



Perceived Quality Evaluation with the Use of Extended Reality

Downloaded from: <https://research.chalmers.se>, 2024-04-26 14:25 UTC

Citation for the original published paper (version of record):

Stylidis, K., Dagman, A., Almius, H. et al (2019). Perceived Quality Evaluation with the Use of Extended Reality. Proceedings of the International Conference on Engineering Design, ICED, 1(1): 1993-2002. <http://dx.doi.org/10.1017/dsi.2019.205>

N.B. When citing this work, cite the original published paper.

PERCEIVED QUALITY EVALUATION WITH THE USE OF EXTENDED REALITY

Stylidis, Kostas; Dagman, Andreas; Almius, Håkan; Gong, Liang; Söderberg, Rikard

Chalmers University of Technology

ABSTRACT

If designers want to communicate quality aspects of the product, there is a need to bring these characteristics into the measurable space of perceived quality (PQ) attributes. To illustrate the solution for designers' dilemma of the "best design choice" in this study we applied the PQ attributes importance ranking (PQAIR) method, with the example of a bread toaster. We choose for evaluation three PQ attributes which can significantly influence visual quality of a product: Gap, Flush and Parallelism. We performed the experiment measuring subjective preferences over the toaster designs of two respondent's groups - "Designers" and "Customers." We used sequentially: (i) web-survey (still images); (ii) desktop system; and (iii) fully immersive head-mounted display system (Virtual Reality).

Consequently, we conducted a post-experiment survey regarding subjective preferences, related to the PQ communication channels that have been implemented during the study. Our results indicate advantages and drawbacks for each PQ communication method that we applied in this experiment and encourage further research in the area of products' perceived quality assessment

Keywords: Perceived quality, Virtual reality, New product development, Robust design, Industrial design

Contact:

Stylidis, Kostas
Chalmers University of Technology
Industrial and Materials Science
Sweden
stylidis@chalmers.se

Cite this article: Stylidis, K., Dagman, A., Almius, H., Gong, L., Söderberg, R. (2019) 'Perceived Quality Evaluation with the Use of Extended Reality', in *Proceedings of the 22nd International Conference on Engineering Design (ICED19)*, Delft, The Netherlands, 5-8 August 2019. DOI:10.1017/dsi.2019.205

1 INTRODUCTION

When speaking about the perceived quality (PQ) of products we have to acknowledge the fact that we are dealing with a complex, multifaceted adaptive system. One of the most challenging tasks in product development today is to achieve “optimal” PQ within the given boundaries, i.e., available technology, manufacturing capabilities, design, and financial limitations. The designers’ ability to control PQ during the product development process is crucial for the market success of a product (Garvin, 1984). At this point, designers have to focus on the product attributes that communicate quality to the customers - PQ attributes. PQ attributes can in turn be defined as characteristics that convey functional and psychosocial benefits of a product to the customer (Steenkamp, 1990). It is apparent that even the relatively “simple” consumer product consolidates a large amount of PQ attributes that can be pivotal to the customer’s “willingness to pay” for the product. If designers want to communicate certain quality aspects of the product there is eventually a need to bring these characteristics into the measurable space of PQ attributes. To illustrate a designers’ dilemma concerning the “best design choice” we applied the PQ attributes importance ranking (PQAIR) method (Stylidis *et al.*, 2018) with the example of a bread toaster (hereafter the “toaster”). We choose for evaluation three PQ attributes with significant influence on the visual quality of a product: *Gap*, *Flush* and *Parallelism*. These three attributes primarily form a spatial relationship (split-line) between the mating parts in an assembled product and create visual cues that allow the customer to detect manufacturing variation. Previous studies have demonstrated that the importance of the split-lines in PQ assessment cannot be underestimated (Wickman and Söderberg, 2007; Hoffenson *et al.*, 2015; Söderberg *et al.*, 2017). While we can evaluate the relative importance of the PQ attributes for a product with the PQAIR method, the communication channels and instruments of PQ attributes presentation to the customer for assessment are still debatable. For that reason, we decided to investigate the question, “Which communication channel or technology can provide a higher level of usability for reliable PQ assessment with the use of the PQAIR method?” A secondary objective was to understand, “Which PQ attributes are the most influential on the PQ of a product’s split lines?” To address these questions, we performed experiments measuring subjective preferences over the toaster designs of two respondent groups - “Designers” and “Customers.” The “Designers” group included individuals with the experience in product design and awareness of the split-lines importance. The “Customers” group consisted of individuals with no or small experience in product development and awareness of the split-lines notion. Thus, we simulated a scenario of designer/customer communication process through product design. We used sequentially: (i) a web-survey, where the visualizations of a toaster were presented in a two-dimensional space (still images); (ii) a desktop system (high resolution monitor), using three-dimensional CAD models of a toaster; and (iii) a fully immersive head-mounted display system (Virtual Reality) demonstrating toaster CAD models in a virtual environment. Applying the PQAIR method, we determined subjective choices for the best and worst toaster design and PQ attributes (*Gap*, *Flush*, *Parallelism*) importance ranking distributions over the chosen PQ area (one of the toaster’s *split-lines*) for each group of respondents. Consequently, we conducted a post-experiment survey and obtained data regarding the subjective preferences of the respondent groups in relation to the PQ communication channels that were implemented during the study. The remainder of this paper is structured as follows: Section 2 introduces the background and previous research; Section 3 describes the methodology implemented in this research and the experimental set up; Section 4 presents results of the experiment; Section 5 discusses findings and suggests further research.

2 BACKGROUND

In this section, we present a brief state-of-the-art literature review, which is essential for understanding the motives and context of our study.

2.1 Perceived quality in product development

The multidimensional nature of PQ is recognized by many. A considerable amount of research, including a variety of perceived quality definitions, has been conducted in the past (Olson and Jacoby, 1972; Gilmore, 1974; Crosby, 1980; Garvin, 1984; Zeithaml, 1988; Steenkamp, 1990; Reeves and Bednar, 1994; Mitra and Golder, 2006; Aaker, 2009). This research has mainly been represented either by customer-centric marketing studies or by the manufacturing approach to quality known as a “conformance to requirements.” Both approaches share a common agreement - they see perceived

quality as the antagonistic entity to “real” or “objective” quality (i.e., not quantifiable, imaginary, subjective). Only recently an erratic shift towards “objectification” of the perceived quality concept has been spotted, with some scholars proposing quantification approaches to perceived quality (Golder *et al.*, 2012; Quattelbaum *et al.*, 2013; Amini *et al.*, 2016). Alas, the majority of the proposed models require further development of practical methods and tools for industrial use in product development. In our research we consistently advocate and develop engineering approaches to PQ (Stylidis *et al.*, 2015). In contrast to a rigid, formal definition of manufacturing quality attributes – engineering tradition regarding PQ is to produce events that make a customer aware of how things are done. In engineering science, the notion of PQ appeared as a part of bigger models; i.e., in the field of Robust Design and particularly in the area of Geometrically Robust Design (Söderberg & Lindkvist, 1999). These research methodologies were among the first to consider PQ from the engineering viewpoint (Wickman & Söderberg, 2007; Wagersten *et al.*, 2011). Recent advances in the area of Robust Design have integrated the engineering PQ approach into industrial applications (Pedersen *et al.*, 2017; Howard *et al.*, 2017).

2.2 Robustness and variation

Variability is unavoidable in the realization of products (Morse *et al.*, 2018). All manufactured products are, more or less, afflicted by manufacturing variation. The manufacturer needs to achieve the geometrical dimensions specified by the designer and the common language between the designer and manufacturer is the use of tolerances. The variation will lead to error in the final product that, in the worst case, can make it hard or impossible to assemble the product, if the variation is significant. The variation could also lead to malfunctioning products or affect the visual and tactile characteristics of the product, i.e. the visual or perceived quality. The term “critical” or “functional” dimensions are used in the case when dimensions are possible for the customer to observe or are essential to the assembly or quality of a product (Bjørke, 1989). Non-critical dimensions are also important because their variation can propagate through the design to affect critical dimensions. The area of the Quality Assurance process has the aim of ensuring that critical dimensions stay within the allowed range of variation. Many tolerance analysis techniques and software applications have been presented that calculate or simulate variation propagation to predict how geometric variation will affect assembly, functionality, and aesthetic perceptions of quality (Löf *et al.*, 2007; Shah *et al.*, 2007). A clear understanding of consumers perception of quality as it relates to variation is needed to determine appropriate levels of permissible variation in critical dimensions. The spatial relationship between the mating parts in an assembled product is a visual cue that allows the customer to detect manufacturing variation. The spatial relationships are defined as split-lines. Two measuring directions, gap and flush, which are expected to be small, parallel and equal (Figure 1) are used in critical measures.

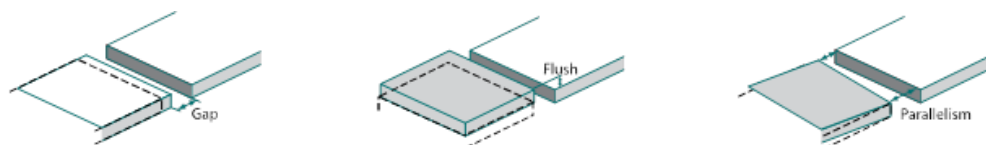


Figure 1. Gap, flush and parallelism.

The area of split lines has been assessed in many different ways using, for example, eye-tracking studies as well as interviews and surveys to understand how split lines influence consumer perceptions of quality (Forslund & Söderberg, 2010). The indication from these assessments has been that misaligned, or improperly spaced split lines, have a negatively influence on consumer perception of products in several cases (Forslund *et al.*, 2013; Hoffenson *et al.*, 2015). In Wickman & Söderberg, (2007) it was investigated the amount of variation in a split-line variation that humans can visibly detect in virtual environments.

2.3 Perceived quality framework (PQF)

The quality perception process is a physical and cognitive event, usually triggered by a physical signal received by our sensory apparatus. The information obtained through the human senses forms the basis of human experience. Thus, it is possible to communicate PQ-related technical elements associated with the customer’s sensorial experience. The vast majority of PQ attributes can be described by one of these sensory categories, or by several in combination. In essence, the Perceived Quality Framework (PQF) reflects human perceptual processing to delineate, test and explore product designs (Stylidis *et al.*, 2015).

The PQ attributes within the framework are organized into categories concerning the primary human senses involved in their assessment; visual, tactile, auditory, olfactory and gustatory (see Figure 2). In our case, quality perception based on the primary senses forms the first level of PQ attributes; *Visual Quality*, *Tactile Quality*, *Auditory Quality*, *Olfactory Quality*, and *Gustatory Quality*.

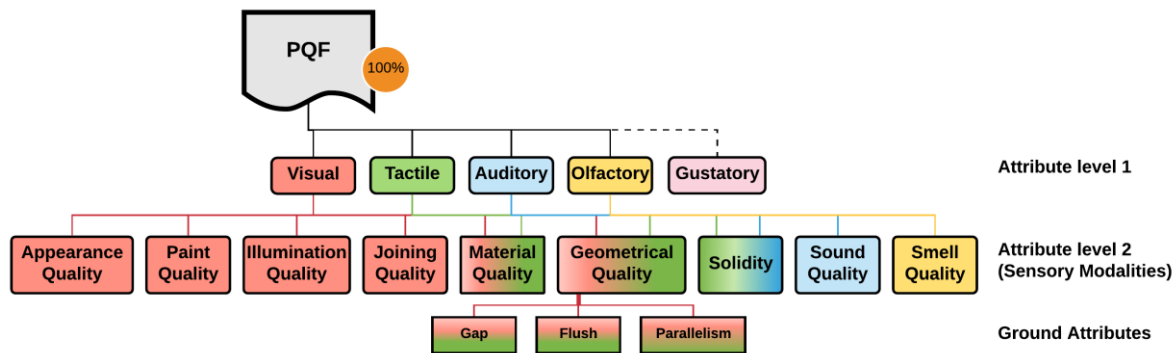


Figure 2. Perceived Quality Framework, adopted from (Stylidis et al., 2015). Sensorial Modalities and Ground Attributes involved in this study.

The second level of PQF attributes is organized into sensory modalities. In our case, sensory modalities are the nine distinctive sets of product attributes encoded for presentation to customers. The baseline of PQ “ground” attributes (GA) is the “lowest point” where designers are still able to communicate technical details to customers and receive meaningful feedback. To avoid ambiguity, every GA has to be coherent to a customer’s experience so that the PQF can stand as a meaningful and accessible frame of reference for both the designer and customer. In our case, only three GA (*Gap*, *Flush* and *Parallelism*) related to the sensorial modality *Geometrical Quality* were involved in the study.

2.4 Perceived quality attributes importance ranking (PQAIR) method

The PQAIR method (Stylidis et al., 2018) was created to assist designers in the decision-making process regarding the evaluation of the PQ attributes relative importance for the final product design. The core of the new method for PQ attributes evaluation is that all identified GA are ranked with regard to their importance (see Figure 3), using either knowledge obtained within the company and/or customer data (e.g., surveys, customer clinics, interviews, internal customer feedback systems, and large datasets).

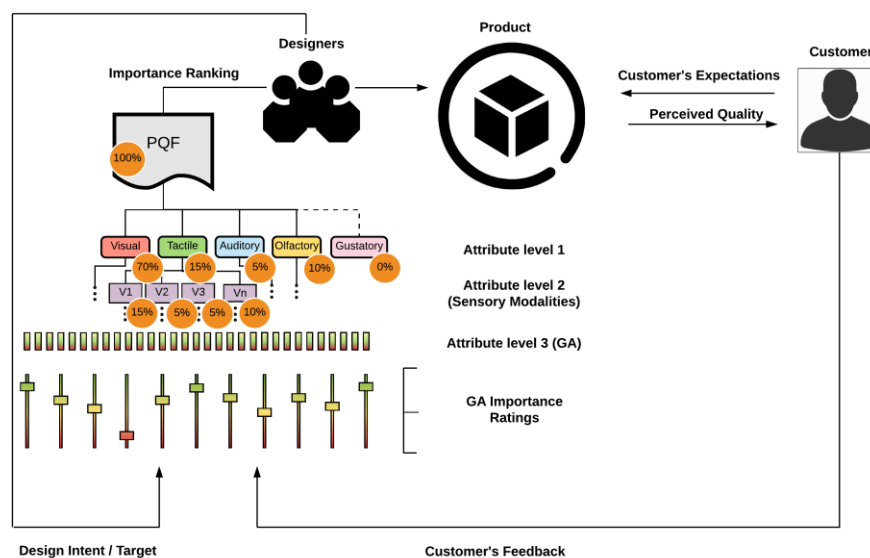


Figure 3. Each GA is mapped onto the universal PQ framework and importance ratings can be calculated per attribute at every level.

These rankings, in combination with the PQF, contribute to the importance score for each branch of the structure at all levels. As a result, designers can obtain an importance score for each PQ attribute,

considering the PQF as a reference model for PQ assessment. In the present study to receive GA importance scores we used the Best-Worst Scaling Method, explained in the subsequent section.

2.5 Best-worst scaling method

To understand the relative importance of GA involved in this study (*Gap*, *Flush* and *Parallelism*), we used a quantitative survey technique called Best-Worst Scaling (BWS). The alternative naming for this method, sometimes used in the literature and in practice, is Maximum-Difference Scaling (or simply MaxDiff). The BWS method was originally developed by [Louviere \(1993\)](#) to understand a respondent's or respondent group's relative valuations of different products or product attributes. According to [Marley and Louviere \(2005\)](#), best-worst tasks also positively affect the consistency of the responses and can be easily understood by respondents. We designed a BWS exercise for three GA (*Gap*, *Flush* and *Parallelism*) of PQF. Consequently, each presented attribute depicted three different conditions: ideal design (i.e., "nominal" - no manufacturing variation), and two clear illustrations of manufacturing variation (see Section 3). This formed experimental BWS design of twelve sets with three items (the PQ attributes) per set.

2.6 Virtual reality technology for product design

The concept of virtual reality (VR) is not new and the first attempt dates back to the 1960s ([Sutherland, 1965](#)). However, thanks to the advancement of both hardware and software, the concept of using computer-generated reality to support daily practice has become ever more mature over the past 10 years. In the area of product design, many previous studies have tried to adopt VR technologies into the various stages of product development. [Tseng et al. \(1997\)](#) proposed a framework that uses VR technology for prototyping and assessing product variants in order to meet the increasing demand on product customization. [Zwingmann et al. \(2002\)](#) reported a system that can virtually assess the product reliability and [Maxfield et al. \(2002\)](#) reported a study that used virtual reality to predict the aesthetic quality of automotive designs, however, some argue that the predictive value of VR assessment should only be limited to the conceptual design phase ([Reuding and Meil, 2004](#)). Nevertheless, in the studies using VR in product design it is generally agreed that the most significant advantages of VR technologies are the realistic visualization of virtual content, nature human-machine interaction and relative ease of examining "what-if" scenarios ([Smith & Heim, 1999](#); [Heragu, 2016](#)). Those unique features will benefit the PQ evaluation process as the realistic model visualization and interaction associated with the VR system will help the agents to make a better judgment on different design variants. Thus, higher quality of PQ evaluation results can be expected to ensure the delivery of customer-satisfaction products.

3 METHODOLOGY

The primary goal of the study was to capture the subjective preferences for the *toaster* designs of two respondent groups "Designers" (N=5) and "Customers" (N=10) and evaluate (i) the usability of each of the communication channels applied in this study (ii) the discrepancy between two groups in PQ attributes ranking. The post-experiment feedback regarding the usability of communication channels was a focus of our interest since the obtained information is considered important to the evaluation of accuracy and appropriateness of PQAIR method. Three different communication channels were picked for the *toaster* designs assessment and these were sequentially used for GA (*Gap*, *Flush* and *Parallelism*) evaluation. Those communication channels were:

- Still images of the *toaster* designs derived from the high-end visualization tool VRED (Autodesk, 2018), presented in the form of a web-based survey.
- Evaluation of the *toaster* designs using VRED-generated 3D models on a desktop system (high- resolution monitor).
- Evaluation of the *toaster* designs using VRED-generated 3D models presented in a VR head-mounted display system (HTC Vive Pro).

As a result, the experiment consisted of four main stages: (i) a web-based BWS study of *toaster* designs; (ii) a desktop system, operator-guided BWS study of *toaster* designs; (iii) a fully immersive VR, operator-guided BWS study of *toaster* designs; (iv) a web-based survey with semantic-differential ratings and open question tasks regarding experience obtained during three previous stages. VRED was chosen based upon being a competitive 3D-visualization and virtual prototyping software used to

create photorealistic renderings. To determine subjective choices for best and worst *toaster* design, we partially applied the PQAIR method. Consequently, we received importance ranking distributions over the chosen PQ area of the *toaster* for each group of respondents with the use of the BWS method. To put it another way, we asked respondents to determine the best/worst *toaster* design with the altered (or nominal) appearance of one specific split-line of the *toaster*. We have chosen three different levels of manufacturing variation in the *toaster* designs. The levels (see Table 1) were decided based upon the experience of the authors in product design and development.

Table 1. Ground attributes manufacturing variation levels.

Ground Attribute Label	Nominal	Min	Max
Gap	X	X + 3 mm	X + 5 mm
Flush	0	X + 3.5 mm	X + 5 mm
Parallelism	0	X + 1.5 mm	X + 2.5 mm

The choice of product for this study was based upon the idea of not being too geometrically complex, as well as being something that many people have seen and used often. As a result of a brainstorming session, it was decided to use a bread toaster as the product for examination. The CAD software Alias Autostudio (Autodesk, 2018) was used to model the product, since Alias Autostudio allows the creation of CAD models with high surface quality. The CAD model was exported into VRED to apply the correct materials to the toaster as well as to add surroundings and light to create as photorealistic renderings of the products as possible. To simulate the geometrical variation with the different toaster designs, VRED was connected to the variation simulation tool RD&T (RD&T Technology, 2018). This enabled manufacturing variation to be simulated, and the result of that variation simulation to be visually presented in VRED. Only one part of the toaster was moved, based upon the variation simulation, to isolate the GAs for evaluation. Figure 4 illustrates examples of the toaster designs with the applied variation.



Figure 4. The white plastic part on the left is afflicted by the following variation: left image - 3 mm Gap, middle image - 3.5 mm Flush and right image - 1.5 mm Parallelism.

A posteriori, we conducted a web-based survey to obtain data regarding feedback from both respondent groups, and their subjective preferences regarding the PQ communication channels (still images, 3D models on a desktop system, 3D models in VR). We were able to assess the level of usability for each of the communication channels, evaluating the rankings of six usability-related attributes: *Learnability*, *Ease of use*, *Usefulness*, *Reliability*, *Holistic appearance*, *Understandability* (see Table 2).

Table 2. A description of the usability attributes evaluated in the study.

Usability Attribute	Description
Learnability	This is the ability to learn how to use the system.
Ease of use	The measure of how easy it is to use the system once it has been learned.
Usefulness	The inclusion of system features necessary for functionality to satisfy the respondent's view on interaction. We look at the degree to which the functionality and features are used by specific groups of respondents.
Reliability	The ability of the system to perform specified functions under the stated conditions for a specified period of time.
Holistic appearance	The holistic perception of a system as a whole regarding interface, feedback and performance.
Understandability	The capability of a system to enable the respondent to understand how it can be used for specific tasks and conditions of use.

3.1 Experimental setup

Each group of participants (both “Designers” and “Customers”) carried out the BWS online study based upon still images of a toaster design as the first step. The VR and desktop system studies were carried out in a laboratory environment. The participants started the exercise in random order, either experimenting with VR or using a desktop system to avoid any learning pattern. Since VRED was used as the visualization tool throughout the whole laboratory experiment, the 3D modeled visual representations of *toaster* designs for VR and the desktop system were the same.

3.1.1 Hardware used in the experiment

The experiment was conducted in a laboratory environment at the premises of Chalmers University. The base stations, used to track the VR-glasses as well as the hand controllers for the HTC Vive Pro, were mounted on the walls to ensure stability in the immersive experience. The computer used for the VR experiments was equipped with a GTX108 graphics card, 32 GB RAM and Core i7 7820HK CPU. The 2D-screen desktop system visualization was carried out with a similar computer and a 1920x1080 Dell 32” screen. Steam was used to control the HTC Vive Pro VR device.

3.1.2 VR experiment

The participants observed a virtual table with three *toasters* positioned next to each other on top of the table (see Figure 5). VR-goggles (HTC Vive Pro) were used as the visualization tool. A participant could freely move around the virtual “table” and explore *toaster* designs. The BWS study was performed with the help of an operator who ran the BWS design exercise created with the help of Sawtooth Software ([Sawtooth, 2018](#)). The operator performed the task of defining manufacturing variation values for each toaster design comparison stages, guided participants, and captured the participant’s subjective choices regarding best and worst *toaster* designs.

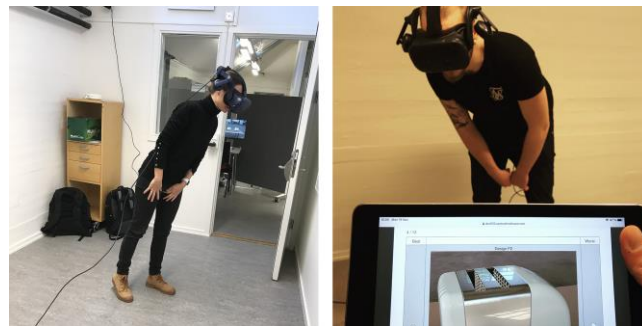


Figure 5. Fully immersive head-mounted display system exercise.

3.1.3 Desktop system experiment

The desktop system experiment set up was the same as for the VR experiment, except for the use of a 2D high-resolution screen and a computer mouse to navigate in the environment. Since not all of the participants had previous experience of operating of CAD systems, the basics regarding the navigation were described at the beginning of the exercise. Thus, the participants were able to train and explore the visualization tool for a short period to familiarize themselves with the virtual environment.

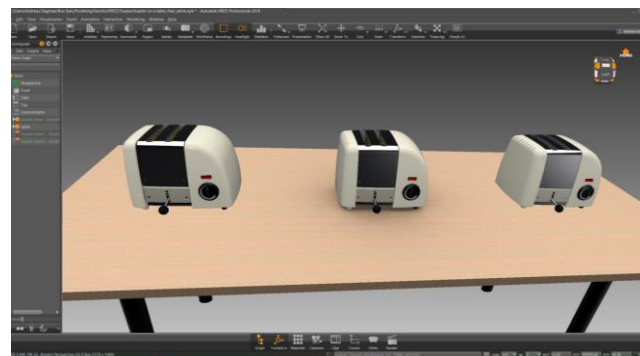


Figure 6. Scene set up for the desktop system with the VRED visualization.

4 RESULTS

The post-experiment survey indicated that the majority of respondents preferred the desktop system or VR as a communication channel for *toaster* design evaluation (see Table 4).

Table 4. A description of usability attributes evaluated in the study, both groups (N=15)

<i>Usability Attribute</i>	<i>Mean Score Still Images</i>	<i>Mean Score 3D Images</i>	<i>Mean Score VR</i>
Learnability	3.21	4.50	4.71
Ease of use	3.43	3.93	4.07
Usefulness	3.57	3.93	3.57
Reliability	3.93	4.07	3.64
Holistic appearance	3.36	3.93	4.07
Understandability	4.07	4.50	4.79

The results of BWS rank-order exercise for three communication channels for the respondent group “Customers” is summarized in Table 5. The GAs with the highest importance score correspond to the “best” *toaster* design and GA with the lowest importance score to the “worst” design. It is apparent that the *toaster* design with the “nominal” parameters for GA was favored by the respondents as the “best” design. The results also indicate that the GA “Gap” was primary determinant for “bad” *toaster* design choice.

Table 5. Ranking results of still images (Web Survey), desktop exercise and VR exercise, best to worst design; team “Customers” (N=10)

<i>Ground Attribute Label</i>	<i>Importance Score Still Images</i>	<i>Ground Attribute Label</i>	<i>Importance Score 3D Images</i>	<i>Ground Attribute Label</i>	<i>Importance Score VR</i>
Gap Nominal	18.98	Parallelism Nominal	21.10	Parallelism Nominal	19.29
Parallelism Nominal	17.91	Gap Nominal	20.44	Gap Nominal	18.53
Flush Nominal	16.44	Flush Nominal	16.64	Flush Nominal	16.85
Flush Max	12.93	Flush Max	11.51	Flush Max	12.95
Flush Min	12.87	Flush Min	10.37	Flush Min	11.33
Parallelism Min	7.80	Parallelism Min	8.12	Parallelism Min	8.80
Parallelism Max	6.42	Parallelism Max	5.45	Parallelism Max	7.23
Gap Max	3.37	Gap Min	3.77	Gap Min	3.74
Gap Min	3.29	Gap Max	2.59	Gap Max	1.27

The respondents from the “Designers” group were generally aware of the split-line notion as well as of design principles. The results of BWS rank-order exercise for the respondent group “Professionals” are summarized in Table 6.

Table 6. Ranking results of still images (Web Survey), desktop exercise and VR exercise, best to worst design; team “Designers” (N=5)

<i>Ground Attribute Label</i>	<i>Importance Score Still Images</i>	<i>Ground Attribute Label</i>	<i>Importance Score 3D Images</i>	<i>Ground Attribute Label</i>	<i>Importance Score VR</i>
Parallelism Min	20.31	Parallelism Nominal	21.41	Flush Nominal	17.76
Parallelism Nominal	18.54	Gap Nominal	19.15	Gap Nominal	16.51
Parallelism Max	15.10	Flush Nominal	16.15	Parallelism Nominal	14.01
Gap Nominal	12.32	Flush Max	13.07	Parallelism Min	13.56
Flush Min	11.11	Flush Min	12.55	Flush Min	12.72
Flush Nominal	10.53	Parallelism Min	8.76	Flush Max	12.36
Flush Max	9.61	Parallelism Max	7.22	Parallelism Max	10.96
Gap Max	1.48	Gap Min	1.04	Gap Min	1.24
Gap Min	1.00	Gap Max	0.65	Gap Max	0.88

5 DISCUSSION AND CONCLUSIONS

Perceived quality and the possibility to assess it is of great importance for companies to stay competitive. Although the area of PQ is considered subjective by many, attempts to make it quantitative have been proposed (e.g., PQAIR method). When assessing products, if no physical examples are available the use of virtual representation is needed.

The objective of the research presented was to try to answer: “Which communication channel or technology can provide a higher level of usability for reliable PQ assessment with the use of PQAIR method?” Based on the results, both the toaster designs evaluation and web-based post-survey, VR and desktop system methods were considered to have the highest usability level for assessing PQ evaluations. The results pointed in the same direction for both respondent groups “Customers” and “Designers”. The drawbacks of the use of still images is that the context of the product is not revealed, i.e., just one viewing angle and exploration of the product is possible. As a designer of a PQ-attribute assessment experiment, one needs to be very careful and accurate to provide the “best” picture for the specific PQ attribute to ensure that the right thing is assessed. The desktop system was considered to have high visual quality and easy to assess the product. It should be stated that the professionals had previous experience in navigating in CAD-software, such as VRED.

VR was considered very easy to learn and to understand since it allowed respondents to move around freely. The drawbacks with VR were the resolution of details and sometimes a perceived blurriness. Gap and Parallelism were easier to distinguish and judge than Flush. This could be improved using different materials and different contrasting background while setting up the scene. VR was also considered more useful to assess holistic experiences and less useful for details with the current setup. To make VR more useful the hardware/software needs to be further refined, aiming at assessing details.

Still images lack the context and meaning of GA, although the participants did achieve similar results regarding the second question: “Which PQ attribute is the most influential for PQ of a product’s split line?” Both respondent groups “Customers” and “Designers” did consider the nominal (without production variation) *toaster* design as the best design regardless of the technology used (with one exception regarding the still images for the designers). The results also indicate that variation in “Gap” direction is perceived as influencing to “bad” design most and “Flush” least. However, “Flush” was also considered hard to identify, which makes the result biased.

Future research suggests some important improvements in this type of study. Firstly, the real product must be included in the study. Secondly, some of the respondents must form a control group that will have a single task, to assess PQ of a real product. Results from this control group must be compared with the other respondent groups evaluating virtual representations of the real product. This set up can allow us to measure the information loss for each of the assessment methods involved in the study.

6 ACKNOWLEDGEMENT

This work was carried out at the Wingquist Laboratory VINN Excellence Centre at the Chalmers University of Technology, in Gothenburg, Sweden, and was supported by the Swedish Governmental Agency for Innovation Systems (VINNOVA). That support is gratefully acknowledged.

REFERENCES

- Aaker, D.A. (2009), *Managing brand equity*, Simon and Schuster.
- Alias Autostudio. [Computer program]. Available at <https://www.autodesk.com/education/free-software/alias-autostudio>. (Accessed 14 December 2018)
- Amini, P., Falk, B., Hoth, N.C. and Schmitt, R.H. (2016), Statistical Analysis of Consumer Perceived Value Deviation. *Procedia CIRP*, Vol. 51, pp. 1–6.
- Bjørke, Ø (1989), *Computer-Aided Tolerancing*, 2nd edition. ASME Press, New York.
- Crosby, P.B. (1980), *Quality is free: The art of making quality certain*, Signet
- Forslund, K. and Söderberg, R. (2010), “Aesthetic consequences of making car exteriors visually robust to geometrical variation”, *J. Des. Res.* Vol. 8 No. 3, pp. 252–271.
- Forslund, K., Karlsson, M. and Söderberg, R. (2013), “Impacts of geometrical manufacturing quality on the visual product experience”, *Int. J. Des.* Vol. 7 No. 1, pp. 69–84
- Garvin, D.A. (1984), “Product quality: an important strategic weapon”, *Business horizons*, Vol. 27 No. 3, pp. 40–43.
- Gilmore, H.L. (1974), “Product conformance cost”, *Quality progress*, Vol. 7 No. 5, pp. 16–19.

- Golder, P.N., Mitra, D. & Moorman, C. (2012), "What is quality? An integrative framework of processes and states", *The Journal of Marketing*, Vol. 76 No. 4, pp. 1–23.
- Heragu, S. S. (2016), *Facility Design*. 4th edn. Boston: CRC Press.
- Hoffenson, S., Dagman, A. and Söderberg, R. (2015), "Visual quality and sustainability considerations in tolerance optimization: A market-based approach", *International Journal of Production Economics*, Vol. 38, pp. 168. <https://doi.org/10.1016/j.ijpe.2015.06.023>
- Howard, T.J. et al. (2017), "The variation management framework (VMF): A unifying graphical representation of robust design", *Quality Engineering*, pp. 1–10.
- Louviere, J.J. (1993), "The best-worst or maximum difference measurement model: Applications to behavioral research in marketing". In *The American Marketing Association's Behavioral Research Conference Phoenix, Arizona*.
- Lööf, J., Hermansson, T. and Söderberg, R. (2007), "An efficient solution to the discrete least-cost tolerance allocation problem with general loss functions". In: Davidson, J.K. (Ed.), *Models for Computer Aided Tolerancing in Design and Manufacturing*. Springer, Dordrecht, pp. 115–124.
- Marley, A.A. & Louviere, J.J. (2005), "Some probabilistic models of best, worst, and best–worst choices", *Journal of Mathematical Psychology*, Vol. 49 No. 6, pp. 464–480.
- Maxfield, J. et al. (2002), "A virtual environment for aesthetic quality assessment of flexible assemblies in the automotive design process", *SAE Technical Papers*. <http://doi.org/0.4271/2002-01-0464>.
- Mitra, D. & Golder, P.N. (2006), "How does objective quality affect perceived quality?" Short-term effects, long-term effects, and asymmetries. *Marketing Science*, Vol. 25 No. 3, pp. 230–247.
- Morse, E., Dantan, J.-Y., Anwer, N., Söderberg, R., Moroni, G., Qureshi, A. J., Jiang, X. and Matheieu, L. (2018), "Tolerancing: managing uncertainty from conceptual design to final product", *CIRP Annals*, Vol. 68, No. 2, pp. 695–717.
- Olson, J.C. & Jacoby, J. (1972), "Cue utilization in the quality perception process". In *SV-proceedings of the third annual conference of the association for consumer research*.
- Pedersen, S.N., Howard, T.J. and Eifler, T. (2017), Perceptual Robust Design.
- Quattelbaum, B., Knispel, J., Falk, B. and Schmitt, R. (2013), "Tolerancing subjective and uncertain customer requirements regarding perceived product quality", *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, Vol. 227 No. 5, pp. 702–708.
- RD&T. [Computer program]. Available at <http://www.rdnt.se/tool.html>. (Accessed 14 December 2018)
- Reeves, C.A. & Bednar, D.A. (1994), "Defining quality: alternatives and implications", *Academy of management Review*, Vol. 19 No. 3, pp. 419–445.
- Reuding, T. and Meil, P. (2004), "Predictive value of assessing vehicle interior design ergonomics in a virtual environment", *Journal of Computing and Information Science in Engineering*, Vol. 4 No. 2, pp.109–113.
- Sawtooth Software [Computer program]. Available at <https://www.sawtoothsoftware.com> (Accessed 14 December 2018)
- Shah, J.J., Ameta, G., Shen, Z. and Davidson, J. (2007), "Navigating the tolerance analysis maze", *Computer Aided Design Applications*, Vol. 4 No. 5, pp. 705–718.
- Smith, R. P. and Heim, J. A. (1999), "Virtual facility layout design: the value of an interactive three-dimensional representation", *International Journal of Production Research*, Vol. 37 No. 17), pp. 3941–3957.
- Söderberg, R. and Lindkvist, L. (1999), "Computer aided assembly robustness evaluation", *Journal of Engineering Design*, Vol. 10 No. 2, pp. 165–181.
- Söderberg, R., Wärmefjord, K., Carlson, J. S. and Lindkvist, L. (2017), "Toward a Digital Twin for real-time geometry assurance in individualized production", *CIRP Annals*, Vol. 66 No. 1, pp. 137–140
- Steenkamp, J.-B.E. (1990), "Conceptual model of the quality perception process", *Journal of Business Research*, Vol. 21 No. 4, pp. 309–333.
- Stylidis, K., Wickman, C. and Söderberg, R. (2015), "Defining Perceived Quality in the Automotive Industry: An Engineering Approach", *Procedia CIRP*, Vol. 36, pp. 165–170.
- Stylidis, K., Wickman, C. and Söderberg, R. (2018), "Perceived Quality Attributes Framework and Ranking Method", *engrXiv*, Vol. 27.
- Sutherland, I. E. (1965), *The ultimate display, Multimedia: From Wagner to virtual reality*, pp. 506–508
- Tseng, M. M., Jiao, J. and Su, C.-J. (1997), "Framework of virtual design for product customization". In *IEEE Symposium on Emerging Technologies & Factory Automation, ETFA*, pp. 7–14.
- Wagersten, O., Forslund, K., Wickman, C. and Söderberg, R. (2011), "A framework for non-nominal visualization and perceived quality evaluation". In *ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (pp. 739–748). American Society of Mechanical Engineers.
- Wickman, C. & Söderberg, R. (2007), "Perception of gap and flush in virtual environments", *Journal of Engineering Design*, Vol. 18 No. 2, pp. 175–193.
- Zeithaml, V.A. (1988), "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence", *The Journal of Marketing*, pp. 2–22.
- Zwingmann, X. et al. (2002), "Product reliability assessment using virtual samples". In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 624–629.