Please Make Me Warm: Combining Product Emotions and Usability to Redesign Microwave Products

Thedy YOGASARA[a] and Yutta Verena HERMANTO[b]
[a,b]*Parahyangan Catholic University*
Jl. Ciumbuleuit 94, Bandung 40141, Indonesia

ABSTRACT
Microwave oven is one of everyday products with many types and brands offering similar functionalities and capabilities. These similarities create tough competition, and in this case, product usability and users' emotional responses to product appearance often become key determinants of purchasing decision. This research combines product emotions measurement and usability testing to evaluate and redesign existing microwave products. Modified PrEmo was used to measure positive and negative emotional responses of thirty participants to three different microwaves. Usability testing involving eight participants was then conducted to the product receiving the most negative emotional reactions. Through these procedures, usability problems and product features that evoked positive emotions were identified and analyzed. The redesign of the microwave encompassed its physical form, color, control panel, and display. Three redesign alternatives were generated, each of which was evaluated using the previous methods. It was found that all alternatives resulted in better usability and product emotions scores, where alternative 3 produced the highest total (combined) score.

KEYWORDS
Modified PrEmo; product design; product emotions; usability.

INTRODUCTION

The growing number of manufacturers producing products that have similar technology and functionality lead to tighter competition among them. To survive in this condition is, among others, to fulfil users' needs by offering products with appealing physical appearance and high level of usability.
Since many products nowadays are often similar in terms of quality, technical characteristics, and price, the experiential or emotional quality of the products is becoming more and more important for enhancing their differential advantage in the market. In some buying decisions, users' emotional responses to the products can even be a decisive factor (Desmet, 2003a). Zhang and Li (2005) argue that affective quality positively impacts on users' cognitive evaluations of a product, which, in turn, can influence their behavioral intention to use it. Moreover, the importance of understanding and fulfilling user emotional needs in product design has been related to the success of a product in the marketplace (Khalid, 2006; Khalid and Helander, 2006).
In addition to affective quality, usability – as a product design criterion – is essential to ensure that products are easy to learn and to use, are satisfying to use, and provide the functionality and utility that are highly appreciated by their target users (Rubin, 1994). Usability is commonly measured by efficiency (e.g. error rate, time on task, mental workload), effectiveness (task completion, output quality), and satisfaction (qualitative and quantitative attitudes) (ISO 9241-11, 1998; Jordan, 1998).
This paper presents the redesign of microwave products based on the combination of product emotions and usability criteria. Microwave was selected because this product category generally has similar basic functions and features from one brand to another, so that when making a buying decision, their users are likely to focus more on usability and aesthetic pleasure offered.

PRODUCT EMOTIONS

Product emotions is a set of emotions that are elicited by product design. Desmet (2003b) establishes a basic model of product emotions (Figure 1), in which four main elements – appraisal, concern, product, and emotion – interact in the process of eliciting emotions. This model describes that all emotional responses result from a process of appraisal in which an individual perceives a product as supporting or conflicting one or more of his/her concerns (Desmet, 2003b). These emotional reactions can be measured using the Product Emotion Measurement Instrument (PrEmo), a non-verbal self-report instrument that assesses 14 emotions that are often evoked by product design (Desmet, 2003a). These 14 emotions comprise 7 pleasant emotions (i.e. desire, pleasant surprise, inspiration, amusement, admiration, satisfaction, and fascination) and 7 unpleasant emotions (indignation, contempt, disgust, unpleasant surprise, dissatisfaction, disappointment, and boredom). Each of the 14 emotions is represented by an expressive cartoon, based on which respondents can report their felt emotions

rather than using their words. Each cartoon is accompanied by a three-point scale that is used for quantifying the evoked emotions. The validity of PrEmo has been assessed through the cross-cultural test of the expressive cartoons and the comparison between PrEmo and a verbal scale in measuring product emotions (Desmet, 2003a).

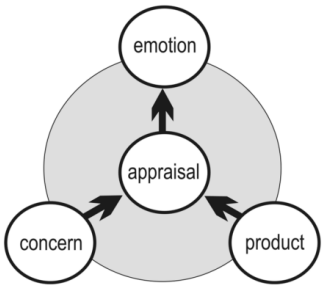

Figure 1. Basic model of product emotions
(Desmet, 2003b, p. 5)

USABILITY

Usability has been an established criterion for determining the goodness of a product or system design. Formally, the International Organization for Standardization defines usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" (ISO 9241-11, 1998). The usability concept can also be defined by several factors: effectiveness, efficiency, satisfaction, likeability, learnability, flexibility, guessability, experienced user performance, usefulness, and task match (Jordan, 1998; Rubin, 1994; Stanton, 1998). Furthermore, Jordan (1998) proposes ten design principles related to product usability: consistency, compatibility, consideration of user resources, feedback, error prevention and recovery, user control, visual clarity, prioritization of functionality and information, appropriate transfer of technology, and explicitness.

Usability testing is a common method used to evaluate a product. This testing is a process that employs participants representing the target population to measure the degree to which a product meets specific usability criteria (Dumas and Redish, 1999; Rubin, 1994). The overall goal of usability testing is to identify and rectify usability problems existing in product design prior to release.

Task scenarios are usually used in usability testing to represent the real tasks that users have to complete in using a product. In relation to the accuracy of task completion, three types of statistics can be measured (Rubin, 1994):

1. Percentage of participants performing successfully within the time benchmark.
2. Percentage of participants performing successfully, regardless of the time benchmark.
3. Percentage of participants performing successfully, regardless of the time benchmark, including those who require assistance.

METHODS AND RESULTS

Product Emotions Measurement

For the subject of study, three different microwaves were selected (Figure 2): sample A (Panasonic), sample B (Ariston), and sample C (Metrowealth). They were used as stimuli to evoke participants' emotional responses.

Figure 2. Three microwave products used in the study (from left to right: Panasonic, Ariston, Metrowealth)

Thirty participants (25 females and 5 males, aged 21-45 years) were recruited using convenience sampling technique based on three screening criteria: participants were adult users, they had used a microwave before, and they had never owned and used the product samples before. Fifty percent of them were expert microwave users (i.e. they had used several brands of microwaves and used them daily or frequently), and the rest were novice users (i.e. they were not really familiar with microwaves and rarely used them). During the experiment, they were shown each of the product samples, one at a time. After observing each sample, participants were asked to fill out a PrEmo questionnaire to report their emotions that were felt when looking at the product.

The PrEmo questionnaire (Figure 3) is a modification of the original one developed by Desmet (2003a). Instead of animation, the questionnaire uses stills of the 14 expressive cartoons describing pleasant and unpleasant

emotions. Each still is accompanied by a three-point scale, representing the following ratings: 0: "I do not feel the emotion expressed by this cartoon", 1: "to some extent I feel the emotion", and 2: "I do feel the emotion". Each of the 14 emotions portrayed by the stills must be scored, since the product may elicit mixed emotions, that is, more than one emotion experienced simultaneously. In addition, each emotion is given a certain weight as follows: desire = 7, pleasant surprise = 6, inspiration = 5, amusement = 4, admiration = 3, satisfaction = 2, fascination = 1, boredom = -1, disappointment = -2, dissatisfaction = -3, unpleasant surprise = -4, disgust = -5, contempt = -6, and indignation = -7. These weights are determined subjectively based on the significance or impact level of the emotions. For example, desire is considered more impactful than pleasant surprise, as desire is likely to generate more influence on users' buying decision.
The total emotion score of a product can be calculated using Equation 1.

$$ES = \sum_{i=1}^{7} P_i \,.\, W_i + \sum_{i=1}^{7} U_i \,.\, W_i \qquad \text{(Equation 1)}$$

Where: ES = emotion score of a product, P_i = point of scale of pleasant emotion i, U_i = point of scale of unpleasant emotion i, and W_i = weight of pleasant/unpleasant emotion i.

Figure 3. PrEmo questionnaire for pleasant emotions

Table 1. Example of emotion score calculation for a product sample

Pleasant Emotions (P)	Scale (P_i)	Weight (W_i)	$P_i.W_i$
Desire	2	7	14
Pleasant surprise	1	6	6
Inspiration	0	5	0
Amusement	1	4	4
Admiration	0	3	0
Satisfaction	1	2	2
Fascination	0	1	0
		$\Sigma P_i.W_i =$	**26**

Unpleasant Emotions (U)	Scale (U_i)	Weight (W_i)	$U_i.W_i$
Indignation	0	-7	0
Contemp	0	-6	0
Disgust	0	-5	0
Unpleasant surprise	0	-4	0
Dissatisfaction	0	-3	0
Dissapoinment	1	-2	-2
Boredom	1	-1	-1
		$\Sigma U_i.W_i =$	**-3**

Table 1 illustrates the calculation of emotion score of product sample A from respondent 1 (ES = 26 + (-3) = 23). The total emotion scores of product sample A, B, and C from the 30 respondents are 351, 258, and 407 respectively. Thus, it can be inferred that product sample B (Ariston) is the lowest in eliciting positive emotional responses from the users. Therefore, this product went through usability testing and redesign process. Besides the product emotions measurement using the PrEmo questionnaire, interviews were conducted with the participants to explore the reasons of positive and negative emotions they felt toward the product samples. Based on the interview, product characteristics that could evoke positive and negative users' emotional reactions were identified. The followings are several examples of participants' comments: "the use of rotary knob seems impractical and out of date", "I like its light color, it looks clean", "symbols on control panel of product sample B are hard to understand", and "the boxy shape of this product looks dull and old-fashioned".

Usability Testing

Usability testing was conducted to product sample B (Ariston), since it had the lowest emotion score. Eight participants (5 females and 3 males, 4 expert users and 4 novice users, aged 21-45 years) participated in this test. The sampling method and screening criteria used were the same as in product emotions measurement. It was ensured that each participant had never owned and used the tested product before.

The usability testing collected objective and subjective data. The objective data consisted of success level in completing tasks and task completion time. Meanwhile, the subjective data comprised participants' opinions and impressions of the tested product, particularly of its usability.

The type of usability testing used was validation test, that is, the objective is to evaluate how the tested product compares to predetermined usability standards. In this research, maximum completion time (MCT) and successful completion criteria (SCC) were used as the standards. MCT is useful for evaluating participant performance and for establishing upper time limit for the completion of each task. MCT of each task was estimated based on mean completion time of the task from 5 expert users, adjusted by an allowance factor of 2.5. Meanwhile, SCC is a description of successful completion of the task. It was established based on usage procedure of the product as written on the manual instruction of product sample B. Table 2 provides examples of task, SCC, and MCT.

Table 2. Examples of task, successful completion criteria (SCC), and maximum completion time (MCT)

Task Description	Task Details		
Set the microwave oven to defrost frozen sausage for 10 minutes	SCC	1	The participant presses button no. 5, the time indicator displays blinking 00:00
		2	The participant sets the microwave timer by using rotary knob no. 11 until the time indicator displays 00:10
	MCT		25 seconds

Instruments for the test included orientation script, task scenarios, and questionnaire. After receiving a brief orientation about the test, participants were asked to skim the product's manual for about 1 minute to grasp the basic functions of the tested microwave. Then, they were prompted to perform a series of tasks (10 tasks) using the microwave B while being observed and video recorded. During the test, the participants were allowed to read the product's manual. The elapsed time and errors were recorded for each unique task on the task scenarios. Relevant participant behavior, comments, and any unusual circumstances that might affect the result were also noted. Having completed all tasks, each participant was debriefed for collecting their opinions about the tested product. These included comments about their performance, difficulties, and specific errors or problems during the test. Moreover, a questionnaire was employed to gather their subjective perceptions of usability of the microwave.

Results of usability testing is presented in Table 3, encompassing the average time to complete each task and the success level of task completion. On average, only 50% of tasks can be completed successfully within the time benchmark. Also, overall, only 38.75% of participants can complete the given tasks successfully within MCT and without errors (meeting SCC). This score can be considered as the usability level of microwave B. Additionally, about 5 out of 10 tasks can not be completed successfully by 50% or more of the participants. These results indicate that microwave B has usability problems, and redesign is required to improve its usability.

According to the questionnaire result, the percentage of respondents who rate the usability of microwave B very positive, positive, neutral, negative, and very negative are 12.5%, 40%, 22.5%, 22.5%, and 2.5% respectively. These ratings to some extent disagree with the outcomes of objective measures. Finally, based on the observation during the test and on the interview during the debriefing session, specific problems and difficulties experienced by the participants when using the product sample B can be identified. The problems mainly relate to such factors as efficiency, likeability, learnability, and guessability. For example, respondents found it very difficult to understand the symbols on control panel, because there was a lack of information labels. Furthermore, the symbols did not fully represent the actual functions as expected by the users.

Table 3. Results of usability testing

Task No.	MCT (sec)	Average Completion Time (sec)	[1]Success Level 1 (%)	[2]Success Level 2 (%)	[3]Unsuccessful (%)
1	25	23	50	37.5	12.5
2	35	34	62.5	37.5	0
3	40	40	50	37.5	12.5
4	50	**57**	0	37.5	**62.5**
5	40	33	75	25	0
6	25	**28**	25	25	**50**
7	25	22	50	37.5	12.5
8	50	**55**	37.5	0	**62.5**
9	35	**38**	25	12.5	**62.5**
10	60	**76**	12.5	25	**62.5**

[1]Success level 1: percentage of participants successfully completing the task within MCT and without errors (meeting SCC)
[2]Success level 2: percentage of participants successfully completing the task, but beyond MCT
[3]Unsuccessful: percentage of participants who do not successfully complete the task

PRODUCT REDESIGN AND EVALUATION

The redesign of microwave B was focused on its physical form, appearance, display unit, and control panel based on the results of product emotions measurement and usability testing. The objective was to improve the product usability and to create product that can elicit positive emotional responses from the users. Three redesign alternatives, as can be seen in Figure 4, were generated from the redesign process.

Figure 4. Three alternatives of redesigned microwave B

Evaluation of the three redesign alternatives was conducted using the previous methods: product emotions measurement and usability testing. The evaluation involved new respondents who had never taken part in the previous measurement or testing, so that the evaluation result would not be biased.

Thirty respondents (27 females and 3 males, 15 expert users and 15 novice users, aged 21-45 years) participated in the product emotions measurement. The procedure and instrument used were the same as in the previous measurement. However, the respondents only examined 3D pictures (virtual prototypes) of the redesigned product. Result of this measurement is recapitulated in Table 4. Compared to the original product, all redesign alternatives result in higher emotion scores, about twice as high as the original score. Majority of the respondents felt pleasant surprise, inspiration, and satisfaction. They found that the color and appearance of alternative 1 and 3 were attractive. Particularly for alternative 3, most respondents were surprised and inspired by its physical form and unique arrangement of its control panel.

Usability testing of the redesign alternatives employed 8 participants (6 females and 2 males, 4 expert users and 4 novice users, aged 21-45 years). The method and instruments used were slightly different from those in the previous test, because the products to be evaluated were virtual prototypes (3D pictures). The type of usability testing used was exploratory test, and interview technique was used to gather objective and subjective data. The participants were given 12 questions related to the redesign alternatives. Their answers were scored based on predetermined criteria, and the scores became a basis to determine if the redesign alternatives were acceptable to the users and if they provided good usability.

Maximum score that can be achieved by each alternative is 300, which is assumed as the maximum usability level of each redesign alternative. The average usability scores of redesign alternative 1, 2, and 3 from the 8 participants are 208.75, 218.75, and 210 respectively. Table 4 shows the usability level of each redesign alternative in percentage of the average score to the maximum score. It can be seen that redesign alternative 2 has the highest usability level, followed by alternative 3, 1, and the original product. During the test, all participants did not have difficulties in answering questions about the function of control buttons. They recognized the control buttons and their functions easily and correctly. This indicates that symbols, information labels, and control buttons arrangement are understandable and facilitate the users to use the microwave properly. Nevertheless, the test shows that participants frequently made errors related to the sequence of operating procedure due to the unavailability of manual instruction of the redesigned product.

Overall, all redesign alternatives of microwave B result in higher emotion scores and usability levels than those of original design. Selection of the best redesign alternative is determined by giving certain weights for each criterion. Since the good product appearance that can elicit users' positive emotional responses is assumed to be as important as product usability, both criteria are given the same weight, which is 50%. The calculation of total

scores for selecting the best alternative is presented in Table 5. Based on these scores, alternative 3 is selected as the best redesign of microwave B.

Table 4. Emotion score and usability level of original product and redesign alternatives

	Original Product (Microwave B)	**Redesign Alternative 1**	**Redesign Alternative 2**	**Redesign Alternative 3**
Emotion Score	258	584	548	**602**
Usability Level	38.75%	69.58% (208.75 of 300)	**72.92% (218.75 of 300)**	70% (210 of 300)

Table 5. Total score calculation for selecting the best redesign alternative

Redesign Alternatives	**Emotion Score (ES)**	**ES Weight 50% (W1)**	**ES x W1 (S1)**	**Usability Score (US)**	**Usability Weight 50% (W2)**	**US x W2 (S2)**	**Total Score (S1 + S2)**
1	584	0.5	292	208.75	0.5	104.38	396.38
2	548	0.5	274	218.75	0.5	109.38	383.38
3	602	0.5	301	210	0.5	105	**406**

CONCLUSION AND FUTURE DIRECTIONS

The use of the combination of product emotions measurement and usability testing is useful for evaluating existing product design, and for providing a basis for redesigning the product. This paper demonstrates the application of both techniques to evaluate and redesign microwave products. The redesigned product is proven to be better than the original design, and is expected to be able to attract consumers by evoking their positive emotional responses and by offering good usability.

In using PrEmo instrument to measure users' emotional reactions to microwave design, this paper introduces the use of different weights for each emotion felt by the users. These weights are considered important as each emotion may produce different level of impact or influence on the users. Currently, magnitude of the weights and their allocation to the 14 emotions in PrEmo are determined subjectively based on certain consideration. Different researchers, thus, may create different weights and ranking of the emotions. The use of these emotional weights requires further research to objectively determine their values and assignment to the emotions.

At the time this research was conducted, PrEmo was a validated and well-known instrument for measuring product emotions. Recently, however, Laurans and Desmet (2012) propose PrEmo2, which is claimed to be an improved version of the original PrEmo. Instead of three-point scale, PrEmo2 uses a five-point scale ranging from "I do not feel this" to " I do feel this strongly". Furthermore, PrEmo2 replaces several previous emotions with the new ones, such as social-context emotions (e.g. pride and shame). Another obvious difference is the use of new character style that maximizes expressivity. In future research, this instrument can be applied to the same products used in this study to investigate if its use produces different or better results.

REFERENCES

Desmet, P. M. A. (2003a). Measuring emotions: Development and application of an instrument to measure emotional responses to products. In M. A. Blythe, K. Overbeeke, A. F. Monk and P. C. Wright (Eds.), *Funology: From usability to enjoyment* (pp. 111-123). Dordrecht: Kluwer Academic Publishers.

Desmet, P. M. A. (2003b). A multilayered model of product emotions. *The Design Journal, 6*(2), 4-13.

Dumas, J. S., and Redish, J. (1999). *A practical guide to usability testing*. Exeter, England: Intellect Books.

ISO 9241-11. (1998). Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability. Switzerland: International Organization for Standardization (ISO).

Jordan, P. W. (1998). *An introduction to usability*. London: Taylor & Francis.

Khalid, H. M. (2006). Embracing diversity in user needs for affective design. *Applied Ergonomics, 37*(4), 409-418.

Khalid, H. M., and Helander, M. G. (2006). Customer emotional needs in product design. *Concurrent Engineering: Research and Applications, 14*(3), 197-206.

Laurans, G., and Desmet, P. (2012). Introducing PrEmo2: New directions for the non-verbal measurement of emotion in design. In *Proceedings of the 8th International Design and Emotion Conference*. London: Central Saint Martins College of Arts & Design.

Rubin, J. (1994). *Handbook of usability testing: How to plan, design, and conduct effective tests*. New York: John Wiley & Sons, Inc.

Stanton, N. A. (1998). *Human factors in consumer products*. London: Taylor & Francis, Ltd.

Zhang, P., and Li, N. (2005). The importance of affective quality. *Communications of the ACM, 48*(9), 105-108.