A|Z

Evaluating emotional response to products: The case of dishwashers

Beyza DOGAN¹*, Hatice Humanur BAGLI²
¹ doganbey@itu.edu.tr • Industrial Product Design PhD Program, Graduate School, Istanbul Technical University, Istanbul, Türkiye
² hbagli@marmara.edu.tr • Department of Industrial Design, Faculty of Fine Arts, Marmara University, Istanbul, Türkiye

Abstract
As the online shopping market proliferates, designers are pushed to start designing for pre-use user experience, focusing on how the user experiences the product during online shopping, namely before actual use. The goal of this study is to provide input to design professionals, helping them to design for pre-use user experience, in other words, design for products sold in the online shopping market. This study sheds light on pre-use user experience by proposing a questionnaire-based method investigating the "affect" dimension of UX. Throughout the study, following two pilot studies and an optimization process, 7 dimensions, which can be rated on a 5-point Likert scale, are generated to assess the emotional response to the dishwasher samples. While proposing a self-reported, practical, cost-efficient, quantifiable method to evaluate the "affect" dimension in UX, this study also posits a framework that can be used to adapt the method to other product genres.

1. Introduction

As the online shopping market is proliferating continuously and e-commerce penetration in Europe is predicted to grow to 67,1% in 2025, well exceeding traditional, in-store brick-and-mortar sales (Statista, 2021). This proliferation brings new challenges to designers who have been designing products that were once mostly sold through in-store shopping. Unlike the in-store shopping experience, the online shopping experience offers little to no room for exploring the actual use of the product, scaling user experience (UX) down to pre-use user experience. Interestingly, while there are many methods and tools to investigate UX during the actual use of the product, only a few of them explores pre-use UX (Oelfin & Schrepp, 2021). As pre-use UX increasingly becoming critical to design successful products due to the proliferation of the online shopping market and the pre-use impression of the user tends to stay mostly the same even after-use, this study focuses on pre-use UX by shedding light on the "affect" dimension of the user experience (Tractinsky et al., 2000). The "affect" dimension is one of the widely accepted key dimensions of UX, which can be defined as the emotional response of the user that is evoked by the interaction with a product (Park et al., 2013) (Park et al., 2018). This paper introduces a method for evaluating emotional response to dishwashers by rating the photographs of the appliance samples, thus focusing solely on pre-use UX.

To investigate emotional response, three Miele dishwasher samples are selected from the existing Miele product portfolio. The selected appliance samples can be seen in Figure 1 with product visuals.

The motivation for developing evaluation methods for the "affect" dimension originates from the fact that the most significant mission of a product designer is seen as stimulating affective responses in the user's mind through the shape of the product (Hsiao & Chen, 2006). So, to support the product design process, several methods in various product genres have been developed. Throughout the literature review, it has been observed that the "affect" dimension has frequently been used interchangeably by various researcher. Similar issue has also been observed by Picard in the context of affective engineering and AI. Picard has observed that the terms "emotional" and "affective" have been used interchangeable in the context of affective (Kansei) engineering and AI. (Picard, 2000) Most of the recent studies that are providing fundamental basis to our study are observed to be using the term "emotion" and its generated phrases like "emotional response" mostly. Therefore, this study makes conscious decision to use "emotion" as an equivalent term for "affect", for the sake of avoiding further confusions in the academic field. The emotion evaluation methods can be

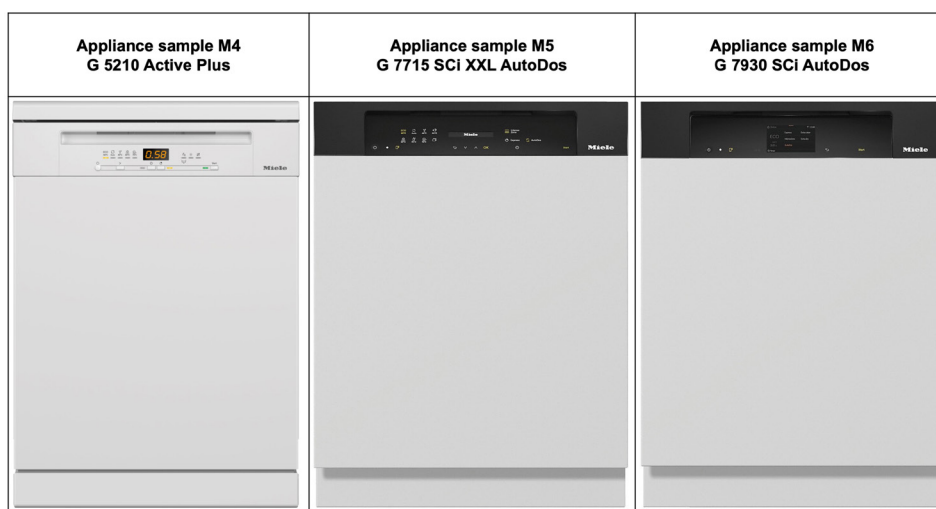| Appliance sample M4 G 5210 Active Plus	Appliance sample M5 G 7715 SCi XXL AutoDos	Appliance sample M6 G 7930 SCi AutoDos

Figure 1. *Appliance samples. Appliance samples from current product portfolio (Miele Gmbh, 2018).*

classified into two main groups: verbal and non-verbal. The non-verbal emotion evaluation tools like Emocards (Desmet et al., 2001), and Self-Assessment Manikin (SAM) (Morris, 1995) (Bradley & Lang, 1994) use visuals or pictograms instead of verbs. Non-verbal methods are largely implemented by many academicians due to the common belief about their high potential to capture unconscious emotional responses (Agarwal & Meyer, 2009). As they, in a way, claim to remove the language barrier between the emotions of the user and the survey. On the other hand, despite their language dependence, verbal methods for evaluation of the emotional dimension in UX are broadly recognized by academicians, as they are relatively easy to implement and adapt. The verbal methods usually use Likert scales and Semantic Differential scales to measure the emotional response (Agarwal & Meyer, 2009). As verbal methods are largely implemented and adapted to home appliances and electronic products, they are identified as the main focus for assessing the emotional dimension of UX in dishwashers. Throughout this study, the verbal emotion assessment method has been optimized to compensate for the language barrier that verbal methods bring due to their nature. On home appliances and electronic products, the evaluation of the emotional dimension in UX is conducted through various methods like Semantic differential scales (Macedo & Silva, 2014), Kansei Engineering (Delin, 2013), and Attrak-Diff (Bevan et al., 2016). Some of these methods use standard scales which can be rated on Likert Scales, while others offer custom scales specific to product genre by finding the adjectives describing the products, then using the adjectives to rate the products through Likert scales or semantic differential scales. For example, AttrakDiff offers one standard adjective set to define any product category leaving less room for adaptability (Bevan et al., 2016). On the other hand, the Kansei Engineering method is more flexible to adapt in any product category as the Kansei words need to be derived specifically for each product. Due to its adaptable nature, the Kansei Engineering method is cho-

sen to assess the dimension of emotion for the field research.

The main inspiration for this study, Kansei Engineering has initially been developed as a product development methodology by Mitsuo Nagamachi which analyses the user's feelings, emotional responses to products in order to stimulate intended emotional response through product design (Schütte, 2002). Kansei Engineering has been widely used in the industry to improve and compare products (Schütte, 2002). In the home appliance industry, it has been implemented in many appliances like dishwashers (Delin, 2013), washers (Ishihara et al., 2014), and vacuum cleaners (Skogman, 2002). Kansei engineering uses "Kansei words" to describe the emotional response generated by the product (Delin, 2013). The collection of Kansei words has been performed in various ways. While some researchers like Delin (2013) preferred collecting the Kansei words through questionnaires conducted on users and designers, some researchers use media materials like marketing materials, blogs, catalogs, etc. For example, Yodwangjai and Pimapunsri (2011) collected 130 Kansei words from various sources like magazines and websites, etc. about furniture through keyword analysis (Yodwangjai & Pimapunsri, 2011). Kansei engineering method has been re-interpreted throughout this study to assess and compare the appliance samples. Especially evaluative adjective (Kansei word) collection process is mainly inspired by Kansei engineering method and uses both data sources namely questionnaires to collect keywords from users and media materials to derive keywords.

While the quantity of Kansei words represents the comprehensiveness of the study, it also brings complexity to the process which would require intense filtering or grouping. At this point studies like Hsu et al. (2000) and Hsiao and Chen (2006) apply keyword analysis to reveal which Kansei word or adjective has the higher distribution frequency (Hsiao & Chen, 2006) (Hsu et al., 2000). Hsiao and Chen (2006) have worked on multiple product genres, namely automobile, sofa, and kettle, to reach overarching infor-

mation helping product designers to convey intended emotion through the product. To do this Hsiao and Chen (2006) derived a set of descriptive adjectives which they name "affective adjectives" from representative product visuals by conducting surveys. Hsiao and Chen (2006) collected 719 adjectives for the automobile, 510 adjectives the sofa, and 645 adjectives for the kettle product categories. After the manual elimination of duplicated, irrelevant, confusing adjectives they reached a list of 100 adjectives per product category. Through a series of card sorting, clustering, and ranking sessions, the adjective list is optimized and adjective pairs ("bipolar adjective pairs") are created to generate a semantic differential scale featuring 28 adjective pairs which can be used on all three product categories (Hsiao & Chen, 2006). Afterward, these 28-item adjective pairs are rated on a 7-point Likert Scale to assess the affective responses towards each product category.

Similarly, in this study keyword analysis is also used to collect and prioritize the evaluation adjectives defining emotional response to dishwashers. In addition to the keyword analysis, the sample adjective sets from existing studies in relevant product genres are also reviewed to support prioritization decisions. A series of pilot studies are conducted to optimize and improve the evaluation format. Consequently, to improve the practicality and efficiency of the rating process and the evaluation process, the adjective pairs are converted into nouns which can be rated on 5-point Likert Scale.

2. Method
The study seeks to investigate the "affect" dimension of the user experience and meanwhile proposes a framework that can be used to develop custom evaluation scales to assess the "affect" dimension of the UX in other product genres. The study consists of two main stages which are "Constructing the tool" and "Final Empirical study".

2.1. Stage 1: Constructing the tool
Stage 1 is the "Constructing the tool" phase which prepares and collects necessary background information for tool design and then tests the tool iteratively through a series of pilot studies to improve and finalize. Stage 1 has four steps; the first step is selecting the appliance samples to evaluate. Second step mainly aims to identify evaluation adjectives for the emotional perception of the appliance sample. The third step is conducting pilot studies to test the rating scale with adjective sets and optimize them. The fourth step is "Constructing the evaluation format" which aims to finalize the rating scale and the rating format through the evaluation of pilot studies.

2.1.1. Step 1: Selecting the appliance samples
The appliance samples are selected from the current product portfolio, aiming to include disparate interface technologies, from disparate price classes, thus with disparate forms in the study. Miele dishwasher portfolio is mostly consisted of three main categories: Semi-integrated dishwashers, Fully integrated dishwashers, Freestanding dishwashers. (Miele & Cie, Miele dishwasher range, 2021) A fully integrated dishwasher practically invisible among the kitchen cabinets. A fully integrated dishwasher has a completely hidden control panel and is designed to be flush with the kitchen cabinets. Due to their integrated invisible nature, they are not seen in the scope of this study. Similarly, semi-integrated dishwasher is a dishwasher that is designed to be integrated into the cabinetry of a kitchen, but with the control panel and display visible on the top of the dishwasher door. In the end semi-integrated dishwashers are also invisible to the naked eye except during a loading or unloading process. Due to it visibility, freestanding dishwasher category is chosen to focus on in this study, as Freestanding Dishwashers Category bring the clearest differentiation between the models. The appliance sample selection process in Freestanding Dishwashers Category is carried out with the aim of including wider scope of the line-up and making sure that the differentiation features among the samples are clear enough to support rating process in the empirical

study. In the end G 5210 Active Plus, Miele G 7715 SCi XXL AutoDos, Miele G7930 Sci Autodos are selected and within the study named as M4, M5, M6 respectively. After defining the appliance samples, a review of the appliance sample visuals quickly showed that there are many similarities between the models. As an example, all three appliance samples have an integrated handle. While Appliance sample M4 has a smaller integrated handle, Appliance sample M4 and M5 feature identical integrated handles. With the intention to find differentiation features between the appliance samples detailed user manual review is conducted revealing similarities and disparities between the models clearly. The control panels as well as the control panel technologies are found to be the sharpest differentiation feature between the models. While appliance sample M4 is featuring a control panel, swamped up with physical buttons and LED indicators, appliance sample M5 features a control panel which is awash with printed symbols with text. On the other hand, appliance sample M6 stands out from the rest with its touch based TFT panel, which considerably decreases the need for printed icons and results in a layout with less interaction elements.

Appliance sample M4 (Miele G 5210 Active Plus)

Miele G 5210 Active Plus is called as appliance sample M4 throughout the study. The appliance sample M4 is chosen from the current product portfolio to represent the lower segments of the dishwasher line-up. The price of the dishwasher at the time of this study was documented as 919,00 €, marking the bottom-line of the pricing in dishwasher category. Appliance sample M4 has full white front door, a fairly crowded control panel with an integrated handle. Appliance sample M4 features no touch panel. Instead, the control panel technology is solely based on push buttons which are occupied by LED indicators, satisfying feedback purposes. In terms of the control elements, there are 5 push buttons which are responsible from functions like on/ off, program

selection, timer, start and express. On the other hand, the feedback about the system status is given through 11 LED indicators. Besides, the control panel also features a small display which shows only a timer. Feedbacks like information about the active program and the status of the ongoing program are not shown on the display, instead the LED indicators are responsible from delivering such feedbacks. In appliance sample M4, aligned with its price segment and control panel technologies, the number of the available dishwashing programs can be found limited in comparison to rest of the Miele dishwasher category line-up. User is supposed to scroll through these programs by a push button. (Miele & Cie, Operating instructions for dishwashers- G 5210, 2018a)

Appliance sample M5 (Miele G 7715 SCi XXL AutoDos)

Miele G 7715 SCi XXL AutoDos is called as appliance sample M5 throughout the study. The appliance sample M5 is specifically chosen from the dishwasher line-up as the representative of the mid-segment models in the Miele dishwasher product category. At the time of the study the price of the appliance sample M5 is recorded as 2.159,00 €, fitting into mid-segment product class definition of the Miele brand. Appliance sample M5 has a discreetly integrated handle which houses the control panel on its back panel, making the control panel the most prominent form element on the dishwasher face. Unlike the appliance M4, M5 features a touch-based technology along with a context based clear text display. The display can show informative text and time. The available controls are formed through backlit icons with descriptive text which are printed behind the glass. On the control panel there are 19 touch buttons, except on/off button all touch buttons are backlit, indicating status information. Compared to the appliance sample M5, available dishwashing programs and the customization options are considerably broader. Hence, the number of the interface elements seems to be considerably more than

the appliance sample M4. The feedback system of the control panel relies on the clear text slim display screen and the backlit touch buttons performing as status indicators. Unlike the appliance sample M4, user does not need to scroll through the options by pushing the same button multiple times, instead every program had its own dedicated touch button, that provides a direct access. (Miele & Cie, Operating instructions Dishwasher- G 7715, 2018b)

Appliance sample M6 (Miele G7930 Sci Autodos)

Miele G7930 Sci Autodos is called as appliance M6 throughout the study. G7930 Sci Autodos is specifically selected from the dishwasher product category to represent the highest product segment. At the time of the study the price of the appliance sample M6 is recorded as 2.919,00 € marking top price segment of the dishwasher product category. The front door features a discreetly integrated handle which houses the control panel on its own back panel, identical to the appliance sample M5. The discreetly integrated handle makes the control panel center of attention. On the other hand, the number of interface elements is considerably less than the other appliance samples. The touch display is accompanied by 4 backlit printed touch buttons. The appliance M6 features more than 13 predefined programs with the options to create custom programs. (Miele & Cie, Operating instructions Dishwasher- G 7930, 2018c).

2.1.2. Step 2: Identification of evaluation adjectives

Due to the adaptable nature of Kansei Engineering, the method development is inspired by Kansei Engineering rather than AttrakDiff. As AttrakDiff offers a standard word pair set to assess any product genre, adaptability of the scale was found to be limited (AttrakDiff, 2015). On the other hand, in the Kansei Engineering method, the evaluation adjective list is not a standard list that claims to fit in every product genre, instead the evaluation adjectives are defined through keyword analysis of relevant sources. So Kansei Engineering method mostly starts with definition of the input sources which would be analyzed to define scale items through set of research activities like keyword analysis, clustering, grouping, etc. In this study, aligned with the Kansei Engineering method, relevant sources are defined as online dishwasher buying guides, blogs, the official Miele website dishwasher page, and product manuals of the appliance samples. While the official Miele website dishwasher page, and product manuals of the appliance samples brings model specific input, online dishwasher buying guides, blogs brings in data from customer perception, expectation by introducing marketing terms in the keyword analysis process. Vast amount of online blog material in form of buying guides and reviews required the study to focus on recent and most influential materials. At the time of the keyword analysis the most recent articles about dishwashers were found at Cnet, which is the oldest and one of the most influential online platform covering tech news, product reviews, how-to guides, product prices. (Cnet, 2018) The most comprehensive article on the Cnet platform has been picked by the researchers for further analysis. (Gebhart, 2018) On these prementioned source a keyword analysis is conducted to collect the evaluation adjectives defining emotional responses to dishwashers. The scope of the defined data sources inevitably affects the accuracy of rating as complete scale is derived from them. According to keyword analysis, an initial set of adjectives are defined. Figure 2 shows the initial set of adjectives with the result of the keyword analysis through expression count, Frequency, and Source. Top ranking twenty-four evaluative adjectives are picked. By doing so, keywords which are not directly describing the product are ignored.

The collected evaluation adjectives are grouped and organized through a cluster analysis featuring hierarchical organization, binary correlation matrix, and antonym pairs generation to form an initial set of adjectives. Throughout the process, researchers

Evaluation Adjectives	Expression count	Frequency	Source
automatic	2	0.10%	Appliance buying guides, blogs*
efficiency	2	0.10%	Appliance buying guides, blogs*
economical	5	0.50%	Miele Webpage**
integrated	2	0.10%	Appliance buying guides, blogs*
quiet	4	0.10%	Appliance buying guides, blogs*
quick	2	0.10%	Appliance buying guides, blogs*
light	3	0.30%	Miele Webpage**
flexible	2	0%	Product manual G6920 Sci***
bulky	2	0.20%	Miele Webpage**
convenient	2	0.20%	Miele Webpage**
comfortable	2	0%	Product manual G6480***
	2	0%	Product manual G6920 Sci***
versatile	2	0.20%	Miele Webpage**
performance	10	0.30%	Appliance buying guides, blogs*
cleaning power	3	0.10%	Appliance buying guides, blogs*
wide	3	0.10%	Appliance buying guides, blogs*
nice	3	0.10%	Appliance buying guides, blogs*
easy	3	0%	Product manual G6920 Sci***
	2	0%	Product manual G6480***
large	3	0.10%	Appliance buying guides, blogs*
smart	2	0.10%	Appliance buying guides, blogs*
small	2	0.10%	Appliance buying guides, blogs*
fashioned	2	0.10%	Appliance buying guides, blogs*
loud	4	0.10%	Appliance buying guides, blogs*
less space	2	0.10%	Appliance buying guides, blogs*
fit	6	0.20%	Appliance buying guides, blogs*

Figure 2. *Keyword analysis results. Initial evaluation adjectives set (Kansei Words) collected from various sources through Keyword analysis (Cnet, 2018; Miele Gmbh, 2018; Miele & Cie, Miele dishwasher range, 2021).*

omitted certain evaluative adjectives that bring incomprehensibility and irrelevancy to the process like versatile, wide, etc. Besides researchers also decided to test the performance of an additional adjective pair: "cheap X expensive". As a result of cluster analysis, an initial set of adjective pairs (antonyms) is generated, and accordingly, the semantic differential scale is formed.

2.1.3. Step 3: Pilot studies
Two sets of Pilot studies are conducted to implement and optimize the initial set of adjective pairs, transforming them into an optimized scale that can evaluate the emotional response on a 5-point Likert Scale (see Figure 3). The details of each Pilot study through sample size, material, evaluation scale, and task are shown in Figure 4. Pilot study 1 included an additional open-ended option for the "How does it look"

question with the aim of collecting new adjectives which might come up during the study. This additional data collection method has been removed for the second round of the Pilot study and Pilot Study 2 is fully focused on testing the scale items.

Throughout this phase, to optimize the evaluation scale, follow-up interviews are conducted and the adjective pairs are compared with adjective pair sets from existing studies from Home Appliances Market. The adjective pairs from existing studies are shown in Figure 5.

2.1.4. Step 4: Constructing the evaluation format
Follow-up interviews are conducted with the participants of Pilot Study 2. Pilot Study 2 showed that the adjective pairs have been still not completely optimized. During the field study, it is observed that

Figure 3. *Question format - pilot study 1. A screenshot from the online questionnaire showing the questionnaire format (Dogan, 2021).*

	Subject	Material	Evaluation Scale	Task
Pilot Study 1	12 users 5 males and 7 females between 18-40 years old with a college degree	Visuals of M4, M5, M6	Each subject is asked to rate the appliances on a 5-point Likert Scale. The rating scale is formed by positioning 5-Point Likert Scale in the middle of the adjective pair. While the adjectives describing negative feelings are positioned on left the adjective describing positive feelings is positioned on right.	The visuals of the appliance samples are shown to the participant one by one with the question of "How does it look?" The participant is expected to rate appliance samples on a 5-point Likert scale.
Pilot Study 2	20 users, between 18-45 years old with college degrees.	Visuals of M4, M5, M6		

Figure 4. *Structure of the pilot studies. Pilot study 1 and pilot study 2 (Dogan, 2021).*

Silva & Macedo (2014)	Neves (2011)
cheap, clean (aesthetics), compact, complex, durable, easy, elegant, expensive, fast, flexible, fragile, functional, heavy, intuitive, large, light, noisy, reliable, robust, spacious, useful	pleasant, neat, controllable, dynamic, stimulating, easy to clean, easy to use, honest, strong, multifunctional, organized, surprising
Automatic, Beautiful, comfortable, difficult, ecological, economical, efficient, ergonomic, innovative, smart, clean (cleaning), modern, practical, resistant, secure, silent, simple, technological	

Figure 5. *Adjectives collected by Silva & Macedo (2014) and Neves (2011). Adjective sets from Silva & Macedo (2014) and Neves (2011) (Macedo & Silva, 2014; Neves, 2011).*
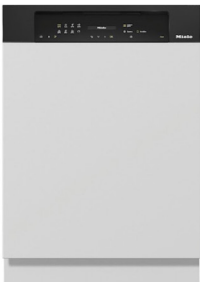
participants were inclined to question the scale items, struggling to comprehend, most importantly getting occasionally bored as follow-up interviews have also validated this observation. It can be concluded that the adjective antonym pairs which are mapped on the semantic differential scale are found to be impractical, time-consuming, hard to comprehend, and in the end thickening the language barrier. More concrete input coming from the follow-up interviews showed that to be able to make sense of the scale, participants are required to comprehend, develop an understanding of 2 evaluative adjectives per dimension and then rate it. This inevitably extends the response time and putting extra pressure on the shoulders of the participants. Consequently, to optimize the rating process, the adjective pairs are converted into nouns, forming dimensions that can be rated on 5-point Likert Scale, bringing in practicality and comprehensibility as seen in Figure 6. Figure 7 shows the comparison between initial and final evaluation scale format. The overall evolution of the evaluation scale format for each keyword can be seen in Figure 9. In its optimized version, the participants are required to comprehend one evaluative adjective to rate every dimension. Testimonial collected from the participants of final field study, confirmed that the scales featuring one adjective per dimension were found to be more practical by majority of the participants.

Figure 6. *The final question format. The final question format with derived dimensions of emotional response on a 5-point likert scale (Dogan, 2021).*

2.2. Stage 2: Final empirical study

Stage 2 is the final phase of the study which tests the final online questionnaire-based method to evaluate the emotional dimension of UX in dishwashers. In this phase final rating scale which is formed by one evaluative name instead of two evaluative adjectives is used. In total emotional response to the appliance samples are measured in seven separate dimensions on a 5-point Likert scale, through seven evaluative names defining these dimensions. The dimensions can be listed as, Practicality, Smartness, Elegance, Modernity, Price, Compactness, Performance.

Subjects: 56 users between 18-65 years old with college degrees.

Material: Online questionnaire with six appliance samples selected from Miele Museum and the current product portfolio. Appliance samples are shown in Table 1.

Procedure: Each subject is asked to complete the online questionnaire. The subjects are allowed to proceed at their own pace.

Task: The visuals of the appliance samples are shown to the participant one by one with the question "How does it look?" The participant is expected to rate appliance samples on a 5-point Likert scale. The semantic differential scale which is optimized as a result of Pilot Study 1 and Pilot Study 2 is implemented through the final empirical study.

2.2.1. Results of the field study

As shown in the Figure 8, it appears that the newest appliance sample, M6, received a high emotional response with an average score of 3.9. It received high scores in the dimensions of smartness (4), elegance (3.8), modernity (4), price (4), and performance (4). This suggests that consumers generally view M6 as a smart, modern, and high-performing appliance that is worth the higher price.

The sample M5 received the second highest overall emotional response with an average score of 3.8. It received high scores in the dimensions of practicality (3.9) which is higher than M6, smartness (3.8), elegance (3.6), modernity (4), price (3.8), and performance (3.8). M5 scores higher than M6 in practicality dimension. It suggests that consumers view M5 as being more practical than M6, conceivably because it has smaller display and bigger printed icons, which obviously offers a bigger surface area to navigate.

The oldest appliance sample, M4, received the lowest overall emotional response with an average score of 3.3. It received lower scores in the dimensions of practicality (3.2), smartness (3.4), elegance (3.1), modernity (3.4), compactness (3.2), and performance (3.4). This suggests that consumers view M4 as being less modern, less smart and less elegant than the other two samples. This view would probably stem from multiple reasons. While appliance sample M4 stands out with its considerably lower interaction technology featuring push buttons rather than touch technologies. This also affects the basis CMF which is used in the construction of the appliance. Unlike glass-based appliance samples M5 and M6, the prominent surfaces in appliance sample M4 are based on colored plastic derivatives.

In conclusion, the newer technology and features and CMF which are asso-

Initial Evaluation Scale Format							Final Evaluation Scale Format						
	1	2	3	4	5			1	2	3	4	5	
hard to use	○	○	○	○	○	easy to use	Practicality	○	○	○	○	○	

Figure 7. *The evaluation scale format. Initial evaluation scale format vs Final evaluation scale format showing how the scale is optimized (Dogan, 2021).*

Appliance Samples	Practicality	Smartness	Elegance	Modernity	Price	Compactness	Performance	Average
M4	3.2	3.4	3.1	3.4	3.3	3.2	3.4	3.3
M5	3.9	3.8	3.6	4	3.8	3.7	3.8	3.8
M6	3.8	4	3.8	4	4	3.7	4	3.9

Figure 8. *Results of field study. Ratings for emotional response per dimension on a 5-point likert scale.*

ciated with newer technologies such as touch panel technology, and glossy glass surfaces, tend to induce higher emotional response in general with one exception: practicality dimension. Practicality dimension is found to be an exception where the appliance sample M5 scores higher than M6. While both of the appliance samples feature touch technologies, the appliance sample M6 features considerably more advanced technology than appliance sample M5. But featuring more advance technology has not helped the appliance M6 to induce better emotional response in practicality dimension.

3. Results and discussion
The paper introduces a method for evaluating emotional responses to home appliances and presents a framework that would be used to implement the tool in various product genres. The presented method is a calculated combination of existing evaluation methods with the aim of improving practicality and efficiency of the rating process and evaluation process. To do this, the existing methods found in the literature for emotion evaluation have been reviewed. As the target product genre belongs to home appliances, especially the methods performed in home appliances have been explored. In addition to the methods implemented in home appliances, the methods evaluating emotional response to consumer electronics have also been investigated. The comparative review of main existing methods like Semantic differential scales (Macedo & Silva, 2014)

assessing washing machines, Kansei Engineering assessing dishwashers (Delin, 2013), and AttrakDiff assessing range hoods (Bevan et al., 2016), etc. revealed several criteria to consider to improve practicality and efficiency of the methods. These criteria could be listed as adaptability to diverse product genres, quantifiability of the results, and the practicality of submitting input.

Regarding "adaptability", the literature review of the existing methods showed that the standard scale used by AttrakDiff is being implemented without adaptation into every product genre (Bevan et al., 2016). In this sense, the process of generating the Kansei word for the Kansei Engineering method is found to be bringing in new levels of adaptability along with complexity (Delin, 2013). The adaptable nature of the proposed tool is mainly inspired by Kansei Engineering rather than AttrakDiff, however, the method that has been developed to reach a scale to evaluate the emotional dimension of user experience creates a framework the list of attributes achieved creates a standard that can be used for different product categories as AttrakDiff suggests. The way that researchers arrange the keyword sources ensures the required room for adaptability. While the method is implemented on dishwashers, the framework could easily be implemented in any product genre. Regarding "quantifiability", the verbal methods for evaluating emotional response have been targeted from the beginning due to their easy-to-quantify nature (Agarwal & Meyer, 2009). Regarding the

	Collected Evaluation adjectives	Initial adjective antonyms		Adjective pair set of Pilot Study 1		Adjective pair set of Pilot Study 2		Final Scale
Use	comfortable	uncomfortable	comfortable	uncomfortable	comfortable	practical	impractical	
	easy	hard to use	easy to use	hard to use	easy to use			Practicality
	practical	impractical	practical	impractical	practical			
	versatile	limited	versatile	mono-functional	multifunctional			
	efficiency	less efficient	efficient					
	convenient	inconvenient	convenient					
	automatic	less automatic	automatic	less automatic	automatic	not smart	smart	
	smart	not smart	smart	not smart	smart			Smartness
	flexible	inflexible	flexible	inflexible	flexible			
	less space	low capacity	high capacity					
	elegant	inelegant	elegant	inelegant	elegant	inelegant	elegant	
	nice	not nice	nice					Elegance
	Antique							
	weird							
	fashioned							
	old-fashion	old -fashioned	contemporary modern	old fashioned	modern	old fashioned	modern	Modernity
	integrated	disintegrated nonintegrated	integrated	complex	clean			
Aesthetic	fit	fitted	not fitted					
	bulky	bulky	compact	bulky	compact	bulky	compact	
	wide							Compactness
	light	heavy	light					
	small	small	large					
	large							
	economical	less economic	economic					Price
	expensive	cheap	expensive	cheap	expensive	cheap	expensive	
Performance	quiet	loud	silent quiet	loud	silent			
	loud							Performance
	quick	slow	fast	slow	fast			
	performance	low performance	high performance	low performance	high performance	low performance	high performance	

Figure 9. The evolution of the scale for evaluating emotional response. The evolution of rating format and scale items for the tool for evaluating the emotional response throughout the study.

"practicality" of submitting input, the method development has started with a semantic differential scale which is developed with the Kansei method and the semantic differential scale has been through multiple optimization processes throughout multiple pilot studies. The pilot studies are followed up by interviews through which the ratings are discussed with participants. During follow-up interviews, two main problems are enunciated by the participants and observed by the researcher: (a) the excessive number of scale items to rate and (b) the language barrier posed by the adjective pairs. To explain the first problem further, the rating process was mostly found lengthy. So, the scale items are reduced to seven, implementing Miller's Law (Miller, 1956). The second problem, the language barrier, was actually stemming from the nature of the semantic differential scale. To rate only one scale item, the subjects were required to comprehend two adjective pairs, then rate. This requirement extends the response time and requires additional cognitive effort. To fix this problem, the adjective pairs are converted into simple nouns, which would later be called as dimensions. The performance of the final evaluation scale is reviewed through another round of follow up interviews, confirming that converting adjective pairs into simple single nouns simplified the rating process for the participants.

In the end, the study synthesizes the method development processes from Kansei Engineering to bring in more adaptability and scale optimization processes from Semantic Differential Scale methods to bring in the practicality of evaluation and input collection to reach more efficient, quantifiable, and adaptable tools to evaluate emotional response. The evolution of the scale for evaluating emotional response starting from collected evaluation adjectives, generated adjective pairs, and optimized semantic differential scale to the final scale featuring dimensions can be seen in Figure 9.

Consequently, the study presents a framework that could be used by researched who would like to evaluate emotional responses in products. The framework can be implemented in other product genres as well. The simple steps presented in the framework to follow can be listed as:

• Stage 1: Constructing the tool

o Step 1: Selecting the appliance samples

o Step 2: Identification of evaluation adjectives

o Step 3: Conducting Pilot Studies

o Step 4: Constructing the evaluation format

• Stage 2: Conducting final empirical study

4. Conclusion

Even if the sample set is focused on a specific product genre, the study aims to propose a method that can be

adapted to other product genres and help designers to test the emotional response to products in the most practical way possible. The method provides advantages compared to the existing methods in the sense of adaptability, quantifiability, and practicality.

Firstly, the adaptable nature of the tool development helps to reach more accurate results by ensuring product genre-specific rating scales.

Secondly, considerably simplified rating format and dimensions help the researchers to easily quantify data for comparison and evaluation of the emotional dimension in user experience, improving quantifiability.

Thirdly, the refined rating scale of the tool which is distilled down to 7 dimensions rated on a 5-point Likert Scale brings practicality and agility to the input submission process and motivates participants.

In the end, the results which this tool provides cannot only shed light on the emotional response but also, can pose solid research questions or hypotheses to explore the link between the emotional response (affect) and various factors like the level of technology, certain interface technologies, focus of the design approach, CMF design, etc. Bringing such perspectives into discussion would further help product management professionals to form more educated design briefs and design practitioners to take informed design decisions.

5. Limitations and further studies

As this tool is developed as a smaller part of an ongoing Ph.D. thesis, the pilot studies, and the final empirical study are performed on appliance samples representing major milestones in the evolution of kitchen technologies. Still, the study details the tool development process with the aim of providing a framework that can be implemented in other product genres and performed on other product models. The choice of appliance samples provided a wide scope of diversity in terms of features and technologies, ensuring a high level of differentiation among the samples. It can be said that the final

result of the field study is also affected by this specific appliance sample choice making the comparison-based evaluation of the results sharper. So, it can be concluded that the comparison capacity of the tool needs to be further assessed by implementation in products with similar technologies. Moreover, the final scale, which originated from the adjective pair sets derived from product genre-related data sources, is not designed to be used in other product genres directly. But the framework that shows how adjective sets can be generated and then converted into nouns to reach the final scale can be implemented in other product genres and used to evaluate emotional response. As the final scale generated through the study has been formed through a series of optimization processes direct implementation of the 7-dimension scale in similar product genres can also be considered in further studies to verify the adaptability of the scale. Similarly, while the adaptable nature of the tool development method poses a significant advantage, further research on the adaptation of the tool in other product genres needs to be conducted to improve the adaptability and flexibility of the tool. To verify the improvements that the tool brings, the practicality of the input submission process and quantifiability of the evaluation process should be compared with existing methods like SD scales, Kansei-Engineering based SD scales, and AttrakDiff.

References

Agarwal, A., & Meyer, A. (2009). Beyond usability: evaluating emotional response as an integral part of the user experience. *In CHI'09 Extended Abstracts on Human Factors in Computing Systems*, (pp. 2919-2930).

AttrakDiff. (2015). *Evaluation Report for Product "Demo - B"*. User Interface Design GmbH.

Bevan, N., Liu, Z., Barnes, C., Hassenzahl, M., & Wei, W. (2016, May). Comparison of Kansei Engineering and AttrakDiff to evaluate kitchen products. *Human Factors in Computing Systems In Proceedings of the 2016 CHI Conference Extended Abstracts* ,

(pp. 2999-3005).

Bradley , M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1), 49-59.

Cnet. (2018, June 21). *Homepage.* Retrieved from Cnet: https://www.cnet.com/

Delin, A. (2013). *Enhancing the perceived quality of a dishwasher through user-centered design.* Master of Science Thesis.

Desmet, P., Overbeeke, K., & Tax, S. (2001). Designing products with added emotional value: Development and application of an approach for research through design. *The design journal, 4*(1), 32-47.

Dogan, B. (2021). *Measuring Cognitive Gap in Evolution of Kitchen Appliances.* Retrieved from Google forms: https://forms.gle/hgWZy8TAhnET-j9db9

Gebhart, A. (2018, June 21). *How to buy dishwasher.* Retrieved from Cnet: https://www.cnet.com/home/kitchen-and-household/dishwasher-buying-guide/

Hsiao, K. A., & Chen, L. L. (2006). Fundamental dimensions of affective responses to product shapes. *International Journal of Industrial Ergonomics*, 36(6) 553-564.

Hsu, S. H., Chuang, M. C., & Chang, C. C. (2000). A semantic differential study of designers' and users' product form perception. *International journal of industrial ergonomics, 25*(4), 375-391.

Ishihara, S., Nagamachi, M., & Ishihara, K. (2014, June). Electronic Product Development with Kansei Engineering/Kansei Ergonomics. *In International Conference on Kansei Engineering and emotion research* (Vol. 17, pp. 1385-1395).

Macedo, V., & Silva, C. (2014). Building a Semantic Differential Scale as Tool for Assisting UX Evaluation with Home Appliances. *In International Conference of Design, User Experience, and Usability* (pp. 308-317). Springer, Cham.

In *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience: Third International Conference, DUXU 2014,* *Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part I 3* (pp. 308-317). Springer International Publishing.

Miele & Cie, K. (2018a, January). *Operating instructions for dishwashers-G 5210.* Retrieved from www.miele.com: https://media.miele.com/downloads/40/fb/00_96106EC61D0B1EEC-B5EA9CEB038040FB.pdf

Miele & Cie, K. (2018b, January). *Operating instructions Dishwasher- G 7715.* Retrieved from www.miele.com: https://media.miele.com/downloads/cd/ce/02_C7F128283F7D1EEEA3D-366E59444CDCE.pdf

Miele & Cie, K. (2018c, January). *Operating instructions Dishwasher-G 7930.* Retrieved from www.miele.com: https://media.miele.com/downloads/4d/d4/03_C7F128288AE61EEE-A3D31CE943934DD4.pdf

Miele & Cie, K. (2021, January). *Miele dishwasher range.* Retrieved from miele.co.uk: https://www.miele.co.uk/e/dishwashers-1015063-c

Miele Gmbh. (2018, June 8). *Willkommen bei Miele- Immer besser.* Retrieved from Miele Shop: https://www.miele.de/haushalt/shop-3996.htm

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review, 63*(2), 81.

Morris, J. D. (1995). Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. *Journal of advertising research, 35*(6), 63-68.

Neves, A. (2011). Experiência cronológica do design de Eletrodomésticos. Programa de Pós-graduação em Design UFPR. Master Thesis.

Oelfin, D., & Schrepp, M. (2021). Evaluation of Anticipated User Experience. *Mensch und Computer 2021-Workshopband.*

Park, J., Han, S. H., Kim, H. K., Cho, Y., & Park, W. (2013). Developing elements of user experience for mobile phones and services: survey, interview, and observation approaches. *Human Factors Ergon. Manufact. Service Ind., 23*(4), 279–293.

Park, J., Han, S. H., Park, J., Park, J., & Kwahk, J. (2018). Development of a

web-based user experience evaluation system for home appliances. *International Journal of Industrial Ergonomics*(67), 216-228.

Picard, R. W. (2000). *Affective computing.* MIT press.

Schütte, S. (2002). *Designing feelings into products: Integrating kansei engineering methodology in product development.* (Doctoral dissertation, Linköping University Electronic Press).

Skogman, A. (2002). Measurements of user feelings regarding existing vacuum cleaners with Kansei Engineering: an evaluation of Kansei Engineering and other methods.

Statista. (2021, May 20). *Digital buyer penetration in Europe from 2017 to 2025.* Retrieved from Statista: https://www.statista.com/forecasts/891317/digital-buyer-penetration-in-europe

Tractinsky, N., Katz, A., & Ikar, D. (2000, March 24). What is beautiful is usable. *Interacting with Computers, 2*(13), p. 127±145.

Yodwangjai, S., & Pimapunsri, K. (2011). Application of Semantic Differential Technique and Statistical Approach to Evaluate Designer's and Consumer's Perception in Furniture Design. *Applied Science and Engineering Progress, 4*(1), 23-30.