

Development of an instrument to measure emotional responses. Work in progress.

Alexander Rozo-Torres^[0000-0001-9343-2264], Elizabeth Garzón-Morales^[0000-0003-2127-5575] and Wilson J. Sarmiento^[0000-0001-7903-8316]

Abstract. This work shows preliminary results of the design process of a questionnaire that provides an instrument to evaluate the emotional reaction by users' self-report. The proposal involves 13 pairs of verbal concepts and their non-verbal representation, including hedonic and pragmatic valuations and affective responses in the three-emotional dimensional model. i.e., pleasure, arousal, and dominance. The preliminary proposal results from conceptual validation of a selected set of instruments in state-of-the-art and users validation of non-verbal representation.

Keywords: Emotional response, Emotional evaluation, Emotional Design

1 Introduction

Interactive objects and systems aim to provide a service, a functionality, that must be communicated through the components that define it. However, more than being a product with which you can interact, users look for experiences to satisfy their necessities. In this way, users' experiences and emotions had been consolidated as an essential part of interactive technologies' design and assessment process in last years involved fields of study within psychology and human social aspect [14, 4].

Emotional design is a field of design focused on this topic; according to Donald Norman, thus, the emotional design seeks to generate appropriate responses to the stimuli of a product [14]. Recognizing that the design of an object can evoke emotions includes determining the emotional response that a person experiences when interacting with them. Therefore, several authors have developed a series of emotional measurement instruments to assess the emotional responses caused by an object, most of which users self-report. However, these instruments have a limitation, some caused by emotional model assumed, languages dependence of concepts interpretation, and extension of the questionnaire. This paper, looking for a contribution of this problem, shows a work in progress that seeks to propose an instrument that involves both verbal and graphic representation of concepts; this is results from validation and integration of well-known instruments.

The rest of this paper is organized as follows; the section 2 presents selected works that support the development of the proposed questionnaire. Section 3 describes the proposal for the construction of the ideal instrument as an emotional measuring

instrument that provides an evaluation of interfaces that generate an emotional connection with people directly related to the information presented and with those who interact. Emphasis is placed on measuring emotions to generate a social impact through the emotional connection that an interface generates, engaging the senses of the human being. Section 3.3 presents the evaluation process with users to obtain the visual perception of a series of images on a term and detect which are the most appropriate images for the construction of the instrument. Finally, in section 4 the possible applications of evaluation as a measurement instrument and future work are discussed.

2 Related Work

Emotional experiences with objects arise from a sensory action (seeing, hearing, feeling, tasting, or smelling) [14]. Emotional design comprehends the sensory experience and the reactions caused by an object, creating a connection with the person. Donald Norman describes emotional design on three levels. The first is the visceral level, a natural reaction elicited by the appearances of an object. The second is the behavioral level, based on the use and experience of interacting with the object. The third is the reflective level, a rational and intellectual phase that includes the interpretation of meanings, memories, and experiences [14].

Other authors such as Patrick Jordan [9] and Aarron Walter [20] describe the emotional design as a scale of needs according to the user and the object, where the highest level is a pleasure. In a previous work of the authors (in review), a convergence between the concept of Norman, Jordan, and Walter was interpreted into *four elements* that revolve around emotional design focused on generating emotions evoked by an object. The first element is the object's information, according to its context and the reason it was created. The second is the appeal, object appearance, and interest. At the same time, the third is related to the interpretation that a person is based on previous experience with the same object or similarities. Finally, the last element is the emotions related to the object and the memories of a person.

The latter being the most used (about 84% of the studies) according to the authors Camila Loiola and Elizabeth Sucupira in their studies on user experience evaluation instruments [13]. Most of the instruments mentioned in the study focus on the use of the product and not on the user experience, which generates the need to propose tools that allow the evaluation of the user experience in real-time, gathering relevant information of the experience without long-term loss.

The evaluation of the emotional reaction of a user may carry out according to the subjective perception of uses or an observer, or objective-based in physiological user's responses [19, 7]. Various emotional evaluations can be categorized by measuring emotions in discrete, dimensional emotional models or physiological signals. The discrete model proposed by Ekman defining universal emotions (happiness, sadness, surprise, fear, anger, and disgust) [7], and some instruments built to measure this set of emotions represented on Likert scales, graphics, or animated characters. An example is Product Emotional (PrEmo), a non-verbal method to identify emotional responses with a graphic representation of an animated character that expresses 7 negative and 7 positive emotions caused by stimulation [5, 12].

On the other hand, the dimensional model proposes a classification of emotions by characteristics such as valence or pleasure, activation or arousal, and dominance. It is a model that welcomes various emotions without labeling them as in Ekman's model. One of the instruments that evaluate emotions verbally in this model is the Standardized Emotional Profile (SEP) [11], which implies emotional three-dimensionality in 27 elements grouped in 9 indices: faith, affection, gratitude, interest, activation, emergence, sadness, fear and skepticism, using a semantic differential that goes from "very" to "not at all" for each element [11]. Another example is the Mehrabian and Russell semantic differential, with a set of 18 pairs of bipolar adjectives graded along a 9-point Likert scale to measure three-dimensionality.

There are other non-verbal approaches to measuring emotions, such as the Affect Grid, being a self-report scale designed to measure pleasure and arousal according to the distribution on a grid [15]. Another example is the Emotional Cards (EmoCards), representing 16 faces in cartoons with 8 different emotional expressions (8 male and eight female faces) depending on the dimensions of liking and arousal [6]. The Self-Assessment Manikin (SAM) is also found as an instrument that measures pleasure (happy-unhappy), arousal (excited-calm) and dominance (controlled-in control) of a person through a graphic scale for each dimension [2]. Similar to the Affective Slider, being a digital instrument composed of two sliders that measure arousal and pleasure, replacing the graphics with emojis that represents an emotional state for each dimension as an alternative to SAM [1]. Other measurement instruments use discrete emotion families aligned in a two-dimensional circle, such as the Geneva Emotion Wheel (GEW), measuring the control and pleasure of emotional reaction [18].

Another instrument is the Anticipated eXperience Evaluation (AXE), which proposes a method of experiential evaluation of the first concepts of products with users [8]. AXE is a qualitative verbal and non-verbal method that assesses the initial perception of future users regarding the experience of the concept of a product through pairs of images separated by a scale [8]. The terms that can be measured, in addition to perception, are the function, emotional activation, the evocation of meanings, and the aesthetics of an object through an interview that has questions related to different categories.

The User Experience Questionnaire (UEQ) is a verbal tool for the quick and direct measurement of user experiences of an interactive product. It is a semantic differential composed of 26 pairs of items (positive and negative) grouped into 6 scales: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty, rated by a 7-point Likert scale. However, this questionnaire using the scale of attractiveness, a pure term valence dimension, to capture the general impression of a user [16]. A short version of the UEQ questionnaire (UEQ-S) is built to adapt to environments that require rapid measurement. This version reduces the 26 item pairs in 8 pairs, making the user response time faster, focusing on the measurement of pragmatic and hedonic quality as an overall value of the user experience [17]. Similarly, Chin et al. [3] proposed Questionnaire for User Interfaces Satisfaction (QUIS), a semantic differential questionnaire of 27 rating scales with a scale of points from 0 to 9, grouped into 5 categories: general reactions to the system, screen, terminology and system information, learning and system capabilities [3].

3 Questionnaire design process

3.1 Defining verbal concepts

In order to define a suitable set of concepts that allow a complete evaluation of the emotional responses of users, we conduct a process of comparison, reduction and match of the concepts involved in instruments of state of the art. The first step was a comparison of 10 emotional measurement instruments (QUIS [3], UEQ [16], UEQ-S [17], SEP [11], AttrakDiff [10], AXE [8], Emo-Cards [6], GEW [18], AffectGrid [15], SAM [2]). This set was compiled using a literature review that included self-report instruments, excluding those based on a discrete set emotional model to basic emotions. All theoretical assessment concepts from the instruments were paired into similar concepts and synonym reduction; the result is a set of eight categories, i.e., 1) perception, 2) functionality 3) pleasure, 4) arousal, 5) mastery, 6) meaning, 7) curiosity, and 8) aesthetics. The first term is a person's perception and impression of the general characteristics of an object. The second is the object's functionality to tell stories, over and above being artistic, collectible or other similar objects, is a tool to convey messages. The next terms are the three main dimensions of emotion in the PAD model: pleasure, refreshment, and mastery. The sixth term is the meaning that a person can understand based on the given context of the object. The seventh is the curiosity a person has when exploring or immersing themselves in the object's content. The last term, aesthetic, is when an object is understood as attractive based on its appearance.

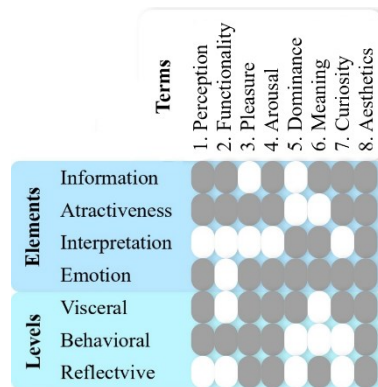


Fig. 1. Conceptual cross-comparison between the set of the 8 terms found and the four elements of the emotional design and three levels of emotional.

Afterward, the process required a conceptual cross-comparison between the set of the 8 terms and the *four elements* of the emotional design and three levels of emotional design. Figure 1 shows this conceptual cross-comparison; in the columns are the 8 terms found, and gray circles highlight conceptual matches.

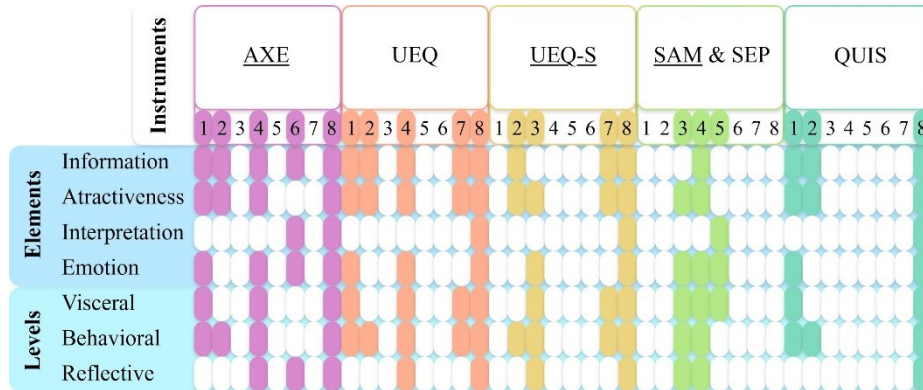


Fig. 2. Conceptual cross-comparison of emotional measuring instruments. This procedure has the assumption that Figure 1 is an ideal template.

A similar comparison was performed involving the main measuring instruments identified in the state-of-the-art review. This comparison of the instruments is presented in Figure 2, which shows in the rows the *four elements* of the emotional design and three levels of emotional design in rows. Each instrument is represented as a set of 8 columns (set of the terms found), where the colored region shows compliance. It is necessary clearly that this procedure has the assumption that Figure 1 is an ideal template.

Table 1. Proposed concepts pair terms.

Obstructive	–	Supportive
Complicated	–	Easy
Inefficient	–	Efficient
Confusing	–	Clear
Boring	–	Exciting
Not Interesting	–	Interesting
Conventional	–	Inventive
Usual	–	Leading edge
Pleasure	–	Unpleasure
Exciting	–	Calm
Controlled	–	Uncontrolled
Relevant	–	Inopportune
Sympathetic	–	Unfriendly

The conceptual cross-comparison process allowed identify a subset of ideal measuring instruments composed by UEQ-S, SAM and AXE. UEQ-S is a short 8-item questionnaire with a 7-point scale for environments that require rapid user experience measurement by measuring hedonic and pragmatic quality based on functionality, satisfaction, curiosity, and aesthetics. SAM compiles graphically the representation of the three emotional dimensions evaluating pleasure, arousal and dominance. AXE evaluates user experience from a graphic and verbal representation through pairs of

images according to their categories, such as perception and meanings. Subsequently was necessary to mix and reduce the concepts that these instruments consider, which got a final set of 13 pairs of concepts proposed, which are presented in the table 1.

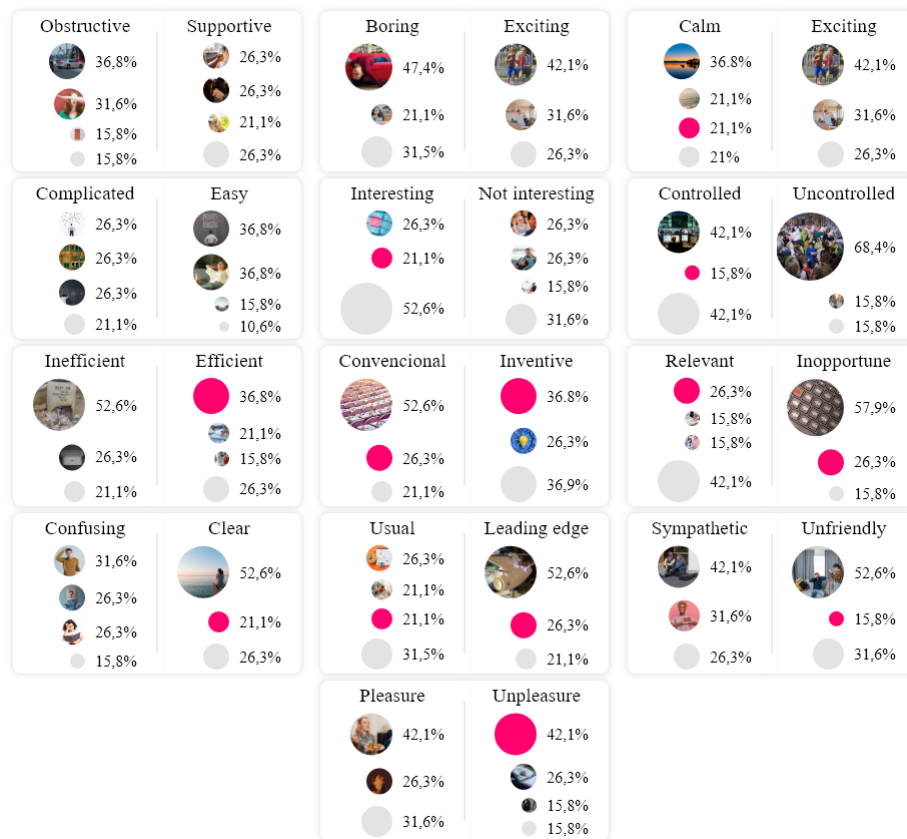


Fig. 3. Results of the questionnaire allowed identifying the better image that represents the verbal concept. From top to bottom is the most voted option. The circled in gray represent the group of images with the least amount of selection by subjects.

3.2 Seeking no-verbal representations

Once defined the set of verbal concepts, it was necessary, searching for images that represent them. Then, a search of 8 images for each concept was performed on the Creative Commons and iStock open-access search engine. As inclusion criteria for this choice were made from the first images out, using as exclusion criteria those that were not related to the definition or were linked to medical concepts, state or political institutions. In order to evaluate image interpretation with meaning correspondence, we

planned a two-level validation. In the first level, a questionnaire was constructed where subjects must mark the image that better represents the meaning according to which Royal Spanish Academy. A summary of the results of the questionnaire is presented in Figure 3 with 19 subjects' responses that selected the image that best matches the term. Most respondents opted for an image, causing 21 of the terms to be represented by one of the images. However, the percentage difference for some images is very similar. The percentages with the least amount of selection were circled in gray.

The second level of validation was an assessment to identify the degree of agreement and disagreement of the previous image choices using a 5 -point Likert scale. In this way, a second questionnaire allowed verifying that the image corresponds satisfactorily and represents the meaning of each concept.

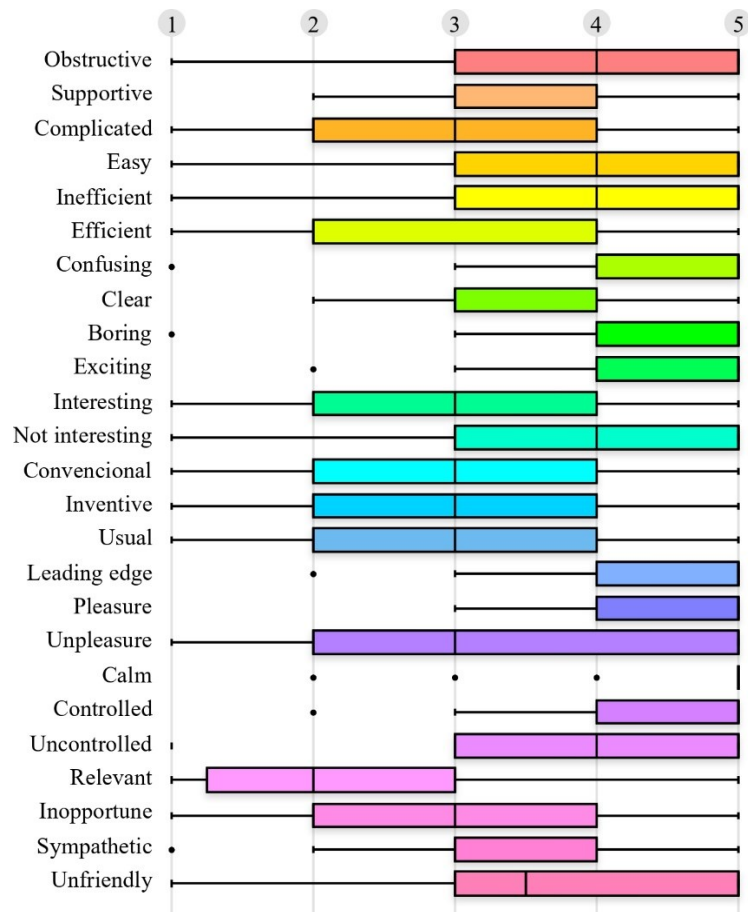


Fig. 4. Results of the questionnaire validate the degree of approval of previously image choice, where 1 is strongly disagreed, and 5 is strongly agreed.

The figure 4 shows the results at correspond to the responses obtained for each concept after applying the questionnaire to a new group of 21 subjects. It is possible to

assert that most concepts observed a desirable level of approval, which is an approval between 3 and 5. However, some concepts had answered with approach levels less than 3 but up to 2. Only a concept, “*relevant*”, had an approval less of 2.

3.3 Preliminary questionnaire approach

The end-stage of the process was to create a graphical representation of a pair of concepts with two images, one by the opposed concept. The graphical representation was inspired by the images chosen and validated.

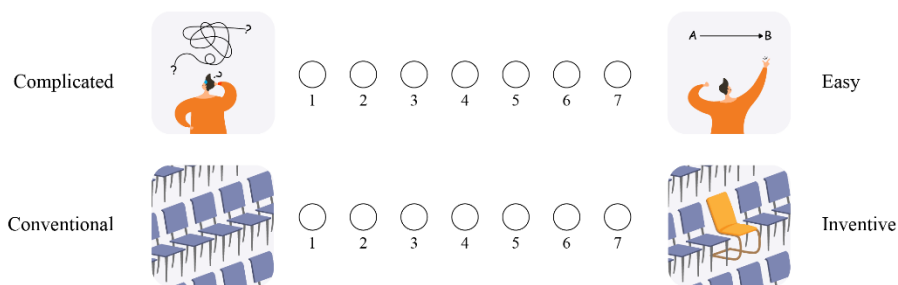


Fig. 5. Preview of our proposal instrument that involves both verbal and graphic representation of concepts to evaluate. It is possible to see two of the pairs of concepts of our proposal, i.e., “*complicated*” - “*easy*” and “*conventional*” - “*inventive*”.

Remember that this is a work in progress, Figure 5 presents the graphical form of two of the pairs of concepts of our proposal, i.e., “*complicated*” - “*easy*” and “*conventional*” - “*inventive*”. It is possible to see a preview of our proposal: an instrument that involves both verbal and graphic representation of concepts to evaluate, that takes main characteristics of the emotional design and is a result of comparison and match of the concepts involved in instruments of state of the art.

4 Conclusions and future work

This paper presents a work in progress that aims to propose an instrument to evaluate the emotional users’ responses. The preliminary results are a set of 13 pairs of concepts to evaluate and a draft of the instrument to include only two evaluation concepts.

This paper describes the process, which starts with a state of the art review and a conceptual cross-comparison with three levels of emotional design and four elements of the emotional design defined by authors in previous work (in review). The process also includes searching images to represent the pairs of and validate these images by two level subjective questionnaires.

A preview of our proposal was presented, including the graphical representation of two pairs of concepts, which showed both verbal and graphic representation in a further evaluation instrument.

Although the proposed instrument is a mixed combination of other instruments, it can be applied and adapted to a wide range of physical and digital interfaces to perform an emotional and experiential evaluation. The results of this evaluation process will help make data-driven instrument design and construction decisions. Future work includes a final version of the instrument with user expert validation and its application in a case study in various areas involving users.

Acknowledgments. This work is a result of the research project INV-ING-3192 funded by the Office of the Dean of Research of Militar Nueva Granada University and Universidad Nacional de Colombia, Bogotá (execution period 2020).

References

1. Betella, A., Verschure, P.F.: The affective slider: A digital self-assessment scale for the measurement of human emotions. *PLoS ONE* 11(2) (feb 2016). <https://doi.org/10.1371/journal.pone.0148037>
2. Bradley, M.M., Lang, P.J.: Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25(1), 49–59 (1994). [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
3. Chin, J.P., Diehl, V.A., Norman, K.L.: Development of an instrument measuring user satisfaction of the human-computer interface. In: *Conference on Human Factors in Computing Systems Proceedings*, vol. Part F1302, pp. 213–218 (1988). <https://doi.org/10.1145/57167.57203>, ISBN: 020114237
4. Desmet, P.M.A., Porcelijn, R., van Dijk, M.B.: Emotional design; application of a research-based design approach. *Knowledge, Technology & Policy* 20, 141–155 (2007). <https://doi.org/10.1007/s12130-007-9018-4>
5. Desmet, P.: *Measuring Emotion: Development and Application of an Instrument to Measure Emotional Responses to Products*, pp. 111–123. Springer Netherlands (2004). https://doi.org/10.1007/1-4020-2967-5_12
6. Desmet, P., Overbeeke, K., Tax, S.: Designing products with added emotional value: Development and application of an approach for research through design. *The Design Journal* 4(1), 32–47 (2001). <https://doi.org/10.2752/146069201789378496>
7. Ekman, P.: An argument for basic emotions. *Cognition and Emotion* 6(3), 169–200 (1992). <https://doi.org/10.1080/02699939208411068>, publisher: Routledge
8. Gegner, L., Runonen, M.: For what it is worth anticipated eXperience evaluation. In: *8th International Conference on Design and Emotion: Out of Control Proceedings* (2012)
9. Green, W.S.: Introduction pleasure with products: Beyond usability 'human factors and desing'. In: Green, W.S., Jordan, P.W. (eds.) *Pleasure with Products Beyond Usability*, pp. 1–6. CRC Press (2002)
10. Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualitat. In: *Mensch & Computer 2003 Interaktion in Bewegung*, pp. 187–196 (2003). https://doi.org/10.1007/978-3-322-80058-9_19
11. Holbrook, M.B., Batra, R.: Toward a standardized emotional profile (SEP) useful in measuring responses to the nonverbal components of advertising. In: *Nonverbal communication in advertising*, p. 304. The Free Press (1987)
12. Huisman, G., van Hout, M., van Dijk, E., van der Geest, T., Heylen, D.: Lemtool: Measuring emotions in visual interfaces. In: *Proceedings of the SIGCHI Conference on Human Factors*

- in Computing Systems. p. 351–360. CHI '13, Association for Computing Machinery, New York, NY, USA (2013). <https://doi.org/10.1145/2470654.2470706>
13. Maia, C.L.B., Furtado, E.S.: A systematic review about user experience evaluation. In: International Conference of Design, User Experience, and Usability. pp. 445–455. Springer (2016). https://doi.org/10.1007/978-3-319-40409-7_42
 14. Norman, D.A.: Emotional Design. Why We Love (or Hate) Everyday Things, vol. 2. Basic Books (2005)
 15. Russell, J.A., Weiss, A., Mendelsohn, G.A.: Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology* 57(3), 493–502 (1989). <https://doi.org/10.1037/0022-3514.57.3.493>
 16. Schrepp, M., Hinderks, A., Thomaschewski, J.: Construction of a benchmark for the user experience questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence* 4(4), 4–40 (2017). <https://doi.org/10.9781/ijimai.2017.445>
 17. Schrepp, M., Hinderks, A., Thomaschewski, J.: Design and evaluation of a short version of the user experience questionnaire (UEQ-s). *International Journal of Interactive Multimedia and Artificial Intelligence* 4(6), 103–108 (2017). <https://doi.org/10.9781/ijimai.2017.09.001>
 18. V., S., Schlegel, K., Scherer, K.R.: Geneva emotion wheel rating study (2012), <https://archive-ouverte.unige.ch/unige:97849>
 19. Van Hout, M.: Comprendiendo, midiendo, diseñando (para la emoción). *Revista Faz* (2), 88–97 (2008), http://www.revistafaz.org/articulos_2/08_emociones_vanhout.pdf
 20. Walter, A.: Designing for Emotion. No. 9 (2011)