TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

OTTO HUITTINEN

MECHANICS DESIGN METHODS AND CUSTOMER EXPERIENCED VISUAL APPEARANCE IN COLOR AND PART INTERFACE COMBINATIONS

Master of Science Thesis

Examiner: prof. Asko Ellman
Examiner and topic approved by the Faculty Council of the Faculty of Engineering sciences
on 14th January 2015

ABSTRACT

OTTO HUITTINEN: Mechanics design methods and customer experienced visual appearance in color and part interface combinations
Tampere University of Technology
Master of Science Thesis, 51 pages, 23 Appendix pages
April 2015
Master's Degree Programme in Mechanical Engineering
Major: Design of Machines and Systems
Examiner: Professor Asko Ellman

Keywords: Part interface, visual quality, variation, tolerance, Six Sigma, Kansei, Voice of the Customer

Visually problematic part interfaces form when manufacturing and assembling variation cause differentiation from nominal dimensions. Big variation in parts and products lead to poor visual quality and unsatisfying design. Visual appearance of part interfaces can be improved by minimizing interface gap and choosing the right colors and design.

Manufacturing and assembling processes can be improved with Six Sigma. Six Sigma is a statistical and scientific problem solving method that aims to reduce process defects to a minimum. Six Sigma is a flexible method that can be used in improving existing processes or designing new processes or products. Six Sigma in improving manufacturing processes consists of five steps: Define, Measure, Analyze, Improve and Control. Variation in manufacturing and assembling processes can be designed in the product with tolerance analysis. Tolerance analysis methods and tolerance management process were studied for this thesis to find out how they support product design process in designing part interfaces. Benefit is found in 3D tolerance analysis, variation in part interfaces can be graphically presented with it before production is started.

Besides gap size, colors in part interface play an important role in how the parts seem to fit to each other. Effect of color in part interfaces was studied in this thesis with customer survey. There are many different product development methods where customer can be included in decision making. Two methods were studied so customer survey could be designed: Kansei Engineering and Voice of the Customer. Kansei Engineering is a method where subjective feelings of the customer can be translated into product properties. Voice of the Customer method was found out to be powerful in making the customer survey. It includes statistical market-research –like methods and data analysis with Six Sigma tools.

13 customers and 15 mechanical design experts participated in the survey to find out the best and worst color combinations in part interfaces. 18 different color combinations were studied with 3 different gap sizes. Main findings from the survey were that the parts seem to fit best with big contrast between parts. Black and white, and red and black combinations managed to hide the gap between parts best. Gap was most visible and fit seemed to be the worst with light color combinations. Gap is sometimes highlighted with different colors in order to reduce its visibility. Black highlighting of the gap did not seem to improve the fitting between parts in black and white combination as suspected.

TIIVISTELMÄ

OTTO HUITTINEN: Mekaniikkasuunnittelun keinot ja asiakkaan kokema ulkonäkö eriväristen osien liitoskohdissa
Tampereen teknillinen yliopisto
Diplomityö, 51 sivua, 23 liitesivua
Huhtikuu 2015
Konetekniikan diplomi-insinöörin tutkinto-ohjelma
Pääaine: Koneiden ja järjestelmien suunnittelu
Tarkastaja: professori Asko Ellman

Avainsanat: Osien liitos, visuaalinen laatu, toleranssi, Six Sigma, Kansei, Voice of the Customer

Visuaalisesti ongelmallinen osien liitoskohta syntyy kun valmistus- ja kokoonpanoprosessien tarkkuus vaihtelee suunnitelluista nimellismitoista. Suuri vaihtelu osissa ja tuotteissa johtaa huonoon ulkonäköön ja epätyydyttävään muotoiluun. Osien liitoskohtien ulkonäköä on mahdollista parantaa pienentämällä liitoskohdan raon kokoa ja valitsemalla oikeat värit ja muotoilun.

Valmistus- ja kokoonpanoprosesseja voidaan parantaa Six Sigma –menetelmän avulla. Six Sigma on statistiikkaan nojaava ongelmanratkaisumenetelmä, joka tähtää prosessin virheiden minimoimiseen. Se on joustava menetelmä, jota voidaan käyttää niin olemassa olevan prosessin parantamiseen, kuin uuden prosessin tai tuotteen suunnitteluun. Valmistusprosessin laadun parantamisessa Six Sigma koostuu viidestä kohdasta: määrittele, mittaa, analysoi, paranna ja valvo. Vaihtelu valmistus- ja kokoonpanoprosesseissa voidaan suunnitella tuotteeseen toleranssianalyysin avulla. Toleranssianalyysimenetelmää ja toleranssienhallintaprosessia tutkittiin tätä diplomityötä varten, jotta saatiin selville miten ne tukevat osien liitoskohtien suunnittelua tuotekehitysprosessissa. 3D-toleranssianalyysi osoittautui hyödylliseksi, sen avulla vaihtelu liitoskohdissa voidaan visualisoida jo ennen osien valmistuksen aloittamista.

Yhteenliitoksen raon koon lisäksi myös värit vaikuttavat siihen, miten osat näyttävät sopivan toisiinsa. Värien vaikutusta tutkittiin tässä työssä asiakaskyselyn avulla. Monissa tuotekehitysmenetelmissä asiakas voidaan ottaa osaksi päätöksentekoa. Kahta menetelmää tutkittiin asiakaskyselyn tekoa varten: Kansei –menetelmää ja Voice of the Customer –menetelmää. Kansei –menetelmässä asiakkaan subjektiiviset tunteet voidaan kääntää tuotteen ominaisuuksiksi. Voice of the Customer –menetelmää käytettiin tutkimuksen tekoa varten. Se sisältää tilastollisia markkinatutkimuksen omaisia menetelmiä ja tietojen analysointia Six Sigma työkalujen avulla.

13 asiakasta ja 15 mekaniikkasuunnittelijaa vastasivat asiakaskyselyyn, jotta parhaat ja huonoimmat väriyhdistelmät osien liitoskohdissa voitaisiin löytää. 18 eri väriyhdistelmää tutkittiin 3:lla eri liitosraon koolla. Tärkeimmät havainnot asiakaskyselystä olivat, että osat näyttävät sopivan toisiinsa parhaiten, kun osien välillä on suuri kontrasti. Väriyhdistelmät musta ja valkoinen, sekä musta ja punainen näyttivät hävittävän raon parhaiten. Rako näkyi parhaiten ja sopivuus näytti huonoimmalta vaaleilla väriyhdistelmillä. Liitoskohdan näkyvyyttä on koitettu hävittää korostamalla sitä korostusvärillä. Musta korostusväri liitoskohdassa ei kyselyn mukaan näyttänyt parantavan musta-valkoisen väriyhdistelmän sopivuutta, kuten oli oletettu.

PREFACE

This thesis was made for Microsoft Mobile between October 2014 and March 2015. All of the resources and funding for this thesis were provided by Microsoft Mobile in Salo. Biggest thanks of all I would like to address to my manager Onni-Matti Halkola for providing me this position. Without providing the topic and valuable guidance, this thesis obviously would not exist. I would also like to thank dimensional quality specialist Juha Niini, for giving me insight and information about tolerance analysis and Six Sigma, as well as guidance with Minitab software. Visual quality specialist Matti Haapala provided me with assistance for creating the survey and helping me focus on correct issues. Six Sigma specialist Ossi Hämeenoja gave me insight and knowledge that I didn't even know to exist. I could not imagine a better way for doing a master's thesis, thanks to all the resources from the great people at Microsoft in Salo.

The examiner and supervisor for my thesis, Professor Asko Ellman, gave his trust in my work and allowed me to build this thesis according to the needs of me and Microsoft Mobile. Thank you also for the guidance and feedback along the process.

Friends and family were supporting me heavily during all the work, either in participating in the survey or giving me feedback. Without their mental and concrete support for this thesis, it would not be what it is.

Salo, 21.4.2015

Otto Huittinen

CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

3D	3-Dimensional
ANOVA	Analysis of Variance
CAD	Computer Aided Design
CAT	Computer Aided Tolerancing
C&E	Cause and Effect
CTQ	Critical to Quality
CTS	Critical to Satisfaction
DOE	Design of Experiments
FMEA	Failure Mode and Effects Analysis
Gauge R&R	Gauge Repeatability & Reproducibility
GD&T	Geometric Dimensioning & Tolerancing
GLM	General Linear Model
QFD	Quality Function Deployment
SIM	Subscriber Identity Module
SPC	Statistical Process Control
VOC	Voice of the Customer
VSM	Value Stream Mapping

1. INTRODUCTION

Phone makers spend a lot of money in manufacturing processes and material chemistry to get the design just right [6]. Despite what kind of materials, finish or color the design uses, quality needs to be perfect. Recently, there has been news about users complaining that their hair is getting stuck in the gap between cover and window in the Apple iPhone 6 plus [2]. Also Samsung's Galaxy Note 4 has gained bad reputation because as popular phone review site GSMarena [25] points out: "The uneven gap between the metal frame and curved screen glass takes a few points away from an otherwise excellent build." Both of these flaws have happened because of real world variation in manufactured parts. Consumers want joints in products to be seamless [6].

When a part is manufactured it is always different from designed nominal measure. How much there is variation depends on the manufacturing process. Products are designed today in a perfect 3-dimensional (3D) computer aided design (CAD) world. However designed parts are never perfect in the real world. Manufacturing variation is something that even the best design cannot avoid. There are however things that can be done in engineering and design that can make the situation much easier. This thesis aims to find the best options.

A way to improve quality of manufactured parts and products is to apply Six Sigma process improving method. Quality improving methods have existed since the beginning of manufacturing, but when mass production was invented, the need for quality improvement grew greatly [1, p. 10]. In mass production, millions of parts need to be manufactured with similar dimensional properties, because the parts need to be interchangeable. Six Sigma relies on statistical and scientific methods to dramatically reduce defect rates and improve process variables [1, p. 8].

In design phase, the best way to ensure that parts in the assembly fit together, is to use tolerance analysis. By doing careful tolerance analysis, it is possible to find out which are the most crucial dimensions in the product. This way it can be known how much dimensions of the part can vary, so that it is still acceptable in the final product.

Dimensional quality in manufacturing and tolerance design affect in size of the part interface. Besides size of gap, also finish of the parts affect how well the parts seem to fit together. Visual gaps are seen differently with different color combinations. It is not straightforward how different colored interfaces are perceived by different people. It depends on the feeling that the user gets from the product and the part interface. There are several methods that address how customer perspective can be taken into account in product design process. A method which addresses how subjective feelings are converted into product parameters, is called Affective Engineering, or Kansei Engineering [26, p. i]. Voice of the Customer (VOC) is a method that uses marketing research techniques to capture consumers’ requirements [16, p. 2]. Voice of the Customer results can be analyzed with Six Sigma tools.

Kansei and Voice of the Customer methods were studied so a survey could be generated. The survey is presented in this thesis to study how part interfaces are seen with different color combinations. The survey targets both normal consumers and mechanical experts from Microsoft. Mechanical experts are accustomed in evaluating visual quality of parts and products. With survey, it is possible to evaluate if visual quality targets are on same level with customers. The main outcome for the survey is to find out the best and worst color combinations in the interface. The target is not to drive the industrial design so that the colors are chosen based on best visual quality. After color designers have chosen the colors based on marketing research, great benefit is in knowing how hard the color combination is, considering visual appearance of interfaces.

2. IMPROVING PART INTERFACE QUALITY

Quality is a subjective term and each person defines it differently. American Society for Quality [23] describes one meaning for quality: “A product or service free of deficiencies.” This definition relates directly to part and part interface quality in a product. It is fairly easy for a customer to see and feel if there is deficiencies in the parts, or part interfaces of a product. Usually this means uneven or too big gaps between parts or big variations in parts or assemblies. This chapter focuses on where the deficiencies in part interfaces come from and how they can be controlled and minimized with mechanical design methods. The problem of visually challenging part interface can be seen in figure 1, where plastic back cover and aluminium ring form a gap.

Figure 1: Part interface in a Lumia 925

How much deviations in part dimensions affect in perceived quality depends on the products visual robustness to geometric variation. [10] Visual robustness is determined by form, colors and materials used in the part interfaces in the product. Visually robust design can make appearance of the product acceptable despite big variations in parts and interfaces. Part interface in figure 1 is not visually robust, as deviation in interface is easily seen and result is low perceived quality. The problem is that appearance in products is usually designed by industrial designer, whose central concern is not variation or visual robustness. [10]

Visually the perfect part interface in a product would be seamless. However, it is impossible for features in different parts to be exactly the same. Dimensional variation in a product comes from manufacturing and assembling processes. Variation in manufacturing can be improved by improving dimensional quality. One statistical quality improving method that is widely used today, is Six Sigma. Six Sigma relies on statistical methods to improve process quality. The way in mechanical design to ensure that gaps are as small as they can, is to use tolerance analysis. With tolerance analysis, variation of the parts can be controlled and designed in the product.

Color combination between parts affects how visual the gap is. If right colors are used, the design becomes visually more robust, and it is much harder to see the gap. A lot of studies have been made about colors, but how they are seen depend highly on the situation. In chapter four, color combination effect in part interface is studied with customer survey.

2.1 Dimensional variation

Gap sizes are determined by the variation in the parts that make the interface. If the parts were perfect and could be manufactured exactly like mathematically precise 3D CAD models, the gaps could be designed to be zero in size. In the real world, manufactured parts are never perfect. The 3D model only represents nominal dimensional state of the part [7, p. 3]. Most CAD programs can detect collision and interference in mechanisms, but they cannot take manufacturing and assembly variation into account [24, p. 78].

Manufacturing variation can happen in both ways, the feature can be bigger or smaller than nominal measure. When two parts are assembled to each other, the gap needs to be the size of the variation in interface feature of both parts. This way it can be ensured that all of the produced parts fit together, or at least it can be designed how big the yield for the assembly is, if manufacturing doesn’t meet design specifications [7, p. 41]. With mass produced products, the assembly needs to be designed so that all of its parts are replaceable.

Variation in the product comes from manufacturing process, assembly process, and inspection process [7. p. 42]. Every manufacturing process have their limits of accuracy and precision they can achieve [7, p. 43]. Variation can be also made smaller with effort for every process. However, there is a limit that is not reasonable anymore, and other processes need to be considered. It is not always clear how tight tolerances can be achieved with certain process, as it varies heavily depending on the situation.

Tool wear is an influencer in manufacturing process variation. All tools wear because of friction in the process, resulting to in-perfect features [7, p. 43]. Also process operator can make a difference, improper usage of materials, tools and operations can easily lead to big variations. Human errors can be big and automated processes are usually better, with less frequency and effect in errors [7, p. 44].

Used material is not perfect either, which causes variation [7, p. 44]. For example sheet material thickness varies, and material properties are not exactly the same always either. Ambient conditions, for example temperature, humidity and vibration, affect in variation. If the machine is operated outside the specified temperature range, it can affect in the process significantly [7, p. 44]. Difference in process equipment influences in variation. Different types of equipment have different qualities. Also if the manufacturing process changes, it has an effect [7, p. 45]. For example, if a hole is made with machining in one process and die casting in the other, it has an effect in variation. Maintenance is related to tool wear, but can also make a big effect. Poorly maintained machinery lose precision and accuracy over time.

Other source of variation in the product is assembly variation. It is essential that the designer knows how the product is assembled [7, p. 46]. Often manufacturers focus only on minimizing the variation in part manufacturing process. This focus easily leads to unreasonably high costs in manufacturing, if the assembly process is not at the same level. Even if the parts are perfect it is possible to produce low quality products with bad assembly process [24, p. 78]. Fixture types and fastening sequences all affect in assembly variation [7, p. 46]. Assembly shift is probably the biggest contributor to assembly variation. For, example if a part is bolted to other part through holes, the variation in holes and bolt form a small clearance. The clearance allows the assembly to shift. Because of this, also gravity has an effect in assembly. Assembly shift is naturally downwards due to gravity.

Manufactured parts and assembled products need to be measured in order to confirm if they are inside specified limits. However, it is good to remember that inspection processes are not perfect either. If measurements are done with poor equipment or in a hurry, it is easy to end up with unreliable results. To overcome this problem, precision and accuracy of measurement equipment need to be verified before making measurements [7, p. 45]. Also same inspection processes need to be used for same measurements so that comparison between variation can be done.

Dimensional variation of products can be minimized by improving manufacturing processes. One way to systematically improve processes is Six Sigma. Dimensional variation can, and should, be also designed into the product, by using tolerance analysis.

2.2 Six Sigma philosophy

Six Sigma is a quality improvement philosophy with systematic problem-solving methods. It aims to reduce process variation and defect rates radically with statistical and scientific process improving methods [1, p. 8]. Six Sigma is a useful way to improve manufacturing processes to ensure best possible quality, it also includes useful statistical tools for data evaluation that are used later in chapter three and four.

Six Sigma can easily be understood only as a goal for minimal variation in process, as sigma refers to standard deviation, σ. Standard deviation states how data is spread around the mean value. Quality of the process can then be described by how well the specification limits fit around the mean value, μ [14, p. 237]. This can be characterized with sigma levels. Upper and lower specification limits have been found to be in ±3 standard deviations from the mean in normally distributed data in figure 2. This process would be considered a "3σ" process [14, p. 237]. In part design, upper and lower spec limits can be translated to tolerance limits.

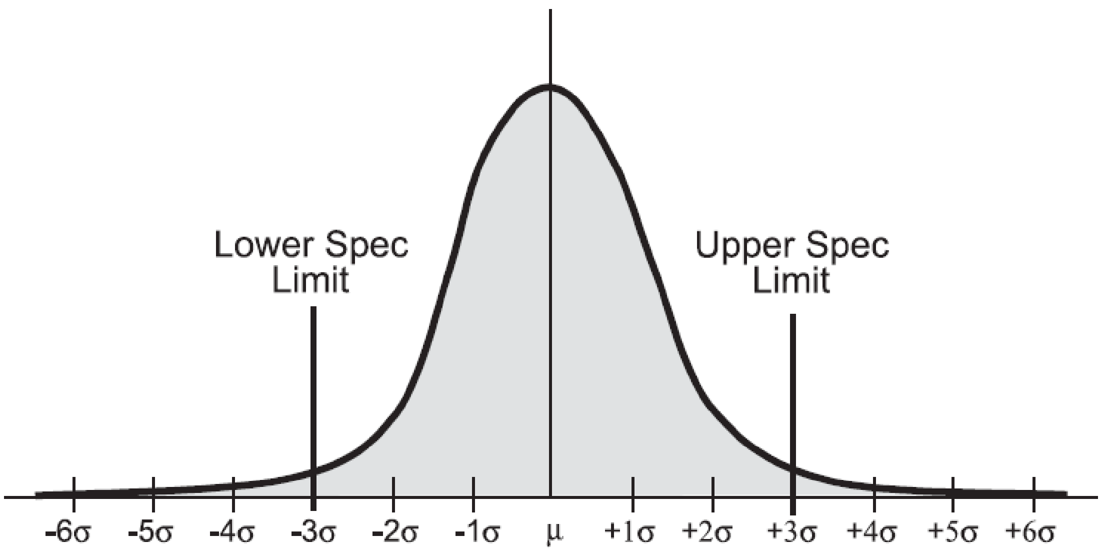


Figure 2: Sigma levels in normal distribution [14, p. 237]

With normally distributed data, 1σ means probability of value being inside 1σ spec limit equivalent of 0,683. In the past, ±3σ was considered acceptable variation, 3σ means probability of 99.73 % of data being inside acceptable limits [14, p. 237]. However, when producing millions of parts 99.73% variation means that 2700 parts in a million are out of specification. Moreover, 1.5σ shift has been observed by companies in long term studies. This means that long term sigma level for 3σ process is 66 803 defect parts in a million. Because of this many defected products, companies have shifted their quality goals to 6σ levels. "Six Sigma quality" in long term translates to only 3.4 defects per million [14, p. 238].

Even though the name Six Sigma suggests a goal of 6σ, less than 3.4 defect parts in a million, it is not often included in the definition of the Six Sigma [1, p. 8]. The emphasis of Six Sigma is in improving quality systematically with statistical methods. Six Sigma project also includes monetary justification, process and quality improvement should be financially reasonable, or it shouldn't be done [1, p. 9]. It should also be doable without statistical experts. The method includes training and certification process for employees, with majority of employees being "green belts" and project leaders being "black belts" [1, p. 9].

Six Sigma in improving existing processes consists of five phases: Define, Measure, Analyze, Improve and Control, shortly DMAIC [19, p. 8]. In define phase the problem and opportunity for the project are characterized: What is important and what can be done? How much money can be saved? Measure phase includes identifying key input variables (KIV's) that have high impact on the project, and assessing what is the current situation. Inputs are then analyzed in the analyze phase, when root causes for current situation are found. Root causes, the most impactful inputs, are then optimized in improvement phase. Finally the process need to be controlled in control phase to ensure that gains are sustained. [19, p. 13]

For new system development, Six Sigma project is a little different, called DMADV: Define, Measure, Analyze, Design and Verify. In this situation a new product, system or solution is designed for the problem in design phase. New design is then evaluated if it is better than original solution in the verify phase. DMADV should be used when existing process is not meeting expectations and there is a demand for a new and better solution. Designing a new system using Six Sigma is also called DFSS: Design for Six Sigma [5].

2.3 Improving process quality with Six Sigma

Improving dimensional quality in manufacturing process is a typical DMAIC process. This chapter introduces a general process for Six Sigma that can be used in improving part dimensional quality. For more detailed example a trial and further study would be needed. The target is to find the optimal key input variable (KIV) settings in the manufacturing process, so that best dimensional quality can be achieved.

First important step after a problematic area, for example a too big part interface, has been found, is to set up a team with all the needed knowledge and skills [21, p. 8]. The commitment form top management for the project is needed, the management needs to be able to see the monetary justification so that all the necessary tools for the project can be had.

The reason for project failure often is "lack of sustained executive sponsorship and commitment." [28, p. 431] This is also the dilemma with Six Sigma projects: the project cannot succeed without management commitment, and management support cannot be had without evidence for successful execution.

The expected profit for the project needs to be clearly stated to management with proper evidence [1, p. 45]. The process of improving the dimensional quality of a part can also be financially justified. Lower variation in parts in the end means that bigger amount of parts can be used in the final product.

Target date for project finalization is needed early in the define phase. Besides organizational and monetary adjustments, the process system needs to be mapped. It includes dividing the project system to subsystems, with input and output variables in each subsystem. For example, part manufacturing can be divided into subsystems for each process, i.e. molding, machining, polishing, painting etc. The end result for define phase is often a project charter. It is a document that clarifies project scope and deliverables, the personnel that are needed, and timing and expected profit for the project [1, p. 47].

In measure phase the current situation for the manufacturing process is evaluated, in order to find out where the variation comes from. The skills and knowledge of the project team is put into test, the process needs to be thoroughly understood. Then key inputs need to be prioritized to find out which inputs have the biggest effect [19, p. 51]. This can be done by doing Failure Mode and Effects Analysis (FMEA). For example, an expert might be able to evaluate if problems in dimensional quality comes from raw material, mold tool problems, or post processing.

Before proper measurements and data collection for the system are made, data collection must be planned and the capability of measurement systems must be evaluated. That way only data that can be trusted is collected. Measurement systems are evaluated using "gauge repeatability & reproducibility" (gauge R&R) methods [1, p. 75]. Then process capability study can be performed and the data can be charted to thoroughly evaluate the data that is collected from different parts of the process.

In analyze phase, the data from measurements is analyzed to find the root causes for the problems in dimensional quality. The objective is to prioritize key system inputs that have biggest effect on key outputs [19, p. 129]. There are several methods for system analysis, including design of experiments (DOE) and statistical process control (SPC) charts and simulation.

Analyze phase includes use of complicated statistical tools and methods. DOE is considered being one of the most challenging methods of Six Sigma. With DOE the system is tested using carefully planned sets of input combinations in random order. For example, a molding machine can be tested using several different control values including temperature, pressure etc. After the tests are performed, the output values are recorded. Interpolation methods are then applied to outputs and the generated model is used to predict output for new possible input combinations [1, p. 123]. DOE is a great method for finding the optimal system input values i.e. when looking for settings to produce dimensionally most accurate parts in a molding machine.

Prioritizing key input variables is a step for applying statistical and graphical tools. They are used for finding the root causes for quality problems so it is known what inputs need to be improved in the next phase. Basic graphical tools are good in identifying relationships between variables. Statistical Process Control (SPC) is needed for identifying if the process is stable or not [19, p. 130]. With hypothesis tests, it is possible to find out if there is differences or relations between variables. Basic graphical tools i.e. boxplots and line plots are in use later in this thesis when analyzing the survey results.

The other important element of analyzing phase is to remove waste from the process [19, p. 178]. It can be done using tools including Value Stream Mapping (VSM), Quality Function Deployment (QFD) and Cause & Effect (C&E) matrices. These methods include brainstorming and analyzing by engineers and customers to streamline and simplify the process and identifying the most critical areas for improvement.

Even though there can be seen two elements in analyzing phase, prioritizing input variables and identifying waste, in Six Sigma project it is often a combined approach [19, p. 129]. The analyzing phase should in the end provide the root causes of defects that are then improved in the next phase. It should also provide prioritized list of key inputs and identified and quantified waste in the process.

DOE is often listed to be implemented in the improve phase of DMAIC project [19, p. 192]. In analyze phase it can be used to clarify input-output relationships, but its most useful use-case is in optimizing outputs. The power of DOE is in testing all factors at the same time, combining and finding the optimal settings for them. It is important to randomize the DOE runs, to avoid noise in the variation. Repetition and replication of the test gives insight of normal variability in the tested system. It should also be considered whether to use narrow or broad studies. Narrow studies are done with small amount of resources and they are used to validate if the results are caused by inputs and not noise. Broad studies include a big amount of resources and time. They are commonly used to validate the results of DOE runs [19, p. 193]. DOE requires a great amount of technical and statistical expertise.

It is important to use the identified waste from analyze phase, and improve the process design [19, p. 207]. There are many techniques that can be applied when removing the waste. These tools help to improve the process flow with for example Kanban, Kaizen and Cell Design and prevent defects with for example Poka Yoke. These are tools that are associated with Lean manufacturing.

Before moving on to control phase, finalized list of key input variables should be available [19, p. 191]. They are the couple inputs that have the biggest effect in the process variables, in this case dimensional quality of the product. Analyze phase should also provide the action plan for improvements, it should give the answer to what are the key input settings that provide the best quality. In the end, the new process design and documentation for it, as well as pilot study for new process, should be done [19, p. 191].

In analyze phase most critical key inputs and process waste were identified. In improve phase the optimal setup for the process inputs was found and identified and waste was removed. The reason for control phase is to maintain the optimal situation, so that during time the process won't slip back to its original state. Control phase requires documenting and detailing the process [19, p. 233]. A final report is needed in the end to provide evidence for success and for validating the business case of the project.

2.4 Designing variation with tolerance analysis

Manufacturing and assembling variation have a big effect in overall low quality, high cycle times and increased cost. Companies that design products containing multiple parts must address this problem by dimensional management and tolerance design. Dimensional management is more than just specifying tolerances to drawings [3, p. 12]. Manufacturing variation is traditionally designed in the product with tolerance analysis.

Tolerance analysis can be divided into two subcategories. It includes methods for determining individual tolerance specifications, as well as so called tolerance stackup, which is the process for determining variation between multiple features [7, p. 47]. Specifying tolerances is not an easy task, and it requires a lot of knowledge in product design to fully understand tolerance specifications in a part. By specifying tolerances, the designer determines how accurately the part must be made to satisfy the needs in the final product. It is possible to use traditional plus/minus tolerances, where the designer allows minimum and maximum dimensions for a feature. Better way to define tolerances is Geometric Dimensioning and Tolerancing (GD&T) [7, p. 15].

GD&T was developed to address the problems in traditional tolerancing. It removes the ambiguity and uncertainty that traditional tolerancing has in tolerancing forms and variation between features [13]. Main difference between GD&T and normal plus/minus tolerancing is that GD&T uses coordinate system that is based on datum reference points of the part [7, p. 131]. All features of the part are tolerated based on the datum reference frame. GD&T is the only way to completely and clearly define allowed variation within features of the parts. Making tolerance specifications is an art of its own, and it does not fully relate to the topic of this thesis. However, tolerance stackup is a tool that can have a big effect in designing part interfaces.

In tolerance stackup, tolerances need to be broken down in order to understand where the variation comes from. Tolerance stackup is an effective way to study for example the variation in a gap between parts. It is a tool for predicting variation and decision making [7, p. 49]. There are several reasons why tolerance stackup should be used: Optimizing tolerances in parts and assemblies, balancing accuracy and precision in manufacturing process to lower costs and determining allowable part tolerances to satisfy final assembly [7, p. 51]. It is also possible to find out if the parts will work in their worst case condition with biggest variation or what the yield will be if they don't work. The result of tolerance stackup can be that the variation is too big in the final product and the design needs to be changed.

Tolerance stackup can be performed in one-, two- or three-dimensional state. Easiest one is one-dimensional analysis, which can be done manually by using pen and paper, others require computer modeling tools because of complexity [7, p. 53]. Basic principle in tolerance stackup is to first find out what is the gap or interface which need to be analyzed. After that, a chain of dimensions and tolerances must be formed through the connecting elements between the parts from interface edge to another.

Tolerance stackup can be further divided to two types: worst case analysis and statistical analysis [7, p. 54]. Like the name says, worst case analysis provides info on how big is the largest possible variation between two features that are studied. Statistical tolerance analysis is more reasonable for a tolerance stackup with multiple parts, with multiple dimensions and tolerances. Statistical tolerance analysis uses mathematical methods for determining the maximum variation in the interface, most commonly root-sum-square method [7, p. 54]. Tolerance stackup can be performed to assemblies as well as single parts, only requirement is that tolerances have been specified to every feature in the tolerance chain.

Tolerance chain represents how the parts are connected together before they form the gap. With shorter tolerance chain, also variation in the gap is smaller. Proper tolerance analysis is needed so smallest possible interface between two parts can be designed. Tolerance analysis and tolerance stackup are tools for specifying and determining if the variation is too big in the designed product.

2.5 Dimensional management

Traditionally tolerances have been applied to drawings late to satisfy drawings based on previous experience [3, p. 12]. The problem with this approach is that the design of the part is already done, and possibility for design changes are limited. Traditionally only 1-dimensional tolerance stackup is done, and it doesn’t necessarily represent the real nature of the part [3, p. 12]. When taking tolerances to design late in the process, it also prevents supplier involvement early. Because of this, it is not possible to take important aspects of product cost, quality, tolerance allocation and assembly process into consideration.

A six step dimensional management process is proposed by Mark Craig [3] where these problems can be beaten. First, dimensional requirements of the product need to be clearly defined [3, p. 13]. Dimensional targets can come from functional requirements or quality improvement goals. Targets need to be clear to all from manufacturing and quality departments to supplier. In the second step, manufacturing and assembly processes are evaluated to see if they meet the product requirements. This is traditionally done with educated guess based on 1-dimensional tolerance analysis. Other option is to build an amount of assemblies and measure them. The best alternative would be to simulate the variation in product using 3D geometry and GD&T.

Third step in the dimensional management process is to ensure that product documentation is correct. Dimensional management documentation includes dimensioning and tolerancing schemes, assembly methods and locating schemes and process control check points. Correct documentation ensures that specifications designed in step two are fully understood in later steps. Step four includes making a measurement plan. Critical features found in simulation in step two must be measured using same references and constrains as in analysis. This is crucial so that comparison of manufacturing and design can be done. From the measurements it can be seen if manufacturing process is capable of producing parts according to design. In step five, measurement plan is put in use and manufacturing capabilities are evaluated. Assembly tools are also evaluated. After this step, it is known if manufacturing and assembling meet design specifications. Final step is to build a feedback loop between production and design. If there are areas in which the product does not meet the design, actual measurement data can be input in the simulation model. After this it can be evaluated if the specification that does not meet the design affect in functions of the product. It can also be evaluated if there are things in design or process that can help to reduce the problem. [3, p. 13-14]

The six step dimensional management process is not easy to implement in product design [3, p. 16]. Organizational structure that supports the process is needed, but the driving force for dimensional management process is a 3D tolerance simulation program.

2.6 3D Tolerance analysis and gap visualization in CAD

Tolerance analysis is often one to suffer because of tight schedules and fast product cycles. Process for tolerance verification is forced to be faster and faster and more flexible [8, p. 7]. Implementing dimensional management process requires 3D tolerance simulation and with CAD being in use everywhere, Computer Aided Tolerancing (CAT) tools are making more sense. CAT is a way to improve the quality of tolerance analysis and potentially speed up the process [8, p. 7]. With 3D tolerance analysis software, it is possible to graphically display the variation in parts and products [7, p. 450], which is especially helpful when designing part interfaces.

The power of 3D tolerance analysis is in complex situations. It is difficult to model long tolerance chains with complex fixtures on a 1-dimensional tolerance stackup. 1-Dimensional tolerance analysis requires often many assumptions and simplifications, and 3D stackup is often much more accurate representation of the variation. In 3D-analysis, it is possible to model translational and rotational effects simultaneously [7, p. 450]. Rotational variation is often hard to analyze with traditional methods. In 3D tolerance analysis, models with dimensions can be brought straight from CAD system, so there is no need for double work. Properly done CAT is linked to 3D model in CAD, so changes in 3D model are reflected in tolerance analysis straight away. Because of these benefits, 3D tolerance analysis can be used for more powerful design optimization than what can be done with linear analysis method [7, p. 450].

There are great benefits of 3D analysis, but also some disadvantages [7, p. 451]. 3D analysis software is more expensive and more complex than linear analysis tools. CAT usually requires 3D CAD models for the analysis, in a simple situation, it might be easier to do linear analysis. 3D CAD models also might be hard to get in some situations. For these reasons it is clever to invest in both, linear analysis and 3D analysis tools.

3D tolerance analysis has a great benefit in being able to visually demonstrate what the variation looks like in 3D model. This way it is possible to include industrial designer in the evaluation of part interfaces, before any real samples are made. In some cases 3D tolerance analysis might be too complicated just for visualizing the variation in a gap. For this kind of situation, it is possible to simulate gap variation using a meshed CAD model [29].

2.7 Effect of color

As stated before, gap size is not the only factor when evaluating how two parts seem to fit each other. Finish, especially colors, of the parts affect as well. Numerous studies have been made about how humans perceive colors. Often there is a tendency to analyze colors independently as separate from anything else. However, the experience of color is heavily influenced by the situation, as many studies have shown. Most of the experiences with colors we observe, is with related colors. This means that in a natural situation, there are many aspects that effect in how color is seen [15, p. 55]. We determine the appearance of color in relation to other lights that are in the same field of view simultaneously [15, p. 51]. Theory of how colors behave in relation to other colors and shapes is a complex one [18]. In part interface, there is often more than one color involved. It is impossible to subtract one color from the combination and confirm that one color is better than other. With different colored backgrounds, colored shape can appear differently. For example, in figure 3 it can be seen that the red square appears larger and more brilliant on the black background [18]. Some colors attract attention away from other objects more easily than others. It is known that color red is this type of color, it is seen to be nearer than other colors [11].

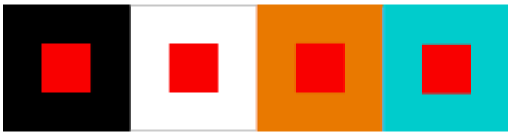


Figure 3: Effect of background color [18]

3. INTERFACE QUALITY FOR A CUSTOMER

Color perception in design is subjective, different colors affect people differently. A lot of thought goes into choosing the right colors and color combinations, involving designers and marketers, so that best colors can be chosen for different market groups. However choosing colors in a product is not only a design and marketing issue, it also affects in mechanical design. When a product includes multiple parts, there are interfaces between parts which are never perfect. The gaps between parts vary because of manufacturing and assembling processes. How these gaps are visualized in the product is affected by the color combinations chosen.

How different color combinations affect different people is not easy to understand. Still it needs to be carefully thought and the result needs to be implemented into the product. Several methods have been developed to understand customer needs and desires in products [16, p. 1]. Kansei Engineering is a product development method, which converts customer's subjective feelings and impressions into concrete design parameters [26, p. i]. Voice of the Customer (VOC) is a process developed to capture consumers' requirements with market research techniques. Kansei and Voice of the Customer methods were studied for this thesis so that best analysis tool could be found for customer survey in chapter 4. Other customer satisfaction analysis methods include Quality Function Deployment (QFD) and Conjoint Analysis [16, p. 1].

3.1 Kansei Engineering method

The word Kansei is a term in Japanese which can roughly be translated to sensitivity or sensibility. The method was developed in Japan and it has mainly been used by Asian companies. The most famous company using Kansei Engineering is perhaps Mazda, they successfully implemented the method in the development of the Miata (MX-5) model. In fact, Mazda Motor Company manager K. Yamamoto was the first one to use the term Kansei Engineering in his speech at Michigan University in 1986 [26, p. 49].

Kansei Engineering is said to consist of three main points: Accurate understanding of customer Kansei, reflecting and translating Kansei into product design and creating a system and organization for Kansei oriented design [26, p. 50].

The process for Kansei Engineering might vary in different research cases, but there are similarities in the procedures and tools used in the evaluation. The first thing in the process is to choose the domain, which consists of selecting the target group and specification of the new product [26, p. 56]. This domain represents the ideal concept, and based on this it is possible to collect samples for further study.

Next in the process is spanning the semantic space. Semantic space includes collection of "Kansei words", which describe the product [26, p. 57]. These can be considered as marketing words for the product. If the product in study would be a tractor, it could be described for example as robust, strong, powerful etc. Number of Kansei words needed depends on the domain, but it is important to collect as many words as possible for a valid result. After collection of words, the most important words must be found.

In the spanning of the space of properties phase, the most important product property features are found. It includes finding the best existing samples that represent different features [26, p. 59]. In the tractor example, these features could be wheel size, color, design of different parts etc. It is also clever to create new concepts so that the process stays innovative, and new solutions can be found.

After the describing words and the products that reflect the descriptions have been found they need to be linked together. This is done in the synthesis phase of the process [26, p. 62]. Often this phase is done with different types of customer surveys. When the best design choices for the domain have been found, it is possible to create new concepts that best reflect the domain. After that the survey can be done again with new concepts, to test the validity of the process so it is possible to find out whether the new concept really reflects the original domain. Kansei Engineering process scheme can be seen in figure 4.

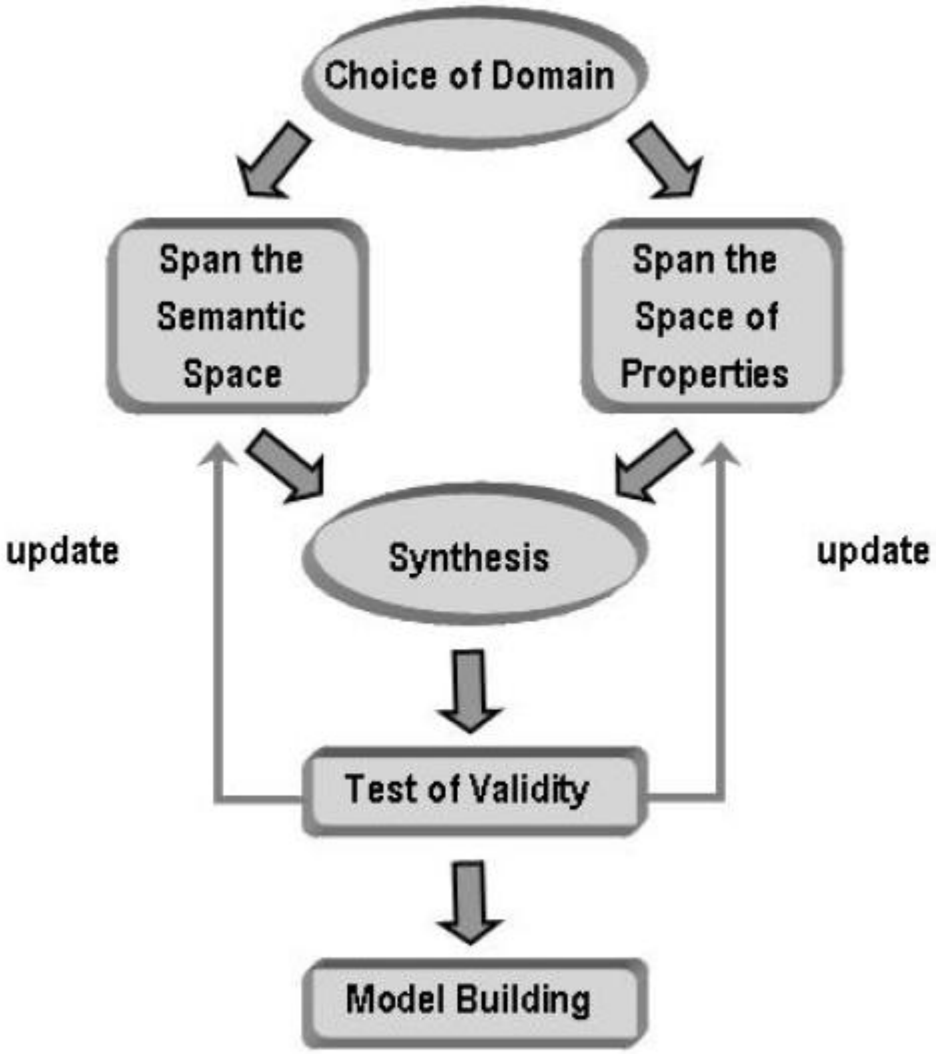


Figure 4: Kansei Engineering process [26, p. 56]

This is shortly the basic process for Kansei Engineering. The extent of the process is mostly dependent on the range of the domain and the process can easily grow very large. However in this thesis the domain is narrowed to a very specific type. The power of Kansei engineering is in designing subjective areas of the product, such as design.

3.2 Voice of the Customer

Voice of the Customer can be described as a consumer-product system. The method builds a dynamic structure between the customer and a company. Voice of the Customer method includes information capturing and analysis for: new product design, customer value study, shifting customer value proposition and product improvements [35, p. 7].

Voice of the Customer data can be divided to two types, qualitative and quantitative, also linguistic and numerical [27, p. 47]. Qualitative VOC includes open ended questions with more subjective information. Quantitative VOC includes more specific and number-oriented questions that are then analyzed with statistical tools. Quantitative VOC uses similar tools as market research, difference is that market research focuses on customer satisfaction, whereas VOC and customer research offers a choice with choice-based questions [27, p. 48]. Quantitative VOC is used in chapter 4 of this thesis to study color combinations in part interfaces.

There are several different techniques to collect the voice of customer, including surveys, interviews, focus groups, panels, brainstorming etc. [27, p. 54]. Best methodology for data collecting depends on the subject. Perfect methodology doesn't exist, all of them have challenges. It is possible to use several techniques to achieve better results. Right data collection technique can be chosen when it is known how much time and money can be used, and what is wanted to know with what accuracy [27, p. 55].

It is easy to think that surveys are just a simple way to collect information, in which you only need to ask and gather answers. However, this kind of attitude leads to unreliable results and even damaging knowledge [9, p. 3]. Besides question forming, target group selection and collection and analyzing of data requires a lot of knowledge. Customer surveys are essential tools for gathering information in VOC [27, p. 55]. They can be done via mail, internet and phone or by comment cards. When performing survey, correct respondent group must have been identified and survey question must be thought out. Question phrasing is extremely important in surveys, survey is designed to produce answers you get [27, p. 55]. Voice of the Customer data is often used as input for Quality Function Deployment to implement detail design specifications in product [16, p. 2]. In the end, Voice of the Customer was the chosen method for color combination survey in chapter 4 and deeper knowledge is needed on the tools that are used with VOC. The power of Kansei engineering would be greater, if it was studied how part interfaces could be used in design.

3.3 Analyzing Voice of the Customer data

Qualitative and quantitative VOC produce different kinds of data, and different analyzing methods are needed. Generally, all VOC data can be divided to three groups: words and notes, attributes and numerical data, and data that can lead to design specifications [35, p. 213]. The data needs to be sufficient and accurate enough to satisfy the needs in product development or competitive position and customer value evaluation.

When analyzing words and notes, the target is to find meanings and patterns from the answers. An affinity diagram can be used. Affinity diagram is a process where one idea is collected from each answer in raw data [35, p. 214-216]. Then ideas are grouped so that similar ideas are in a same group. After that, headers are created for each group and further analysis can be started. After completing the affinity diagram, it is easier to find critical-to-quality (CTQ) and critical-to-satisfaction (CTS) metrics. CTQ and CTS are features that are the most important to product quality and can be transformed to design specifications. Often qualitative VOC data includes answers like: “I want a lightweight mobile phone.” This kind of requirement cannot be a CTQ feature, because it does not include performance standard or specification limit [35, p. 223]. CTQ statement could be for example: I want a mobile phone weighing less than 150 grams.” CTQ statements can later be used as input for QFD analysis.

Data from quantitative VOC is often in form of surveys, benchmarking or interviews. Quantitative VOC data can be classified to two types: attribute and variable data [35, p. 221]. Attribute data can be further divided to categorical and discrete data. Categorical data can be for example color, gender, social class etc. Discrete data can only be integers, numbers without decimal points and fractions. Discrete data can be for example number of defective units, number of scratches in a part etc. Variable, or continuous data can be any real number, it can include for example length, weight, volume or time [35].

Quantitative data can be analyzed with Six Sigma. Basics in Six Sigma were already introduced in lowering manufacturing variation in chapter 2. However, further knowledge in Six Sigma tools is needed for customer survey analysis in chapter 5.

Statistical thinking and use of Six Sigma is based on three principles: “All work occurs in a system of interconnected processes. Variation exists in all processes. Understanding and reducing variation are the keys to success.” [20] Studying Voice of the Customer with a survey can be based on these principles, as well as improving manufacturing processes. Six Sigma tools are used in chapter 4 to understand the variation in survey data. Data is analyzed with basic statistical tools and graphs to help in classification of the data. Also tools like hypothesis testing and correlation are used. Six Sigma data is usually analyzed with statistical software, like Minitab.

Minitab is a program that provides tools for statistical process control, design of experiments, reliability analysis and measurement system analysis. It is often used when performing Six Sigma process improvements, and in fact, Minitab was designed especially for Six Sigma professionals [17].

Descriptive statistics are simple numerical and graphical methods for displaying basic properties in data. Some of the graphical descriptive statistic methods are dot plots, histograms and box plots. Box plot is a useful way to describe data, its median value and variation. In Minitab, upper and lower whisker represents upper and lower 25% of the distribution of data [31]. This is shown in figure 5 where B is upper and D is lower whisker.

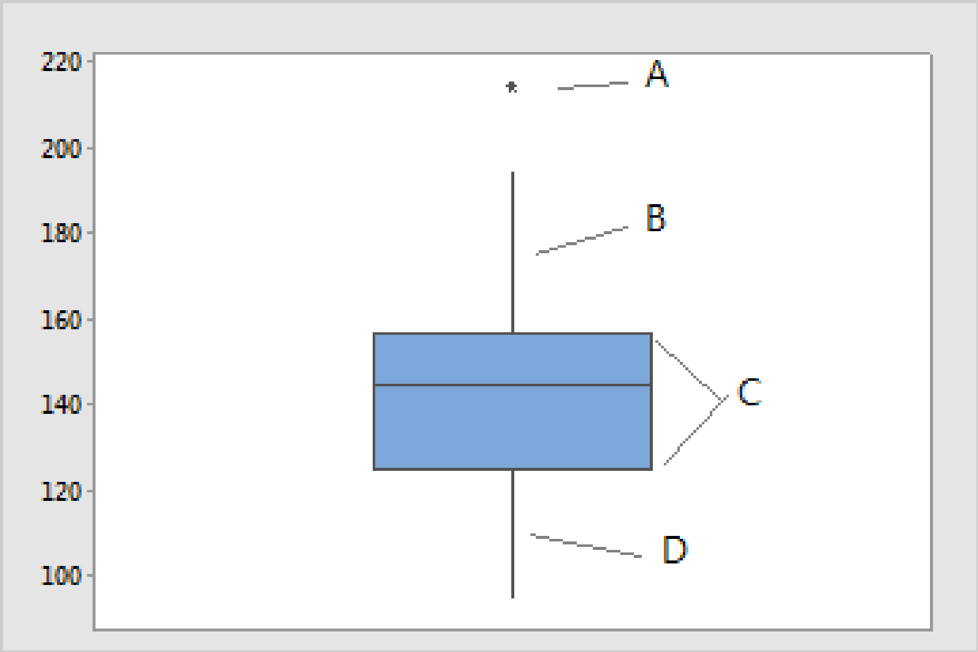


Figure 5: Box plot in Minitab [31]

Letter C in figure 5 represents middle 50% of data and the line in the middle is median. Minitab shows outliers with asterisk symbol, as seen in figure 5 in letter A. These are data points that do not fit in range with other data [31]. In some forms of box plots outliers are not in use and whiskers show the minimum and maximum values of all data.

Interval plot is another way to easily compare groups. It is a graphical summary that at default shows the mean value and 95% confidence interval bar [34]. Confidence interval indicates how likely it is that the sample taken from unknown population contains the value. Often 95% confidence interval is used and for mean value, it states that with 95% confidence the mean value is between the bars. For example, in figure 6 for group number 1 mean is between 10 and 20 with 95% confidence [32].

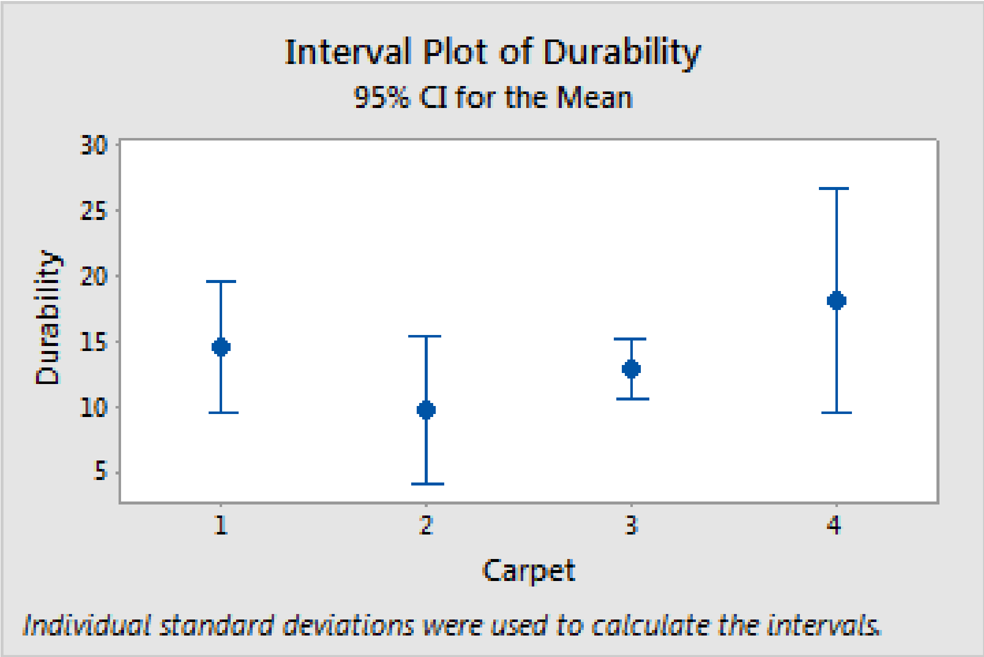


Figure 6: Interval plot in Minitab [34]

In figure 6, the means seem to be different. However with interval plot, difference in means is only significant when the bars do not overlap [34]. Box plot and interval plot are graphical methods for comparing and displaying data properties. Minitab includes also many numerical methods that can be used in analysis. One of them is Item Analysis, which can be used in analyzing reliability of the survey. Item Analysis can be used in evaluating how well multiple questions measure one characteristic. One of the items in Item Analysis is Cronbach's alpha. It is a single factor that tells how well set of items measure one characteristic. Generally, a limit of 0.7 for Cronbach's alpha is used for reliable surveys [12]. In chapter 4, Item Analysis is used for evaluating correlation of answerers to reveal unreliable answerers.

Hypothesis testing is a tool for decision making. It includes several different tests, including mean tests, median tests, variance tests and relationship tests. The hypothesis that is tested in this statistic technique is called null hypothesis. If null hypothesis proves to be false in the test, alternative hypothesis is true. Null hypothesis is accepted or rejected based on probability value called p-value, which is acquired from the test. With Minitab, several different hypothesis tests can be performed. Commonly used hypothesis tests include 1 Sample-t, Paired-t, 2 Sample-t, ANOVA, Chi-square, Correlation and Regression. [19, p. 145]

One type of Analysis of Variance (ANOVA) procedure is General Linear Model (GLM) [33]. General Linear Model is used in determining if the means of groups differ from each other. Several different comparisons can be performed with GLM to find significant differences between means of groups. Null hypothesis in multiple comparisons is that there is no difference between means. Many methods for multiple comparisons have been developed. For survey analysis in appendix B, Tukey and Bonferroni methods from Minitab are used in comparing groups. Tukey method is used in survey evaluation in chapter 4. Power of Tukey is in doing pairwise comparisons, whereas Bonferroni's inequality method is a conservative procedure with larger confidence intervals [33].

4. STUDY WITH DIFFERENT COLORS

The reason why Kansei Engineering and Voice of the Customer were studied in this thesis is to find out how different color combinations affect to visual robustness in part interfaces. Unlike in the normal Kansei Engineering process, it is also important to find out which are the worst color combinations. It is not possible to choose the domain to represent the ideal solution. The Kansei that is needed to find out in this review process could be described as: How different color combinations affect in visual appearance of part interfaces? However, the question is very limited which affects greatly to the review process, and benefit for using Kansei Engineering would be lost in the process. The real benefit from Kansei engineering would be shown, if the survey studied part interfaces as a part of product design. Because of these factors, the information was gathered and analyzed with Voice of the Customer methods in the form of survey.

Designing a survey is a complicated task. It is important to remove all external factors that can affect in the answers when the survey is done with chosen samples [30]. Also the survey needs to be possible to execute in a reasonable amount of time. Number of color combinations and part interface types are unlimited, so the samples that are used need to be carefully considered.

4.1 Designing the survey

Part interfaces that can be affected by choosing the right color combinations can be found in many places. Primary goal in this thesis was to choose multiple different part interfaces for review, but when combined with multiple colors, the survey matrix would grow too large. This is why only one type of interface was chosen for color combination review. The assumption is, that the affected visual quality would not change dramatically in different interface types with same colors, so the conclusion of this review process could be used also in different situations.

In mobile phones, one type of problematic visual gap between parts can be found in between the cover and the SIM-card (Subscriber Identity Module) door. In many mobile phone designs, the cover is not removable, and there still needs to be a way to insert the SIM-card into the phone. Often this is done with a small lid. This is also repeatedly a problematic area in visual quality, because the gap needs to be reasonably sized to allow movement between the parts. Because of this and easy availability of parts, it was decided to study the interface between SIM-card door and a phone cover with different colors.

The target was to use real parts that were manufactured with correct methods. As the number of combinations grew, it became apparent that there was not enough phone covers available. To get as many colors as possible reviewed at the same time, a SIM-door test block was designed for the survey. The test blocks were manufactured with Stereolithographic 3D printing, which uses photo-reactive resin. It is a cost efficient and dimensionally accurate way to manufacture this kind of small sample batch [22]. The blocks were painted with typical spray paints after printing. There were small variations in sample block colors, but it was not seen critical. With test blocks, it was possible to remove all external factors in the survey and also physical gap sizes could be controlled with decent accuracy. Three gap sizes were chosen to represent normal situations, this way it was possible to see where the color combination effect is the most radical.

SIM-door colors in the survey were chosen based on actual colors used on Lumia 920 phone model: black, red, yellow, white, blue and grey. The order of SIM-doors was mixed in sample blocks to minimize the effect of place in the block. Base colors for the test blocks were chosen with the expertise of visual quality specialist to represent the most common and visually hardest color combinations: white, black and white with black background. The reason why only white and black base colors were chosen is that multiple mixed colors usually does not appear in mobile phones. For example, yellow and blue color combination is very rarely seen in design. In the real world, SIM-door color and base color are usually same. In this survey, there is two cases where door and base are same colored, white plus white, and black plus black. Background of the gap is sometimes made black in order to reduce the visibility of gap between parts. White base color with black background was chosen for survey so it could be seen whether background color makes any difference. Sample block with different colored SIM-card doors can be seen in figure 7, where the red arrow highlights the gap between parts.

Figure 7: Part interface between a SIM-Door and a white sample block

Finally the outcome was six SIM-door colors and three background colors. These combinations were all reviewed with three gap sizes: 0.15mm, 0.25mm and 0.35mm. These are gap sizes that are starting to look problematic in real products. Often 0.1mm gap can be achieved. When designing the survey, it seemed reasonable to choose bigger gap sizes so difference between colors could be seen.

In the end, total combination of samples in the review process was 54. One verification block was also designed, so that the answerer needed to answer one situation twice, with same color combination and same gap size. This way reliability of the answerer could be evaluated. Interface samples for the survey can be seen in figures 8, 9 and 10.

Figure 8: Samples with 0.35mm gap

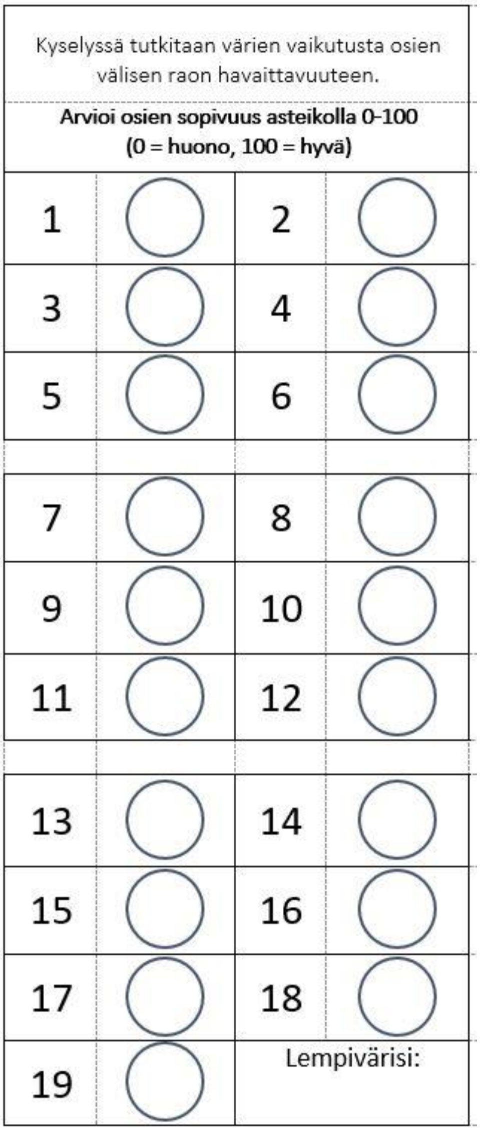
Figure 9: Samples with 0.25mm gap

Figure 10: Samples with 0.15mm gap

As seen from the figures above, sample size is pretty large. With enough time, the survey should have been planned with Design of Experiments (DOE) methods [30]. With DOE, the sample size could have been reduced and more reliable results could have been found. Making DOE experiments can be extremely time-consuming, at least with no previous experience [4, p. 7].

The target of this thesis was to find out the customer experienced visual appearance of part interfaces with different colors. Designing a survey that answers this question reliably was not an easy task. The survey was planned to be answered on paper in a face-to-face type of situation. The samples needed to be presented physically, so there was no need for computer made reply sheet.

After the samples were designed, the next very important task was to create a question that would be in line with original target, and that would also be easy to understand and answer. The question needed to get the answerer focused on the right issue, but should not lead to any specific answer. This proved to be hard, because not many consumers pay attention to these kind of details in products. Finally the question phrase was concluded to be: “Evaluate the fit between parts on a scale from 0 to 100 (0 = bad, 100 = good)” A scale of 0 to 100 was chosen, because there is so many different samples and smaller scale could be too narrow. The question was presented in Finnish, because the target group consisted only on Finnish speaking persons, and it would be easier for them to understand what the survey is looking for. Also a presentation for the survey was made so that the answerer could focus on the right issue: “This survey studies how colors affect in how the interface is visualized.” Survey questionnaire can be seen in figure 11.



Kyselyssä tutkitaan värien vaikutusta osien välisen raon havaittavuuteen.

Arvioi osien sopivuus asteikolla 0-100 (0 = huono, 100 = hyvä)

1		2	
3		4	
5		6	
7		8	
9		10	
11		12	
13		14	
15		16	
17		18	
19		Lempivärisi:	

Figure 11: Answer sheet

Two target groups were decided for the survey: normal customers and mechanics experts from Microsoft Mobile. When surveying both consumers and professionals, it could be seen if there is any difference in how visual appearance of parts is evaluated. Mechanics experts from Microsoft are used to evaluating gaps in products, but it is not clear how different color combinations affect in their evaluation. Customers are not used to evaluating defects and it is not clear how they will see this kind of subject. Still, customer satisfaction is the most important thing in a product and it doesn't actually matter what mechanical experts think about visual appearance of the parts. If there is big difference between answers, experts might need to shift their targets.

The target population from customer side is huge, all of the people in the world that use a smartphone. All of the answerers were Finnish, so it is not seen whether there is differences between people from different cultures. Still, most customers do not have any preconceptions for this kind of subject, and opinions are formed just after seeing the samples.

The samples are graded just by comparing them to each other. After consulting a Six Sigma specialist, it was decided that about 17 answers for the survey is needed on both sides to achieve some kind of reliability [30].

4.2 Conducting the survey

From the beginning of survey creation, it was clear that making up an easy-to-understand question for the survey wasn't easy. Although the question seemed easy and descriptive at first, extra clarifying was needed by many answerers. This is not optimal, because when interfering with the survey answerer, it is easy to lead her to a specific answer. Because of this, spoken guidance was kept minimal and similar to all answerers. Despite question clarification, there was misunderstandings whether to base the grading on how visible the gap is, or how well the parts fit together. This was reported by a couple of answerers. This misunderstanding was significant, because these interpretations lead to totally opposite scores. Favorite color was asked also in the survey. There was seen misunderstandings whether to pick a color from the survey or general favorite color. Many answerers skipped the question entirely. It was not absolutely clear, if favorite color was important information in the survey, so there was not any actions made to improve yield for the question.

Answering to one gap size, total of 19 samples, took approximately five to ten minutes. Because of this, it was impossible to ask multiple gap sizes from different answerers, and decision to ask all three gap sizes from all answerers was dropped already in the beginning. Answering to multiple gap sizes was also seen boring and losing of focus was seen from customers. Multiple gap sizes were evaluated by six customers, who had more time for answering. Most answerers answered only to samples from one gap size. Gap size for each answerer was picked randomly, so that amount of answers to gap sizes would be similar.

The change in survey plan didn't affect the original target. The target of this thesis was to find out how color combinations affects how the interface is seen, not how gap size affects how the interface is seen. Also it is expected that the answer would be similar with different gap sizes and different gap sizes work as a verification to each other.

Feedback was gathered from mechanics experts at Microsoft, who are familiar with the subject and accustomed in analyzing visual quality of products. However, they are not used to evaluating how color affects in the interface. This target group consisted of Finnish men and women from about 30 to 50 years old. The data was collected from mechanics experts during busy working hours. It was seen that some answerers didn't have much time to focus on the subject, which can affect in the data. Answers were collected from all mechanics experts from Microsoft in Salo that were easily available and ready to answer. 15 answerers from mechanics experts participated in the survey, from which 3 were women and 12 were men.

Other target group consisted of normal consumers with different backgrounds, all of them being Finnish and very familiar with smartphone use. Consumer target group consisted of broader age range, consisting of men and women aged from about 15 years to 60 years. 7 answerers were women and 6 were men. For customers this kind of survey topic was very unfamiliar and this caused confusion and lose of focus. Answerers from customer side were found from friends and family that were willing to participate in the survey. Total of 13 different customers participated in the survey, from which one customer answered to all three gap sizes, and five customers answered to 0.35mm and 0.15mm gap sample blocks. Collected survey results from mechanics experts and customers can be seen in Appendix A.

5. SURVEY RESULTS AND ANALYSIS

After the results were collected, they were analyzed with Minitab software, using tools introduced in chapter 3. Whole Minitab report can be found in Appendix B, the most important information from the report is presented in this chapter thoroughly.

First, reliability of the survey was needed to be evaluated. There was misunderstanding with the survey question and with reliability evaluation it could be analyzed if all answerers understood the question correctly. Reliability analysis can be done by analyzing how similar the scores are. The analysis was done in Minitab by using Item Analysis. One factor that is used in evaluating the correlation is Cronbach's alpha. The value was initially low, 0.7323, when comparing the answers from mechanics experts. The values can be seen in table 1 below.

Table 1: Minitab Item Analysis on mechanics experts

Cronbach's Alpha = 0,7323

Omitted Item Statistics

Omitted Variable	Adj. Total Mean	Adj. Total StDev	Item-Adj. Total Corr	Squared Multiple Corr	Cronbach's Alpha
Answerer1	945,0	128,7	0,6366	0,9022	0,6907
Answerer2	958,1	128,0	0,5389	0,9390	0,6955
Answerer3	970,9	117,9	0,6685	0,9782	0,6672
Answerer4	960,9	134,0	0,2534	0,9725	0,7268
Answerer5	950,3	142,1	-0,1609	0,9777	0,7947
Answerer6	935,3	149,9	-0,6938	0,9266	0,7869
Answerer7	933,8	135,1	0,2992	0,9279	0,7224
Answerer8	952,1	121,0	0,6070	0,9726	0,6792
Answerer9	929,4	125,4	0,8650	0,9799	0,6693
Answerer10	943,8	130,4	0,7196	0,9824	0,6929
Answerer11	934,4	138,4	0,1270	0,9023	0,7342
Answerer12	965,6	126,2	0,4646	0,8081	0,7020
Answerer13	960,8	135,3	0,4322	0,9854	0,7169
Answerer14	945,8	139,3	0,1258	0,5720	0,7336
Answerer15	965,0	130,2	0,4706	0,9398	0,7045

In Omitted Item Statistics it can be seen what the values would be if answerer was removed from the study. It can be seen that Cronbach's alpha would be higher if answerer 5 and 6 were removed from the study. Total correlation is also negative for these answerers, which means that there is negative correlation with these two answerers compared to others.

After going through the scores found in appendix A, it was noticed that these answerers had graded the samples in similar order but totally opposite from the others. It became apparent, that the question was misunderstood by these answerers, because when misinterpreting the survey question it leads to this kind of situation. As stated in chapter 4.2, a couple of answerers reported that the question and clarification were misleading. As it seemed obvious that question was misunderstood by answerer 5 and 6, the values were calculated after removing these two answerers from the study. Results can be found in table 2.

Table 2: Minitab Item Analysis on mechanics experts corrected

Cronbach's Alpha = 0,8416

Omitted Item Statistics

Omitted Variable	Adj. Total Mean	Adj. Total StDev	Item-Adj. Total Corr	Squared Multiple Corr	Cronbach's Alpha
Answerer1	794,7	138,2	0,7101	0,8628	0,8172
Answerer2	807,8	140,2	0,4597	0,8652	0,8330
Answerer3	820,6	130,5	0,5975	0,9763	0,8261
Answerer4	810,6	146,2	0,1783	0,9535	0,8517
Answerer7	783,5	144,3	0,3948	0,9270	0,8366
Answerer8	801,8	128,7	0,7293	0,9721	0,8104
Answerer9	779,1	135,8	0,8834	0,9547	0,8067
Answerer10	793,5	140,9	0,7288	0,9788	0,8211
Answerer11	784,1	149,6	0,0856	0,8307	0,8497
Answerer12	815,3	134,9	0,5477	0,7594	0,8278
Answerer13	810,5	144,1	0,6066	0,9669	0,8297
Answerer14	795,5	149,7	0,1550	0,5048	0,8458
Answerer15	814,7	140,2	0,5105	0,9255	0,8293

From table 2 it can be seen that Cronbach’s alpha is much higher after removing answerer 5 and 6 from the study. Data analysis was done including answerer 5 and 6 and also without them to see how big the difference is. There was not seen big differences in the end, which also tells about reliability of the study. As the assumption was that question was understood incorrectly, it was reasonable to remove them from the study for further examination. In table 3 Item Analysis is done to customer scores.

Table 3: Minitab Item Analysis on customers

Cronbach's Alpha = 0,8922

Omitted Item Statistics

Omitted Variable	Adj. Total Mean	Adj. Total StDev	Item-Adj. Total Corr	Squared Multiple Corr	Cronbach's Alpha
Answerer1	652,4	199,6	0,5519	0,6785	0,8868
Answerer2	663,8	189,1	0,7904	0,9483	0,8739
Answerer3	658,5	197,8	0,4084	0,9303	0,8932
Answerer4	667,9	181,6	0,8805	0,9365	0,8667
Answerer5	675,7	200,9	0,4320	0,9529	0,8908
Answerer6	688,6	194,5	0,6625	0,9135	0,8811
Answerer7	690,7	201,3	0,6612	0,8843	0,8862
Answerer8	692,4	192,2	0,7055	0,8974	0,8786
Answerer9	674,3	193,9	0,4905	0,8671	0,8903
Answerer10	700,4	182,5	0,7862	0,9463	0,8731
Answerer11	683,2	196,8	0,3798	0,8601	0,8965
Answerer12	687,4	192,6	0,6462	0,7960	0,8813
Answerer13	660,0	202,8	0,4491	0,8264	0,8907

In table 3 it can be seen that Cronbach's alpha value is reasonably high, and negative correlation is not seen. Based on this we can say that customers have understood the question similarly.

Validity of the answerers was studied by asking the same color combination with same gap size twice. This was grey door on black base color. A line plot was formed for each answerer to see how their answer changed with same combination. Results from mechanics experts can be seen in figure 12. Here each color represents different answerer. Black-Grey and Black-Grey1 are essentially same color combination with same gap size. Steeper angle in the line means more radical change in score to same combination from specific customer.

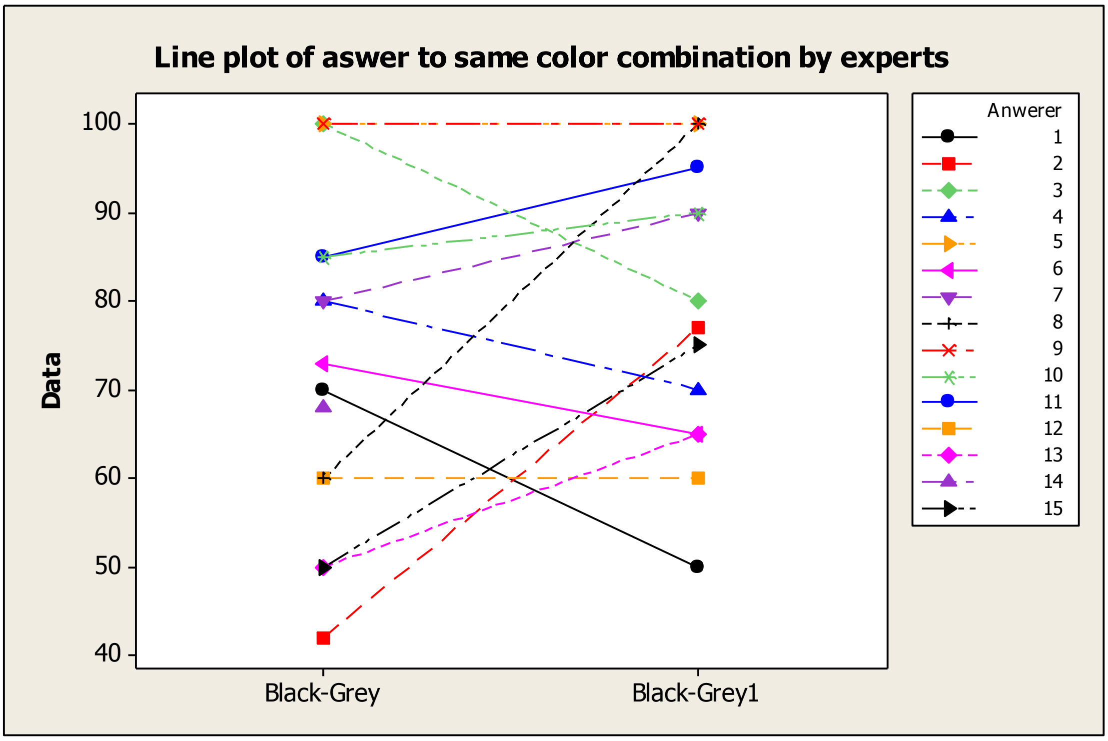


Figure 12: Mechanics expert answer validity study

As seen in figure 12, scores have changed a lot with some answerers. One factor that can explain this kind of phenomenon is focus. Many answerers from mechanics experts didn't have much time to focus on the survey. Survey results were collected during working hours, and many answerers were focused on their own work. Overall the mean value for both black and grey combinations is still similar. From figure 13 the result from customers can be seen.

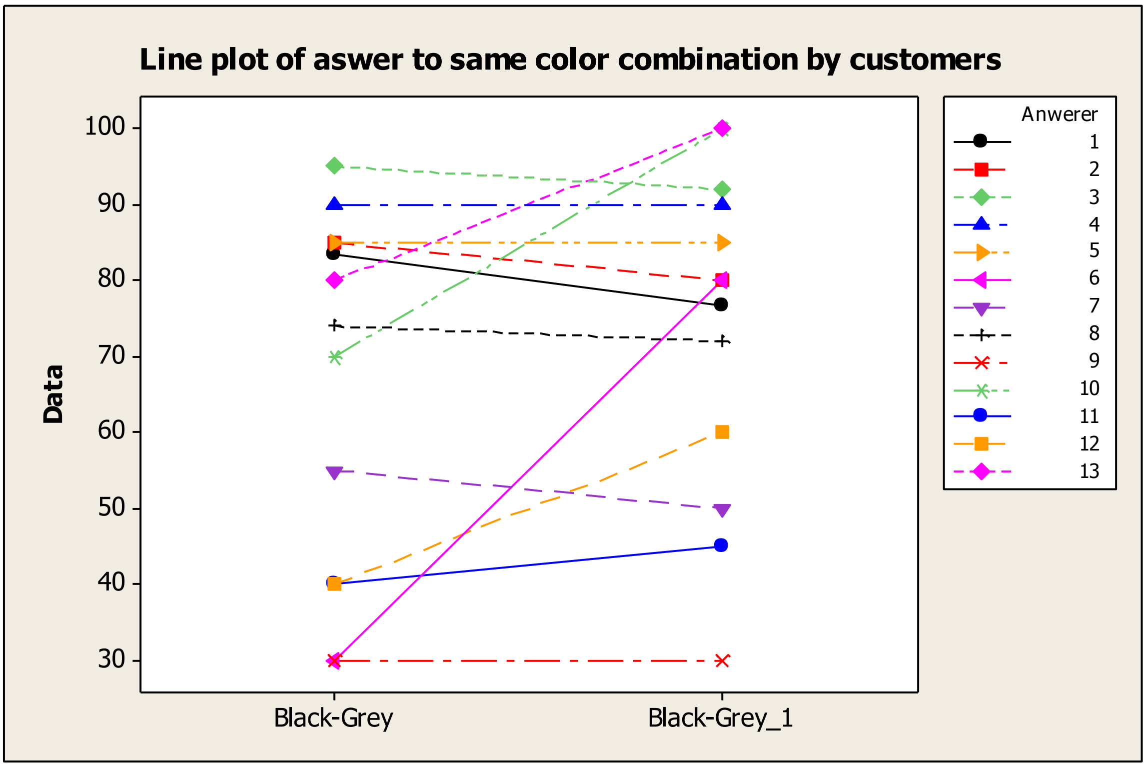


Figure 13: Customer answer validity study

Customers have graded same interface exceptionally similarly, as seen in figure 13. For 9 customer answerers, the grade has changed 5 points or less. Customers had more time to focus and answer on the survey, which can explain the situation. Still, there is especially big change in score from answerer 6. This answerer didn't come up in survey reliability study in the beginning so otherwise the scores from answerer 6 have not changed radically from others. It is clear that the question was not misunderstood by answerer 6.

After reliability and validity of the survey was evaluated, real analysis of survey data could begin. As stated in chapter 4.2, thorough gap size evaluation was dropped in the beginning of data collection. However, for the interest of seeing if there is difference between different answerers scoring different gap sizes, an analysis was done. It must be noted, that sample size is extremely low for this study, only 4 answerers graded 0.15mm sized gaps and 5 answerers graded 0.35mm gaps. Still, it might be seen if answerers have generally scored higher scores on smaller gap sizes. In figure 14, mechanics experts' scores has been divided to different gap sizes.

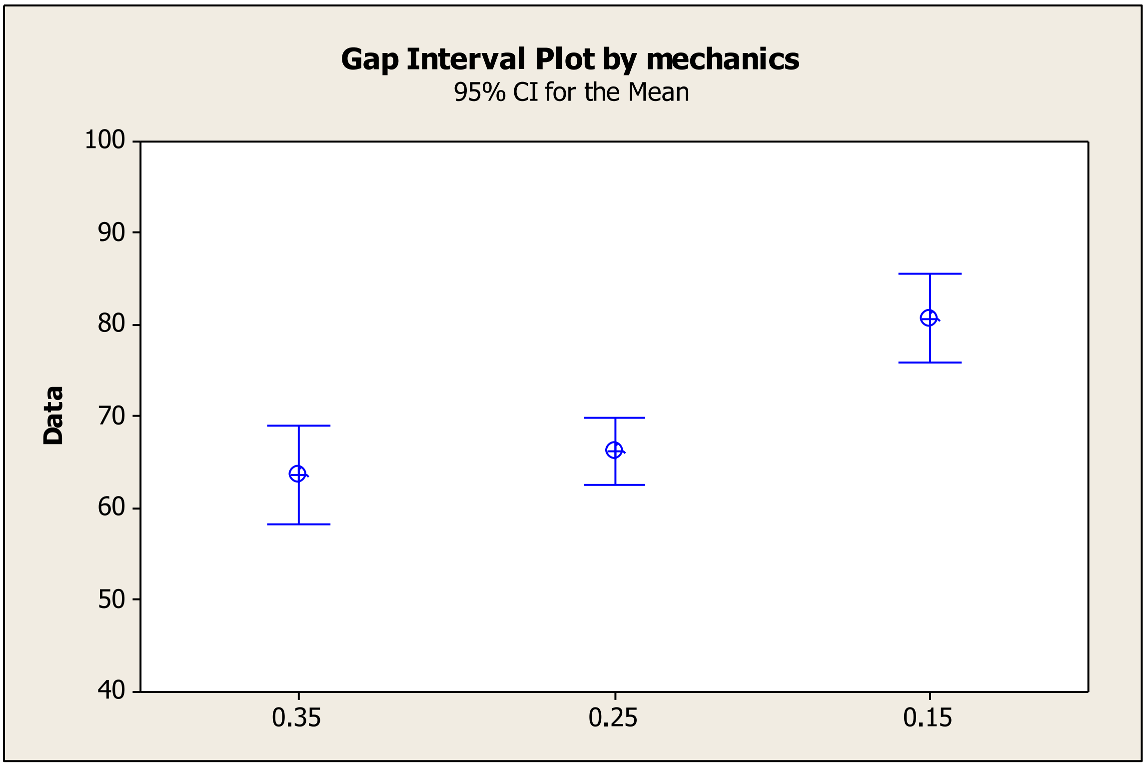


Figure 14: Gap analysis on mechanics experts

From the figure above, we can see that smallest 0.15mm gap size has gotten clearly the best scores. The data is reasonable also otherwise, as smaller gap sizes have gotten better scores. Meaningful difference doesn’t exist between 0.35mm and 0.25mm samples, because confidence intervals for both includes same values. In figure 15 we can see the same data from customers.

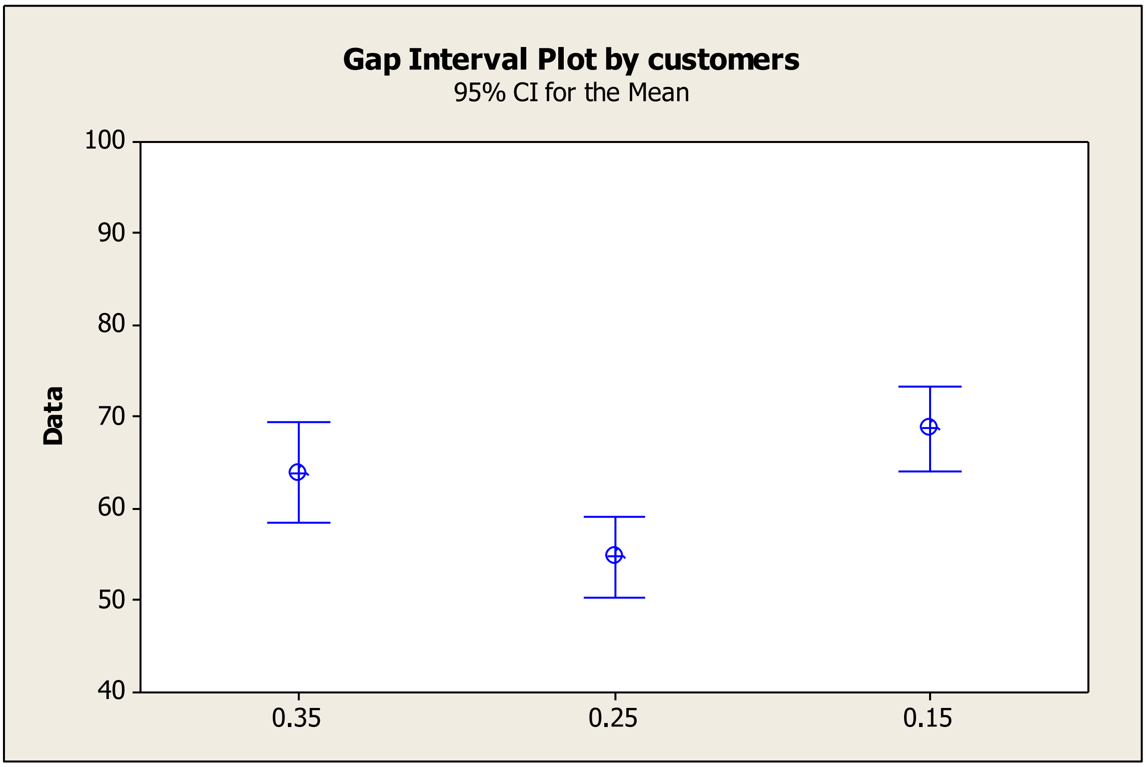


Figure 15: Gap analysis on customers

Gap size scores have been divided randomly in customers' scores. Precise evaluation cannot be made from gap sizes based on this survey data. Mechanics experts are used to analyzing gaps and they are able to tell what kind of gap is good based on previous experience. Customers are not used to evaluating visual qualities of part interfaces. Figures 14 and 15 support this statement.

Even further, two gap sizes, 0.15mm and 0.35mm were evaluated by 6 same customers. There was about 5 minutes break between analyzing gap sizes, they couldn't see the samples at the same time. All except one answerer told that they cannot tell any difference between the samples. This can also be seen in figure 16, as there is not much favor towards smaller, 0.15mm gap size.

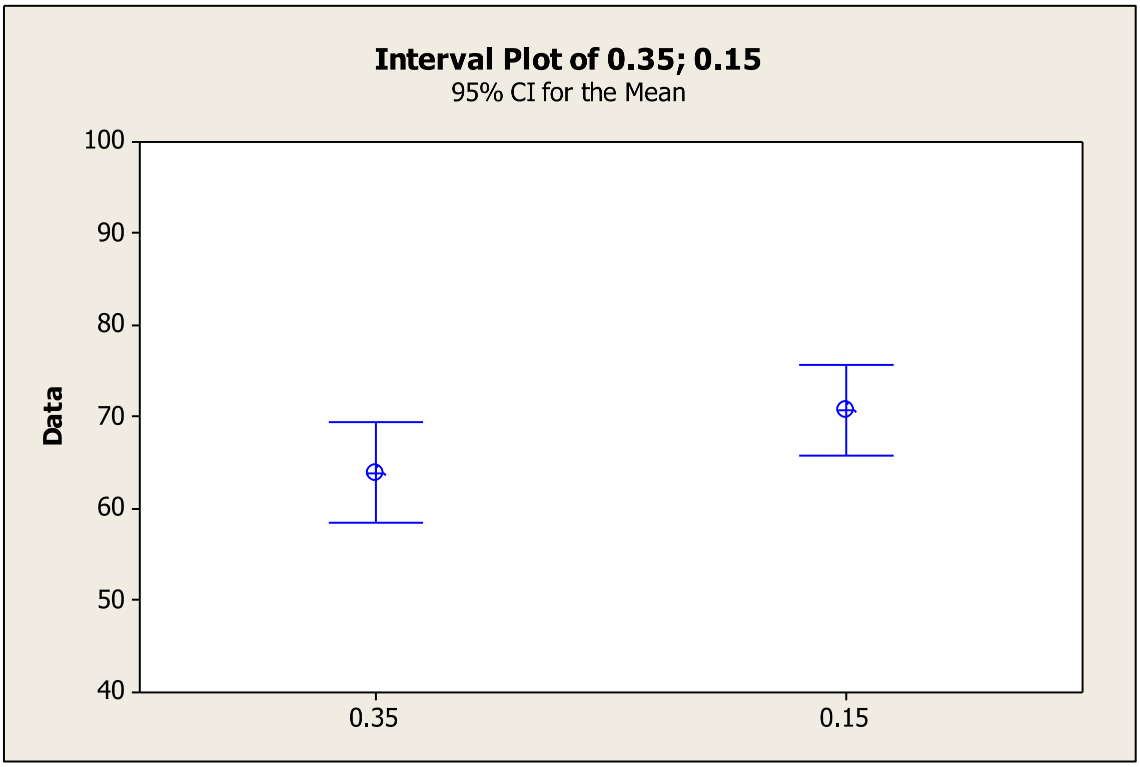


Figure 16: Difference in gap size scores by 6 customers

Base color effect was also studied with different gap sizes. Results from mechanics experts can be seen in figure 17. Here scores are displayed for every base color with different sized gaps.

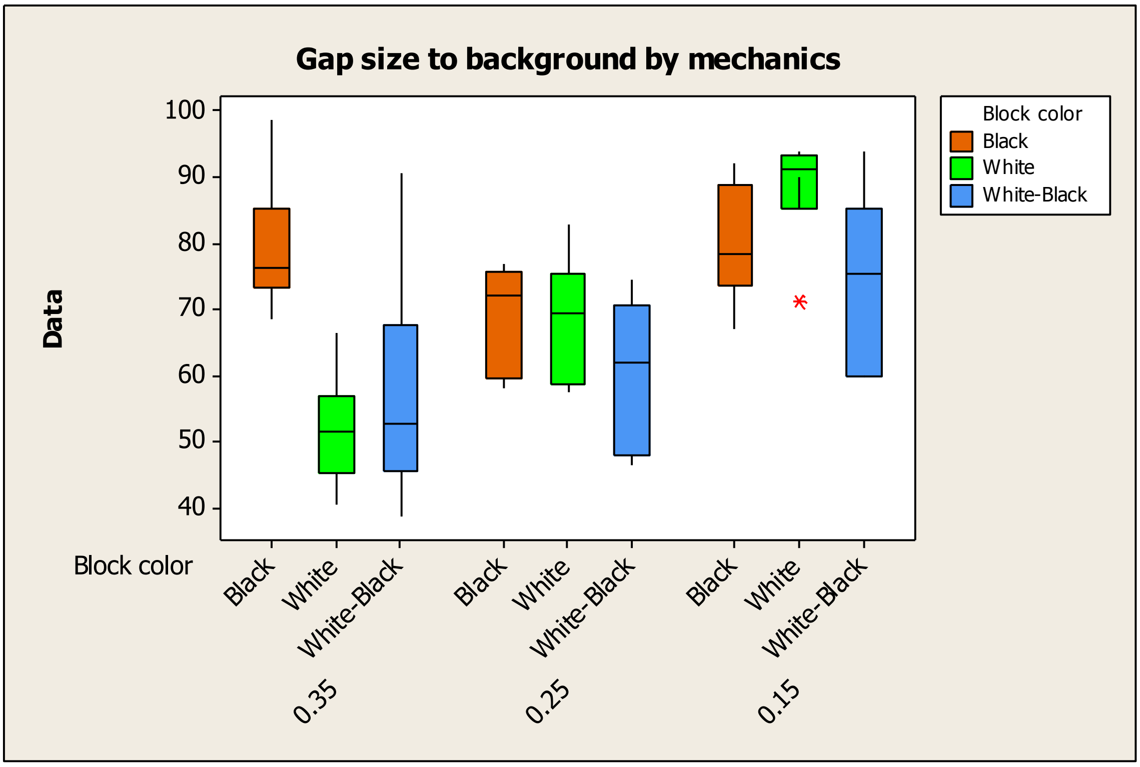


Figure 17: Gap size effect in base color by experts

From figure 17 we can see that with bigger gap size, black base color is clearly the best. Same kind of phenomenon can be seen from customers, as can be seen in figure 18. Black color seems to hide bigger gaps well.

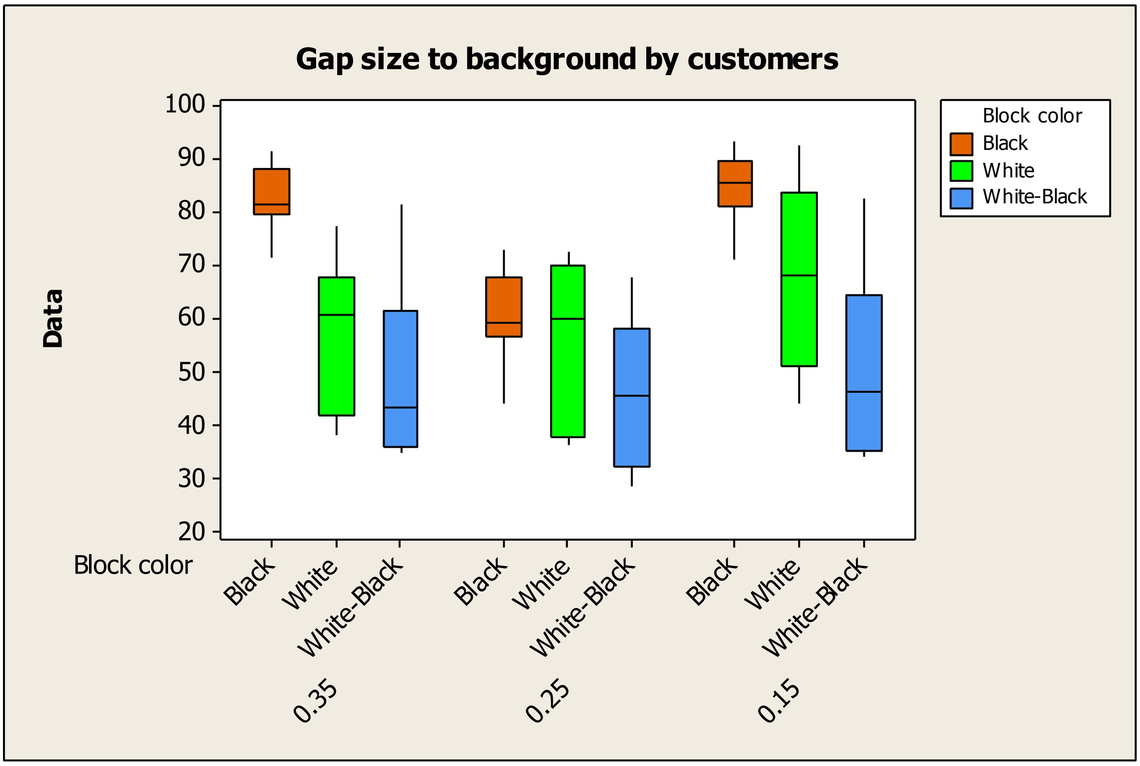


Figure 18: Gap size effect in base color by customers

Finally we can get to the main point of the survey, analyzing color combinations. The survey contained three blocks with different base colors, black, white and white with black background. In figure 19 we can see how the scores divided between them with all base colors, for customers and experts.

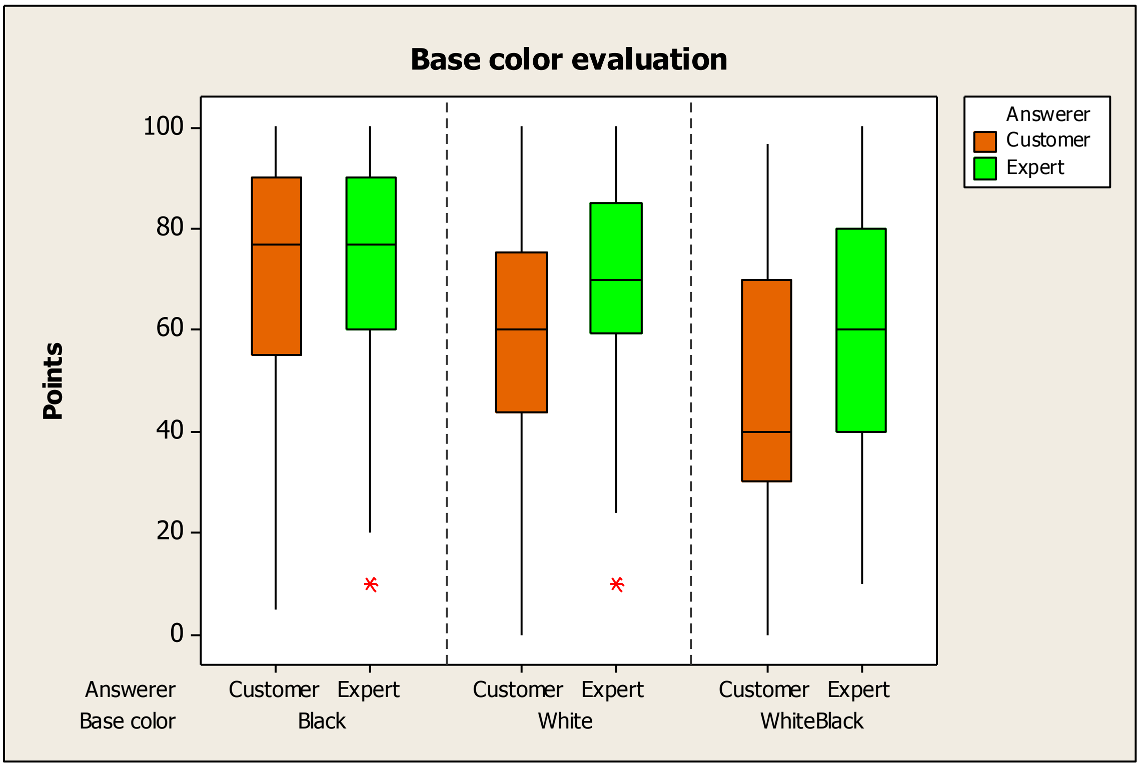


Figure 19: Base color evaluation

It can be seen that black base color has generally gotten better grades, from both customers and experts. The scores are similar between customers and experts also otherwise. Customers have been more critical and graded base colors with bigger difference. Also all six door colors were separately analyzed. The data from both, mechanics experts and customers can be seen in figure 20.

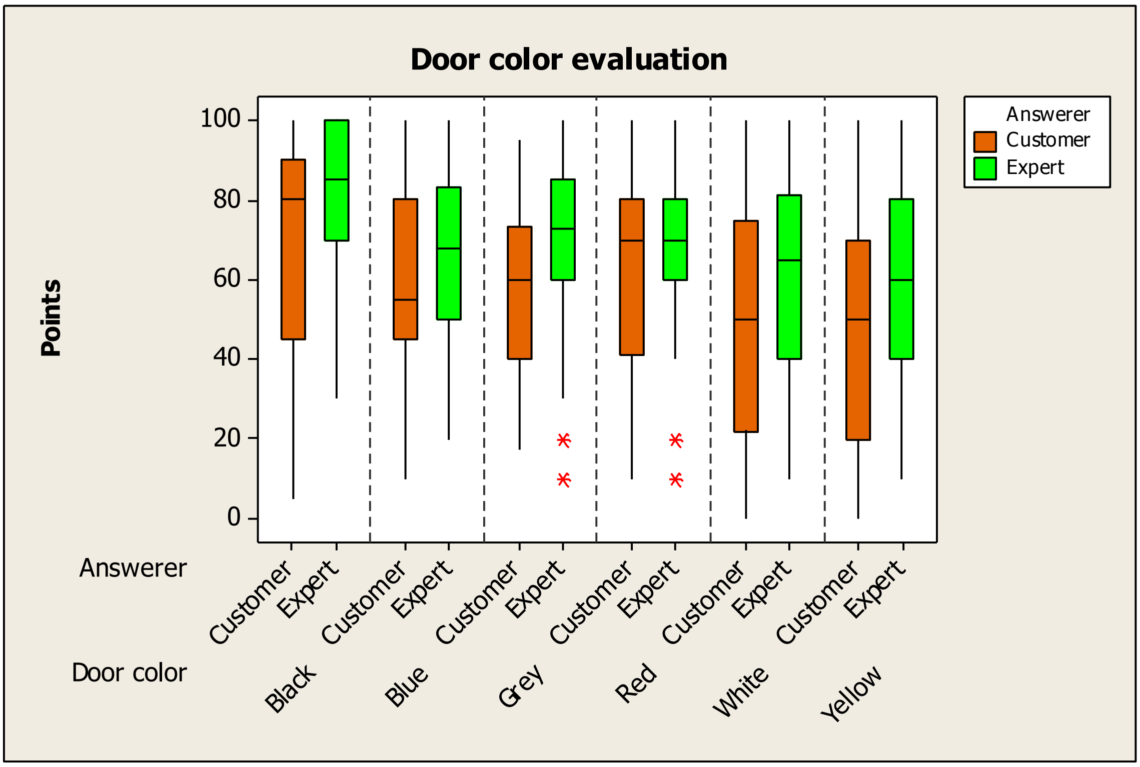


Figure 20: Door color evaluation

In figure 20, door colors from different base colors and gap sizes are combined into same study. From door color evaluation, it can be seen that dark colors are also here better than light colors, black and red being on the top. Variation is much bigger with white and yellow colors. In chapter 2.7, it was stated that we evaluate colors based on the surroundings. Very precise conclusions cannot be made when stripping the other interface color from the study, but dark colors seem to be generally better as a part of interface. From figure 21 we can finally see how color combinations in all cases are seen by both, mechanics experts and customers.

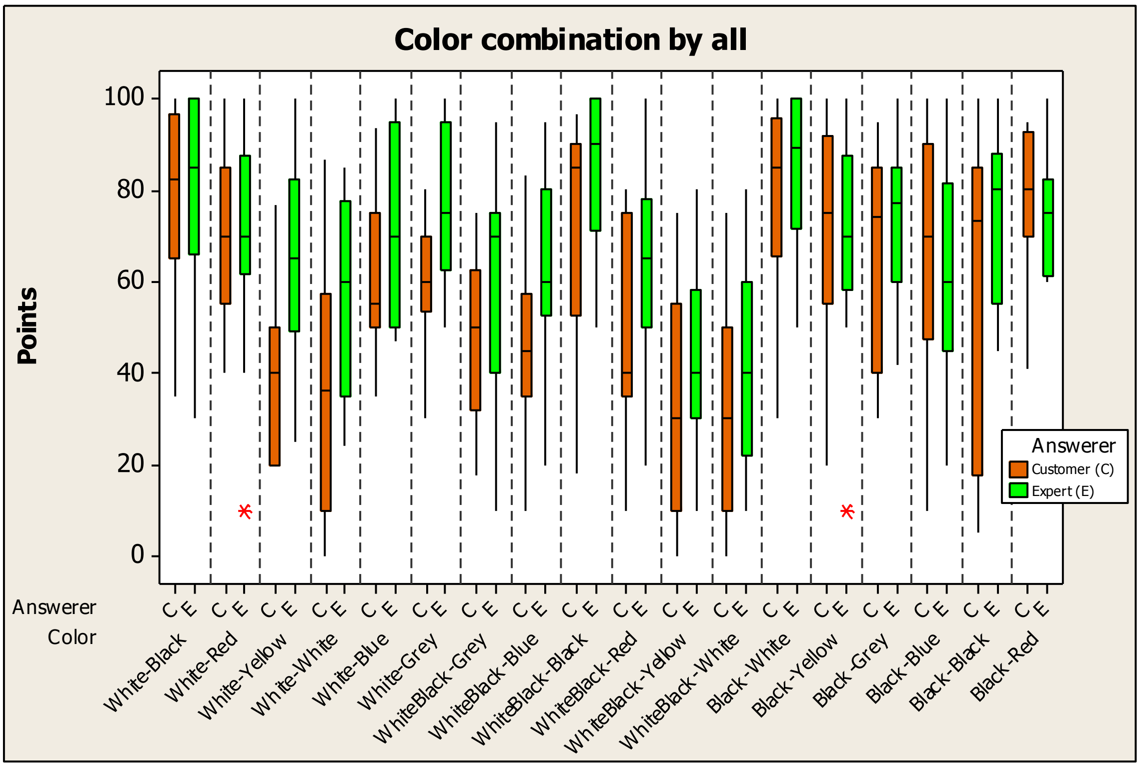


Figure 21: Color combination evaluation by all

From figure 21 we can see the best and worst color combinations and how customers and mechanics experts have evaluated them for all gap sizes. The answers are again in similar order, customers have given generally lower scores. Variation is bigger among customers in most combination cases. Because the combinations were graded in similar order, it seemed reasonable to combine the scores for final evaluation. Box plot is a great way to look at the variation in the data. For comparing groups, interval plot is easier to observe. In figure 22, scores from all three gap sizes and both customers and experts have been combined to same interval plot.

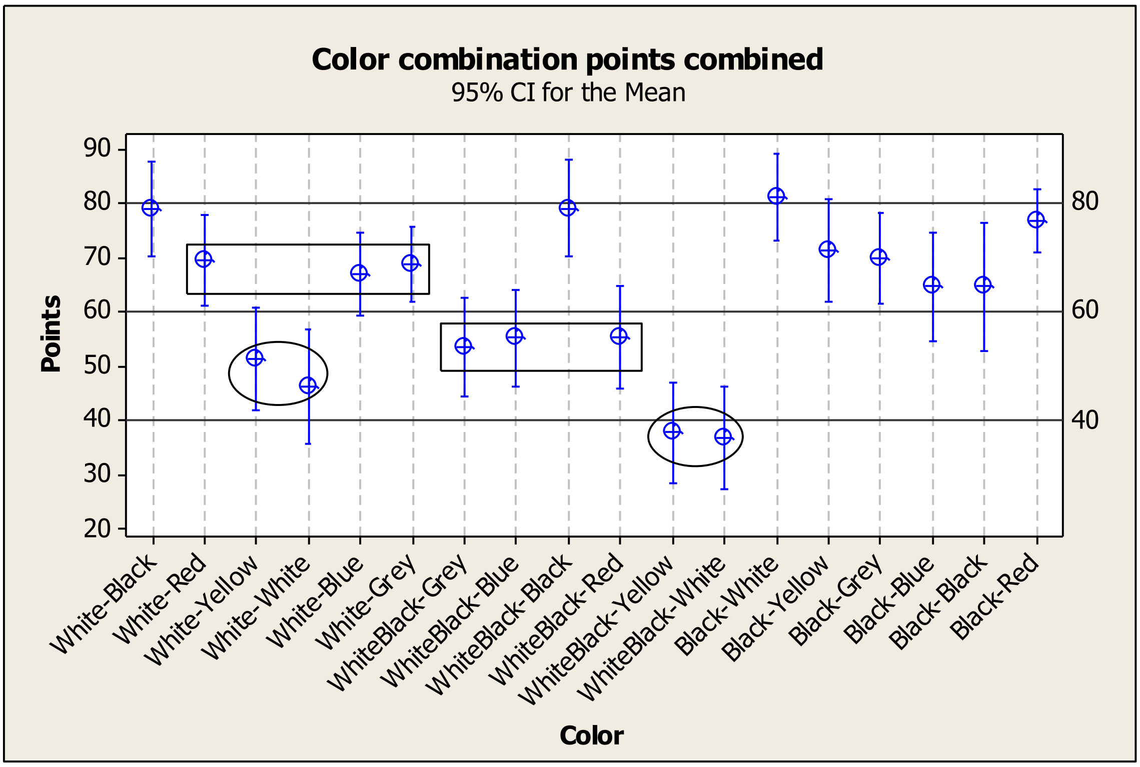


Figure 22: Color combination evaluation combined

From figure 22 it is possible to do final analyzing of color combinations in part interfaces. Best colors seem to be black and white combinations. The grades between white-black and black-white are very similar. Theory presented in the beginning stated that the scores would be similar on different types of part interfaces with same colors and this finding supports the theory. Interestingly black and red combination is ranked among the top. In fact, when performing General Linear Model analysis, Tukey method placed black and red combination amongst three black and white combinations as we can see in table 4.

Table 4: Tukey grouping for color combinations

Grouping Information Using Tukey Method and 95,0% Confidence

Color	N	Mean	Grouping
Black-White	25	81,4	A
WhiteBlack-Black	26	79,1	A
White-Black	26	79,0	A
Black-Red	26	76,8	A
Black-Yellow	26	71,4	A B
Black-Grey	26	70,1	A B
White-Red	26	69,6	A B
White-Grey	26	68,9	A B
White-Blue	26	67,1	A B C
Black-Black	26	64,8	A B C
Black-Blue	26	64,8	A B C
WhiteBlack-Red	26	55,4	B C D
WhiteBlack-Blue	26	55,3	B C D
WhiteBlack-Grey	26	53,6	B C D
White-Yellow	26	51,5	B C D
White-White	26	46,3	C D
WhiteBlack-Yellow	26	37,9	D
WhiteBlack-White	26	36,9	D

Means that do not share a letter are significantly different.

From survey data in appendix A, we can see that color red was clearly the second favorite color after black. It can be that favorite color has effect on how we visualize the interface, whether we like the interface visually or not. Best color combinations according to the study can be seen in figure 23.

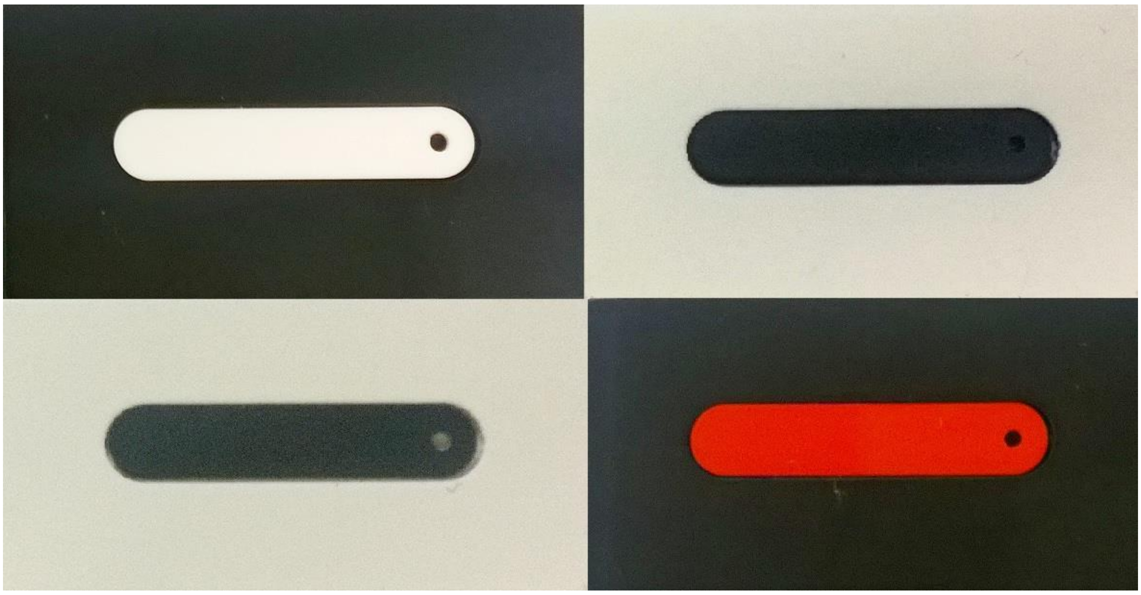


Figure 23: Best color combinations

Light colors, yellow and white, are placed in the bottom with white base colors. Clearly the worst combinations are white base color with black background plus white and yellow door colors. Worst color combinations can be seen in figure 24.

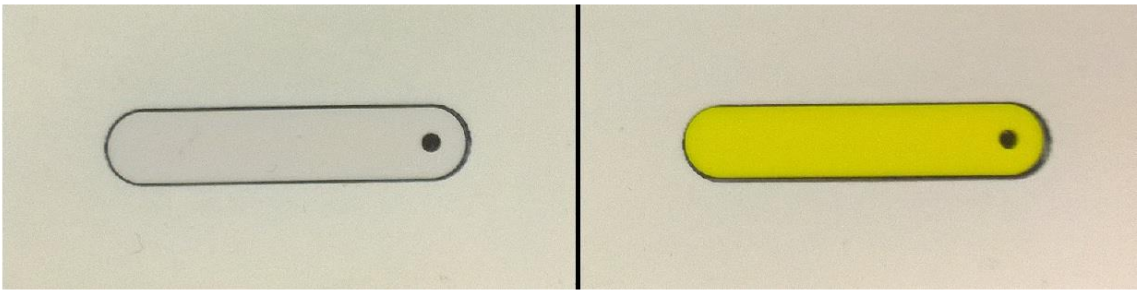


Figure 24: Worst color combinations

Three colors, red, blue and grey are all rated surprisingly similarly with white base color. This is highlighted in figure 22 with boxes. This phenomenon happens also with white base color with black background. This finding supports a theory that the color as such does not really matter, only the brightness of the color matters.

All the scores are similar with white and white-black base colors, white-black being in about 10 points lower level. From this we can see that black background behind interface only helps if black colored part is forming the interface with other color.

Black base color has greatly lower variation between door colors than white color. It can be said that with various colors, black color is the easiest when designing an interface. Black base color with black door is still seen as the worst from black base colors. From this finding it can be concluded that the gap disappears more easily with big contrast between the parts that make the interface.

5.1 Error estimation and survey conclusion

Estimating how big the error in this study is, is not easy. There is no way to calculate a specific measure for survey error [9, p. 48]. Two types of errors can be identified in survey: specification error and processing error [9, p. 39]. Most of the error in this survey comes from specification errors. Questionnaire and survey design are sources for error in this study. Couple of answerers misunderstood the survey question entirely. However, this should not be seen in final scores as these answerers were identified before data analysis. It should also be questioned whether samples used in this survey represent different types of interfaces, and carefulness is needed when drawing conclusions from the survey. The target population from customers is huge, and it is clear that the sample from population does not represent all of the target population. This is known as coverage error [9, p. 42]. For example, all answerers were from the same culture so cultural preferences cannot be seen. Different colors are preferred in different cultures, but human eyes work similarly everywhere. The survey tried to find out how well two parts fit together, and it is assumed that cultural differences does not affect in evaluation.

In the beginning, it was stated that 17 answers was needed to achieve reliability in this study. The number of answerers gathered was 13 from consumers and 15 from mechanics experts. In the end, both rated the combinations in similar order, so the data was combined for final conclusion. This provided more reliable results as the number of answerers for color combination evaluation is 28. Still, carefulness is needed in analyzing the final scores, because there is not big differences in many color combination cases. Conclusions were drawn from figure 22, which shows mean values for color combinations with 95% confidence interval. High confidence interval doesn’t show big differences between colors, because the interval bar for mean value is large. If the confidence for mean value is lowered to 70%, it is possible to see bigger differences, as seen in figure 25. However, if conclusions were drawn from figure 25, the results would not be as reliable.

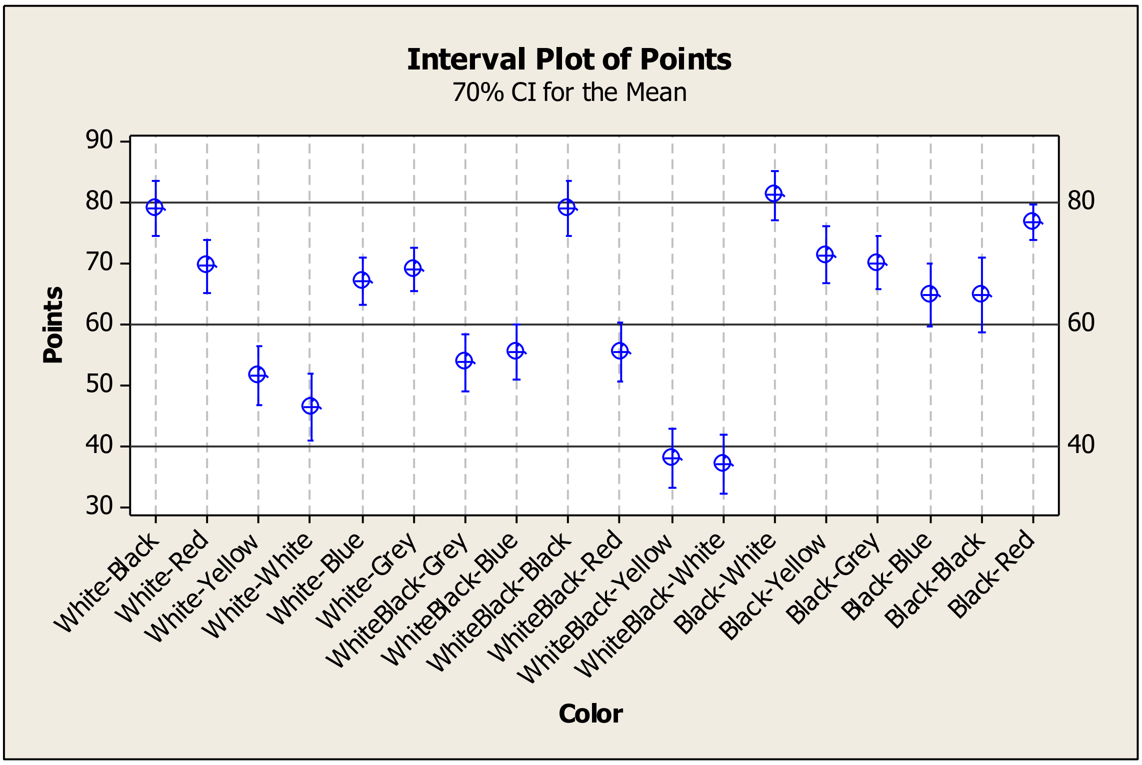


Figure 25: Color combination points with 70% confidence interval

Before building the survey, assumption was that visual interface between parts would disappear more easily with high contrast between colors. Even though the survey created was not perfect, it seemed to produce similar scores from answerers. In the end, it seems that the primary assumption is proven with the survey, black and white combinations got clearly the highest scores. Usually, the color of the gap between parts appears black or dark, because parts form a shade in the interface. Perhaps because of this reason, black color as one part provides a better looking interface and the gap is not easily seen.

In the beginning of survey creation, it was assumed that scores would be similar with same colors in different types of interfaces. Even though black-white and white-black combinations produced similar outcomes, the situation is not as simple. If the interface was bigger, the result could be different. In chapter 2.7, it was found out that red square appears bigger on a black background. It might be that this is the reason why red-black combination in this type of interface hided the gap well. The result might be different if the parts were side by side and not inserted to each other. However, it is assumed that big contrast between parts hides the gap well in all types of interfaces.

Light colors combined with white base color produced the lowest scores in the survey, which was expected. The only way to make this kind of interface visually better, is to reduce size of the gap or try to make the gap appear lighter. In fact, based on the survey it seems that black highlighting of the gap is not really helpful in any kind of situation. It only produced marginally better results with white-black combination. Because the difference is so small, it should be evaluated how much effort should be put into trying to highlight the gap with different colors.

One important finding of the survey is in seeing how little consumers pay attention to these kind of features in products. It seemed that mechanical experts were able to distinguish the difference between smaller and bigger gaps. Consumers only seemed to focus on finding the differences between colors, which admittedly was stated in survey clarification. Quality is subjective, even though six customers in this survey did not find differences between 0.35mm and 0.15mm gap, focus on trying to minimize visual interface features should not be forgotten. As stated in the introduction of this thesis, big variation in part interfaces can lead to big visual and functional problems.

6. CONCLUSIONS

Product development cycles are constantly speeding up, better products need to be developed faster. It is important that products meet customers' expectations, half-finished products cannot be launched in order to compete in the market. Involving quality improvement methods and customer feedback in product development is essential.

Mechanical quality of the product depends highly on dimensional quality of the parts. A way to improve dimensional quality is to improve manufacturing processes with Six Sigma. Six Sigma is a quality improving method that tries to reduce variation in the process with statistical and analytical problem solving tools. The process for Six Sigma consists of five steps: Define, Measure, Analyze, Improve and Control. It requires a lot of knowledge and expertise in the process improved and statistics. The basis of Six Sigma process is in thorough and reliable measuring of the process and analyzing and improving them with Design of Experiment methods. With DOE, it is possible to test and optimize several variables at the same time.

Dimensional variation is something that cannot be designed in the product with 3D CAD programs, as only dimensionally perfect parts can be seen on CAD software. A way to design dimensional variation in the product is to use tolerance analysis. With tolerance analysis, it is possible to calculate how much variation the design can withstand. The drawback of traditional tolerance analysis is that it is usually applied late in the design process and the possibilities for design changes are limited. A way to take dimensional quality better into account is to incorporate dimensional quality management process with 3D tolerances. Using 3D tolerance program allows more accurate and realistic tolerance calculations. The best way is to tie 3D CAD and 3D tolerance software together to allow for real time changes in tolerance analysis. For part interface design 3D tolerance software has a great benefit in being able to show how the variation looks before any real samples are made. This way it is possible to involve industrial designer in the interface design process, and it is possible to verify what kind of variation is allowed.

This thesis is focused on part interface quality. Gap size between parts is determined by dimensional variation in manufacturing and assembling. The visual appearance of an interface is not only determined by its size, the colors of the parts that form the interface influences heavily in how it is seen. A lot of studies has been done in color analysis, how colors are seen varies heavily depending the situation. This is why it needs to be studied, how colors affect customers in the case of part interfaces. When trying to take into account customers feeling of a product, a customer satisfaction analysis method is needed. Several methods have been developed for this problem. The power of Kansei Engineering method is in capturing subjective feelings of customers. It tries to convert feelings that are hard to explain into concrete design parameters. Voice of the Customer is more traditional way to include customers in the design process. It uses marketing research tools, but gives customer a choice, instead of measuring only satisfaction. Qualitative and quantitative data can be collected and analyzed with Voice of the Customer methods. Voice of the Customer is also involved with Six Sigma. Both are based on statistical thinking principles: “All work occurs in a system of interconnected processes. Variation exists in all processes. Understanding and reducing variation are the keys to success.” [20] Quantitative Voice of the Customer process provides numerical results that are analyzed with same tools that are in use in Six Sigma process, including descriptive statistics, hypothesis testing and correlation.

Voice of the Customer was used to find out how part interfaces are seen with different color combinations. Quantitative VOC in the form of a customer survey was developed for this problem. It is easy to think that making a survey is just asking a question and analyzing the answers. However, making a reliable survey is much more than that, it is a time consuming process. Part interface for the survey was found in mobile phone SIM-door gap. It is often a visible and problematic gap in mobile phones. Samples were made with stereolithographic 3D printing, so that all external factors could be removed from the study. 3 different gap sizes were chosen to represent different variation in parts. Question of the survey was tried to keep simple to understand, and in the end it was concluded to be: “Evaluate the fit between parts”. Still, it was shown that this kind of topic was unfamiliar to many, and answering was not always easy. Two target groups were chosen for survey, mechanics experts and normal customers. In the end 13 answerers was gotten from customers and 15 from mechanics experts. When analyzing the survey results it was shown that the question was not always correctly understood, as was also reported by a couple of answerers. Still, clear differentiation for color combinations was found in many cases.

According to the survey, black color as one part usually provides a better looking interface and the gap is not easily seen. When the gap between parts grow, superiority of black color is even bigger. Still, if both parts of the interface are black, the result is not so good. Gap is always a little different colored than the parts, so the colors does not blend as easily. The survey consisted of two kinds of white base colors, where one was highlighted with black background behind the SIM-door. The only case where the highlighted one showed better results than without highlighting, was with black door. It seems that highlighting the gap with black accent in white base color only helps if the other color of interface is black. Even with black and white combination, black accent color behind the SIM-door showed only marginally better results.

The best combinations, where the gap between parts disappeared most easily, was found to be black and white combinations. Interestingly, black and red combination showed to be also in favor of answerers. It was found out that red object in black background actually appears bigger than in white background. Red color was also favorite to many answerers.

These might be the reasons why red and black combination was found to be so good. Otherwise red, blue and grey got very similar results with different base colors. Also similar results were observed from very light yellow and white parts. This supports a theory that the color of the parts itself is not the biggest factor when looking at visual appearance of part interface. Big contributor to appearance is the lightness of the color. When there is a big contrast between the parts that form the gap, it is not so easy to see the gap itself. The worst colors proved to be light colors, yellow and white, with white base colors. When this was still highlighted with black accent behind the SIM-door, the visual appearance turned out to be the worst.

How these results change with different gap sizes would be interesting to study. It can be figured, that with very small gap sizes, the differences between color combinations would disappear, as the gap is not seen with any color combinations.

REFERENCES

[1] Allen, T. Introduction to Engineering Statistics and Six Sigma. London, 2006, Springer. 481 p.

[2] Auchard, E. 'Hairgate': iPhone hit by new social media storm. Reuters. 7.10.2014. [accessed on 6.2.2015]. Available: http://uk.reuters.com/article/2014/10/07/us-apple-iphone-hair-idUKKCN0HW1MA20141007.

[3] Craig, M. Dimensional management versus tolerance assignment. Assembly Automation 16(1996)2, pp. 12-16.

[4] Dean, A. & Voss, D. Design & Analysis of Experiments. Secaucus, NJ, USA 1998. Springer.

[5] DMAIC Versus DMADV. iSixSigma. [accessed on 6.3.2015]. Available: http://www.isixsigma.com/new-to-six-sigma/design-for-six-sigma-dfss/dmaic-versus-dmadv/.

[6] Dolcourt, J. Why smartphone design matters (Smartphones Unlocked). cnet.com. [accessed on 5.3.2015]. Available: http://www.cnet.com/news/why-smartphone-design-matters-smartphones-unlocked/.

[7] Fischer, B. Mechanical Tolerance Stackup and Analysis. 2nd ed. 2011, CRC Press. 472 p.

[8] Germani, M., Mandorli, F., Mengoni, M. & Raffaeli, R. CAD-based environment to bridge the gap between product design and tolerance control. Precision Engineering 34(2010)1, pp. 7-15.

[9] Gideon, L. Handbook of Survey Methodology for the Social Sciences. 2012, Springer New York. 520 p.

[10] Giordano, M., Mathieu, L. & Villeneuve, F. Product Life-Cycle Management: Geometric Variations. 2012, Wiley.

[11] Guibal, C.R.C. & Dresp, B. Interaction of color and geometric cues in depth perception: When does "red" mean "near"? Psychological research 69(2004)1-2, pp. 30-40.

[12] Item Analysis with Cronbach's Alpha for Reliable Surveys. Minitab, Ltd. [accessed on 25.2.2015]. Available: http://www.minitab.com/en-us/Support/Tutorials/Item-Analysis-with-Cronbach-s-Alpha-for-Reliable-Surveys/.

[13] Justification for GD&T. Advanced Dimensional Management LLC. [accessed on 19.2.2015]. available: http://www.advanceddimensionalmanagement.com/justification-for-gdt/.

[14] Koch, P.N., Yang, R.J. & Gu, L. Design for six sigma through robust optimization. Structural and Multidisciplinary Optimization (2004) pp. 235-248.

[15] Kuehni, R.G. Color: An Introduction to Practice and Principles (3rd Edition). Somerset, NJ, USA 2012, John Wiley & Sons.

[16] Lokman, A.M. Design & Emotion: The Kansei Engineering Methodology. Malaysia 2010. Universiti Teknologi MARA (UiTM).

[17] Minitab 17. SixSigma.us. [accessed on 9.3.2015]. Available: http://www.6sigma.us/minitab-training.php

[18] Morton, J.L. Basic Color Theory. Color Matters. [accessed on 19.2.2015]. Available: http://www.colormatters.com/color-and-design/basic-color-theory.

[19] Pereira, M. SBTI Lean Six Sigma Pocket Guide. 2011, SBTI Europe. 257 p.

[20] Phadnis, S. The Importance of Statistical Thinking. iSixSigma. accessed on 23.2.2015]. Available: http://www.isixsigma.com/implementation/case-studies/importance-statistical-thinking/.

[21] Ploytip Jirasukprasert, Garza-Reyes, J.A., Kumar, V. & Lim, M.K. A Six Sigma and DMAIC application for the reduction of defects in a rubber gloves manufacturing process. International Journal of Lean Six Sigma 5(2014)1, pp. 2-21.

[22] ProJet® 6000 HD. 3D Systems. [accessed on 2.3.2015]. Available: http://www.3dsystems.com/3d-printers/professional/projet-6000-hd.

[23] Quality Glossary - Q. American Society for Quality. [accessed on 6.2.2015]. Available: http://asq.org/glossary/q.html.

[24] Rots, R.M. Pinpointing assembly VARIATION. Machine Design 78(2006)16, pp. 78-80.

[25] Samsung Galaxy Note 4 review: Four of a kind. GSMArena. 10.10.2014, [accessed on 6.2.2015]. Available: http://www.gsmarena.com/samsung_galaxy_note_4-review-1147.php.

[26] Schütte, S. Engineering Emotional Values in Product Design -Kansei Engineering in Development. Linköping, Sweden 2005. Linköpings Universitet, Division of Quality Technology and Management, Department of Mechanical Engineering. Linköping Studies in Science and Technology, Dissertation 951. 49 p.

[27] Shillito, M.L. Acquiring, processing and deploying voice of the customer. 2001, CTC Press LLC. 271 p.

[28] Stevenson, W. Operations Management. Tenth ed. New York 2009, McGraw-Hill/Irwin. 906 p.

[29] Stoll, T. & Paetzold, K. Gap and Flush Visualization of Virtual Nonideal Prototypes. International Design Conference, 19-22.5.2008. Croatia 2008, pp. 681.

[30] Survey design interview. Ossi Hämeenoja 10.12.2015.

[31] What is a boxplot? Minitab, Ltd. [accessed on 25.2.2015]. Available: http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/graphs/graphs-that-compare-groups/boxplots/boxplot/.

[32] What is a confidence interval? Minitab, Ltd. [accessed on 25.2.2015]. Available: http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/introductory-concepts/confidence-interval/confidence-interval/.

[33] What is a general linear model? Minitab, Ltd. [accessed on 26.2.2015]. Available: http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/anova/basics/what-is-a-general-linear-model/.

[34] What is an interval plot? Minitab, Ltd. [accessed on 25.2.2015]. Available: http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/graphs/graphs-that-compare-groups/interval-plots/interval-plot/.

[35] Yang, K. Voice of the Customer: Capture and Analysis. McGraw-Hill Publishing.

APPENDIX A: SURVEY RESULTS

Customers

Table 4: Customer scores to 0.35mm gap

0,35mm			Answerer					
Block	Door	No.	1	2	3	4	5	10
White	Black	1	90	100	60	90	60	65
	Red	2	80	70	60	80	60	40
	Yellow	3	70	50	30	10	50	20
	White	4	100	60	20	20	60	0
	Blue	5	90	70	50	70	50	30
	Grey	6	70	80	50	70	60	40
White-Black	Grey	7	40	40	60	50	40	0
	Blue	8	80	30	70	60	50	0
	Black	9	90	90	100	90	30	90
	Red	10	90	40	90	70	40	0
	Yellow	11	50	30	90	10	30	0
	White	12	60	40	80	10	30	0
Black	White	13	100	100	100	90	70	90
	Yellow	14	90	90	100	90	70	90
	Grey	15	**90**	**80**	**100**	**90**	**80**	**40**
	Blue	16	100	80	100	90	80	40
	Black	17	70	100	80	90	90	0
	Red	18	90	95	100	90	80	40
Black	Grey	19	**70**	**70**	**90**	**90**	**80**	
	Favorite color		Black	Red	Black			White

Table 5: Customer scores to 0.25mm gap

0,25mm			Answerer						
Block	Door	No.	1	6	7	8	11	12	13
White	Red	21	90	70	40	50	100	50	85
	Black	22	90	90	35	49	100	75	70
	White	23	90	10	36	22	60	0	50
	Yellow	24	80	20	30	20	40	20	45
	Grey	25	70	59	30	48	70	60	60
	Blue	26	100	50	35	47	80	55	80
White-Black	Blue	27	90	45	40	30	50	50	80
	Grey	28	70	51	35	29	25	45	65
	Red	29	80	40	35	19	35	40	70
	Black	30	100	60	45	18	90	80	83
	White	31	60	30	30	10	10	20	75
	Yellow	32	50	20	30	15	5	10	70
Black	Yellow	33	100	60	60	70	20	50	50
	White	34	90	71	55	76	30	60	85
	Blue	35	100	40	60	65	50	10	90
	Grey	36	**80**	**30**	**55**	**74**	**40**	**40**	**80**
	Red	37	90	41	70	75	60	80	95
	Black	38	60	12	40	13	30	75	80
Black	Grey	39	**70**	**80**	**50**	**72**	**45**	**60**	**100**
	Favorite color		Black			Black	Red		Red

Table 6: Customer scores to 0.15mm gap

0,15mm			Answerer						
Block	Door	No.	1	2	3	4	5	9	10
White	Black	41	100	100	100	90	60	100	100
	Red	42	100	80	95	90	60	60	80
	Yellow	43	80	60	70	30	50	40	45
	White	44	70	50	70	30	60	10	20
	Blue	45	90	55	80	70	50	50	80
	Grey	46	80	60	80	70	60	80	50
White-Black	Grey	47	50	60	80	70	40	75	35
	Blue	48	80	50	60	40	30	30	20
	Black	49	100	90	70	90	40	90	100
	Red	50	70	40	70	30	30	80	20
	Yellow	51	50	40	60	10	30	60	0
	White	52	60	40	60	10	30	40	0
Black	White	53	100	95	90	90	80	100	100
	Yellow	54	90	90	100	90	80	80	100
	Grey	55	**80**	**90**	**90**	**90**	**90**	**30**	**100**
	Blue	56	100	85	100	90	80	45	100
	Black	57	90	100	80	90	90	5	45
	Red	58	90	95	90	90	80	80	100
Black	Grey	59	**90**	**90**	**94**	**90**	**90**	**30**	**100**
	Favorite color		Black	Red	Black			White	White

Mechanics experts

Table 7: Mechanics expert scores to 0.35mm gap

0,35mm			Answerer				
Block	Door	No.	1	2	3	4	5
White	Black	1	100	43	60	30	15
	Red	2	80	63	10	40	10
	Yellow	3	70	58	40	60	40
	White	4	50	24	40	70	50
	Blue	5	90	47	60	50	20
	Grey	6	85	68	60	60	60
White-Black	Grey	7	70	73	10	30	80
	Blue	8	80	74	20	30	60
	Black	9	100	93	90	70	100
	Red	10	65	76	20	60	80
	Yellow	11	60	35	10	40	95
	White	12	40	24	10	40	80
Black	White	13	100	93	100	100	100
	Yellow	14	90	56	80	100	100
	Grey	15	**70**	**42**	**100**	**80**	**100**
	Blue	16	60	83	40	80	80
	Black	17	50	86	90	60	80
	Red	18	80	72	60	70	100
Black	Grey	19	**50**	**77**	**80**	**70**	**100**
	Favorite color		Black	Black			Red

Table 8: Mechanics expert scores to 0.25mm gap

0,25mm			Answerer					
Block	Door	No.	10	11	12	13	14	15
White	Red	21	65	60	90	70	70	75
	Black	22	95	85	90	80	72	75
	White	23	70	85	30	60	85	25
	Yellow	24	60	85	40	65	71	25
	Grey	25	70	100	75	65	78	50
	Blue	26	80	85	50	70	70	50
White-Black	Blue	27	80	95	60	55	75	50
	Grey	28	60	95	20	50	60	75
	Red	29	70	85	70	50	60	50
	Black	30	85	85	95	60	72	50
	White	31	50	80	10	40	75	25
	Yellow	32	50	80	40	40	56	25
Black	Yellow	33	85	85	10	60	68	50
	White	34	95	85	80	55	77	50
	Blue	35	90	50	40	50	68	50
	Grey	36	**85**	**85**	**60**	**50**	**77**	**75**
	Red	37	85	95	60	62	78	75
	Black	38	80	85	50	45	82	75
Black	Grey	39	**90**	**95**	**60**	**65**		**75**
	Favorite color		Red	Black	Yellow	White	Cyan	

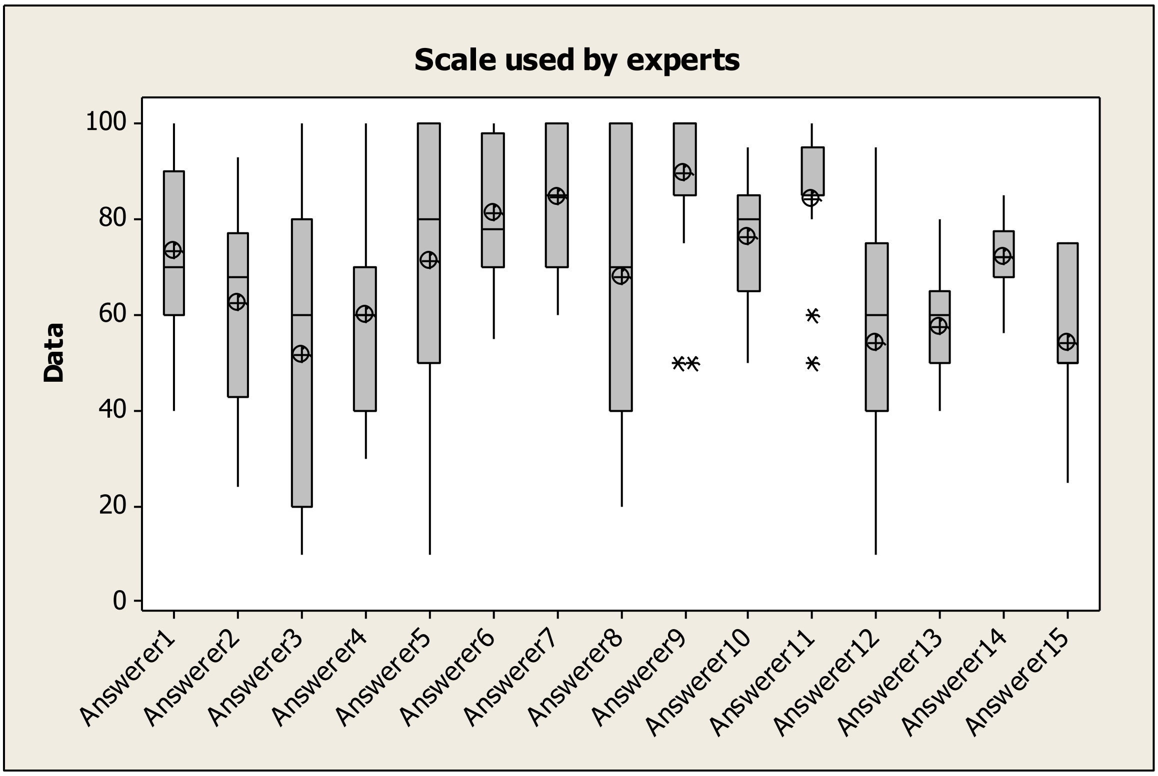
Table 9: Mechanics expert score to 0.15mm gap

0,15mm			Answerer			
Block	Door	No.	6	7	8	9
White	Black	41	73	100	100	100
	Red	42	80	100	100	85
	Yellow	43	100	100	80	85
	White	44	90	80	40	75
	Blue	45	75	100	100	100
	Grey	46	70	100	90	100
White-Black	Grey	47	95	90	70	75
	Blue	48	98	60	60	85
	Black	49	75	100	100	100
	Red	50	80	80	40	100
	Yellow	51	100	70	20	50
	White	52	100	70	20	50
Black	White	53	70		70	100
	Yellow	54	65	70	70	100
	Grey	55	**73**	**80**	**60**	**100**
	Blue	56	78	70	20	100
	Black	57	98	100	70	100
	Red	58	55	60	80	100
Black	Grey	59	**65**	**90**	**100**	**100**
	Favorite color			Black	Blue	

APPENDIX B: MINITAB REPORT

Color questionnaire results

This graph shows how each answerer from mechanics experts has used the scale from 0-100.



Item and Total Statistics

Variable	Total Count	Mean	StDev
Answerer1	17	72,9	17,5
Answerer2	17	59,8	21,1
Answerer3	17	47,1	31,0
Answerer4	17	57,1	20,2
Answerer5	17	67,6	30,9
Answerer6	17	82,6	14,0
Answerer7	17	84,1	15,4
Answerer8	17	65,9	29,0
Answerer9	17	88,5	17,1
Answerer10	17	74,1	13,6
Answerer11	17	83,5	12,2
Answerer12	17	52,4	26,5
Answerer13	17	57,2	11,1
Answerer14	17	72,1	7,8
Answerer15	17	52,9	19,5
Total	17	1017,9	140,5

Internal consistency of the answerers is measured with Cronbachs Alpha.

Cronbach's Alpha = 0,7323

Omitted Item Statistics

Omitted Variable	Adj. Total Mean	Adj. Total StDev	Item-Adj. Total Corr	Squared Multiple Corr	Cronbach's Alpha
Answerer1	945,0	128,7	0,6366	0,9022	0,6907
Answerer2	958,1	128,0	0,5389	0,9390	0,6955
Answerer3	970,9	117,9	0,6685	0,9782	0,6672
Answerer4	960,9	134,0	0,2534	0,9725	0,7268
Answerer5	950,3	142,1	-0,1609	0,9777	0,7947
Answerer6	935,3	149,9	-0,6938	0,9266	0,7869
Answerer7	933,8	135,1	0,2992	0,9279	0,7224
Answerer8	952,1	121,0	0,6070	0,9726	0,6792
Answerer9	929,4	125,4	0,8650	0,9799	0,6693
Answerer10	943,8	130,4	0,7196	0,9824	0,6929
Answerer11	934,4	138,4	0,1270	0,9023	0,7342
Answerer12	965,6	126,2	0,4646	0,8081	0,7020
Answerer13	960,8	135,3	0,4322	0,9854	0,7169
Answerer14	945,8	139,3	0,1258	0,5720	0,7336
Answerer15	965,0	130,2	0,4706	0,9398	0,7045

It can be seen that there are two answerers who have rather different answers compared to the rest.

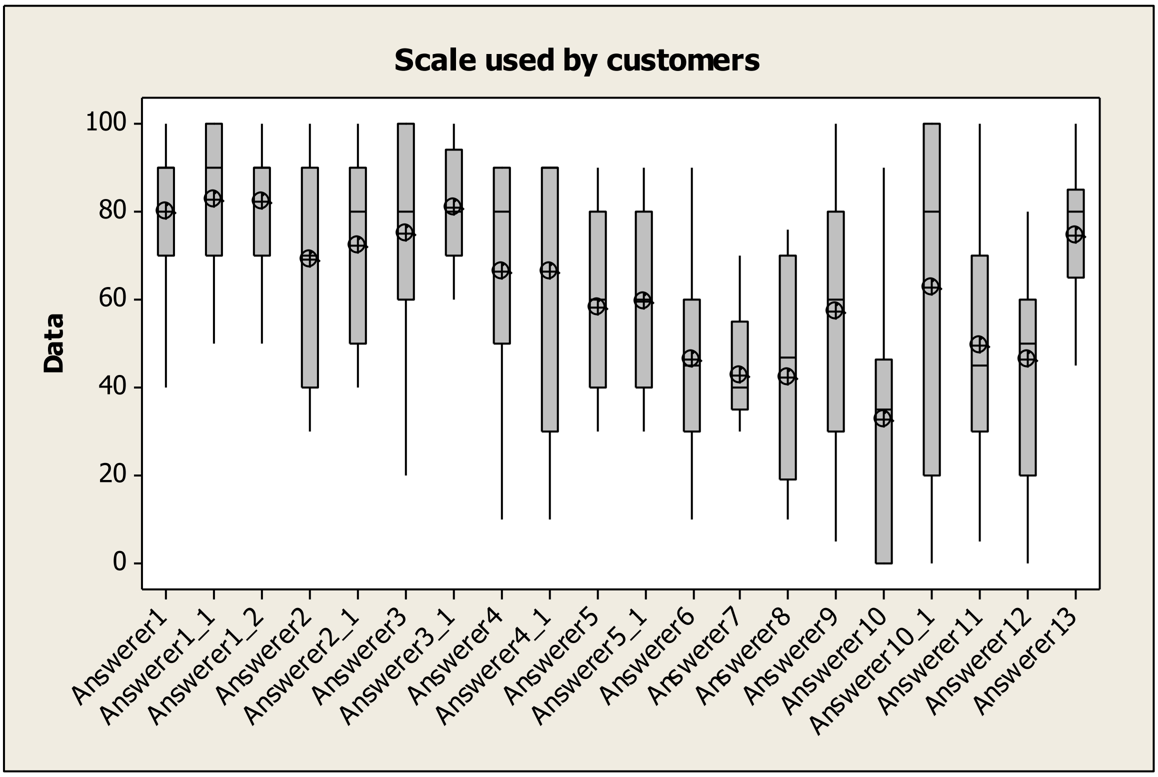
After removing two answerers not in line we have much better internal consistency.

Cronbach's Alpha = 0,8416

Omitted Item Statistics

Omitted Variable	Adj. Total Mean	Adj. Total StDev	Item-Adj. Total Corr	Squared Multiple Corr	Cronbach's Alpha
Answerer1	794,7	138,2	0,7101	0,8628	0,8172
Answerer2	807,8	140,2	0,4597	0,8652	0,8330
Answerer3	820,6	130,5	0,5975	0,9763	0,8261
Answerer4	810,6	146,2	0,1783	0,9535	0,8517
Answerer7	783,5	144,3	0,3948	0,9270	0,8366
Answerer8	801,8	128,7	0,7293	0,9721	0,8104
Answerer9	779,1	135,8	0,8834	0,9547	0,8067
Answerer10	793,5	140,9	0,7288	0,9788	0,8211
Answerer11	784,1	149,6	0,0856	0,8307	0,8497
Answerer12	815,3	134,9	0,5477	0,7594	0,8278
Answerer13	810,5	144,1	0,6066	0,9669	0,8297
Answerer14	795,5	149,7	0,1550	0,5048	0,8458
Answerer15	814,7	140,2	0,5105	0,9255	0,8293

Here we can see how customers have divided their scores.



Item and Total Statistics

Variable	Total Count	Mean	StDev
Answerer1	18	80,56	17,31
Answerer2	18	69,17	25,22
Answerer3	18	74,44	25,72
Answerer4	18	65,00	31,30
Answerer5	18	57,22	18,73
Answerer6	18	44,39	21,84
Answerer7	18	42,28	12,36
Answerer8	18	40,56	23,78
Answerer9	18	58,61	29,04
Answerer10	18	32,50	33,27
Answerer11	18	49,72	29,18
Answerer12	18	45,56	25,02
Answerer13	18	72,94	14,27
Total	18	732,94	209,65

Internal consistency is again measured with Cronbach's alpha.

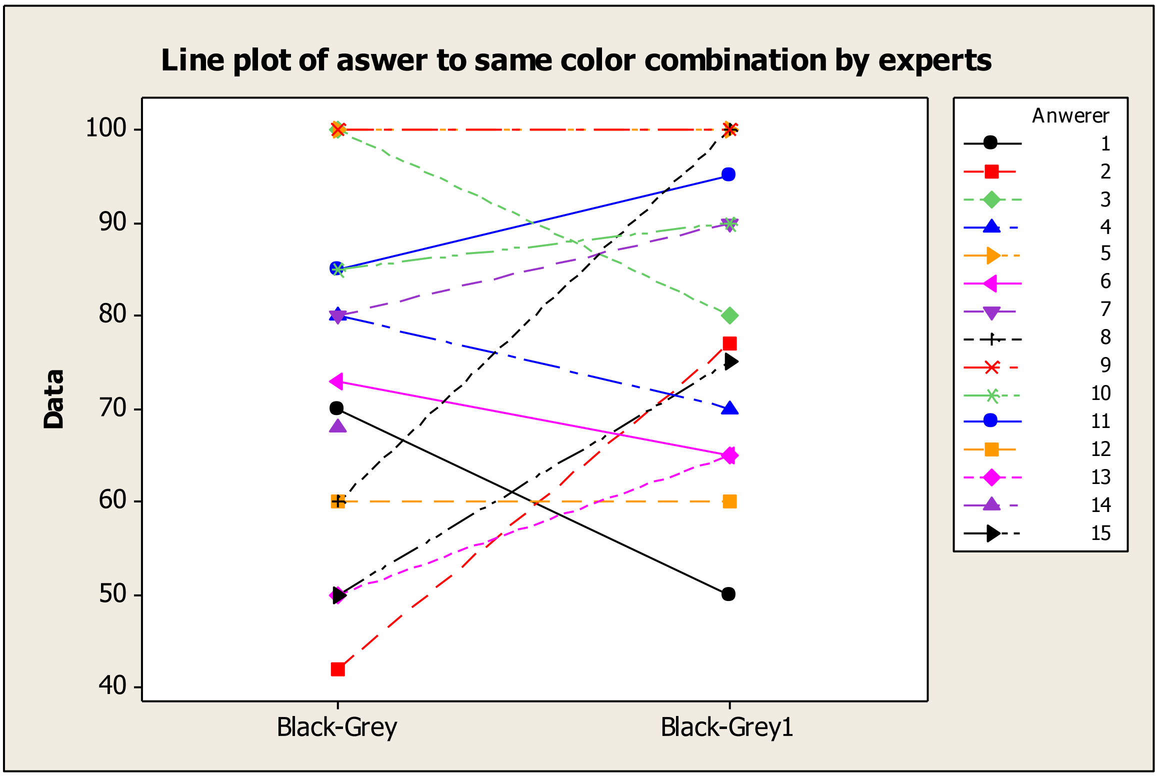
Cronbach's Alpha = 0,8922

Omitted Item Statistics

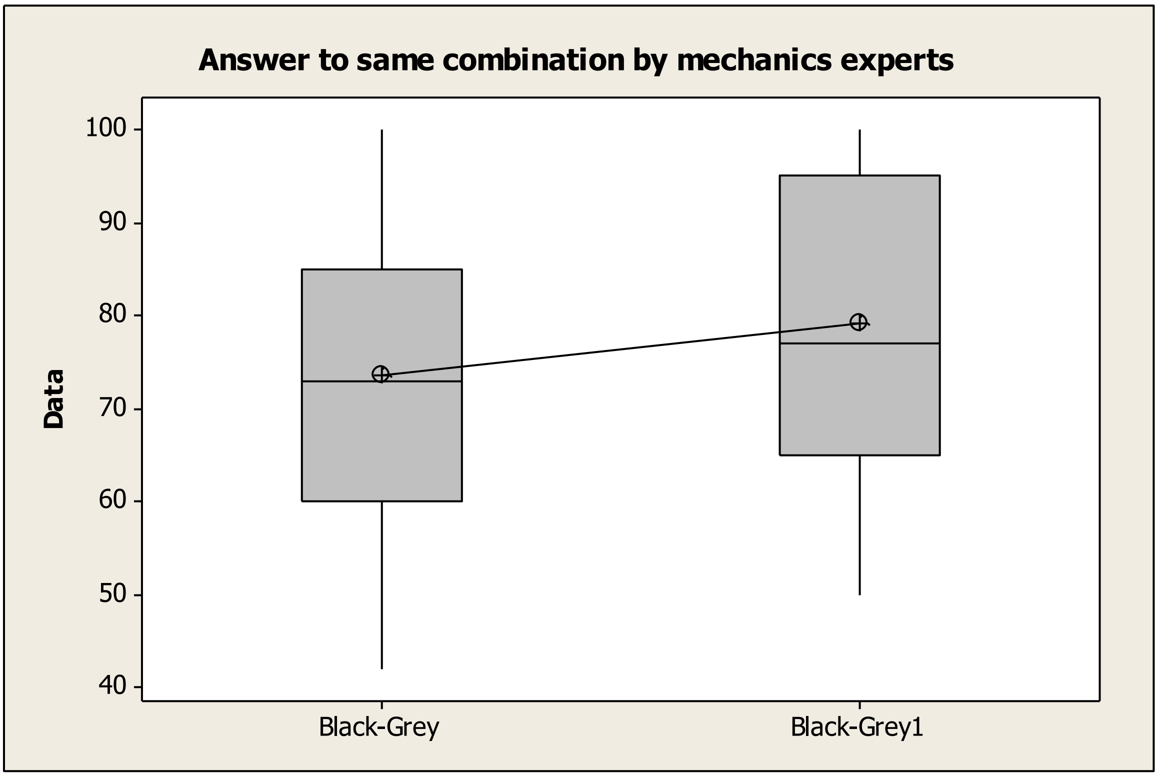
Omitted Variable	Adj. Total Mean	Adj. Total StDev	Item-Adj. Total Corr	Squared Multiple Corr	Cronbach's Alpha
Answerer1	652,4	199,6	0,5519	0,6785	0,8868
Answerer2	663,8	189,1	0,7904	0,9483	0,8739
Answerer3	658,5	197,8	0,4084	0,9303	0,8932
Answerer4	667,9	181,6	0,8805	0,9365	0,8667
Answerer5	675,7	200,9	0,4320	0,9529	0,8908
Answerer6	688,6	194,5	0,6625	0,9135	0,8811
Answerer7	690,7	201,3	0,6612	0,8843	0,8862
Answerer8	692,4	192,2	0,7055	0,8974	0,8786
Answerer9	674,3	193,9	0,4905	0,8671	0,8903
Answerer10	700,4	182,5	0,7862	0,9463	0,8731
Answerer11	683,2	196,8	0,3798	0,8601	0,8965
Answerer12	687,4	192,6	0,6462	0,7960	0,8813
Answerer13	660,0	202,8	0,4491	0,8264	0,8907

It can be seen that customers have quite good internal consistency between their answers.

Validity of the test was evaluated when same color combination was asked twice from the answerers

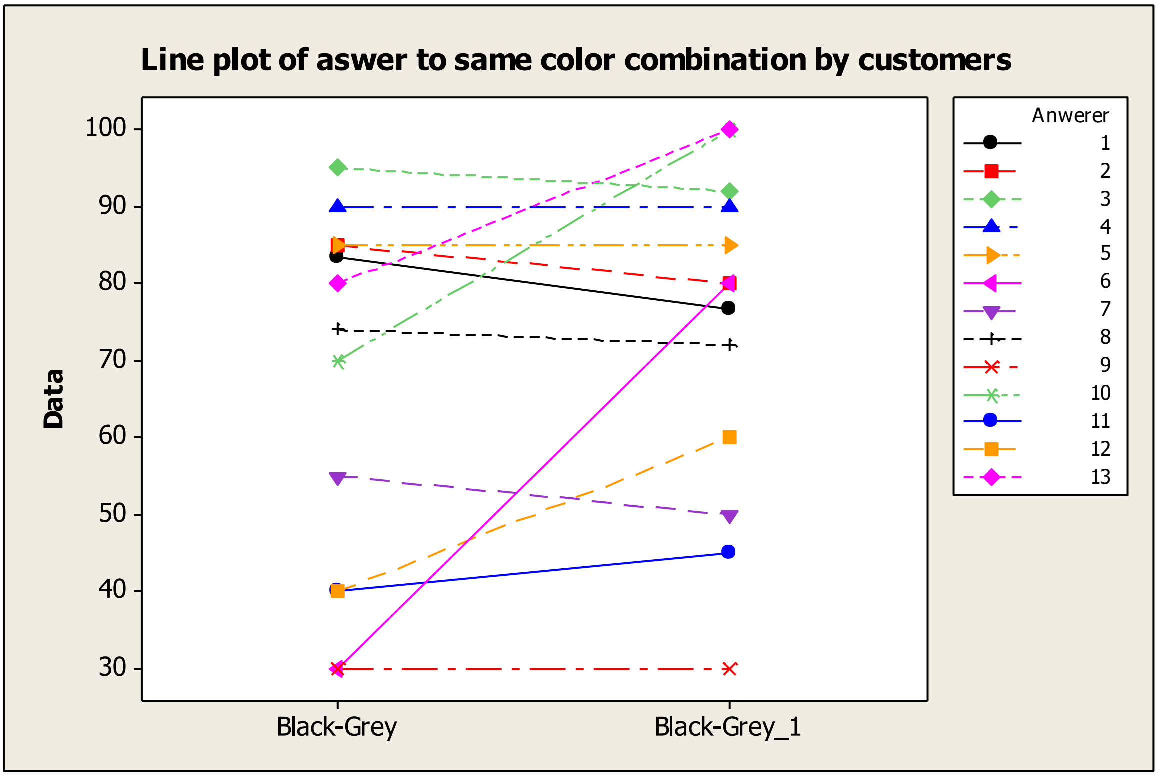


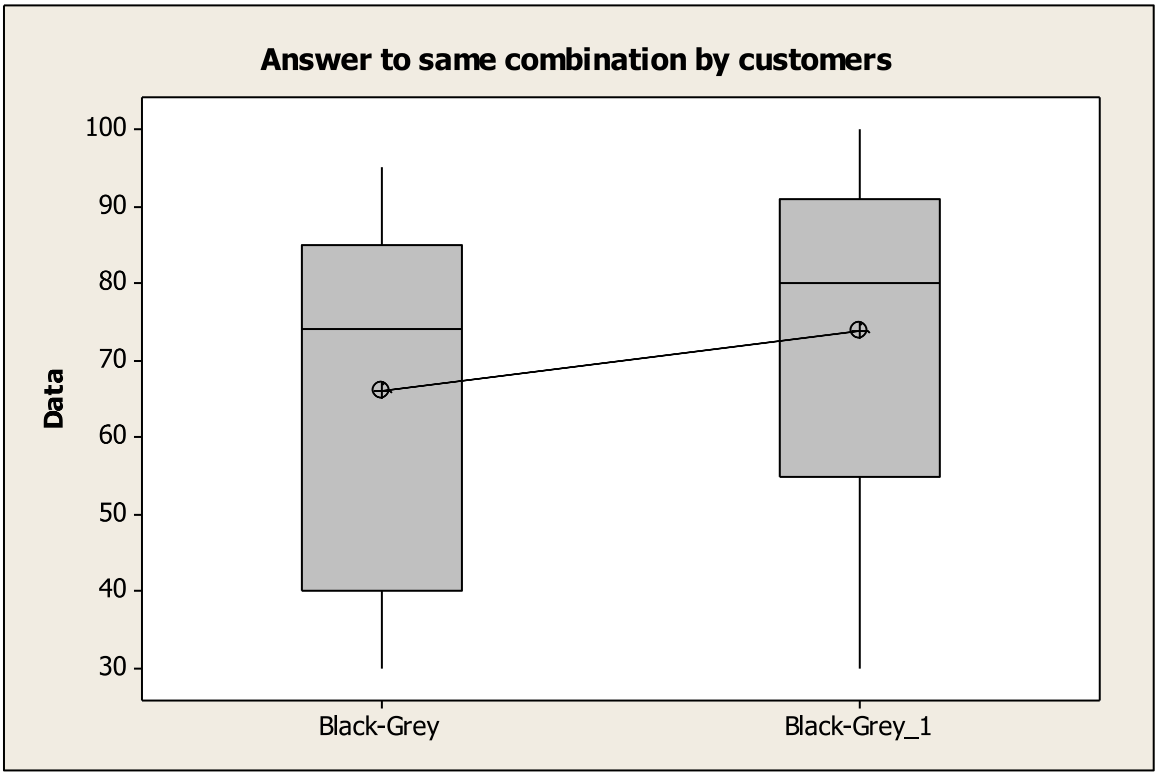
Bigger slope in the line means that the informants answer has a big change to the same color combination. It is not clear where this change comes from.



When combining the answers the change is not that big.

This graphs shows the answer to same combination by customers. Customers have stayed better in line. Here we can see that answerer 6 has had a dramatic change in scoring the same combination. Answerer 6 didn't come up in internal consistency study, so otherwise answerer 6 has stayed in line with others.



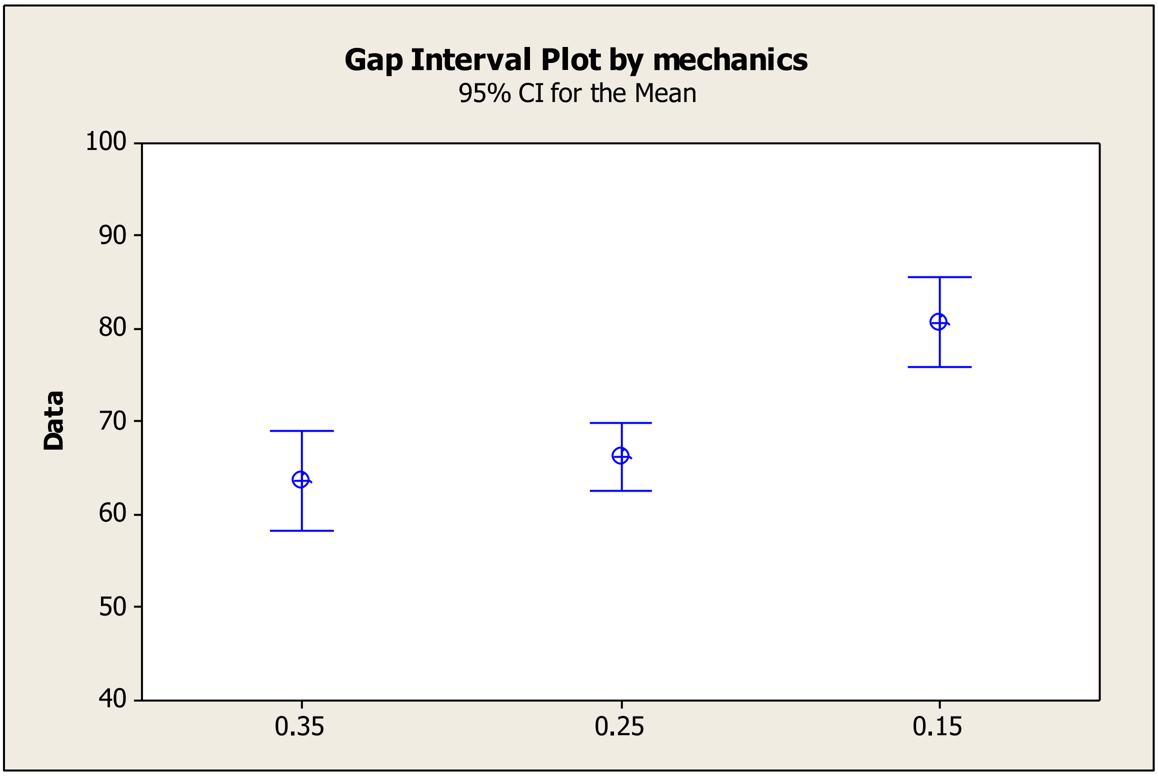


Analyzing gaps

The primary target of the survey was to also include gap analysis in the study. However, when starting the survey its size proved to be too large. Multiple gap sizes were analyzed only by a couple of customers. Gap analysis was still concluded by combining the results from different answerers to see if there is difference between scoring different gap sizes.

Gap evaluation by mechanics experts

The graph below shows how scores from mechanics experts has divided to different gap sizes.



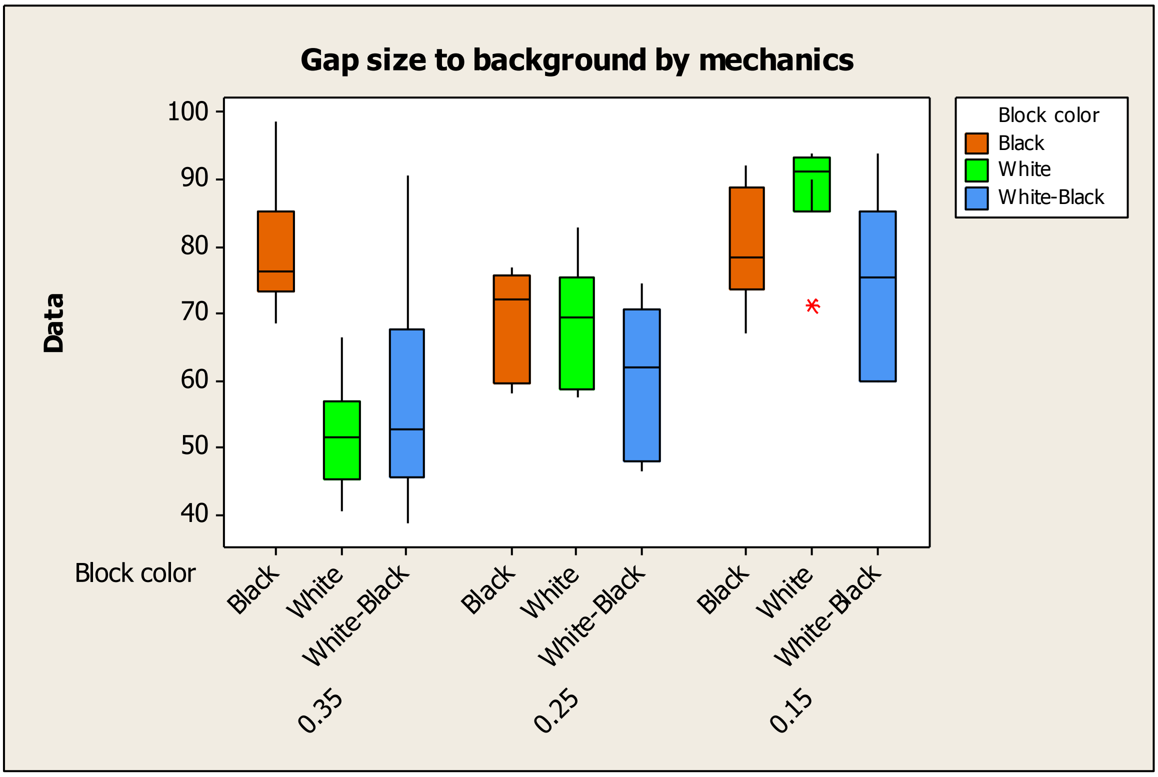
Two-sample T tests were done to compare the gap sizes, to see if there is any real difference.

```
Two-sample T for 0.35 vs 0.25

        N  Mean  StDev  SE Mean
0.35   95  63,7   26,2      2,7
0.25  113  66,2   19,7      1,9


Difference = mu (0.35) - mu (0.25)
Estimate for difference:  -2,60
95% CI for difference:  (-8,88; 3,69)
T-Test of difference = 0 (vs not =): T-Value = -0,81  P-Value = 0,417  DF =
206
Both use Pooled StDev = 22,9068


Two-sample T for 0.25 vs 0.15

        N  Mean  StDev  SE Mean
0.25  113  66,2   19,7      1,9
0.15   75  80,7   20,7      2,4


Difference = mu (0.25) - mu (0.15)
Estimate for difference:  -14,49
95% CI for difference:  (-20,39; -8,58)
T-Test of difference = 0 (vs not =): T-Value = -4,84  P-Value = 0,000  DF =
186
Both use Pooled StDev = 20,0960
```

```
Two-sample T for 0.35 vs 0.15


       N   Mean   StDev   SE Mean
0.35   95   63,7    26,2      2,7
0.15   75   80,7    20,7      2,4


Difference = mu (0.35) - mu (0.15)
Estimate for difference:  -17,08
95% CI for difference:  (-24,39; -9,77)
T-Test of difference = 0 (vs not =): T-Value = -4,61  P-Value = 0,000  DF =
168
Both use Pooled StDev = 23,9725
```

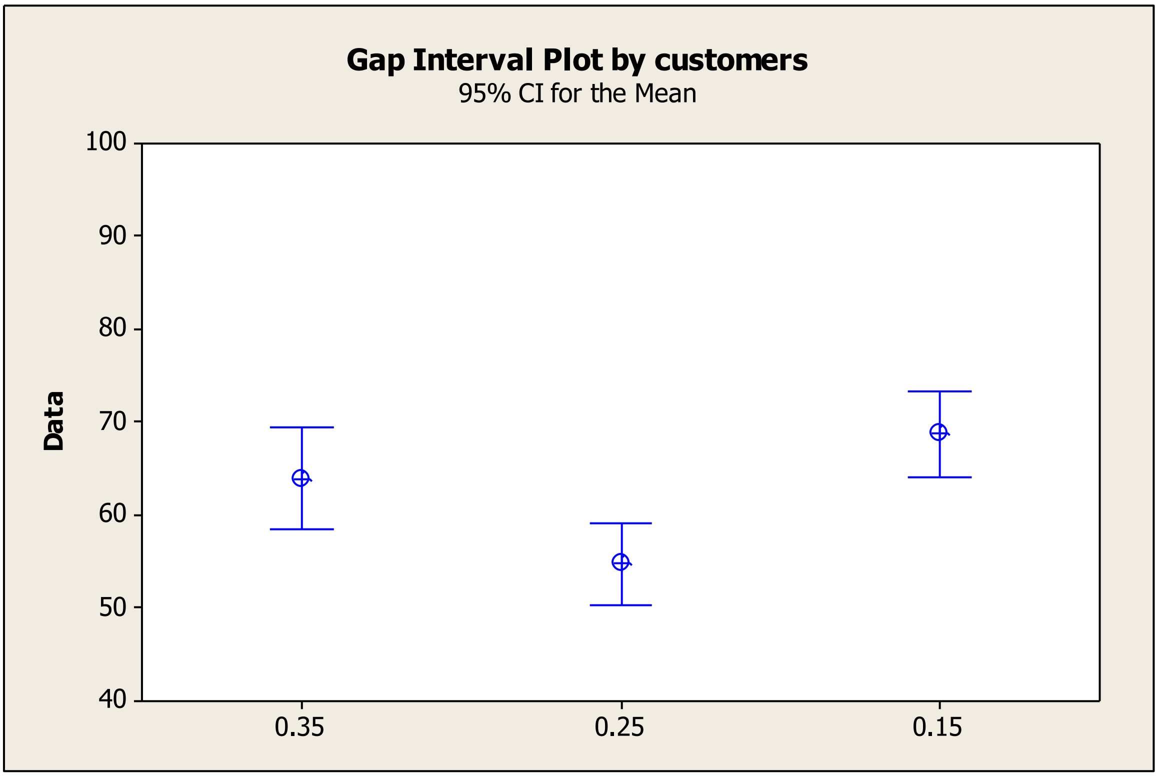
With P-value lower than 0.05, there is statistical difference between the data. So mechanics experts see 0.15mm gap clearly better than the others.

From these results we can see that even though the study wasn't designed with gap evaluation in mind, mechanics experts can tell the difference between different gap sizes. It is good to remember that one answerer only answered to one gap size, so the difference comes from scoring between different answerers.



In this graph it can be seen how scores were divided to different background colors in different gap sizes. Black block color is clearly the best with biggest gap size 0.35.

Gap evaluation by customers



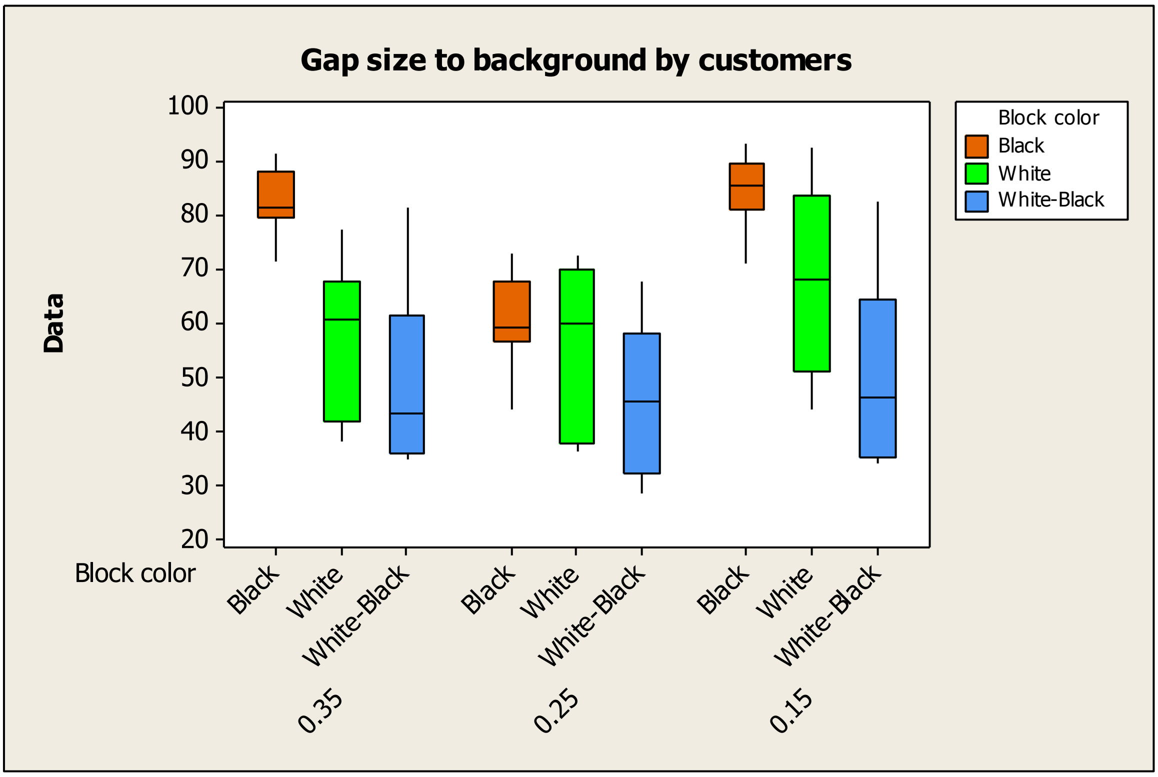
Two-sample T test was done also evaluate how customers saw the difference between gaps

```
Two-sample T for 0.35 vs 0.25

        N  Mean  StDev  SE Mean
0.35  113  63,9   29,3      2,8
0.25  133  54,8   25,6      2,2


Difference = mu (0.35) - mu (0.25)
Estimate for difference:  9,12
95% CI for difference:  (2,22; 16,02)
T-Test of difference = 0 (vs not =): T-Value = 2,60  P-Value = 0,010  DF = 244
Both use Pooled StDev = 27,3855


Two-sample T for 0.25 vs 0.15

        N  Mean  StDev  SE Mean
0.25  133  54,8   25,6      2,2
0.15  133  68,8   27,0      2,3


Difference = mu (0.25) - mu (0.15)
Estimate for difference:  -14,02
95% CI for difference:  (-20,37; -7,66)
T-Test of difference = 0 (vs not =): T-Value = -4,34  P-Value = 0,000  DF =
264
Both use Pooled StDev = 26,3148
```

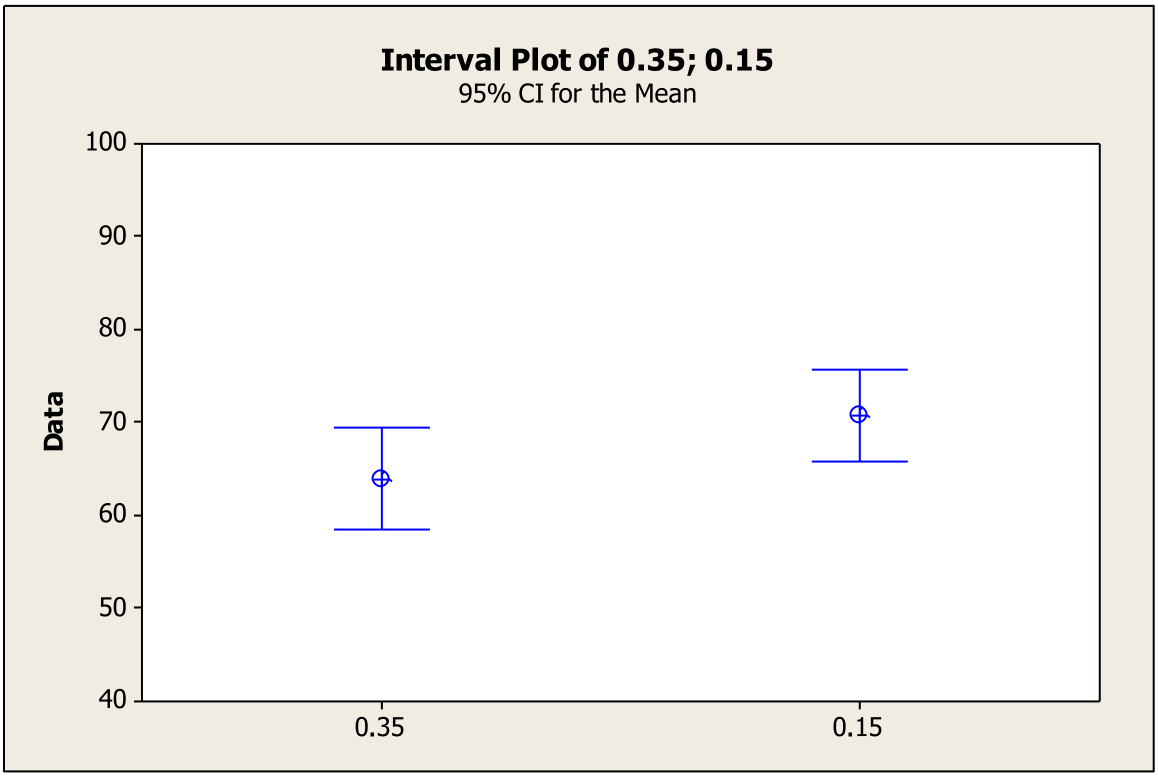
```
Two-sample T for 0.35 vs 0.15

         N  Mean  StDev  SE Mean
0.35   113  63,9   29,3      2,8
0.15   133  68,8   27,0      2,3


Difference = mu (0.35) - mu (0.15)
Estimate for difference:  -4,90
95% CI for difference:  (-11,97; 2,18)
T-Test of difference = 0 (vs not =): T-Value = -1,36  P-Value = 0,174  DF =
244
Both use Pooled StDev = 28,0875
```

With customers there can't be seen similar results than with mechanical experts. There are differences between gap sizes but they do not follow common logic that smaller gap size is better.

The next graph shows how scores were divided to different background colors in different gap sizes by customers. Black block is also clearly the best with biggest 0.35mm gap size.



Gap analysis was done also to compare scores to different gap size by same answerer. Total of 6 customers answered to both 0.35mm and 0.15mm gap size samples.




```
Two-sample T for 0.35 vs 0.15

        N  Mean  StDev  SE Mean
0.35  113  63,9   29,3      2,8
0.15  114  70,7   26,3      2,5


Difference = mu (0.35) - mu (0.15)
Estimate for difference:  -6,84
95% CI for difference:  (-14,12; 0,44)
T-Test of difference = 0 (vs not =): T-Value = -1,85  P-Value = 0,065  DF =
225
Both use Pooled StDev = 27,8369
```

All except one answerer told they can't see any difference between 0.15mm and 0.35mm gaps. It is good to remember that they couldn't evaluate different gap sizes at the same time, there was a couple minutes break between the answers.
Even though the answerers told they couldn't see the difference, there is a minor favor towards smaller 0.15mm gap. This shouldn't still be concluded, because p-value between gaps is bigger than 0.05 at 0.065.

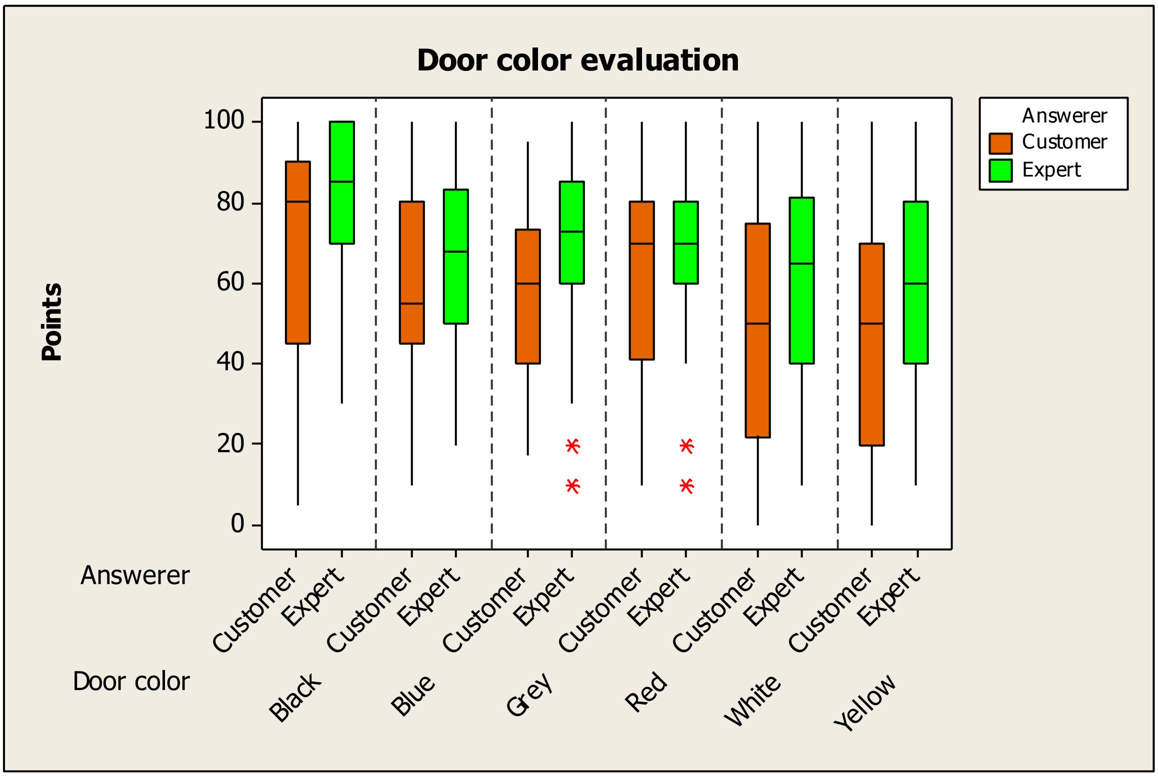
For all gap size evaluations in this study the sample sizes are too low and poorly executed. To overcome this problem the number of block and door color combinations should have been reduced with design of experiments (DOE). Still, the reason for this survey was to analyze the color combinations and it can be executed as planned.

Analyzing colors

This graphs shows how base color affects to mechanics experts and customers.

Black base color is clearly the best for both, customers and experts. It can be said that mechanical design wise black color is the easiest. Variation is big for white block with black background.

This graph shows how answers change depending on door color.



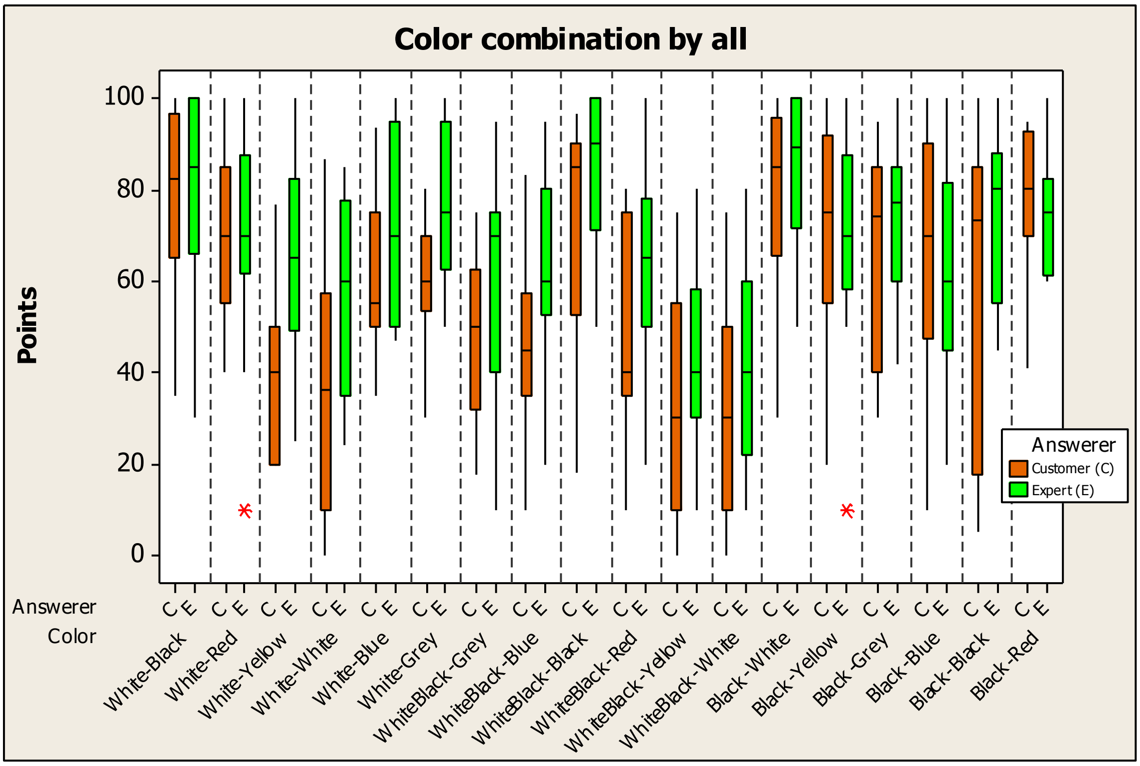
Black color is again seen as the best door color on different colored backgrounds. Light colors like white and yellow are seen as the worst. Again, experts have answered similarly to customers, but they have rated a little higher scores. Also the change between different colors seem to be bigger with customers.

Grouping for color analyzing was done using Bonferroni and Tukey methods. They group the answers based on statistical difference.

```
Means that do not share a letter are significantly different.
Grouping Information Using Bonferroni Method and 95,0% Confidence

Door color   N  Mean  Grouping
Black       78  74,3  A
Red         78  67,3  A
Grey        78  64,2  A B
Blue        78  62,4  A B
White       77  54,5    B
Yellow      78  53,6    B

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95,0% Confidence

Door color   N  Mean  Grouping
Black       78  74,3  A
Red         78  67,3  A
Grey        78  64,2  A B
Blue        78  62,4  A B
White       77  54,5    B
Yellow      78  53,6    B

Means that do not share a letter are significantly different.
```

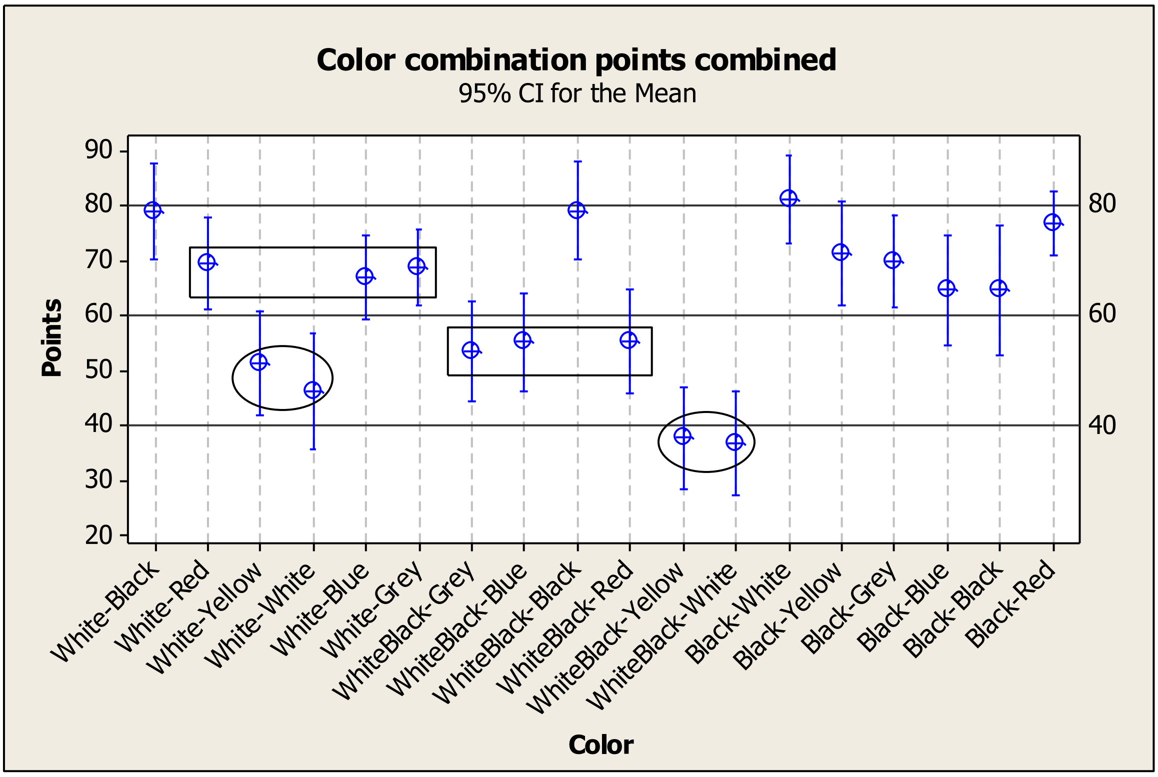
Bonferroni and Tukey method conclude that dark colors like black and red are seen as the best and ligh color such as white and yellow are the worst.

Finally we can see the best color combinations according to mechanics experts and customers.




```
Grouping Information Using Bonferroni Method and 95,0% Confidence

Answerer    N  Mean  Grouping
Expert    245  68,3  A
Customer  247  58,7    B

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95,0% Confidence

Answerer    N  Mean  Grouping
Expert    245  68,3  A
Customer  247  58,7    B

Means that do not share a letter are significantly different.
```

Bonferroni and Tukey method clearly state that there is difference between customer and mechanics expert scores. However when looking at the graph above, the order of color combinations seem similar. Overall mechanical experts have given better scores, but in similar order than customers. Because of this it is sensible to combine the results and look at the color combination results as a whole.



The best combination seems to be white door on a black block. With this combination, the gap is hardest to see. The grade is almost same with inverse situation, where the door is black and block is white. Based on this information, it can be that visual appearance of different types of part interfaces are similar with same color combinations. To prove it there is need for another trial.

Although black and dark colors came up as the best with block and door study, the combination of dark colors are not on top in color combination evaluation. With big contrast between the parts that make the interface, gap disappears more easily.

The scores are similar for white block and white block with black background. WhiteBlack block has lower scores in all except with black door.

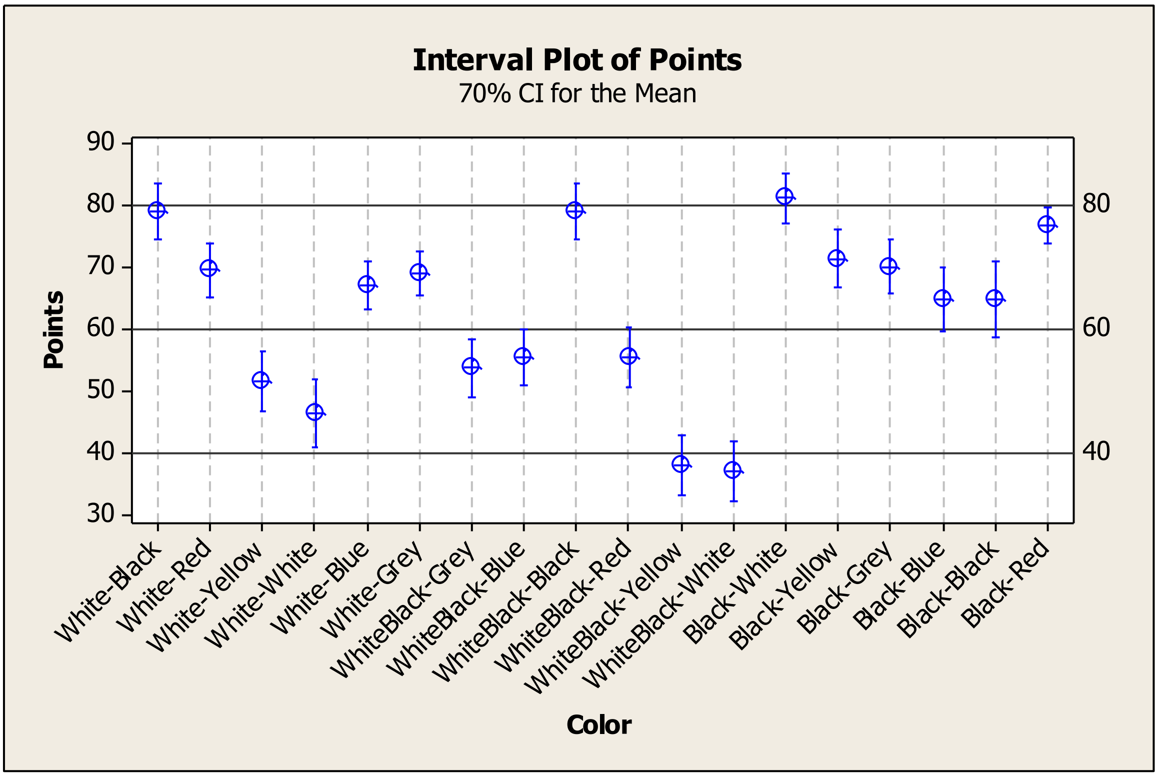
Clearly the worst cases were with yellow and white SIM-door on a white block and black background. In this situation the gap is easiest to see and the interface is visually disturbing.

```
Grouping Information Using Bonferroni Method and 95,0% Confidence

Color               N  Mean  Grouping
Black-White        25  81,4  A
WhiteBlack-Black   26  79,1  A
White-Black        26  79,0  A
Black-Red          26  76,8  A B
Black-Yellow       26  71,4  A B C
Black-Grey         26  70,1  A B C
White-Red          26  69,6  A B C
White-Grey         26  68,9  A B C
White-Blue         26  67,1  A B C D
Black-Black        26  64,8  A B C D
Black-Blue         26  64,8  A B C D
WhiteBlack-Red     26  55,4    B C D E
WhiteBlack-Blue    26  55,3    B C D E
WhiteBlack-Grey    26  53,6      C D E
White-Yellow       26  51,5      C D E
White-White        26  46,3        D E
WhiteBlack-Yellow  26  37,9          E
WhiteBlack-White   26  36,9          E

Means that do not share a letter are significantly different.
```

```
Grouping Information Using Tukey Method and 95,0% Confidence

Color               N  Mean  Grouping
Black-White        25  81,4  A
WhiteBlack-Black   26  79,1  A
White-Black        26  79,0  A
Black-Red          26  76,8  A
Black-Yellow       26  71,4  A B
Black-Grey         26  70,1  A B
White-Red          26  69,6  A B
White-Grey         26  68,9  A B
White-Blue         26  67,1  A B C
Black-Black        26  64,8  A B C
Black-Blue         26  64,8  A B C
WhiteBlack-Red     26  55,4    B C D
WhiteBlack-Blue    26  55,3    B C D
WhiteBlack-Grey    26  53,6    B C D
White-Yellow       26  51,5    B C D
White-White        26  46,3      C D
WhiteBlack-Yellow  26  37,9        D
WhiteBlack-White   26  36,9        D

Means that do not share a letter are significantly different.
```

The result from graph above is confirmed with statistical analysis. It can be seen that Tukey method has ranked black and red color combination as similar to black and white combinations. This is against the assumption that gap disappears with combination of dark and light colors. From the result tables we can see that many of the answerers have ranked red color as their favorite. This might be the reason why black and red combination is liked.

Analysis from color combinations is done with high confidence interval, 95%. Because high confidence interval is used, differences are not seen in many combination cases. When lowering confidence interval to 70%, the differences between means are greater. Result with 70% confidence can be seen in the graph below.


Interval Plot of Points
70% CI for the Mean
Points
90
80
70
60
50
40
30
White-Black
White-Red
White-Yellow
White-White
White-Blue
White-Grey
WhiteBlack-Grey
WhiteBlack-Blue
WhiteBlack-Black
WhiteBlack-Red
WhiteBlack-Yellow
WhiteBlack-White
Black-White
Black-Yellow
Black-Grey
Black-Blue
Black-Black
Black-Red
Color