5	9	9
P9	4	2	1	0	8	1	0	11	0	0	9	4	3	0	0	4	0	0	3	6	7	8	4	5
P10	4	1	1	1	1	0	1	2	2	3	6	2	3	4	5	5	0	3	6	5	5	9	4	7

Contact: Weihua Lu, Jean-François Petiot
LUNAM, Ecole Centrale de Nantes, Institut de Recherche en Communications et Cybernétique de Nantes (UMR CNRS 6597)
1, rue de la Noë, BP 92101, 44321 Nantes, France
Tel: Int +33 (0)2 40 37 69 59
Fax: Int +33 (0)2 40 37 69 30
Email: weihua.lu@irccyn.ec-nantes.fr, Jean-Francois.Petiot@irccyn.ec-nantes.fr
URL: http://www.irccyn.ec-nantes.fr/~petiot/

Weihua Lu is a PhD candidate at IRCCyN (Institut de Recherche en Communications et Cybernétique de Nantes), UMR CNRS 6597. Next to this study, she has been working as a university lecturer in China on Industrial Design. She has a special interest in emotional design and sensory analysis. Her current focus is on design methodology.

Jean-François Petiot is Professor of Engineering Design at the Ecole Centrale de Nantes. He teaches and researches in customer oriented design and user centered design. He is interested by many aspects of design, in particular user's perceptions, subjective assessments, sensory analysis, preference mapping, and psychoacoustics. He has a deep interest for applications in acoustics and in the domain of the perceived quality.