Towards Remote Assessment of User Experiences of Visual Product Representations

Kristav, Per; Warell, Anders

Published in:
[Host publication title missing]

2009

INTERNATIONAL CONFERENCE ON DESIGNING PLEASURABLE PRODUCTS AND INTERFACES, DPPI09
13- 16 OCTOBER 2009, COMPIEGNE UNIVERSITY OF TECHNOLOGY, COMPIEGNE, FRANCE

TOWARDS REMOTE ASSESSMENT OF USER EXPERIENCES OF VISUAL PRODUCT REPRESENTATIONS

Per Kristav[1], Anders Warell[2]
[1]PhD student, Adjunct Lecturer, Department of Design Sciences, Faculty of Engineering, Lund University, Lund, Sweden. per.kristav@mkon.lth.se
[2]PhD, Associate Professor, Department of Design Sciences, Faculty of Industrial Design, Lund University, Lund, Sweden. anders.warell@design.lth.se

ABSTRACT
Understanding consumer experiences of products is becoming increasingly important for producers acting on the global market. This paper presents a pilot study done as part of the development of a VIPET (Visual Internet Product Experience Tool), a future tool for remote user assessment of product experiences. The aim of the tool is to provide users with a convenient and intuitive approach to assess visual representations of products by allowing respondents to place visual representations of products in the form of images in relation to each other on a type of bipolar visual analogue scales in the form of 2x2 charts. The purpose of the study was to evaluate a prototype version of the visual interface of the tool using international respondents. The prototype tool was emailed to respondents in five countries, asking them to assess an everyday type of product, food packaging containers, against six parameters, measuring types of perceptual experience. The results of the study indicate that remote assessment of visual representations of products using a tool of this type is a promising approach. Even though some test results are presented the major interest of this paper is not in the assessments of the specific products as such, but rather the characteristics and quality of the assessment method itself and the implications for the development of a tool for assessing consumer experiences.

Keywords: Product assessment, international respondents, evaluation tool, user experiences, remote assessment.

1 INTRODUCTION
In product design it is important to capture and assess subjective user perceptions and consumer experiences of products [1], [2]. However, perceptions and experiences of products are not always sufficiently understood or considered in the product development process. As a consequence, product requirements used during the design process may not sufficiently consider perceived values of actual consumers, which are increasing in importance for subjective evaluation of products. As designers' conceptions of product experiences, such as symbolism and user experience do not necessarily match those of the user group [3], products may become more attractive to design practitioners than to target consumers, thereby missing important market targets.

Approaches for eliciting and understanding user experiences of products in 'objective' ways, thereby potentially increasing the likelihood that experience related requirements are communicated effectively within organisations and design processes, have appeared during the last decade [4]. Sperling and Eriksson [5] and Sperling [6] developed a tool, the User Compass Chart (UCC), inspired by the sector chart used by Russell [7] for measuring affections. Sector charts have been used in industrial strategic product planning of future products in relation to existing competitors and to direct design towards a target. They are also frequently used by practising industrial designers as a research tool. One example is a cross-cultural study were 30 subjects from the United kingdom and Finland were asked to sort 36 photographs of living room chairs into piles of semantically similar chairs and then to describe each pile. The results of this study were then presented as a chart with two intersecting parameters [8].

Although effective in understanding perceptual experiences of users, such tools often require very specific resources in terms of facilities, hardware, personal guidance, and training. For example, the procedure of using the UCC tool was resource demanding and required continuous assistance by test leaders in order to guide respondents through the complex multistep procedure of using the tool. As such, it is limited in use due to constraints in resources, time and place. Furthermore, it requires that participants are physically present in order to provide guidance during use, which makes reaching users on international markets a costly and time consuming activity.

The original UCC consists of a physical 'game board' with two intersecting vertical and horizontal axes forming a 2x2 grid measuring distinct dimensions. The UCC tool was developed to measure user perceptions, feelings and mental projections of visual product qualities. Each axis of the game board features a bipolar scale, which is labelled at each endpoint with opposite adjectives. Each adjective is presented together with synonyms provided in order to guide respondents in the understanding of the dimensions used. The axes represent two intersecting rating scales with the central part of the UCC featuring a 'neutral' (neither/nor) zone. Adjectives with positive connotations are located in the north-east (or upper right) sector (see example of a UCC-chart in Figure1).

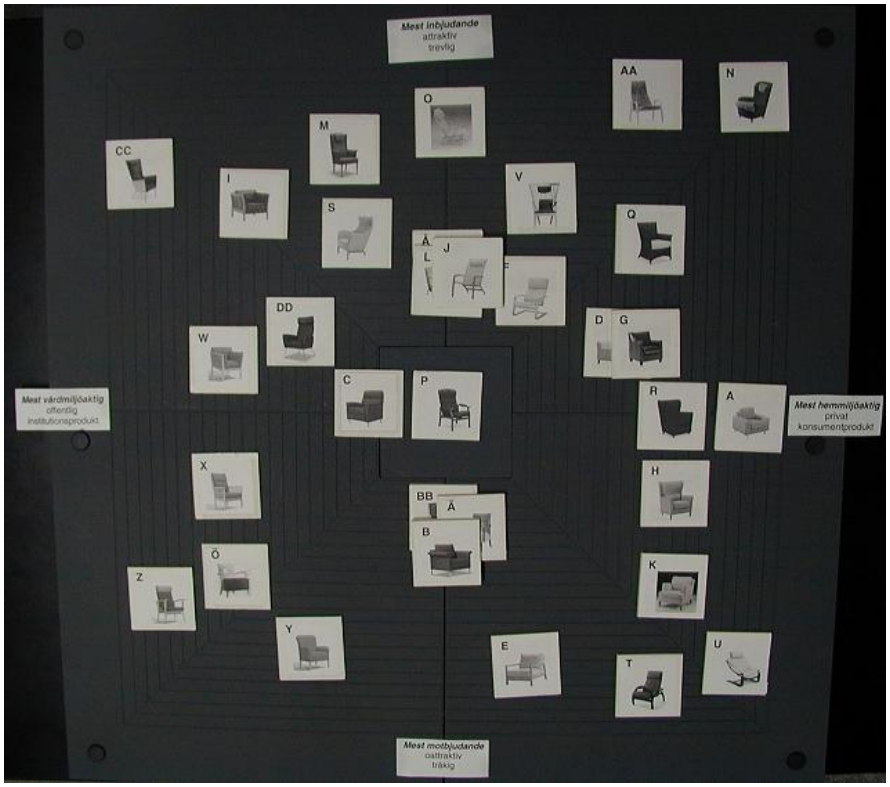


Figure 1. Example of a 2x2 UCC chart with axes featuring bipolar parameters for assessment, showing pieces of furniture placed in relation to two bipolar scales by a respondent in a usability lab setting.

In the UCC tool, all images of products or design details are displayed simultaneously at the time of assessment, which gives the respondent a favourable opportunity to compare and relate them to each other. The tool also gives the respondent an easily accessible opportunity to readjust all the placements in the test. In the design process, establishing such relative relations between assessments of products or details are often as important as giving each product assessment an exact value.

Sperling and Eriksson [5] used the UCC tool to categorize user experiences of material samples in a study of automotive interior materials, while Sperling [6] used pictorial product representations to study perceptions of future easy-chairs (see Figure1). The UCC tool has been used in a usability lab, where respondents are asked to position representations such as material samples or product images in

the sectors in relation to how they rate the sample with respect to the dimensions of the two axes. While positioning the samples, respondents are asked to think aloud about their reasoning for each placement. The experiment and the completed boards are audio-visually documented for further analysis. The UCC tool was positively received among participants, partly as it provided a 'play-like' interaction with the game board, adding a level of enjoyment to the test itself.

Studies with some similarities to UCC have been carried out in usability laboratory environments; however they did not involve remote users or customers [9], [10], [11], [13], [14], nor have they used the internet to assess consumer experiences of products as in this study. Reaching out to large numbers of geographically dispersed respondents at a reasonable cost and effort is potentially interesting for global manufacturers and marketing researchers within industry. There is thus an opportunity for a tool which alleviates some of these shortcomings by enabling less resource demanding assessment procedures, thus also facilitating access to geographically dispersed markets.

2 OBJECTIVE

It is essential to assess how respondents understand the tool with respect to how perceptions are elicited. The purpose of this research is to support the development of VIPET (Visual Internet Product Experience Tool), an emerging internet based tool for quantitative analysis of visual data [12]. This tool is intended to meet a need by manufacturers to understand consumer perceptions of non-utilitarian perceived values of products. In order to facilitate access to larger numbers of respondents through remote communication using the internet, the tool is intended to be self-instructional and suitable for use in a web based application.

The objective of this paper is to evaluate a preliminary version of the VIPET tool for remote product assessment. While the VIPET has some similarities to the UCC tool, such as simultaneous presentation of multiple product representations and the relative placement of representations in relation to each other along analogue scales, it also presents some distinct differences in terms of the remote interaction aimed for in order to achieve more effectively assessments with larger numbers of respondents. Is it possible to design a tool that alleviates the cognitive challenges for respondents experienced with the physical UCC tool? What is the nature of an interface and procedure, which allows respondents to interact with the tool in a self-guided assessment without the support required in the usability lab environment used for UCC assessments?

One final objective of the evaluation of the prototype tool in this study is thus to determine whether respondents are able to place visual representations of products on two-dimensional 2x2 grids featuring bipolar axes without the support of research leaders being present during the test. What sort of problems have to be negotiated to improve the performance of the tool for future internet based investigations?

3 METHOD

For the research study presented in this paper, a prototype tool was created in order to test the method of inquiry with international respondents. As the object of assessment, food containers for jam, marmalade, herring and similar food products were chosen. This choice of this product type was made since food jars are common everyday products, which are well known among people in all markets, and do not require any technical, contextual or use-related previous understanding.

The present study involved three specific steps: (1) the elicitation of parameters for assessment, (2) the selection of study objects for the assessment, and (3) assessment using international respondents. The method is described below.

3.1 Elicitation of parameters for assessment

The purpose of the first step of the study was to elicit the parameters to be used on the bipolar assessment scales during the test. The elicitation of parameters was done using a group of master level industrial design students. In a class of 29 students, a series of 13 images of a broad range of packaged food products were shown on a large screen using a computer projector. The images were taken from a

combination of two types of sources: from the selection of 32 food containers purchased for the purpose of this research study; and from a selection of images of packaged food products sourced from the World Wide Web. The food containers purchased represented a typical cross section of containers for jam, herring and marmalade, widely available on the domestic grocery market in Sweden.

The selection of images used for the elicitation of parameters for assessment, represented a broader range of packaged food products than the images used for the test with international respondents. This was done in order to ensure that a sufficiently broad range of descriptive words (parameters) were generated, as the selection of jars purchased for the study was relatively homogenous in terms of type and appearance.

The students were given the task to write down what they came to think of when they were shown the images (associations, thoughts and reflections). In groups of three to four students, words were generated during a period of two minutes for each image. In all, 719 words, predominantly adjectives, were generated during the session. Subsequently, the words were categorised into six categories. The categories were 'natural' in terms of representing 'natural' clusters of meanings of words as interpreted during analysis. Eight categories of words were generated, representing associations to origin, quality of the food product, period of time, type of meal, typical occasions of use in terms of traditions and festivities, style and expression of packaging, general associations of the food product, and health/diet associations. After further reduction, six parameters were chosen as dimensions for assessment. The finally chosen parameters were: Reuse-worthy, Scandinavian, Gift-worthy, Expensive, Authentic and Appetising.

3.2 Selection of study objects

The objective of this step of the study was to reduce the total number of containers from the 32 purchased, to a suitable number of study objects to be rated by respondents. The original group of containers consisted of a range of materials and types, including glass jars, metal cans, and a plastic bottle type of container.

The reduction of containers was done using the same group of students as in the first step. All containers were photographed against a neutral white background and the images arranged, in four rows and eight columns, in a random order as a collage on an A3 size paper (see Figure 2). Instructions on how to rank the containers where attached. Individually, the students were asked to indicate the seven most 'characteristic' (distinctive, eye-catching) containers by circling seven images of their own preference. The students were then asked to rank these containers from 1-7 and, in their own words, to describe the 3 most characteristic features of these containers. Following this, the procedure was repeated for the seven least attractive containers, which were ranked 8-14 and verbally qualified in the same manner. Subsequently, all ratings were summarised and the seven most and seven least characteristic containers, as rated by the student group, were selected as study objects for the study.

3.3 Assessment using international respondents

Based on the findings from the first two steps, a Microsoft PowerPoint-based version of the tool was developed and sent out in a survey format to participating international respondents using e-mail. The PowerPoint format was chosen as it is a widely available application suitable for visual-verbal presentations, accessible to most home computer users.

The test consisted of a total of eleven pages, including test procedure instructions, an explanatory example describing the image positioning procedure, the main study featuring three 2x2 charts with two bipolar axes, a page to be completed with personal details (age, gender, place of birth, place of living and occupation), and submission instructions. The test was distributed to participants as a PowerPoint file attachment in an e-mail send out. In the body of the e-mail message, instructions were provided for downloading a free PowerPoint viewer application for participants who did not have a copy of the full software package installed. Apart from personal details, no qualitative data was requested.

Study participants were recruited from five countries; Sweden (45 participants), Iceland (20), Czech Republic (20), New Zealand (14) and Estonia (10). The Swedish participants were part of the Industrial Design User Group; a group of individuals recruited from three Swedish medium sized cities who have volunteered to occasionally participate in research studies. The New Zealand participants were industrial design students at Massey University in Wellington, while the remaining participants were acquaintances of other foreign students studying at the industrial design programme at Lund University in Sweden.

Figure 2. The original assortment of 32 food containers acquired for the study.

In the main study, three 2x2 charts with bipolar axes were used to allow respondents to assess a total of six parameters, elicited in the first step of the study design. Each axis exhibited one of the six parameters (reuse-worthy, Scandinavian, gift-worthy, expensive, authentic and appetising, respectively), and each chart displayed two of the parameters. Each of the three bipolar charts exhibited all 14 containers, elicited in step two of the study design procedure, to place in relation to each other and the two parameters featured on the bipolar axes. Respondents were instructed to choose, drag and drop each image into a position on the grid which they felt best matched the words indicated on the axes. In all, thus, each respondent had to place a total of 42 images against the six parameters.

In letting respondents moving images of products on an analogue scale the process differs from other scales that have fixed steps for assessment such as Likert scales [15] and semantic differential scales [16]. Furthermore, as in the UCC tool, respondents were given the opportunity to reposition visual representations after initial placement. Displaying the images of all products simultaneously provides the respondent with the opportunity to compare and relate all product representations with each other.

4 RESULTS

In all, 109 tests were emailed to participants. Of these, 35 responses were returned. Due to time limitations, non-responding participants were not reminded or invited to participate in the study a second time. Of the returned tests, five were not valid as they were not completed correctly.

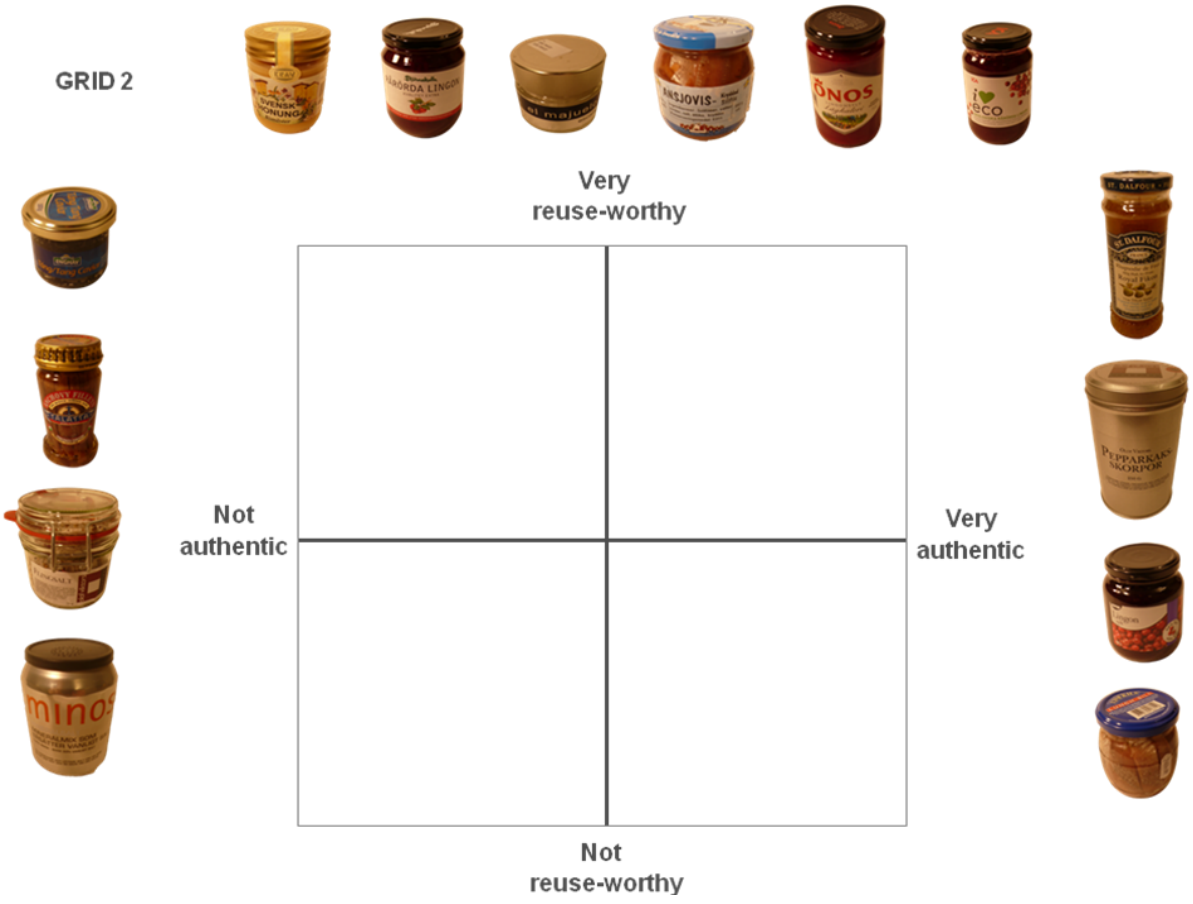


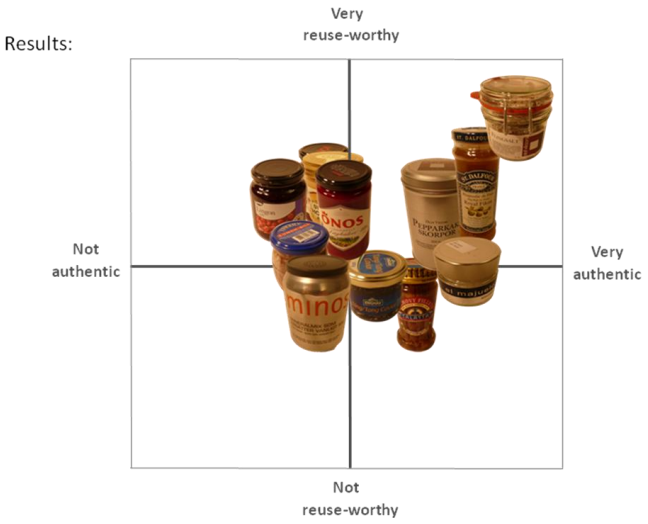
Figure 3. A screen shot of one of the three 2x2 charts of the test, with two bipolar parameters as assessment dimensions ('authentic' and 'reuse-worthy', respectively).

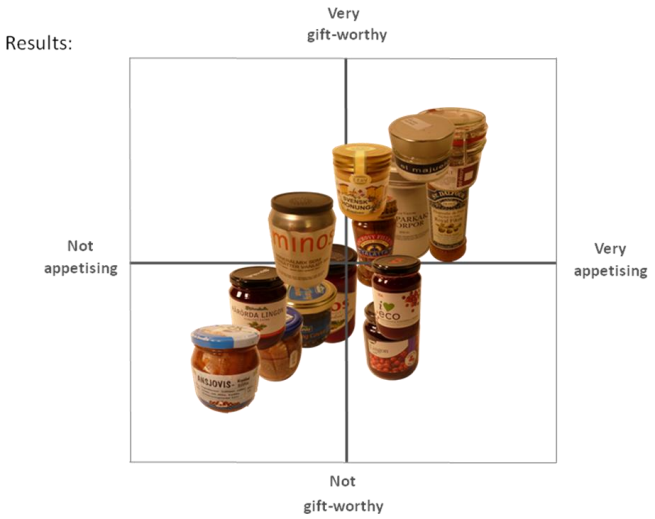
The positions of the images, as placed by the respondents, were manually extracted from print-outs of the returned PowerPoint slides. In compiling the results, the placements of each specific container (one of the fourteen assessed in the study) from all of the respondents were plotted on top of each other. For each container, the visual mean of all placements was plotted and the image representation of each container was placed at the centre of this area of placement markings. This was repeated for each of the 14 images. The visual mean positions of all placements are illustrated in Figure 4 below. These figures can therefore be regarded as a visual account of the total results.

5 DISCUSSION

The objective of this paper was to evaluate a prototype version of a tool intended for remote assessment of user experiences of visual product representations. In recent years, untraditional and exciting visual and game-like product emotion tools have entered the scene [17]. Much is in favour of the idea that people would rather mediate something about themselves while experiencing play rather than just filling in a traditional survey form. Game type tools for product assessment are more likely to motivate respondents to accept and complete more time-consuming tasks than less engaging alternative survey designs.

This study investigated the effectiveness of the physical UCC tool (Sperling and Eriksson [4] and Sperling [5]) transformed into a non-assistive PowerPoint application as a tool for remote assessment rather than as a one-on-one, guided evaluation tool for usability lab settings (a less extensive application has previously been tested and presented for direct communication [13]). This was done with focus on issues arising concerning visual representation, remote interaction and multiethnic





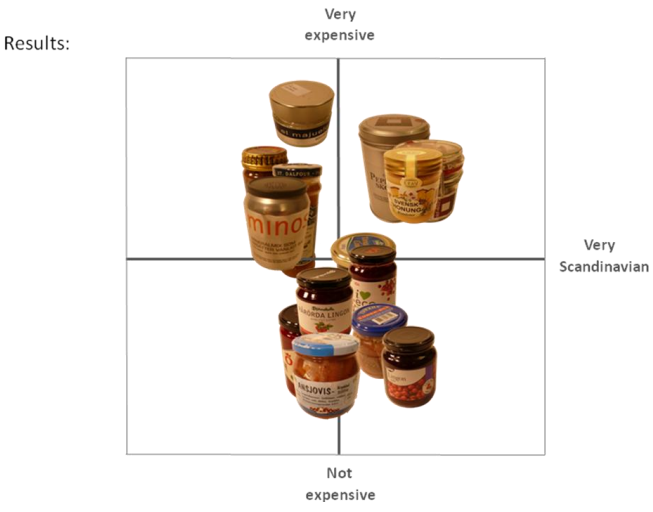


Figure 4. The results from the test visualised in the three 2x2 charts used in the study (with bipolar parameters reuse-worthy and authentic; appetising and gift-worthy; and Scandinavian and expensive, respectively), indicating the visual mean of all image placements.

variety of respondents. Particularly, the focus was on identifying new types of problems associated with remote assessment, that do not affect the physical UCC tool in the usability lab environment. While some implications were expected; important aspects were revealed that will guide the development of future versions of the VIPET-tool.

Studies of the physical UCC tool have indicated that it is favourable to provide respondent's with the opportunity to readjust their placements of visual samples on the game board once their initial positioning is completed. As each sample is placed one by one, respondents often adjust the positioning of samples as they become assessable in relation to each other. As a unique feature of the UCC tool, thus, the repositioning of samples further stimulates respondents to verbally comment and motivate the inter-relatedness of the placements, providing a high level of qualitative content in addition to the quantitatively measurable placements of samples on the UCC board. In this study the output generated was strictly quantitative in nature, as respondents were not asked to motivate their placements. However, as with the UCC board, respondents could adjust the positioning of samples as they become assessable in relation to each other during the placement of images in the test.

Although an internet based study has benefits in terms of scale, for example with respect to number of respondents and their geographic distribution, researchers do not have the same control over and knowledge of the situation as when respondents are physically present in, e.g., a usability laboratory. An inherent restriction in an internet based tool is that it can only provide the respondent with a limited amount of feedback and assistance, and this in a rather un-customised and standardised manner. Thus, a tool for remote assessment can never provide the richness of face to face interaction. This must be considered when designing a future tool relying on remote interaction.

Approximately six months after the survey, a post-study evaluation was performed in the form of an email asking respondents to the test to submit their lasting impressions of using the tool. Numerous participants responded that they enjoyed performing the test, expressed in positive terms such as: *"It was fun and different", "...it was fun to do and rather interesting",* and *"I liked it".* Additional aspects of the tool in relation to feedback from the post-survey evaluation is discussed in the following.

5.1 Remote interaction and interface design

Regarding the test procedure, participants indicated that the procedure was experienced as being too "open-ended" and sometimes confusing to execute. Also, some respondents expressed a concern that the test procedure was time-demanding to learn and understand, having to read all the textual instructions word by word. One respondent replied: *"It took me too long until I got the hang of how it worked, but then it became fun and a bit different, and I mean this in a positive way."* Future development of the tool will have to consider more self instructive and visual guidance, rather than text instructions, which were used to compliment the visual example provided in the test tool.

Some respondents found the interface awkward to use and it was sometimes misunderstood. One respondent commented: *"The test could be improved if it could be done online. Downloading to PowerPoint seemed like barrier".* This is a natural drawback of having to download a file attachment; future development of the tool will consider the possibility of online login access to a platform independent assessment application, in order to eliminate these shortcomings.

The ability to review images through scaling (possible only through a conventional cursor based drag and enlarge feature) to enable a closer inspection of the containers was limited. As noted by another respondent: *"One thing that made it difficult was that the images of the jars were quite big, and this made it difficult to lay them out precisely compared to other jars. Perhaps the jars could be made a little smaller, or have a large image of the jar to look at, and then a smaller one to move."* Future versions of the tool may include scaling and place pointer features to remedy these difficulties.

5.2 Visual representation

As indicated by the five invalid responses, some of the respondents completely misunderstood how to execute the test. A reoccurring error was that respondents, instead of placing the provided images on the 2x2 charts, placed one product in each quadrant and left the rest of the images unconsidered. It seems that the multiple tasks of (1) choosing images one by one, (2) placing them on a two-dimensional grid, (3) considering each of the two parameters on the axes simultaneously during positioning, and (4) repositioning if desired in relation to other images, is too complex a procedure for some users. The importance of guiding the respondent in the task of negotiating only a limited amount of visual information at each step has been confirmed in this study. A step in this direction may be achieved through using bipolar assessment scales of one dimension at a time, rather than 2x2 charts. Additionally, splitting up the task of positioning images in several defined steps, each having a limited degree of freedom, may be a solution to consider in future developments of the tool.

5.3 Multiethnic respondents

Some problems arose that where related to the fact that the interface was used by respondents of different nationalities [18]. Some respondents found the containers represented in the images too similar to assess. In future investigations, the familiarity of users on different markets of products to be assessed must be considered, and more information obtained prior to such studies [19]. How do respondents relate to, and what do they already know about, the products to be assessed? Two respondents elucidate these circumstances in the following: *"The products you had chosen to be placed where awfully similar to each other in appearance and design"*, and *"...in my opinion some of the products were in exactly is the same group, the same level of aesthetics..."*.

Obviously, factors such as product awareness, and cultural and market related aspects with respect to specific product types, must be considered in future studies. Familiarity with some of the food products and food containers might also have caused some bias in this study. A large share of the examples would have been familiar to Swedish participants, but probably not to respondents of other nationalities. This might have influenced, and even facilitated, the positioning of images for some respondents, as they would have had previous knowledge of and associations to the products.

Furthermore, language problems were encountered in this study, which were apparent when complex text instructions were employed. For example, some non-native English speaking respondents misinterpreted the English text instructions, resulting in invalid or unusable responses. Future versions of the tool will most likely benefit from providing respondents with customised versions of the tool featuring instructions in their own language.

In this pilot study 109 tests where sent out and 35 where returned. Five of the returned tests were not correctly executed as the questions were misunderstood by the respondents. The number of completed tests can normally be related to how time consuming or stimulating the respondent considers the task to be. How devoted the respondent is to complete a test depends also on the commitment the respondent has towards the distributor of the test. Such conditions must always be considered in each unique test situation.

5.4 Future development

Analysing the position data manually was very time consuming even though the number of respondents was quite limited. A tool using an integrated database application would be able to store data automatically, more easily present differences between respondent groups and also help identify correlations between background data and specific product placements. A database with all product placements registered as numerical coordinates would also provide an effective starting point for statistical analysis. However, the findings from this study indicate that the interface needs further development in order to make numerical analysis efficient and valid.

This study has contributed to the development of a new form of remote evaluation tool for subjective assessment of experiential product qualities. Specifically, it has investigated the characteristics of a prototype version of the interface of a tool for elicitation of consumer experiences with respect to how participants of a test respond to a proposed design of the tool. Future work will address the problems encountered in this study. The experiences made vouch for that the future VIPET *(Visual Internet*

Product Experience Tool), through further development, is likely to become a useful tool for future web assessment of consumer perceptions of products.

REFERENCES

[1] Jones, T. *New product Development*, 1997 (Butterworth-Heinemann, Oxford).

[2] Pine II, B. J. and Gilmore, James H. Welcome to the experience economy. *Harvard Business Review*, 1998, 76(4), pp. 97-105.

[3] Keitsch, M. Hjort af Ornäs, V. Meaning and Interpretation: An analysis of two theoretical perspectives in product design. In *Proceedings of the 7th International Conference on Design & Emotion*, Hong Kong, Hong Kong Polytechnic University, School of Design, October 3-6, 2008 (Hung Hom, Hong Kong).

[4] Jordan, P. W, Green, W. S. *Pleasure with products – Beyond Usability*, 2002 (Taylor and Francis, London).

[5] Sperling, L. Eriksson, P. Drivers' experiences of interior materials in present & future interiors of cars and trucks, In PD. Bust (Ed.) *Contemporary Ergonomics* 2006 (Taylor & Francis, London).

[6] Sperling, L. Kristav, P. Olander, E. Lekeberg, H. Eriksson, J. Exploring emotions for design of your future chair. In *Proceedings of the International Scientific Conference Design & Emotion*. Chalmers University of Technology, Göteborg, September 2006.

[7] Russell, J.A. A circumplex model of affect. *Journal of Personality and Social Psychology*, 1980, 39(6), pp 1161-1178.

[8] Miller, H. Kälvinen, M. Objects for an enjoyable life: social and design aspects. In *Proceedings of the International Conference on Affective Human Factors Design*, Singapore, June 27-29, 2001 (Asean Academic Press, London).

[9] Fenech, O. Borg, J. A Sensation Model of Product Elicited Emotions, *Proceedings of the Design & Emotion Conference*, Göteborg, 27-29 September 2006.

[10] Fisher, T. Nordli, H. Has that shape got a weight problem? Understanding human/object relationships through product elements: studying shape emotion and discourse using a Kelly's grid, *Proceedings of the Design & Emotion Conference*, Göteborg, 2006.

[11] Skogen, M. Visual White Space and Emotional Exclusivity: A Student Exercise, Proceedings of the Design & Emotion Conference, Göteborg, 27-29 September 2006, pp.7-10.

[12] Kristav, P. Sperling, L. Warell, A. Towards An Internet Based Visual Tool for Communication with Consumers in Early Phases of the Product Development Process. Full paper at the 10th QMOD Conference, First European Conference on Affective *Design and Kansei Engineering*, Lund University & Linköping University, Campus Helsingborg, June 2007.

[13] Sperling L, Rosenblad T. A tool for communication with consumers – an experiment concerning vacuum cleaners. Presentation at a national event of the European ENGAGE project, Linköping University, Chalmers University of Technology and Lund University, at the University Collage of Craft and Design, Stockholm, December 2006.

[14] Yaliang C, Lin-Lin, C. How to Rate 100 Visual Stimuli Efficiently, in International *Journal of Design* 2008, 2(1).

[15] Likert, R. A Technique for the Measurement of Attitudes. Archives of Psychology 1932, 1(140)

(Woodworth,R.S. ed.)

[16] Gould, D J, Kelly, D, Goldstone, L, Gammon, L. Examining the validity of pressure ulcer risk assessment scales: developing and using illustrated patient simulations to collect the data, *Journal of Clinical Nursing*, 2002, 10(5), pp697-706.

[17] Desmet, P.M.A. Designing Emotions. Delft Technical University, 2002.

[18] Galdo del, E. Nielsen, J. International User Interfaces, 1996. John & Sons, New York.

[19] Castillo, C, José. The user-reported critical incident method for remote usability evaluation. Thesis submitted to the Faculty of Virginia Polytechnic Institute for Master of Science in Computer Science, 1997.