NTNU

Norwegian University of
Science and Technology

Prototyping for technically trained audiences vs. financially trained audiences

Jonas Neraal Jakobsen

Master of Science in Entrepreneurship
Submission date: July 2017
Supervisor: Roger Sørheim, IØT
Co-supervisor: Martin Steinert, MTP

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

This page is intentionally left blank.

Abstract

Start-ups and new product development projects are fuzziness and ambiguity set in motion. In the chaos of these early phase projects, more and more research is turning activities like prototyping, marketing, funding and production into a science, but there is still much left to chance. This master thesis is an attempt to add to the science of early stage development activities and take them out of the black box, trying to understand how to successfully create communication type prototypes for technically trained audiences vs. financially trained audiences. During the work on this problem, I discovered that a more interesting relationship is between those who possess knowledge about the technical aspects of a prototype, and those who do not. This is due to the fact that one's education and work experience does not provide a complete picture of one's knowledge. Also, in order to understand what determines an evaluation one has to understand what aspects of a prototype affects the judgement, and in what degree. Thus, in addition to background, I wanted to understand how design resolution and function affects the judgement of the audience.

To understand the evaluations of prototypes I designed an experiment using real start-up prototypes: Wiral, a lightweight cable cam system and MovieMask, a face accessory for mobile devices to simulate a cinematic experience. The experiment was used to gather data on the effects of background, design resolution and function on prototype evaluations. The experiment was then run with 18 participants with different professional backgrounds at different locations in Norway and the US. I then performed a statistical analysis on the data registered from the experiments. The experiment setup, participant requirements and time restrictions resulted in a low sample size, reducing the impact of the statistical analysis in this thesis. However, the statistical analysis did reveal some statistically significant results for background, design resolution and function. There was not enough evidence to conclude the effect of these variables, but there was an interesting correlation between the tests showing that the evaluations of MovieMask and Wiral had large differences in distributions, when testing for the effects of background, design resolution and function indicating that the concepts themselves are affected differently. Lastly I present interpretations of results, implications, limitations of the study and a recommendation for further work on the subject.

This page is intentionally left blank.

Sammendrag

Start-ups og prosjekter for ny produktutvikling er usikkerhet satt i bevegelse. I kaoset til disse tidligfaseprosjektene dukker det stadig opp mer og mer forskning som gjør aktiviteter som prototyping, markedsføring, funding og produksjon til vitenskap, men det er fortsatt mye som er styrt av tilfeldigheter. Denne masteroppgaven er et forsøk på å tilføre til forskningen på aktiviteter i tidligfase produktutvikling og forsøke å forstå hvordan man skaper gode prototyper for kommunikasjon for teknisk trent publikum og finansielt trent publikum. I løpet av arbeidet på dette problemet fant jeg at et mer interessant forhold er mellom de som har kunnskap om de tekniske aspektene av en prototype og de som ikke har det. Dette er fordi kunnskap ikke er godt nok definert av utdanning og arbeidserfaring. I tillegg ønsker jeg å forstå hva det er som påvirker en evaluering, og dermed også hvilke aspekter ved prototyper som påvirker denne evalueringen. Derfor har jeg, i tillegg til bakgrunn, valgt å forsøke å forstå hvordan design og funksjon påvirker evalueringen til publikumet.

For å forstå evaluering av prototyper designet jeg et eksperiment som tar i bruk prototyper fra ekte start-ups: Wiral, et lettvekt kabelkamerasystem og MovieMask, et ansiktstilbehør for smarttelefoner som simulerer en kino-lignende opplevelse. Eksperimentet ble brukt for å samle inn data på effekten av bakgrunn, design og funksjon på evaluering av prototyper og ble kjørt med 18 deltakere med forskjellig profesjonell bakgrunn ved forskjellige lokasjoner i Norge og USA. Deretter utførte jeg statistisk analyse på dataen samlet inn fra eksperimentet. Oppsettet til eksperimentet, krav til bakgrunnen til deltakere og begrenset tid resulterte i et lite datagrunnlag, som reduserer styrken til den statistiske analysen i denne oppgaven. Likevel avslørte den statistiske analysen flere statistisk signifikante resultater for bakgrunn, design og funksjon. Det var ikke nok bevis til å konkludere med den nøyaktige effekten av disse variablene, men den viste en interessant korrelasjon mellom testene hvor evalueringen av MovieMask og Wiral hadde store forskjeller i distribusjon når de ble testet for effekten av bakgrunn, design og funksjon. Dette indikerer at konseptene også blir påvirket i forskjellig grad av bakgrunn, design og funksjon. Til slutt legger jeg fram en tolkning av resultatene, implikasjoner, begrensninger og et forslag for videre arbeid på problemet.

This page is intentionally left blank.

Acknowledgements

I would like to thank the companies and people who generously took the time participate in my experiments. A special thanks to my advisors at TrollLabs, Jørgen and Stephanie, for all your help.

I would also like to thank my mother and my girlfriend, Linn for their constant worry and support.

This page is intentionally left blank.

Table of Contents

List of figures

List of tables

This page is intentionally left blank.

1 Introduction and Background

This goal of this thesis is explore how background, design resolution and function impact prototype evaluations. There has been little research into the perceptions of people with different backgrounds when it comes to prototypes and product development. Bryan-Kinns and Hamilton (2002), discuss the importance of considering your audience when building a specific prototype, but does not include what make people evaluate prototypes differently. How evaluation is done is complex and often regarded as a black box to avoid over-complication. To make it more complicated, the factors like design resolution and function make it harder to predict the evaluation of a specific prototype.

The setting of this thesis early stage projects like start-ups and New Product Development (NPD) in companies. These cases are special in that they are often dependent on support from other instances to succeed, or proceed. This means that an essential part of their early work is focused on making *communication* material/prototypes to present to potential stakeholders, to show their progress, skills, competitive advantage and so on. As start-ups and new product development teams juggle between product development activities and other activities like getting strategic partners and financial support, investors attempt to create science out of choosing which projects to support. This thesis is an attempt to provide useful insight to both the product developers and the investors in order to help make good decisions and informed prioritization.

Ulrich and Eppinger (2012) divides the general purpose for prototyping into four categories: Learning, communication, integration and milestones. There are multiple purposes for developing a prototype, but in this thesis I will focus on what Ullrich and Eppinger (2012) describes as communication prototypes, specifically for stakeholders external to the product development team. By eliminating all other information about the case but the product and the prototype, I will attempt to perform statistically significant analyses on the effects of design, functionality and personal background on the subjective concept evaluation. This will hopefully lead to a better understanding of how to communicate correctly to external parties.

To evaluate the impact of these factors I designed an experiment, described in chapter 3, using real cases from the start-up environment surrounding the Norwegian University of Science and Technology (NTNU) and NTNU's School of Entrepreneurship. I then ran the experiment with

18 professionals of different backgrounds at different locations in the US and Norway. The participant recorded their evaluations of prototypes I provided in a questionnaire. I then ran several statistical analyses on the data to investigate the impact of the experiment, explained in chapter 4.

2 Theory

2.1 Decision making and absolute judgement

We are all decision makers. Our lives are formed by the decisions we make, consciously or unconsciously. To help our judgement in these decisions, we gather information to base our decisions on. In a study of decision making Saaty (2008) argues that to make a decision, we need to know the problem, the need and purpose of the decision, the criteria of the decision, their subcriteria, stakeholders and groups affected and the alternative actions to take. We will from this knowledge attempt to make a picture of the problem and determine the best alternative. In the same study, Saaty (2008) also propose that there is two ways to learn about anything; be that an object, a feeling or an idea. Firstly, we can examine or study it to learn about its various properties, synthesize the findings and draw conclusions. The second is to compare it to other similar entities and relate them by making comparisons. The two ways of learning can be related to Blumenthal's (1977) Absolute and Comparative Judgement. Absolute Judgment is, as the first way of learning, based on the relation between a single stimulus and some information held in short-term memory and information gained through previous experience. Comparative Judgement is the identification of some relation between two stimuli both present to the observer. The Absolute Judgement is of interest in the upcoming experiment, as it is what makes up the participant's subjective evaluation. I will attempt to quantify the participants Absolute Judgement through the experiment.

2.2 Backgrounds

As one of the goals of this thesis and experiment is to examine the effect personal background has on a concept evaluation based on a prototype, it is important to understand both what *background* is in this context, and how it could impact the results. First of all, I had to define what backgrounds I was going to use, and how I believed that would affect the results of the experiments. Starting out, I hoped to conduct the experiments with *investors* with technical and non-technical education. An investor in this context will from here on be defined as any person external or internal to a company or start-up, that has the opportunity to either support with, or deny a project the resources controlled by a potentially supportive instance. Externally this can mean people like Venture Capitalists, Business Angels and External Company Managers, who can assist the inventors in some way that help them achieve some strategic or project goal. Internally this would typically be decision makers such as project managers and upper management who can allow a project to continue and provide funds, or discontinue and shut

down. In addition, the experiment participants should have a position where they are typically presented with prototypes or ideas of concepts and be in a position to affect the progress of the concept in a large degree.

2.3 New Product Development

This thesis is set to early stage product development scenarios, often described as New Product Development (NPD). NPD is a well discussed topic and one model for NPD that was created by Booz et al. (1982) is presented here. The model contains 7 stages of product development and are shown in figure 2.2.

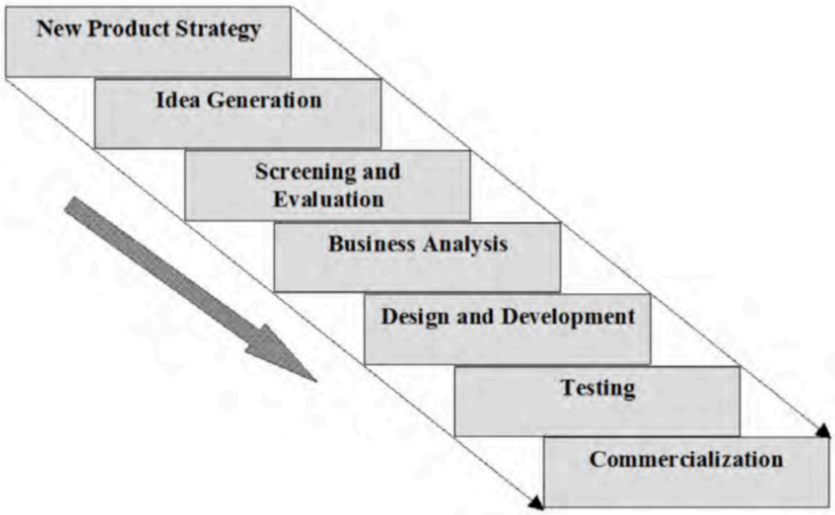


FIGURE 1 STAGES OF NPD (BOOZ ET AL. 1982)

These stages describe a typical process of product development seen from a management perspective. However, the model is heavily focused on internal actions, excluding important activities such as stakeholder communication. Stakeholder activities are often excluded from product development models as they are viewed as mainly financial activities, performed to acquire resources from an instance such as a Venture Capitalist firm or a company manager and are not always required. These activities, however, are critical for NPD projects, such as start-ups, in that they often cannot complete the development process and commercialization without an influx of both financial resources and outside competencies. It is therefore important to understand what criteria stakeholders and investors affect their evaluation and decision.

When trying to understand what criteria the investors utilize to make up an opinion, one has to take into consideration that concept evaluation is made up of an accumulation of factors that affects the final opinion. In a normal venture capitalist concept pitch situation, the start-up will present ideas, team, prototypes, IP, strategy, financials, competencies and market numbers in order to persuade the investor to support the project. The investors final decision is thus

dependent on multiple criteria. Tyebjee and Bruno (1984) collected existing literature on venture evaluation criteria into table 1 shown below. It describes the diversity of factors that make up the evaluation of venture capital firms. Factors such as management commitment/quality/skills, market size, product, financial history, marketing plans and rate of return rank among the most important when making a funding decision. These are carefully designed evaluation criteria that the venture capitalists utilize to reduce their risk of funding an unfit venture.

Venture Evaluation Criteria

WELLS (1974) Sample: Eight Venture Capital Firms (Personally interviewed)		POINDEXTER(1976) Sample: 97 Venture Capital Firms (Mailed questionnaires)	TYEBJEE & BRUNO (Study I) Sample: 46 Venture Capitalists (Telephone survey, open-ended questions)	
Factor	Average Weight	Investment Criteria by Rank Order Of Importance	Factor	% of Respondents Mentioning
Management Commitment	10.0	1. Quality of Management	1. Management Skills & History	89
Product	8.8	2. Expected Rate of Return	2. Market Size/Growth	50
Market	8.3	3. Expected Risk	3. Rate of Return	46
Marketing Skill	8.2	4. Percentage Equity Share of Venture	4. Market Niche/Position	20
Engineering Skill	7.4	5. Management Stake in Firm	5. Financial History	11
Marketing Plan	7.2	6. Financial Provisions for Investor Rights	6. Venture Location	11
Financial Skill	6.4	7. Venture Development Stage	7. Growth Potential	11
Manufacturing Skill	6.2	8. Restrictive Covenants	8. Barriers to Entry	11
References	5.9	9. Interest or Dividend Rate	9. Size of Investment	9
Other Participants in Deal	5.0	10. Present Capitalization	10. Market/Industry Expertise	7
Industry/Technology	4.2	11. Investor Control	11. Venture Stage	4
Cash-Out Method	2.3	12. Tax Shelter Consideration	12. Stake of Entrepreneur	4

TABLE 1 – COLLECTION OF VENTURE EVALUATION CRITERIA (TYEBJEE AND BRUNO 1984).

In addition to the multiple evaluation criteria, there lies a more fundamental judgement in the less conscious sense. The evaluation criteria described in the table by Tyebjee and Bruno (1984) can be considered objective evaluation criteria, in the sense that these criteria should be applied to any venture presented to the venture capital firm. Meanwhile, there is also a subjective evaluation performed by any of the persons being presented to, based on the person's cognitive capabilities and existing knowledge, namely the Blumenthal's (1977) Absolute Judgement. This subjective evaluation will depend on previous experiences and knowledge, personal competencies, mood and other day to day impressions. Understanding this subjective evaluation based on experience is what I will attempt to shed light on in this thesis.

2.4 Prototyping

Houde and Hill (1997) define the prototype as any representation of a design idea, regardless of medium. Ulrich and Eppinger (2012) claim that prototypes serve four purposes in new product development: Learning, communication, integration and milestones. Different types of prototypes are necessary in the different stages of a NPD process. It is therefore important to gain an understanding of various types of prototypes and how to take advantage of them. In this setting, communication type prototypes are in focus, trying to understand how to optimize these prototypes for different audiences. These prototypes are used to engage external parties to convey a message. Depending on the message and the audience, the designers must figure out to present their ideas and wishes through prototypes.

Houde and Hill (1997) has created a model showing the dimensions of what a prototype prototypes. The purpose of the model is to make it easier to develop and subsequently communicate about this kind prototyping strategy. The three dimensions are shown in Fig. 2.4. The dimensions are *role; look and feel;* and *implementation*.

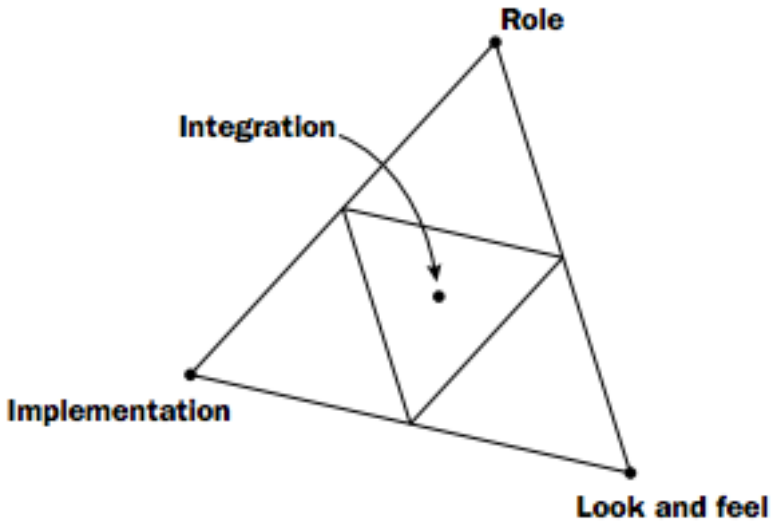


FIGURE 2 DIMENSIONS OF PROTOTYPES (HOUDE AND HILL 1997)

Role refers to the role the product should serve in the user's life. This requires a setting and a clear understanding of what the product is meant to represent for the user. This dimension is important in that it requires the designer to understand the user's need. To communicate accurately, the designer must attempt to understand how the user will react to a prototype and its features.

The *look and feel* is the dimension questioning the actual experience of using the product. When focusing on this dimension, the designer must build a prototype which simulates the user experience. This dimension is heavily related to the design resolution of a prototype. To communicate the wanted impression of a prototype, the designer must consider what fea8tures a final product is intended to include to create the desired value. This thesis will try to create an understanding of how the user receives and interprets the message. What feelings does the designer want the product to produce in a user, and how does he communicate that in a prototype?

Implementation refers to the mechanisms that allow the product to perform its function. To prototype this, a more or less completely functioning model must be built, which is often resource demanding. This dimension should display the intended function of a prototype and is therefore a relevant dimension in this setting. According to Houde and Hill (1997), implementation prototypes are often used by the design team to demonstrate to their organization the technical feasibility of the artefact. Again, I will attempt to look at the investors perspective, and try to understand how they perceive the implementation dimension of the prototype.

Houde and Hill (1997) also identify a fourth prototype; an *integration* prototype, which is a combination of all three dimensions. Integration prototypes are built to represent the complete user experience as an artefact. This fourth prototype is important to balance and resolve constraints across the design dimensions. This also verifies that the design is complete and coherent. The prototypes used for the experiment can all be regarded as integration prototypes, as they are built to simulate the complete user experience, not only one of the previous three dimensions. In many ways, this thesis will discuss the balance in these integration type prototypes and which of the dimensions cause what reactions in the audience.

2.5 Function

Roozenburg and Eekels (1995) describes function of a product as the intended and deliberately caused ability to bring about a transformation of a part of the environment of the product. As such, function is referred to as the intended function of the prototype, i.e. the function(s) the inventors meant for the final product to have (e.g. a clock is made to be able to discern the time of day). There might be several key functions assembled for the prototype to actually function

as intended (e.g. the battery, mechanics, buttons and so on) but the function is regarded as the effect that is produced of the combined efforts of the sub-functions.

According to Roozenburg and Eekels (1995) a product usually has several properties and each property, or group of properties, represents the possibility to function, but argues that we usually only notice a few. This is due to the fact that products must be used in a specific manner to produce a certain function. This means that even though the designer may have incorporated many functions into a product, the functions stays a "hypothetical statement" until the we bring the product into certain conditions. The product does only behave as intended if it is used as the designer prescribed. The instructions for use are not given facts for the designer, like the function, but are thought up – together with the form of the product – and thus form an essential part of the design (Roozenburg and Eekels 1995). This presents a major issue in design; will the audience be able to see the same values in a product as the designers, and how do one communicate these values?

A design process is usually initiated by the realization of a problem, or, in other words, the realization of a missing function. Hence, the design specification is the detailed list of the properties needed to realize the intended function of the product. Roozenburg and Eekels (1995) argues that by fulfilling functions a product satisfies needs, and this gives people the possibility to realize one or more values. Put in the context of Houde and Hill (1997), the function greatly affects the *Role* dimension. In the experiment presented in this thesis I will be comparing the effect on participant's evaluation of prototypes with and without function. I will also attempt to see if there is a difference in effect between participants with and without a technical understanding of the concept.

2.6 Design Resolution

Wiklund et al. (1992) examined the relationship between what Virzi et al. (1996) describe as the aesthetic refinement (which is referred to as design resolution in this thesis) of prototypes and perceived usability of a prototype. The aesthetic refinement are aspects of the product the does not directly influence functionality (i.e., graphic design and color). Wiklund et al. (1992) found that the aesthetic refinement did not affect the rating of scales such as ease of use and ease of learning. They argued that prototype fidelity does not affect the kinds of problems one can detect. The prototypes used in their experiments were digital tools on a computer of varying

closeness to the actual product in refinement. Although using physical prototypes for the experiment in this master thesis, one could expect to see similarities in the results. Another experiment by Walker et al. (2002) also found little difference in the number, type and severity of usability issues found with different prototype fidelities (Fidelity describes how easily prototypes can be distinguished from the final product).

The result of a design process is what we commonly regard as the design of a product. As the design process is in reality an iterative process, described in the basic design cycle (figure 3), a product will very likely have taken several forms before ending up with a final design, hopefully satisfying not only the designer's assumptions, but also the user's needs. These several designs will, if the design process is successful, get closer and closer to a product that possess the features intended for the final product. These iterations that gradually approach the final design can be viewed as different resolutions of one product, the resolution increasing the closer to a commercialized product it gets. Design is inevitably one of the most important factors that affect the impression of a prototype. Design comprise both the aesthetic and functional, making up the first impression as well as affecting the experience as a whole. In this thesis I will be referring to the design resolution of the prototypes as the degree of completeness. Meaning high resolution prototypes should display better aesthetics and functional properties than that of a low resolution prototype.

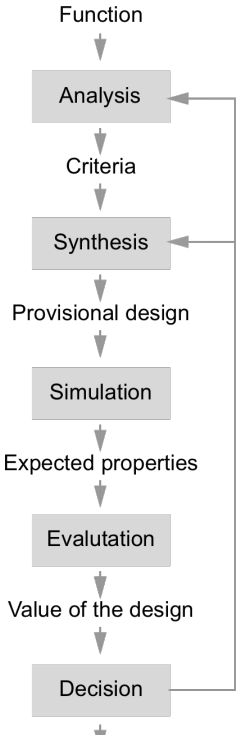


FIGURE 3 THE BASIC DESIGN CYCLE (ROOZENBERG AND EEKELS 1995)

Figure 4-7 shows prototypes of a piece from a boardgame using magnets with respectivly high and low resolution. The high resolution piece is made in cast plastic made for mass production. The low resolution piece is made of laser cut Medium Density Fiber (MDF) to simulate a lower refinement. The prototypes clearly shows a difference in refinement, though their function is close to identical.

Figure 4 High Resolution Top - Cast in Plastic

Figure 5 High Resolution Bottom - Cast in Plastic

Figure 6 Low Resolution Top - Laser cut MDF

Figure 7 Low Resolution Bottom - Laser cut MDF

In the experiment I will be comparing the effect on participant's evaluation of prototypes with low or high design resolution. I will also attempt to see if there is a difference in effect between participants with and without technical understanding of the concept.

3 Experiment

To understand the evaluation criteria of people with different backgrounds, I decided to set up a research experiment. The experiment was specifically created to explore the effects of exposing participants of different backgrounds to varying degrees of design resolution and function. The following section will explain in detail how the experiment was set up and conducted in order to find answers to the hypothesis.

3.1 Hypotheses

The method used to understand how investors evaluate prototypes requires several considerations and factors. I want to alter background, design resolution and function as stimuli to explore the evaluation criteria of investors. With the three independent variables being controlled in the experiment, design resolution, function and background, three hypotheses are formulated: Background Hypothesis, Design Resolution Evaluation Hypothesis and Function Evaluation Hypothesis.

3.1.1 Background Hypothesis

As described in the introduction, I set out to understand how people of different backgrounds perceive prototypes, and how it affects their evaluation. This implicates that there is a form of difference in the evaluations of people of different background.

Thus, the null hypothesis is stated as follows:
Participants of different backgrounds will consistently evaluate prototypes equally.
With the corresponding alternative hypothesis stated as follows:
Participants of different backgrounds will consistently evaluate prototypes differently.

This hypothesis is based on the argumentation that one's background and previous knowledge make up a substantial part of one's perception, evaluation and decision making (Saaty 2008; Blumenthal 1977). More knowledge and relevant background gives an initial base for evaluation and could affect how they perceive both design resolution and function.

3.1.2 Design Resolution Evaluation Hypothesis

As a manipulated factor for understanding evaluation criteria, Design Resolution is believed to create some difference in the participant's evaluation.

This means that the null hypothesis is stated as such:

Participants will evaluate the prototypes of higher design resolution with equal scores as the prototypes with lower design resolution.

The corresponding alternative hypothesis is stated as follows:

Participants will evaluate the prototypes of higher design resolution with different scores than the prototypes with lower design resolution.

This hypothesis is based on the argumentation that design resolution affects the impression of a product and will thus affect the evaluation of the product. A higher design resolution is expected to yield higher evaluation scores.

3.1.3 Function Evaluation Hypothesis

The other manipulated factor, Function is believed to create a reaction with the participants and their evaluations.

The null hypothesis is stated as follows:

After seeing the function of the prototype, the participants will not change their scores in the re-evaluation.

The corresponding alternative hypothesis being:

After seeing the function of the prototype, the participants will change their scores in the re-evaluation.

This hypothesis is based on the argumentation that function is an important part of a product, and not seeing the function versus seeing the function will produce different results.

3.2 Experimental Setup

To get a clear understanding of how people perceive prototypes, I decided to use prototypes from existing start-ups in the university environment to conduct a controlled experiment. The experiment used a total of 4 prototypes from 2 concepts.

I wanted to see how people of different backgrounds evaluate prototypes of different resolution and function. Thus, the experiment was setup to allow participants to see and interact with prototypes, then fill out their evaluations in a comprehensive questionnaire. To increase the amount of data points, each participant was issued two prototypes; one from each case. In addition, to see the effect of seeing a prototype with and without function, each of the two prototype evaluation processes were divided into two: evaluation of prototype without function and re-evaluation of the same prototype with function. Each questionnaire was ended with a demographic survey, providing general demographics, and information about educational and professional background.

The questionnaire was designed to fit most backgrounds, meaning the questions generally required no prerequisite knowledge. Before the evaluation part, the questionnaire contained instructions and a description containing the bare essentials about the concepts. The questions regarded the technical aspects of the prototypes to capture the participants feeling about the design and function, as well as questions about the respective markets to understand the overall impression of the concept.

3.3 Evaluation Cases

As mentioned, the prototypes used in the experiments are all from actual start-ups, originating from NTNU. The start-ups have achieved various degrees of success, but are both more than one year old, and have, or is about to launch Kickstarter campaigns.

3.3.1 Wiral

Wiral is a light weight cable cam system created to allow for more creative filming in challenging environments. Wiral propels a camera down a wire using an electric motor and a remote control. The concept is created to be ideal for filming in challenging areas like forests, indoors or in crowded areas, and is lighter, cheaper and easier to set up than its competitors.

FIGURE 8 WIRAL CONCEPT

The two prototypes of Wiral used for the experiment both hold the intended function of the inventors in that they can both be propelled over a wire by a motor with a remote control. To allow the participants to see both the prototype with and without function, they were first allowed to interact with the prototype only without placing it on a wire. After they evaluated the prototype, I placed the prototype on the wire and used the remote control to propel it across the wire.

The high resolution prototype body is 3D-printet plastic. The motor is a brushless motor situated in the top wheel. The electronics can be found inside the body in a separate container. The battery is revealed by opening the lid of the body. The prototype has no ON/OFF button.
The low resolution prototype body is also 3D-printet plastic but in a lower resolution. The motor is situated beneath the drive wheel and connected by a rubber belt. The mechanics are found behind the battery in the body. The prototype has a simple ON/OFF button.

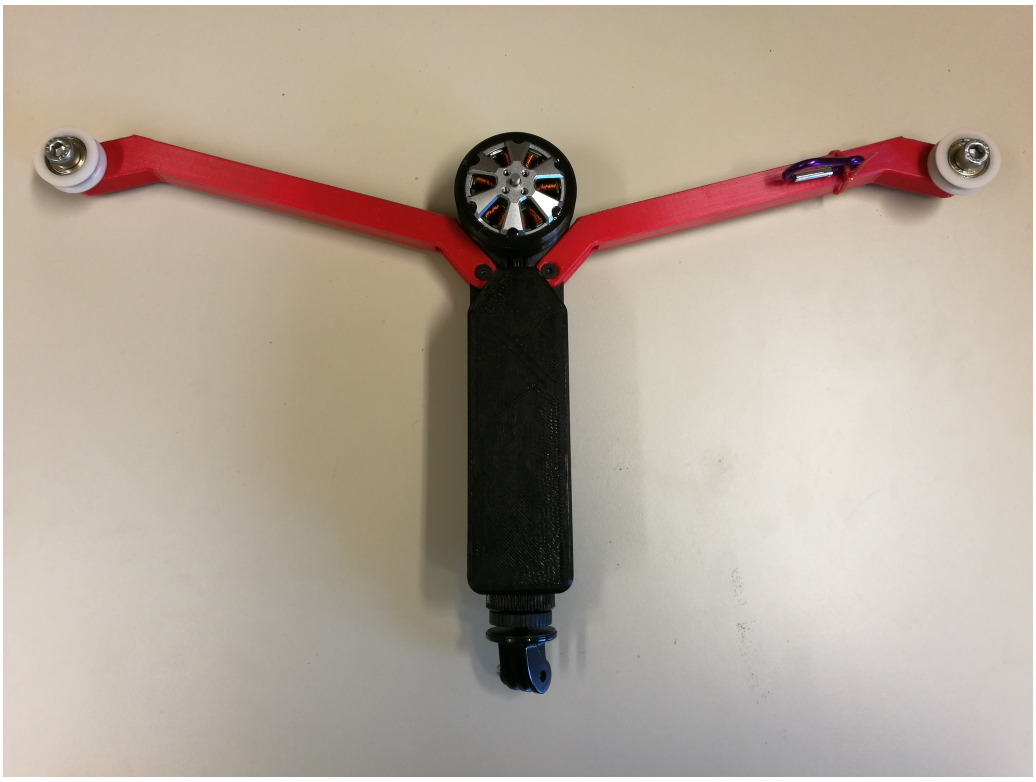


FIGURE 9 HIGH RESOLUTION WIRAL FRONT

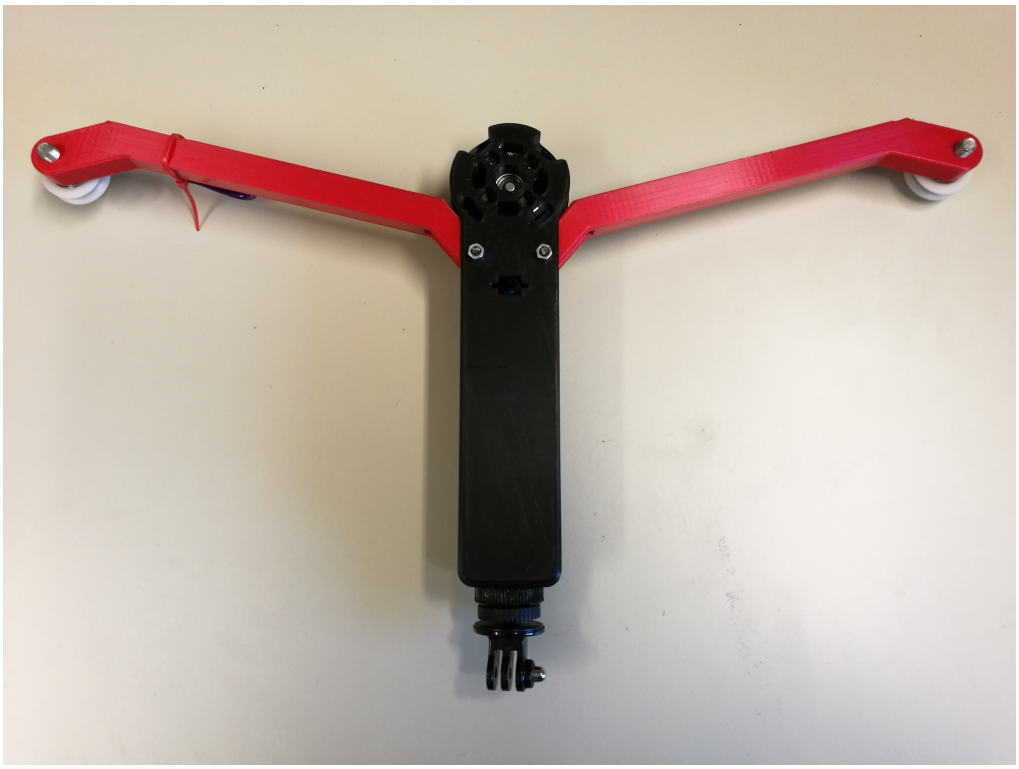


FIGURE 10 HIGH RESOLUTION WIRAL BACK

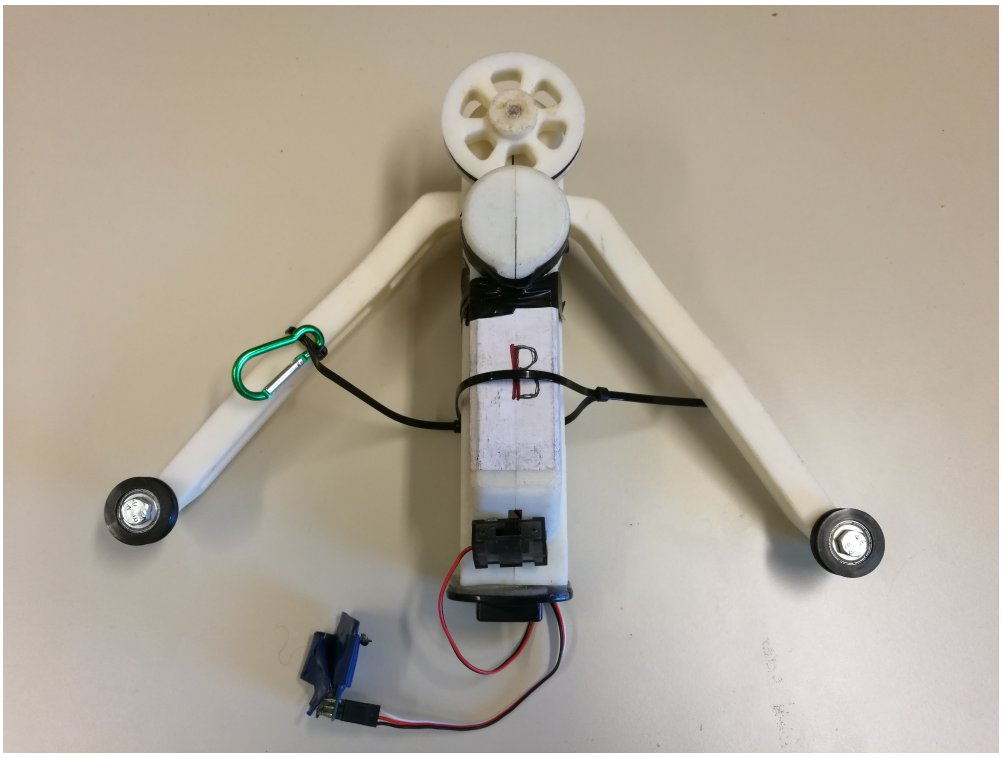


Figure 11 Low Resolution Wiral

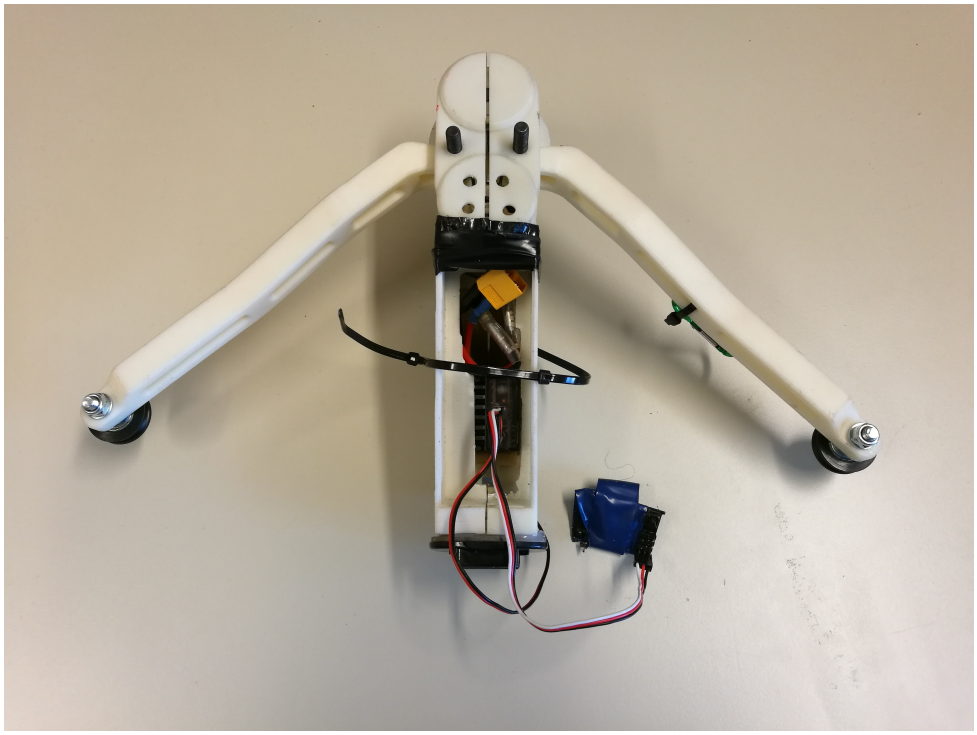


Figure 12 Low Resolution Wiral Back

3.3.2 MovieMask

MovieMask is a set of head mounted glasses using a smartphone to view content similar to VR glasses. The lenses and case are made to simulate a cinematic experience from a smartphone video, rather than to view VR content.

FIGURE 13 WIRAL CONCEPT

The two prototypes used in the experiment both hold the intended function. To allow participants to view both prototypes with and without function, they were first allowed to interact with the prototype without inserting a phone into the case. After evaluating the prototype, I inserted my phone with one specific movie playing from the start.
The high resolution prototype has molded plastic body, covered in high quality fabric. The head strap has a strap around your head as well as an overhead strap. For the head mount, the plastic is covered with a foam cushioning and soft fabric for the nose. The prototype is closed with a zipper and the mounting mechanism consists of two elastic bands with mounts on the ends.

The low resolution prototype has a 3D-printed plastic body covered in tape. The head straps around your head. The head mount has simple foam cushioning. The prototype is closed with a zipper and the mount consists of a spring loaded plastic mount.

FIGURE 14 HIGH RESOLUTION MOVIEMASK

FIGURE 15 LOW RESOLUTION MOVIEMASK

3.4 Tools, Equipment and Materials

The experiment was setup at different locations, using different rooms, to facilitate for the participant's location. All rooms were equipped with a table, a chair and a wire for Wiral to travel on. To minimize outside input and biases, participants performed the experiment individually. There was also a facilitator in the room with the participants, providing instructions, prototypes and display of function.

Figure 16 General Experiment Setup. Paricipant was seated in the closest chair.

Each participant received a consent form, a questionnaire and a pen. The participants were exposed to two prototypes from two separate cases and were allowed to interact with them during the experiment. All participants were also shown the function of each prototype.

3.5 Questionnaire

The questionnaire consisted of 5 parts:

1. Evaluation of Wiral
2. Re-evaluation of Wiral
3. Evaluation of MovieMask
4. Re-evaluation of MovieMask
5. Demographics

Parts 1 and 2 were introduced with a short instruction of the experiment procedure. The questionnaire also provided a bare essential description of the concept, to ensure understanding of the concept while keeping unnecessary information at a minimum. Each of the 4 evaluation parts consisted of 2 parts, one concerning the technical issues of the prototype, and one about the market. This was to get a complete image of their opinion and their knowledge areas.
The questionnaire was designed to capture the unbiased opinion of each participant. It contained both scales, where the participants were asked to evaluate the concept based on a statement and a provided scale, yes and no questions, and text based questions to allow the participants to fill in their opinions to a larger extent than in the scaled parts. Because I wanted to capture the essence of each participant's opinion, the questions were largely based on their subjective opinion, rather than a qualified answers based on hard facts. By choosing participants that were less likely to know the concept beforehand and providing only essential information about each concept, I wanted to limit their ability to perform a biased analysis of the concept and base their answers on the prototype presented and their personal judgment.
A complete questionnaire can be found in appendix 1.

3.6 Experiment Procedure

Participants were greeted and introduced to the room where the experiment was to be held. When seated they received a short description of the procedure of the experiment, asked to sign a consent form and asked to turn their phones silent. The participants were then given the questionnaire form and given the first prototype. To reduce the learning effect in the experiment, the order and type of prototypes were randomized. As there were 2 prototypes from each concept, with 1 of each resolution there was a total of 8 orders/groups:
High Resolution = HR, Low Resolution = LR

1. HR Wiral – HR MovieMask
2. HR Wiral – LR MovieMask
3. HR MovieMask – HR Wiral
4. LR MovieMask – HR Wiral
5. LR Wiral – HR MovieMask
6. LR Wiral – LR MovieMask
7. HR MovieMask – LR Wiral
8. LR MovieMask – LR Wiral

Participants were asked to notify the facilitator between each part in the questionnaire (Part 1 through 4). After part 1 and/or 3, participants were shown the function of the prototypes. When

the participant had seen the function, he/she was asked to re-evaluate the prototype in part 2 or 4. After finishing the re-evaluation, the previous prototype was removed and the participants were given a prototype from the other concept. The experiment was concluded when participants finished part 5, the demographic part of the questionnaire. The whole experiment took from 20-30 minutes for each participant.

3.6.1 Wiral function

Wiral's function was shown by the facilitator placing the prototype on a wire, already strung inside the experiment room, as shown in figure 17. The facilitator controlled the prototype with a remote, bringing it 4 times back and forth across the wire, to a halt at the center of the wire. The prototype was then left on the wire for the participants to view for the re-evaluation of the prototype.

FIGURE 17 WIRAL ON WIRE TO DISPLAY FUNCTION

3.6.2 MovieMask function

MovieMask's function was shown by the facilitator placing a Huawei P9 into the prototype with a specific movie, playing nature scenes without sound. The movie was 53 seconds long and participants were allowed to view till the end if they wished. After the video was done, the prototype was left on the table for the participants to view for the re-evaluation of the prototype.

3.7 Variables for Evaluation of Hypotheses

3.7.1 Independent variables

For the following statistical analyses I used three Independent variables: Design Resolution, Function and Background.

Design Resolution was divided into groups based on the prototypes they were given, either high or low resolution. The test was run for Wiral and MovieMask separately, to remove the issue of one participant being in both groups as participants received one prototype of each concept in randomized order. This meant that each groups had 8-10 participants.

Function was divided into groups based on which part of the experiment their evaluations were given, namely before and after being shown the function. All participants saw two prototypes without, and then with, function, meaning this test ran with all 18 participants in both groups.

Background was divided into groups based on the participants' backgrounds. Background was evaluated on four different levels: Education, Work experience, Claimed Knowledge and Combined Knowledge. As with Design Resolution, Background had to be separated based on the individual participant (as one participant cannot have more than one background). Thus, each group had 8-10 participants.

3.7.2 Dependent variables

The dependent variables in these analyses were the individual evaluation points in the questionnaire. The variables were mostly ordinal values from the evaluations using the Likert scale. There were 10 such ordinal variables and 4 continuous variables from evaluations of time to market, product cost, production cost and size of market. The continuous variables had many missing values, as many participants chose to refrain from answering. Also, most of these variables failed tests of normality. The few data points and non-normal distribution meant the analyses of these variables would be less powerful, and thus I decided to treat the continuous variables as additional assessment information, like the textual answers in the questionnaire.

Disagree ☐ *Somewhat disagree* ☐ *Neutral* ☐ *Somewhat agree* ☐ *Agree* ☐

A1. The inventors have identified and addressed the technical issues of the problem.
Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A2. The concept is likely to work as described above.
Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A3. The prototype is well made.
Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A4. You do **not** like the design of the prototype.
Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

FIGURE 18 EXCERPT FROM THE QUESTIONNAIRE USING LIKERT SCALE

The dependent variables will be referred to as their respective label in the questionnaire (A1, A2, A3…. A8 and so on). The variables will also have some identifiers providing information of what group the prototype belongs to: W = Wiral prototype, M = MovieMask prototype and R = Re-evaluation).

This page is intentionally left blank.

4 Results

The experiment and questionnaire was designed to collect data for statistical analyses. The data was analyzed using SPSS Statistics™ by IBM. Eighteen participants were recruited to understand the effects of design resolution, function and background on prototype evaluation and the whole population was used for the analyses. In the following chapter I will present the tests used and results of the statistical analyses.

4.1 Statistical Tests

4.1.1 Mann-Whitney U Test

Mann-Whitney U Test, which is commonly regarded as a nonparametric alternative to independent-samples t-test. The test allows comparison between independent groups without normal distribution and ordinal variables (the Likert scale used for the evaluation provides ordinal variables that are not necessarily normally distributed). The test ranks the individual evaluations, independent of the group it is in, and then compares the ranks. To run Mann-Whitney U Test successfully 4 assumptions must be met. The first is that you only have one dependent variable measured on the continuous or ordinal level. The second requires that you have one independent variable with two categorical, independent groups. Third, the test requires independence of observations, meaning there should be no relationship between the observations in each group of the independent variable. The final assumption is that one must determine whether the distribution of scores for both groups of the independent variable have the same shape. This is assessed through vis8ual inspection.

4.1.2 Wilcoxon Signed-Rank Test

Wilcoxon Signed-Rank Test, commonly regarded as a nonparametric equivalent to paired samples t-test. Like the Mann-Whitney U test, Wilcoxon signed-rank test allows comparison of non-normally distributed and ordinal data. The test is used to determine whether there is a median difference between paired or matched observations, e.g. participants tested on two occasions or under two different conditions. Wilcoxon signed-rank test has three assumptions for successful testing. The first is identical to the first assumptions for Mann-Whitney U test: continuous or ordinal dependent variable. The second is that you should have one independent variable with two categorical related groups or matched pairs. The final assumption requires that the distribution of the differences between the two related groups to be symmetrical in shape, which is also assessed through visual inspection.

4.1.3 Statistical Significance

The statistical significance of the tests performed in the following section is dependent on the reported significance level (Sig.). The tests used report two values of significance: Exact and Asymptotic significance level. The exact p-value does not correct for ties in the data (i.e., when two or more participants have identical values on the dependent variable). Obviously, this happens more frequently with ordinal dependent variables with few possible values. This means the exact p-value will be inflated and the asymptotic significance level is a better measurement of significance. Thus, all significance levels reported in the analyses are asymptotic. The significance level for all the tests are $p < .05$.

4.2 Test Result Walkthrough

4.2.1 Design Resolution

To compare design resolution evaluations of low and high resolution prototypes, I used Mann-Whitney U Test. The Design Resolution hypothesis, stated previously, was as follows:

Null hypothesis:

H_0: Participants will evaluate the prototypes of higher design resolution with equal scores as the prototypes with lower design resolution.

The corresponding alternative hypothesis is stated as follows:

H_A: Participants will evaluate the prototypes of higher design resolution with different scores than the prototypes with lower design resolution.

The corresponding hypothesis in terms of the Mann-Whitney U Test is as follows:

Null hypothesis:

H_0: the distribution of scores for the two groups are equal

Alternative hypothesis is:

H_A: the distribution of scores for the two groups are not equal

However, another way to express the alternative hypothesis is as follows:

H_A: the mean ranks of the two groups are not equal

4.2.1.1 Wiral results

A Mann-Whitney U test was run to determine if there were differences in evaluation score between high and low resolution prototypes of Wiral for all ordinal dependent variables (A1-8 & B5-6 for both initial evaluation and re-evaluation). For all but A5, RA1 and RA2, distributions of the evaluation scores for high and low resolution were not similar, as assessed by visual inspection. Only B5 had evaluation scores for high resolution (mean rank = 11.55) that were statistically significantly higher than for low resolution (mean rank = 6.94), U = 60.500, p = .050.

Figure 19-29 displays the distribution of evaluations for the individual questionnaire statements.

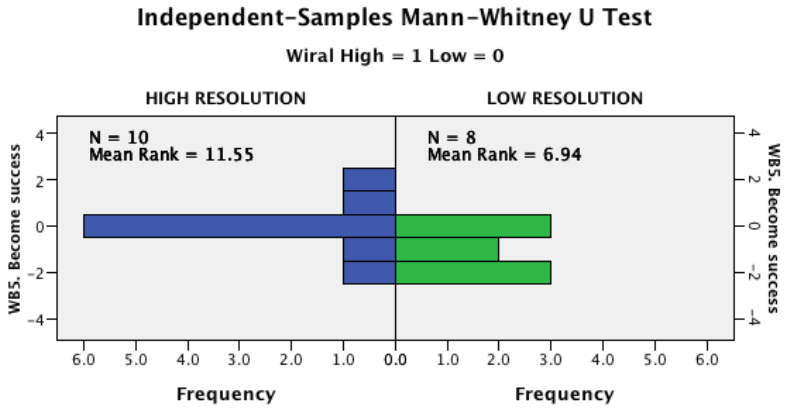


FIGURE 19 EVALUATION STATEMENT B5: YOU BELIEVE THE PRODUCT WILL BECOME A SUCCESS.

For A5, RA1 and RA2 distributions of the evaluation scores for high and low resolution were similar, as assessed by visual inspection. Median evaluation score for high resolution (respectively -1.50, 1.00 and 2.00) and low resolution (respectively -2.00, 0.50 and 1.00) was not statistically significantly different.

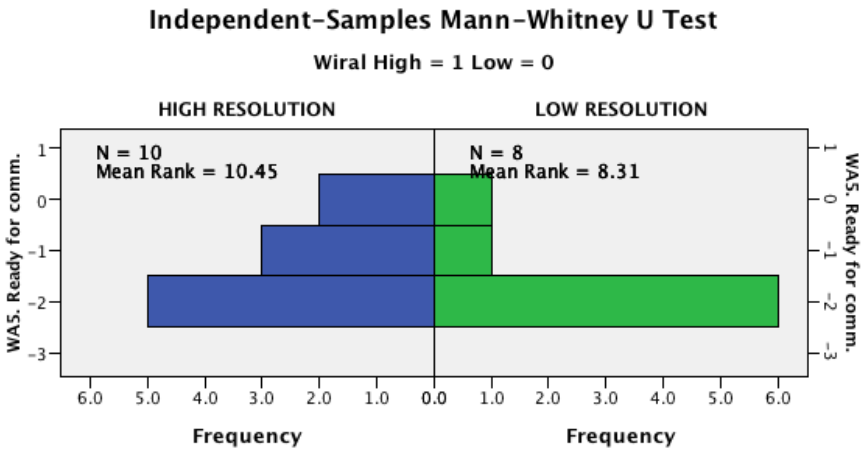


FIGURE 20 EVALUATION STATEMENT A5: THE PRODUCT IS READY FOR COMMERCIALIZATION (I.E READY FOR SELLING TO CONSUMERS).

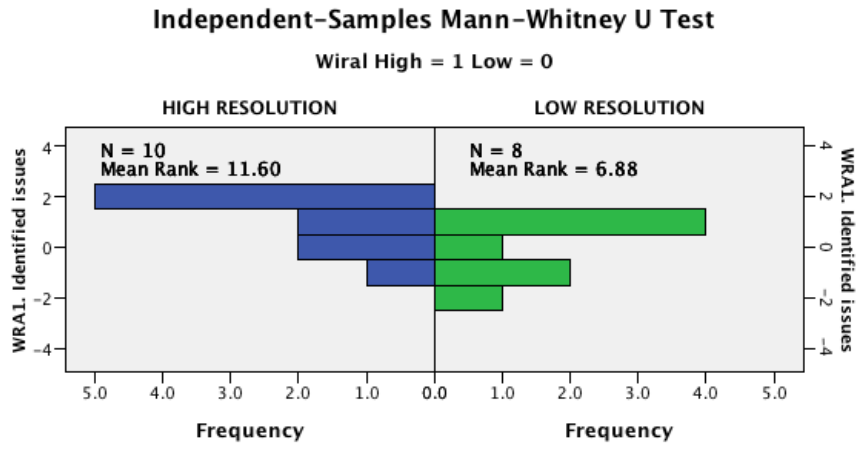


FIGURE 21 EVALUATION STATEMENT A1: THE INVENTORS HAVE IDENTIFIED AND ADDRESSED THE TECHNICAL ISSUES OF THE PROBLEM.

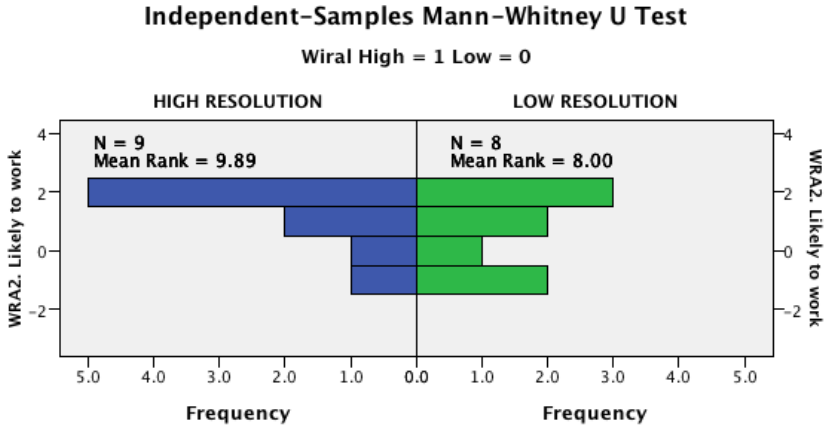


FIGURE 22 EVALUATION STATEMENT A2: THE CONCEPT IS LIKELY TO WORK AS DESCRIBED ABOVE.

Although there was only one statistically significant value, and some similar distributions, the high resolution prototype Mean Rank was consistently higher than that of the low resolution prototype, indicating that people tend to evaluate higher resolution with higher scores.

4.2.1.2 MovieMask

Another Mann-Whitney U test was run to determine if there were differences in evaluation score between high and low resolution prototypes of MovieMask equal to the test run for Wiral. For all but B5 and RB6, distributions of the evaluation scores for high and low resolution were not similar, as assessed by visual inspection. A3, A6, A8, RA3 and RA8 had evaluation scores for high resolution (respectively mean rank = 12.81, 12.44, 13.56, 12.19 and 12.69) that were statistically significantly higher than for low resolution (respectively mean rank = 6.85, 7.15, 6.25, 7.35 and 6.95), U = (respectively 66.500, 63.500, 72.500, 61.500 and 65.500), p = (respectively .012, .032, .002, .044 and .017).

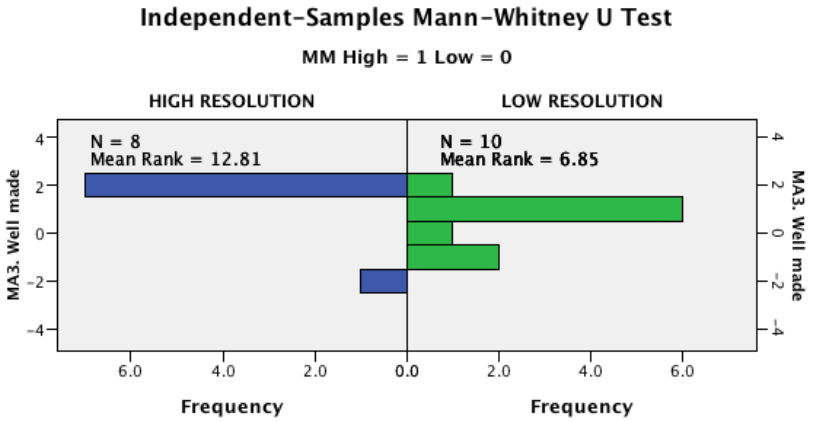


FIGURE 23 EVALUATION STATEMENT A3: THE PROTOTYPE IS WELL MADE.

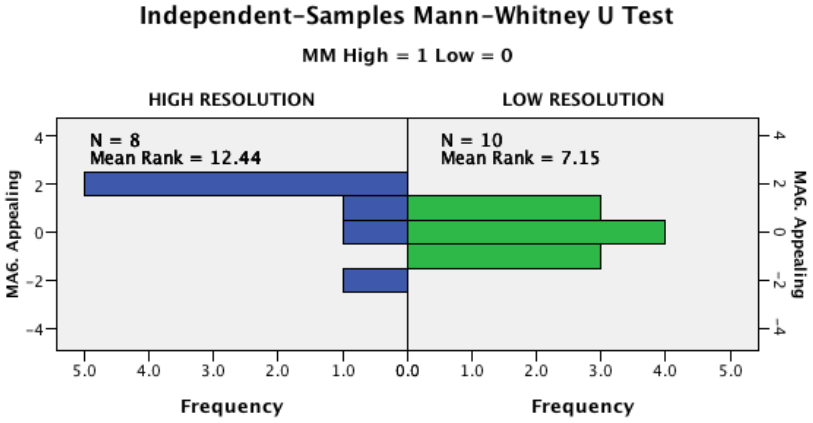


FIGURE 24 EVALUATION STATEMENT A6: THE PROTOTYPE IS APPEALING.

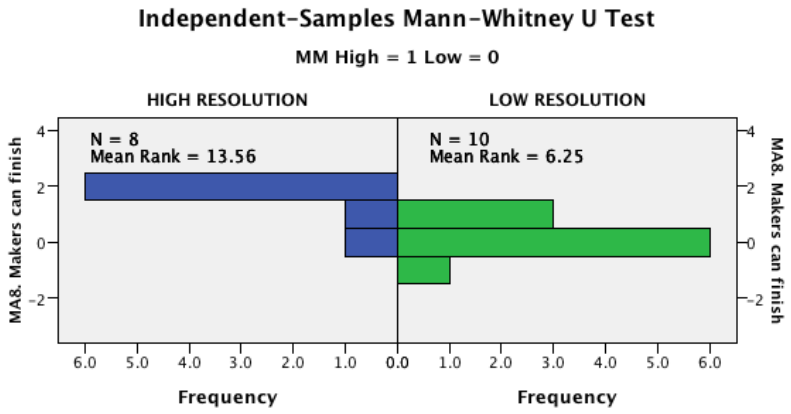


FIGURE 25 EVALUATION STATEMENT A8: THE PROTOTYPE INDICATE THAT THE MAKERS POSSESS SKILLS AND KNOWLEDGE TO MAKE THE FINISHED PRODUCT.

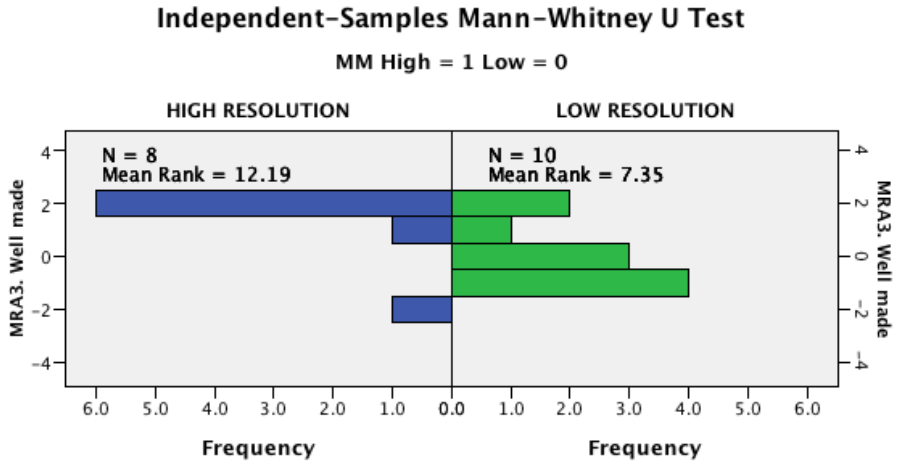


FIGURE 26 EVALUATION STATEMENT RA3: THE PROTOTYPE IS WELL MADE.

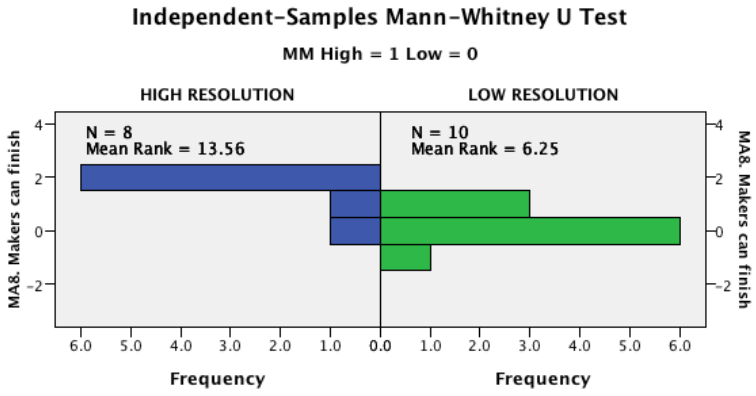


FIGURE 27 EVALUATION STATEMENT RA8: THE PROTOTYPE INDICATE THAT THE MAKERS POSSESS THE SKILLS AND KNOWLEDGE TO MAKE THE FINISHED PRODUCT.

For B5, and RB6 distributions of the evaluation scores for high and low resolution were similar, as assessed by visual inspection. Median evaluation score for high resolution (respectively -

1.50, 1.00 and 2.00) and low resolution (respectively -2.00, 0.50 and 1.00) was not statistically significantly different.

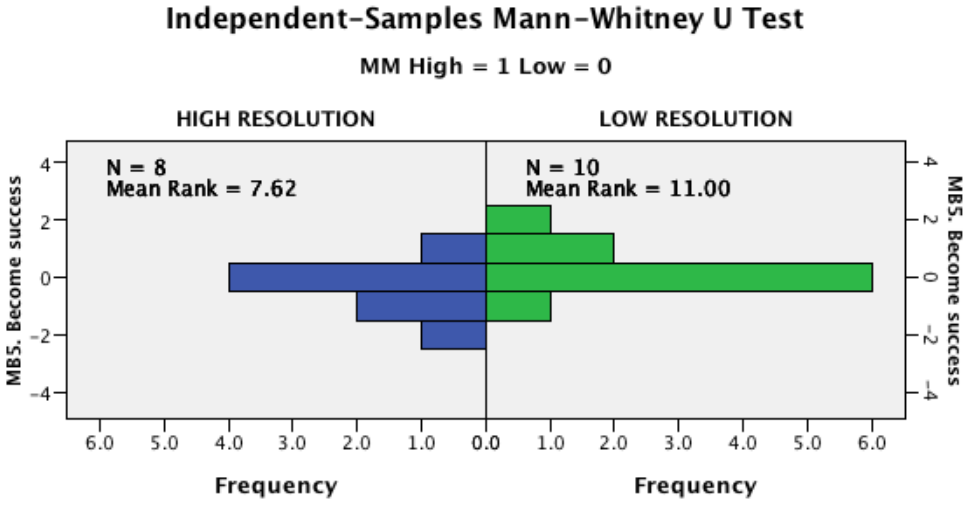


FIGURE 28 EVALUATION STATEMENT B5: YOU BELIEVE THE PRODUCT WILL BECOME A SUCCESS.

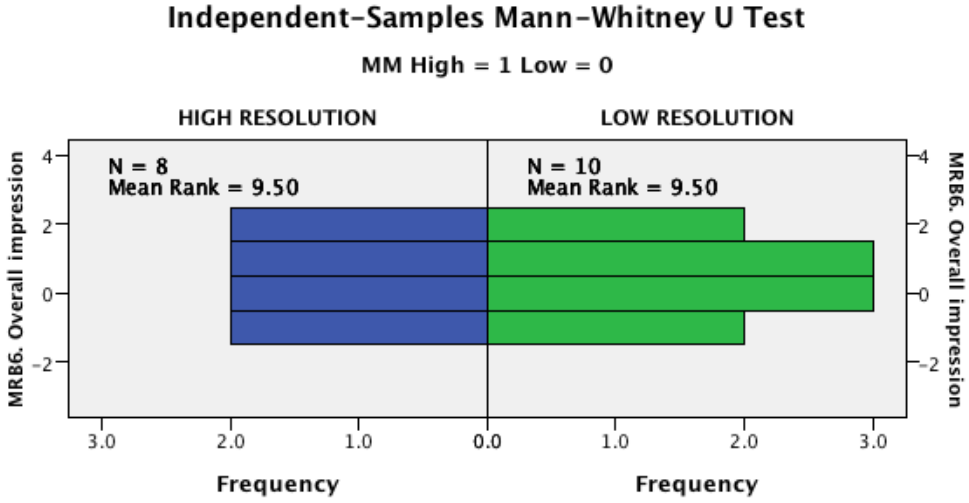


FIGURE 29 EVALUATION STATEMENT RB6: WHAT IS YOUR OVERALL IMPRESSION OF THE PROTOTYPE?

4.2.2 Function

To compare the samples of evaluations before and after being show function, I used Wilcoxon Signed-Rank Test. The Function Evaluation Hypothesis was stated previously like so:

H_0: After seeing the function of the prototype, the participants will not change their scores in the re-evaluation.

The corresponding alternative hypothesis being:

H_A: After seeing the function of the prototype, the participants will change their scores in the re-evaluation.

The corresponding hypothesis in terms of the Wilcoxon Signed-Rank Test is as follows:

Null hypothesis:

H_0: median difference = 0 (or alternatively, $\theta = 0$)

Alternative hypothesis:

H_A: median difference $\neq 0$ (or alternatively, $\theta \neq 0$)

A Wilcoxon Signed-Rank Test was run for each of the ordinal variables (A1-8 & B5-6) for both low and high resolution prototypes of Wiral and MovieMask. Out of the 20 tests, only one (A6 for Wiral, figure 30) returned a statistically significant median increase in evaluation. Most tests also showed a non-symmetrical distribution of differences. A general trend for both Wiral and MovieMask prototypes was that the differences tended to be more positive (the test subtracting the re-evaluation from the pre-evaluation), indicating that people tended to evaluate higher after being shown the function. Below the results with symmetrical distribution and/or statistically significant median increase are displayed.

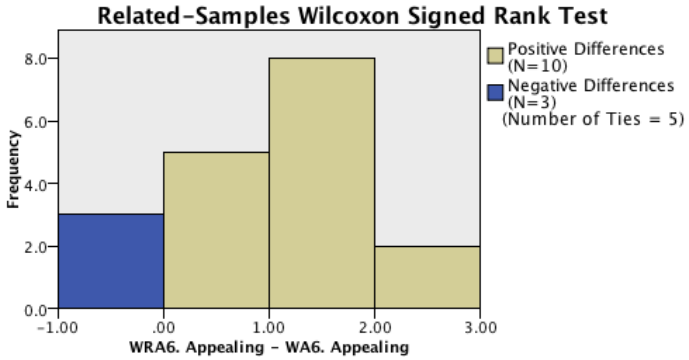


FIGURE 30 DISTRIBUTION OF DIFFERENCES FOR EVALUATION STATEMENT WA6: THE PROTOTYPE IS APPEALING.

As seen on figure above. Out of 18 participants, 10 evaluated higher after function, 5 had no change and 3 evaluated lower. Data are medians unless otherwise stated. Distribution of differences appears to be symmetrical as assessed by visual inspection. The Wilcoxon signed-rank test determined that there was a statistically significant median increase in evaluation (1.00) for subjects having seen the function (1.00) compared to subjects before seeing the function (.00), p = .039, z = 2.066.

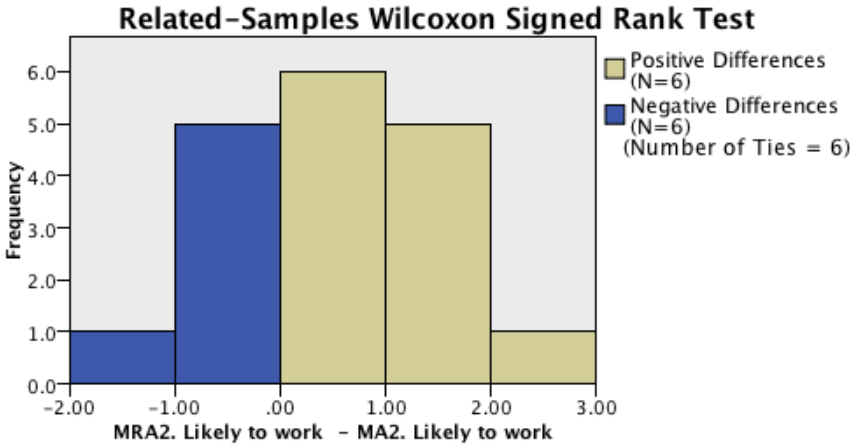


FIGURE 31 DISTRIBUTION OF DIFFERENCES FOR EVALUATION STATEMENT MA2: THE PROTOTYPE IS LIKELY TO WORK AS DESCRIBED ABOVE.

As seen on the figure above, out of 18 participants, 6 participants evaluated higher after function, 6 had no difference and 6 evaluated lower. Looking at the figure, it is truly symmetrical. The test was not statistically significant, p = 1.000, z = .000, Median difference = 0.

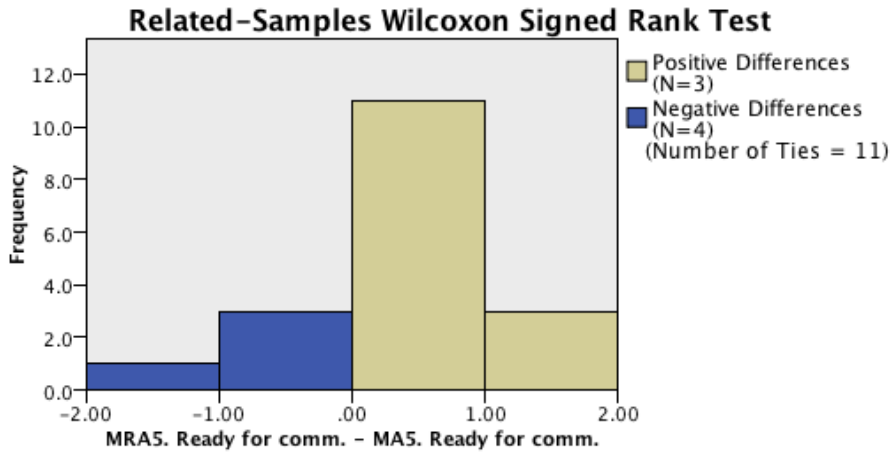


FIGURE 32 DISTRIBUTION OF DIFFERENCES FOR EVALUATION STATEMENT MA5: THE PRODUCT IS READY FOR COMMERCIALIZATION (I.E. READY FOR SELLING TO CONSUMERS).

As seen on the figure above, out of 18 participants, 3 participants evaluated higher after function, 11 had no difference and 4 evaluated lower. The distribution of differences appears to be symmetrical as assessed by visual inspection. There was no statistically significant increase in median in evaluations, p = .527, z = -.632, Median difference = 0.

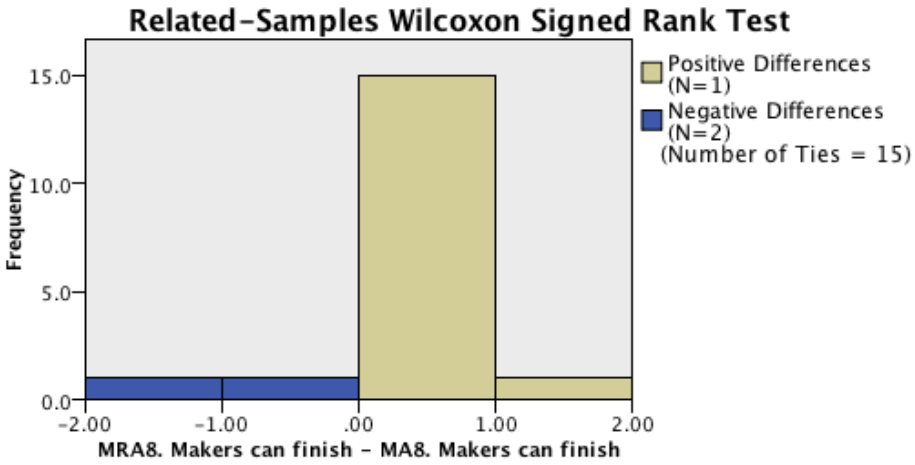


FIGURE 33 DISTRIBUTION OF DIFFERENCES FOR EVALUATION STATEMENT MA8: THE PROTOTYPE INDICATE THAT THE MAKERS POSSESS SKILLS AND KNOWLEDGE TO MAKE THE FINISHED PRODUCT.

As displayed on the above figure, out of 18 participants, 1 participant evaluated higher after function, 15 had no difference and 2 evaluated lower. The distribution of differences appears to be symmetrical as assessed by visual inspection. There was no statistically significant increase in median in evaluations, p = .414, z = -.816, Median difference = 0.

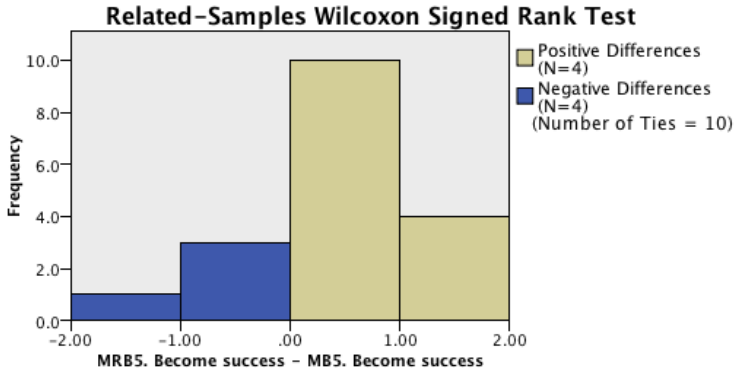


FIGURE 34 DISTRIBUTION OF DIFFERENCES FOR EVALUATION STATEMENT MB5: YOU BELIEVE THE PRODUCT WILL BECOME A SUCCESS.

As seen on the figure above, out of 18 participants, 4 evaluated higher after function, 10 had no difference and 4 evaluated lower. The distribution of differences appears to be symmetrical as assessed by visual inspection. There was no statistically significant increase in median in evaluations, p = .763, z = -.302, Median difference = 0.

4.2.2.1 Signed Rank test

As most of the tests returned non-symmetrical distribution of differences, Signed Rank Test is proposed as an alternative to Wilcoxon Signed-Rank Test. The (paired-samples) sign test tests whether the median differences of the paired observations are equal to 0 (zero) in the population. It upholds to the same assumptions as Wilcoxon Signed-Rank Test, save only for the symmetrical distribution of differences. It is recommended to perform the Wilcoxon Signed-Rank Test if your data pass the assumption of symmetrical distribution as it is a more powerful test. Since most of my data did not pass this assumption, I performed a Signed Rank Test on the data that did not pass.

14 Signed Rank Tests were performed for all ordinal variables for all prototypes. Only one (A4 for MovieMask, table) returned a statistically significant median increase in evaluation.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between MA4. Not like and MRA4. Not like equals 0.	Related-Samples Sign Test	.039[1]	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

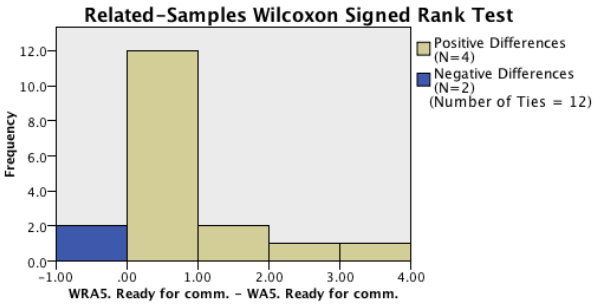
[1]Exact significance is displayed for this test.

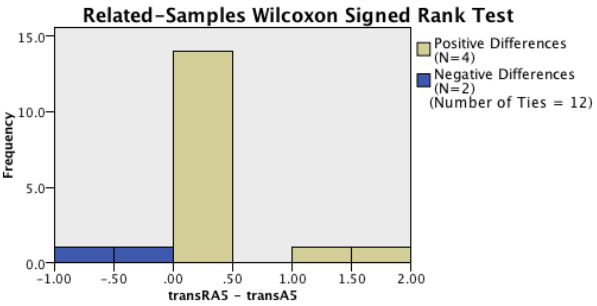
TABLE 2 SIGNED RANK TEST RESULTS FOR EVALUATION STATEMENT MA4.

The test returned 8 positive differences, 1 negative differences and 9 ties. Data are medians unless otherwise stated. The Signed Rank test determined that there was a statistically significant median increase in evaluation (.00) for subjects having seen the function (-2.00) compared to subjects before seeing the function (.00), p = .039, z = 2.066.

4.2.2.2 Distribution skewedness

One of the assumptions of the Wilcoxon signed-rank test is that the distribution of differences is symmetrical. Many of the distributions for the dependent variables were symmetric, but many were not. To correct for skewed distributions, there are options to transform the data. To show this, I attempted to transform some of the data points A5 for to see whether I could achieve symmetry. To do so, the scale [-2 -1 0 1 2] had to be converted to a positive scale [0 1 2 3 4], in order to find the square root of the data (used to convert moderately positively skewed data). As the figures below show, the transformation did appear to correct the skewedness somewhat and make it more symmetrical. The transformation did not change the number of positive, negative and ties.





4.2.3 Background

To examine the effect of backgrounds, a series of Mann-Whitney U Tests was run to determine if there were differences in evaluation scores between different definitions of backgrounds. The first definition, Combined Knowledge, was combined of the participant's educational background, work experience, expressed knowledge in the questionnaire and in part my subjective impression of the participant's knowledge (formed by talking, observing and interacting with the participants before, during and after the experiment). The group had N = 9 in each group. The data set was then split on two Combined Knowledge variables (Wiral and MovieMask) and analyzed using the Mann-Whitney U Test.

Report

Median	A1.	A2.	A3.	A4.	A5.	A6.	A7.	A8.	B5.	B6.	RA1.	RA2.	RA3.	RA4.	RA5.	RA6.	RA7.	RA8.	RB5.	RB6.
COMBINED KNOWLEDGE WIRAL																				
COMMON KNOWLEDGE	1.00	1.00	1.50	.00	-2.00	.00	.00	.00	.00	.00	1.00	.50	1.00	.00	-1.00	.00	1.00	1.00	.00	1.00
EXTENSIVE KNOWLEDGE	1.00	1.00	1.00	-2.00	-2.00	1.00	1.00	1.00	.00	1.00	1.00	2.00	1.00	-2.00	-2.00	1.00	1.00	.00	.00	2.00
Total	1.00	1.00	1.00	.00	-2.00	.00	1.00	.00	.00	1.00	1.00	1.00	1.00	.00	-2.00	1.00	1.00	.00	.00	1.00
COMBINED KNOWLEDGE MOVIEMASK																				
COMMON KNOWLEDGE	1.00	.00	1.00	-1.00	-1.00	.00	1.00	.00	.00	1.00	1.00	.00	1.00	.00	-1.00	1.00	1.00	.00	.00	.00
EXTENSIVE KNOWLEDGE	1.00	1.00	2.00	-2.00	1.00	1.00	2.00	1.00	.00	1.00	1.00	.00	2.00	.00	1.00	1.00	2.00	1.00	.00	1.00
Total	1.00	.50	1.00	-2.00	-.50	.50	1.00	1.00	.00	1.00	1.00	.00	1.00	.00	-1.00	1.00	1.50	1.00	.00	.50

TABLE 3 MEDIAN CALCULATIONS OF EVALUATIONS SORTED BY COMBINED KNOWLEDGE.

None of the results from the test were statistically significant, but as seen from table 3, participants placed in *extensive knowledge* group generally had a higher median score than participants in the *common knowledge* group, indicating that they tended to rate prototypes higher.

The second category of backgrounds was educational background. These groups were divided between *engineering* and *others,* whereas others included studies like business development, architecture and industrial design, with N = 9 in each group. Another Mann-Whitney U test was run for the groups. As with the Combined Knowledge groups, none of the tests showed statistically significant differences in evaluations.

Report

Median	A1.	A2.	A3.	A4.	A5.	A6.	A7.	A8.	B5.	B6.	RA1.	RA2.	RA3.	RA4.	RA5.	RA6.	RA7.	RA8.	RB5.	RB6.
Wiral																				
OTHER STUDY	1.00	1.00	2.00	.00	-2.00	.00	.00	.00	.00	1.00	1.00	.50	1.00	.00	-2.00	.00	1.00	.00	.00	1.00
STUDIED ENGINEERING	1.00	1.00	1.00	.00	-2.00	1.00	1.00	.00	-1.00	1.00	1.00	2.00	1.00	-1.00	-2.00	1.00	1.00	.00	.00	2.00
Total	1.00	1.00	1.00	.00	-2.00	.00	1.00	.00	.00	1.00	1.00	1.00	1.00	.00	-2.00	1.00	1.00	.00	.00	1.00
MovieMask																				
OTHER STUDY	1.00	1.00	1.00	-1.00	.00	.00	1.00	1.00	.00	1.00	1.00	.00	1.00	.00	-1.00	1.00	2.00	1.00	.00	.00
STUDIED ENGINEERING	1.00	.00	1.00	-2.00	-1.00	1.00	1.00	.00	.00	1.00	.00	1.00	.00	.00	-1.00	.00	1.00	.00	.00	1.00
Total	1.00	.50	1.00	-2.00	-.50	.50	1.00	1.00	.00	1.00	1.00	.00	1.00	.00	-1.00	1.00	1.50	1.00	.00	.50

TABLE 4 MEDIAN CALCULATIONS OF EVALUATION SORTED BY EDUCATIONAL BACKGROUND.

The third category of backgrounds was work experience. This variable was split between those who had worked in an engineering company and those who had not, also with N = 9 in each group. The Mann-Whitney U test revealed that the re-evaluation of statement A2 for the MovieMask prototypes had evaluation scores for participants who has worked for an engineering company (mean rank = 11.07) that were statistically significantly higher than for those who had not (mean rank = 4.00), U = 50.000, p = .016. The distribution of evaluations was, however, not similar as assessed by visual inspection.

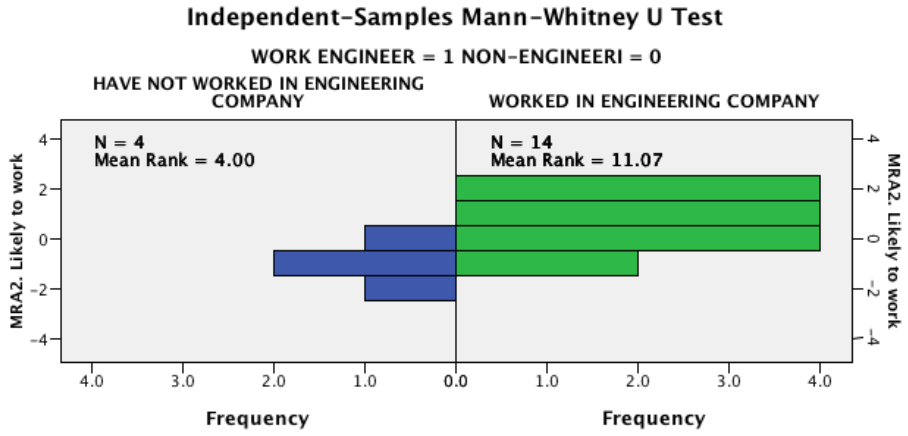


FIGURE 35 EVALUATION STATEMENT RA2: THE CONCEPT IS LIKELY TO WORK AS DESCRIBED ABOVE

A fourth category was defined by their answers to statement A12 in the questionnaire, regarding the participants' knowledge of the technology used in the prototypes. For Wiral there was equally many participants in each group, while MovieMask had 8 participants who claimed knowledge of the technology, while 10 did not. For the tests run on Wiral, none were statistically significant, while for MovieMask there were 7 statistically significant differences. Evaluations of A5, A6, A7, A8, RA5, RA7 and RA8 had evaluation scores for participants who claimed knowledge (respective mean rank = 13.12, 13.19, 12.75, 12.94, 12.50, 13.25, 13.19) that were statistically significantly higher than for those who did not (respective mean rank = 6.60, 6.55, 6.90, 6.75, 7.10, 6.50, 6.55), U = (respectively 69.000, 69.500, 66.000, 67.500,

64.000, 70.000, 69.500), p = (respectively .008, .007, .015, .010, .026, .004, .006). A6 also had similar distribution for both groups, although different location.

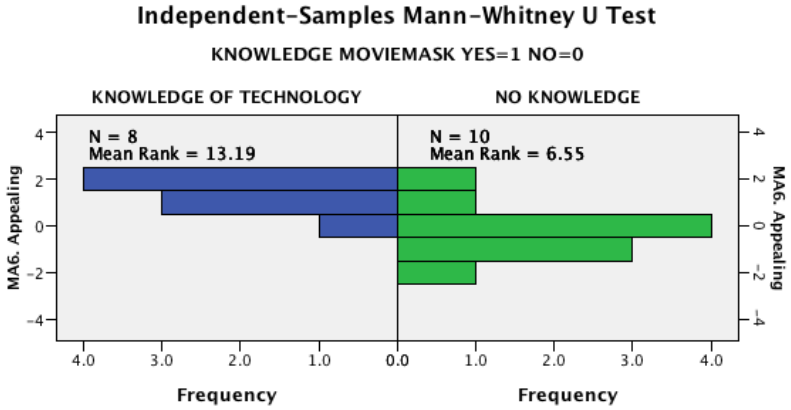


FIGURE 36 EVALUATION STATEMENT A6: THE PROTOTYPE IS APPEALING.

4.3 Demographics and Other Observations

The participants were a mix of men and women with nationalities from Norway, Denmark, Italy, Ireland, Portugal, USA, India and France with an average age was 38. 2 participants held Doctorates, 14 held a Master's degree, 1 a bachelor's and 1 had a college degree. 14 participants had held a job at an engineering company.

A Mann-Whitney U Test was run on the basis of gender, but returned no statistically significant results, indicating that there is no significant difference between evaluations based on gender.

The questionnaire contained evaluation points that were not included in the statistical tests above. This was in part because of the data itself, and partly because it was registered purposely as additional information to provide context to the evaluations. The following section will do a short evaluation of this data. The data includes estimations of months to finished product, sales price, size of market and production cost. The table below show the calculated means of the additional data. Keep in mind that the analysis of this data is not statistical, and is purely descriptive for the sake of discovering trends in the data. The data is lacking in sample size and should be evaluated as such, whereas one evaluation will have a large impact on the outcome.

Evaluation points	High resolution Wiral Pre-function	High resolution Wiral Post-function	Low resolution Wiral Pre-function	Low resolution Wiral Post-function	High resolution MovieMask Pre-function	High resolution MovieMask Post-function	Low resolution MovieMask Pre-function	Low resolution MovieMask Post-function
MEAN A9 - Estimate months to final product.	10,00	9,80	12,83	12,33	10,56	14,00	6,00	7,13
MEAN B1 - Estimate sales price.	168,33	187,38	716,67	693,33	103,63	139,88	120,86	121,71
MEAN B3 - Size of market	1600,00	1400,00	633,33	683,33	10125,00	7634,00	5500,00	5059,50
MEAN B4 - Production costs	38,14	82,00	120,75	127,67	21,79	85,17	48,42	49,42

TABLE 5 TABLE OF MEANS FOR THE ADDITIONAL VARIABLES.

Figure 37 displays the means for the participants' estimation of time to market. One interesting, perhaps counterintuitive point is that participants seem to believe that the high resolution prototype of the concept MovieMask needs more time to market than its low resolution counterpart. An interpretation could be that participants tend to have higher expectations for higher resolution prototypes, and thus expected more of the prototype. The low resolution prototype, on the other hand, could have performed closer to the expectations of the participants. As showed in the figure, this is not the case for the prototypes of Wiral. Another notation is that the participants increase their estimates after seeing the function of MovieMask, while they decrease their evaluation after seeing the function of Wiral (although very little).

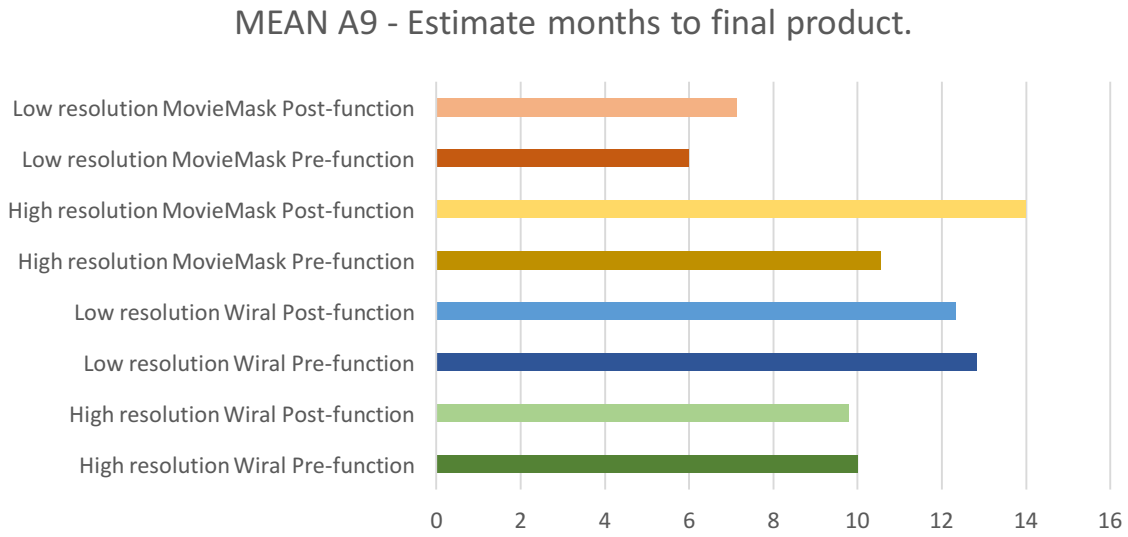


FIGURE 37 MEANS OF VARIABLE A9: PLEASE ESTIMATE HOW MUCH TIME THE INVENTORS NEED TO MAKE A PRODUCT THAT CAN BE SOLD TO CONSUMERS.

Figure 38 shows the means for the participants' estimates of a finished product's sales price. Immediately we can notice that he low resolution prototype of Wiral has a much higher estimation overall than its high resolution prototype. When looking at the raw data from this evaluation, it's made clear that this is due to some extreme outliers, heavily affecting the mean. When removing the outlier (3000 USD) the mean comes down to 260 USD, which is not as big

a difference, but is still higher than that of the high resolution prototype. One interpretation of this could be that the low resolution prototype revealed much more of the electronic setup making it seem more complicated than the high resolution version, where the only electronic to be viewed was the battery.

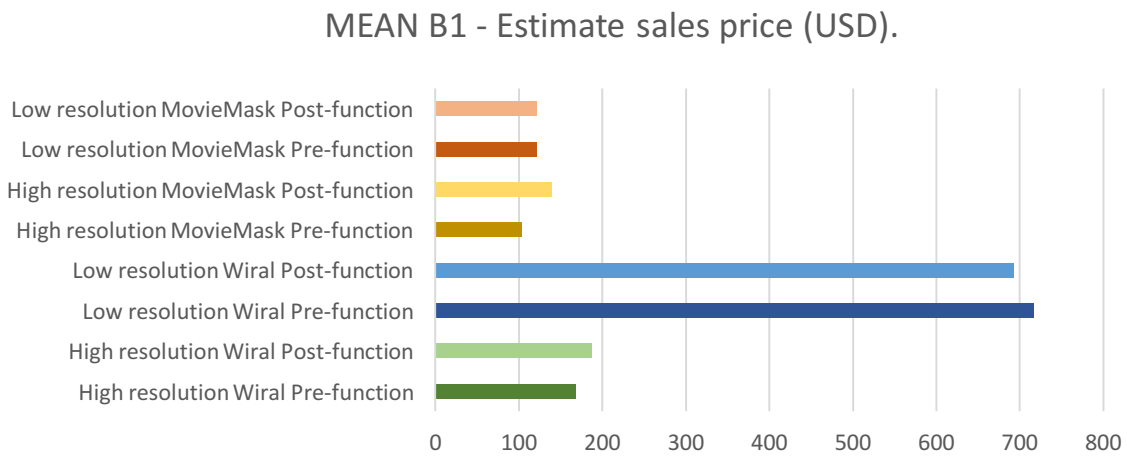


FIGURE 38 MEANS OF VARIABLE B1: PLEASE ESTIMATE THE SALES PRICE FOR A COMMERCIALIZED VERSION OF THE CONCEPT IN USD.

Figure 39 show the means for the participants' estimation of the size of the market. As is clear from the participants believe the market to be fairly larger for MovieMask, than for Wiral. Note that this evaluation point suffers from small data size and should be in no way be interpreted as a truth before more data has been established. However, the general impression from the participants was also that they believed MovieMask to have a broader market than Wiral (also gathered from interactions with participants after the experiment).

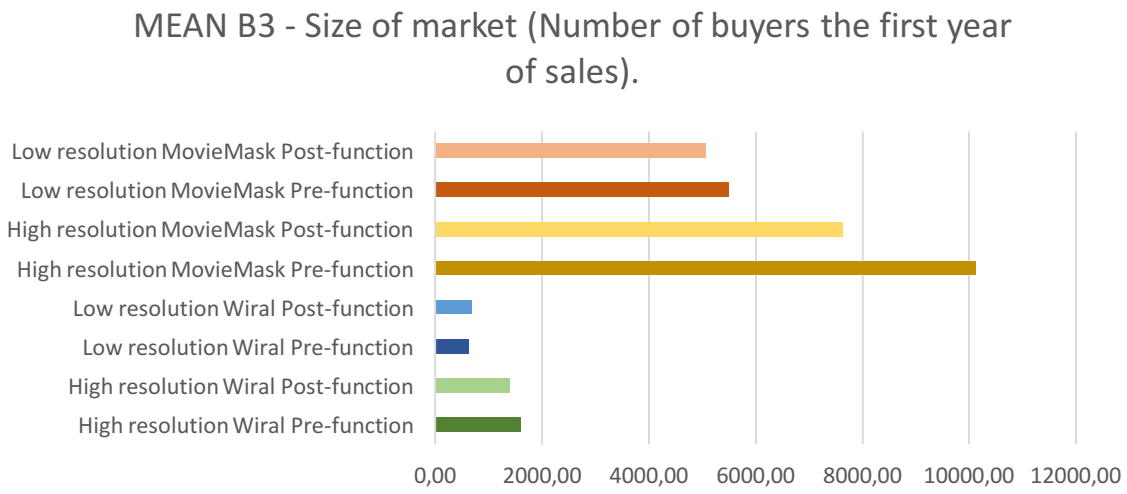


FIGURE 39 MEANS OF VARIABLE B3: PLEASE ESTIMATE THE SIZE OF THE MARKET THE FIRST YEAR.

Figure 40 display the means of the participants' estimation of production costs. Like the previous figure, this evaluation point is also lacking in data size and should be evaluated as such. Again we can see the same trend for Wiral as with the sales cost, namely that the low resolution prototype is believed to have higher production costs than the high resolution prototype. Another interesting notation is that participants increased their estimations for both the high resolution prototypes after seeing the function. This was also the case for the estimation of sales costs, indicating that they feel the products are more complicated after seeing the function of the prototypes, and require both more resources to finish and fetch a higher sales price.

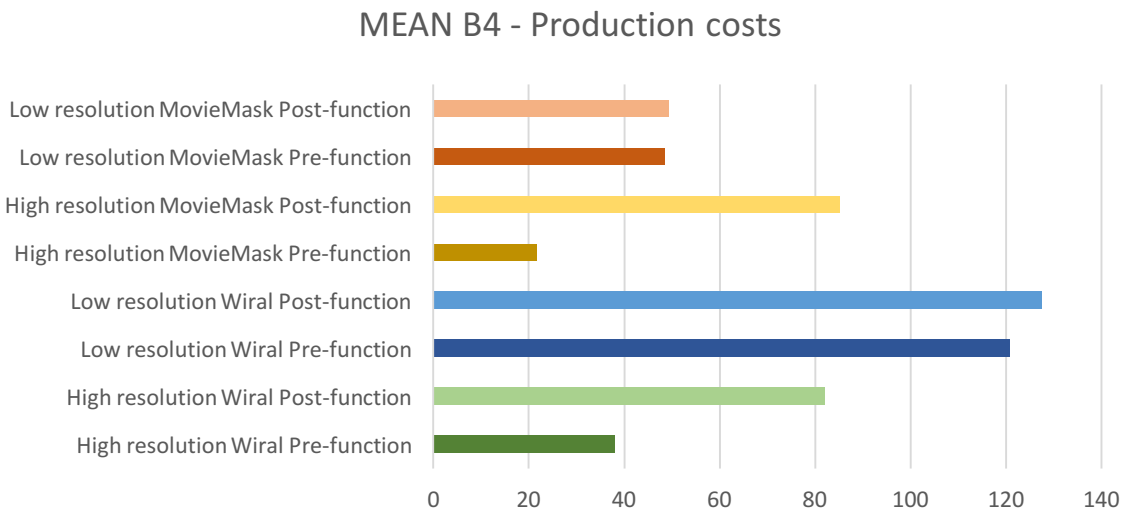


FIGURE 40 PLEASE ESTIMATE THE PRODUCTION COSTS FOR THE FINAL PRODUCT IN USD.

4.4 Evaluation of Hypotheses

This section is dedicated to the evaluation of the hypotheses stated in section 3, as was the goal of the experiment. The evaluation is based on the statistical evidence from the former tests in this chapter.

4.4.1 Background Hypothesis

This hypothesis was evaluated with the Mann-Whitney U Test with four different definitions of the independent variable. I will discuss the hypothesis individually for these four definitions. The first was based on a combined variable of several inputs to build groups of backgrounds. The statistical test revealed no statistically significant results, indicating that there is no difference in evaluations based on this combined variable. For the second definition of the variable, based on education, also had no statistical evidence that any of the groups evaluated differently than the other. The statistical analysis of the third definition yielded one statistically

significant result. However, the distribution of evaluations for the two groups were not similar, meaning the test should be disregarded as valid results. The fourth and final definition of the background variable was based on the participants' claimed knowledge of the technical aspects of the prototypes. This test yielded 7 statistically significant results. However, only one of these had a similar distribution, meaning the other 6 should be disregarded. The one significant result showed that participants who evaluated the MovieMask concept, and claimed to have knowledge of the technology tended to evaluate the evaluation statement (A2. The concept is likely to work as described above) higher than those who did not claim knowledge. This was not the case for the evaluation of the same statement for Wiral, which is interesting as one would perhaps expect the two have similar results. To summarize, the majority of results did not support the alternative hypothesis when grouping by background, although some results did lead to rejection of the null hypothesis, accepting that the groups did evaluate differently based on their background.

4.4.2 Design Resolution Evaluation Hypothesis

The design resolution evaluation hypothesis was evaluated with Mann-Whitney U Tests with the groups split on high or low resolution prototypes. The tests were run individually for the two concepts. For Wiral, there was only one statistically significant result, although the distribution was not similar, meaning the result should be disregarded. The test of the evaluations of MovieMask revealed 5 statistically significant results, but again none of the distributions were similar. To summarize, although having some statistically significant tests for design resolution, the distributions indicate that the tests should be regarded as faulty, meaning there was no statistical evidence that participants evaluated the prototypes differently based on resolution.

4.4.3 Function Evaluation Hypothesis

The function evaluation hypothesis was evaluated using Wilcoxon Signed-Rank Test, and also Signed Rank Test for the test with non-symmetrical distribution of differences. The Wilcoxon tests revealed one statistically significant result with symmetrical distribution for Wiral, leading to rejection of the null hypothesis and accepting the alternative hypothesis for the evaluation statement (A6. The prototype is appealing). The statistical evidence showing that participants evaluated higher after seeing the function of the prototype. Running the results that failed the assumption of symmetrical distribution in the less powerful Signed Rank Test revealed another statistically significant result for MovieMask leading to rejection of the null hypothesis for the evaluation statement (A4. You do **not** like the design of the prototype). Again, the statistical

evidence showed that participants evaluated higher after seeing the function of the prototype. To summarize, there were two cases of statistically significant increases in evaluations after seeing the function, indicating that there is in fact a difference in evaluations based on the function of the prototype.

4.5 Interpretation of Results

As discussed in section 4.2, much of the results show statistically insignificant results. Some did, however, show significant results although many of these failed assumptions of either identical or symmetrical distribution. Because of the difficulty of declaring similarity or symmetry due to the low amount of data, I have chosen to discuss the statistically significant results regardless of the distribution to spot trends and find points of interest for further research. I recognize that these tests are to be regarded with skepticism, but can be indications of large variations in evaluations to be checked with an appropriate amount of data.

4.5.1 Background tests results

Most of the results from the tests were statistically insignificant, nudging me towards accepting the null hypothesis of zero change based on background. However, there were several statistically significant results, especially for the MovieMask prototypes (although most failed the assumption of identical distribution). The most interesting definition of background seemed to be claimed knowledge, looking at the number of statistically significant results. The participants claiming knowledge tended to evaluate the prototypes higher than those who did not. This could be because they already have some idea of the need for such a concept, and thus evaluate the need higher than those who have no knowledge of similar products, their use and the need for said products.

Another interpretation is that people without extensive knowledge about the technical aspects of the prototype have to rely on their subjective opinion of the product and how it could fit into their own and other's lives. Meanwhile, participants with extensive knowledge about the technical aspects could perform a more technical analysis, judging how the product would work, technically and in a consumer setting.

The difference between the different concepts is interesting with the definition of claimed knowledge (MovieMask having 7 statistically significant results, while Wiral had none). One

interpretation of this could be that the products' imagined role in the participant's life made a differentiation between the concepts. Looking at the market number estimations, and the participants' general attitude towards the concepts, they estimated a far larger market for the MovieMask prototypes than the Wiral prototypes. This could have been caused by the intended use of the prototypes. Most people have some relation to watching movies (in one format or the other), while fewer have experience with filming and producing movies (MovieMask being a tool for the former and Wiral a tool for the latter). Thus, it might be easier for people to imagine the role of MovieMask in their life and judge it based on the role. Another interpretation is that MovieMask belongs to a product category that is more familiar to the participants. VR-goggles and mobile phone face accessories have had a recent boost in popularity fronted by companies like Google and Samsung. This allows participants to perform a more complete judgement basing the evaluation on existing knowledge on comparable products. Indeed, looking at the textual answers of the participants, most was aware of similar products to MovieMask, and some had even some working experience with similar concepts, making a more complete Absolute judgement (Blumenthal 1977). Meanwhile, few had any previous knowledge about similar products to Wiral, making it harder to evaluate a concept they had no experience with, and perhaps no immediate personal need for.

4.5.2 Design Resolution tests results

The test with design resolution as independent variable returned 1 statistically significant value for Wiral and 5 for MovieMask. Meanwhile, none of these results passed the assumption of identical distribution. These results indicate that there is actually a difference in evaluation between high and low resolution prototypes. This might not be very surprising as one might expect that a higher resolution prototype will get a higher evaluation. This is also reflected in the mean rankings of the groups. Throughout the tests, the high resolution prototypes tended to have a higher mean ranking than those of the low resolution prototypes. This was, however, less true for MovieMask than for Wiral. The evaluations of the high resolution Wiral prototype was consistently higher than the low resolution prototype, while MovieMask's ranks varied somewhat. One factor affecting the results could be the fact that the high resolution prototype of MovieMask was actually a commercialized product (mass produced). This could have manifested in the participant's evaluation in that they felt that the prototype had come too far without correcting their concerns so that it might not be corrected at all.

Looking at the results, none of the evaluation statements that were statistically significant (A3, A6, A8 and B5) was about the usability or functionality of the prototypes. According to the research of Wiklund et al. (1992), the design resolution should not affect the participants ability to find issues and thus evaluate the problems equally, independent of resolution, which seems to also be the case in this experiment. It did however affect the impression of the prototype, and their belief in success for the product. This factor may be important to the developers in that a potential investor may believe a higher resolution prototype will have a higher rate of success.

Again, like with background, there is a substantial difference in statistically significant results between Wiral and MovieMask. In the case of design, this could be related to the commonness of the concepts. Like discussed earlier, products like MovieMask are not uncommon and most people have some opinion of such concepts one way or the other. Wiral, on the other hand can be placed in a less common product group, and could even be said to address a niche market. The difference in results could also, as mentioned, be caused by the fact that the high resolution prototype of MovieMask was indeed a commercialized product, looking much more like a finished product. One could argue that the difference in resolution between the two prototypes of Wiral was less (they both had 3D-printed components), while MovieMask's prototypes 3D-printed frame vs. molded plastic frame.

4.5.3 Function tests results

The statistical tests for function only returned two statistically significant values, whereas one was from the less powerful Signed Rank Test. The Wilcoxon Signed-Rank Test that returned a statistically significant difference was, however, also symmetrical, meaning the value should be fairly trustworthy. The test was of the evaluation statement A6. (The prototype is appealing). The statistical evidence indicate that the function of the prototype has an impact on the prototype's appeal. Interestingly, the same test for MovieMask did not reveal the same significance, which creates uncertainty. One could argue once again that the function of MovieMask was somewhat expected based on previous experiences with similar products, and thus did not make the same impact on the test results, while how Wiral would function was more uncertain to the participants. In that respect one could argue that the less familiar the function of a product is to the audience, the more impact it will have on their evaluation.

5 Limitations

During the process of building, running and analyzing the experiment there was discovered some limitations of the study. The following section is dedicated to discuss said limitations and how it might have affected the results in order to prepare for future work.

The first and most obvious limitation of this research is the lack of data. Due to the experimental setup and choice of participant requirements, it was difficult collecting enough data to run significant statistical tests. The prototypes used for the experiment were the only prototypes of the same resolution and function, meaning I could only run the experiment with one participant at once, and not spread the prototypes out for more participants. Meanwhile, the physical prototypes meant they could be broken or not work equally for all participants. The setup also meant I had to be present for all the experiments in order to both set up and conduct the experiment. Although I attempted to keep my influence to a minimum, this is unavoidably adding a factor of influence for the participants. The participant requirements were set for project decision makers in a professional setting, who are generally harder to find and get to participate in experiments, as opposed to students who are in abundance at a university and often happy to be a part of research.

The lack of data obviously affects the statistical analyses of the experiment by reducing the impact and increasing uncertainties. This is especially true for analysis of ordinal variables using visual inspection of distributions to evaluate the assumptions of the test. The way the experiment was set up, more participants are required to produce truly statistically significant results. The sample sizes for the comparisons of low and high resolutions yielded about half of the total number of participants in each groups (N = 8-10), which is very low. Both the Mann-Whitney U Test and the Wilcoxon Signed-Rank Test required visual comparison of distribution of data. This approach is, however, not well suited for low sample sizes. Small sample sizes make comparison difficult as the individual results will have a large impact on the distribution of data. This is especially true for the Mann-Whitney U Tests in these analyses as it was used on groups with only 9 participants (total in both groups N = 18). The lack of data can be attempted corrected with transformation, like shown in section 4.2.2.2, but the best way is to acquire more data.

To see the effect of function on evaluations, the participants were first asked to evaluate a prototype without function. They were then shown the function and asked to re-evaluate the prototype. The reason this was the only order, and not both *no function – function* and *function – no function*, is that one can argue that one cannot unlearn what you have seen. Once you have seen the function, one cannot evaluate a prototype based on the assumption that they do not know the function of the prototype. Ideally, one would test these in opposite order to eliminate learning effects of the previous. In a future experiment this can be done by showing only one of the two to each participant. This would mean half the amount of data points, but would eliminate the learning effect. Also, in the experiment the participants did in some ways not experience the "full" function of the prototype in that they were not allowed to control the prototypes themselves. To create identical conditions for the participants, I chose to control the showing of function for the participants. Additionally, in the case of Wiral a large part of the function is not only the ability to propel a camera down a wire, but also the result of the capturing (the movie recorded using Wiral). Without seeing the result of the filming, they could only speculate on the final result. This could have had an impact of the results making it hard to compare the two concepts in terms of functionality.

Some participants had design background, with extensive knowledge of prototyping methodology and design thinking. Two participants mentioned that they had evaluated the prototypes as Learning types (Ulrich and Eppinger 2012), trying to understand what questions the designers wanted the prototype to answer. Prototypes were not introduced as communication type prototypes and the participants were free to put it into whatever category they felt appropriate. In order to link the results from this experiment to real investment cases, it should have been made clearer that participants evaluated the prototypes in an investment setting. This distinction is especially important when addressing participants with different backgrounds. If introducing a prototype to a product developer, he may have one view on for what purpose the prototype was made, while a person in a management position may have another.

The learning effect in this experiment was attempted removed by randomizing the order of the prototypes. However, the order should have been restricted to one group (order high – low or low – high), rather than two. Having resolutions of low – high and high – low means one can evaluate and discard the learning effect between these two groups because of the cross examination. Meanwhile, if another category is added, such as low – low and high – high

resolutions, one has to check whether these have a different effect on the evaluations than the other group, complicating the statistical analysis. I chose to disregard this difference because of the already small sample size, in order to conduct the statistical analysis, but this should be considered in future research.

In order to collect more data points for each participant, they were each given two prototypes, one from each concept. After being in the room with the participants I noticed that some deliberately compared their evaluations of the concepts instead of individually evaluating the two. This is not necessarily bad, but it is something to note as a possible source of difference (as not all participants seem to do the same).

This page is intentionally left blank.

6 Implications

The following section is a discussion of the implications and how I believe the results can affect future research. As discussed in the previous chapter, the analyses have several limitations which must be taken into account when discussing the implications of the study.

Several tests with different definitions of background was run, but only some returned statistically significant results. It seems, perhaps not surprisingly, that education and work experience are not accurate measure of a person's knowledge by themselves. A person's evaluation is after all made up of impressions and experiences from all parts of their life, not only education and work. The definition that returned the most statistical evidence was self-claimed knowledge, where the participants who claimed knowledge gave higher scores than those who did not claim knowledge. This is an interesting indication that people who feel they know and understand the technical aspects of a prototype gives higher evaluation, regardless of the prototype. This could also be interpreted from other definitions as well, where participants who had studied engineering tended to give higher scores than non-engineers. More data would be required to confirm this, but should be taken into account in further research. As with the tests for the other two independent variables, the tests for background returned different results for Wiral and MovieMask. It is not surprising that two concepts are evaluated differently, but one could expect the distribution of evaluations to have the same impact from an independent variable. As mentioned, more data is needed to confirm anything, but I find this difference interesting as it was not only present for one of the tests, but for all.

As discussed earlier, many of the evaluation statements returned insignificant differences in the statistical tests. Although some of the tests were statistically significant while upholding the assumptions of the tests, many did not show the same difference. For example, the statistically significant tests for design resolution all had higher mean rank for the high resolution prototypes. This indicates that there is indeed a difference in evaluations based on the resolution of the prototypes. This may seem contrary to the research done by Wiklund et al. (1992) but it is worth noting that the discussed evaluation statements were not regarding the usability of the prototype, but rather what Houde and Hill (1997) refer to as the *look and feel* of the prototype. It is perhaps not surprising that people would rate a higher resolution prototype higher in aesthetic refinement than its low resolution counterpart. It is also interesting that MovieMask seemed to have more significant differences than Wiral. As discussed, this could be due to

familiarity of the concept or the difference in resolution between the prototypes of the same concept. Either way, there seems to be an interesting relation between the prototype and its resolution that should be further explored in future research.

Looking at the results from testing the impact of the function of the prototypes, there is little statistical evidence of change in evaluation. It is difficult to say anything certain without testing with an appropriate amount of data, but it seems from the test results that function has less impact on the evaluations than design resolution and background. However, the fact that some tests returned statistically significant results for one concept and not the other, indicate that there is a connection between the concepts prototype and the effect of displaying function.

7 Conclusions

The goal of this thesis was to examine the effects of background, design resolution and function on prototype evaluations to try to assist both early stage developers and *investors* in understanding what to prioritize and how to make good decisions. This was done by the creation of an experiment design, running 18 experiments and statistical analyses of the acquired data. Because of the lack of data, the statistical analyses were limited, but did provide some insights that should be further explored. The tests uncovered that people tend to evaluate prototypes differently based both on their background and the design resolution of the prototype. With the current data, function seemed to have little effect on the evaluations. However, the tests run on function did share the same indications as the tests on background and design resolution of difference in distributions based on concept. For further research, it would be interesting to uncover what makes this difference in distribution.

For further work on the subject on what affect prototype evaluations, I suggest a preliminary experiment in addition to the completion of data collection for the experiment presented in this thesis. As this subject deals with human evaluations there should be a high amount of data points in order to provide sufficient statistical background for analysis. For this I suggest a preliminary experiment using a digital version. By making a digital version, that can be issued by email, using video and illustrations to provide design resolution and function and an online questionnaire to capture the evaluations, one can reach out to a vast number of participants in a much shorter timeframe. This would provide valuable insight and allow quick uncovering of trends. As a digital version cannot replace a physical experiment, I propose that the results from the preliminary study is used to create a design to test the trends more thoroughly.

This page is intentionally left blank.

Bibliography

Blumenthal, Arthur. L. 1977. *The Process of Cognition*. Englewood Cliffs: Prentice Hall.

Booz, Allen, and Hamilton. 1982. "New Product Management for the 1980's." New York.

Bryan-Kinns, Nick, and Fraser Hamilton. 2002. "One for All and All for One?: Case Studies of Using Prototypes in Commercial Projects." In, 91. ACM Press. doi:10.1145/572020.572032.

Houde, Stephanie, and Charles Hill. 1997. "What Do Prototypes Prototype." In *Handbook of Human-Computer Interaction*, 2:367–81. Amsterdam: Elsevier Science BV.

Roozenburg, Norbert FM, and Johannes Eekels. 1995. *Product Design: Fundamentals and Methods*. Vol. 2. Wiley Chichester. http://tocs.ulb.tu-darmstadt.de/34922482.pdf.

Saaty, Thomas L. 2008. "Decision Making with the Analytic Hierarchy Process." *International Journal of Services Sciences* 1 (1): 83–98.

Tyebjee, Tyzoon T., and Albert V. Bruno. 1984. "A Model of Venture Capitalist Investment Activity." *Management Science* 30 (9): 1051–66.

Ulrich, Karl T., and Steven D. Eppinger. 2012. *Product Design and Development (5th Ed.)*. 5th ed. McGraw- Hill.

Virzi, Robert A., Jeffrey L. Sokolov, and Demetrios Karis. 1996. "Usability Problem Identification Using Both Low- and High-Fidelity Prototypes." In , 236–43. ACM Press. doi:10.1145/238386.238516.

Walker, Miriam, Leila Takayama, and James A. Landay. 2002. "High-Fidelity or Low-Fidelity, Paper or Computer? Choosing Attributes When Testing Web Prototypes." In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 46:661–665. SAGE Publications Sage CA: Los Angeles, CA. http://journals.sagepub.com/doi/abs/10.1177/154193120204600513.

Wiklund, Michael E., Christopher Thurrott, and Joseph S. Dumas. 1992. "Does the Fidelity of Software Prototypes Affect the Perception of Usability?" In *Proceedings of the Human Factors Society Annual Meeting*, 36:399–403. SAGE Publications Sage CA: Los Angeles, CA. http://journals.sagepub.com/doi/abs/10.1177/154193129203600429.

APPENDIX 1

Evaluation Questionnaire

NTNU
Faculty of Economy and Management
Department of Industrial Economics and Technology Management

Our date 26.04.17 | Our referanse JOJA
Your dato | Your referanse

1. Concept Specific Questionnaire

This questionnaire is designed to collect your impression of the concept presented.

Case: Cablecam

The prototype is for a light-weight "cablecam" system that allows you to move a camera along a line/wire propelled by an electric motor, controlled with a remote.

Instructions

Please take a few minutes to make yourself familiar with the prototype. When you feel ready, please continue with the questionnaire you have been given.

Part A. Technical

The questions in this section are designed to collect information about your impression of the technical aspects of the prototype. Please indicate your degree of agreement with the following statements. The scale used is described below.

Disagree	*Somewhat disagree*	*Neutral*	*Somewhat agree*	*Agree*
☐	☐	☐	☐	☐

A1. The inventors have identified and addressed the technical issues of the problem.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A2. The concept is likely to work as described above.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A3. The prototype is well made.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A4. You do **not** like the design of the prototype.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A5. The product is ready for commercialization (i.e. ready for selling to consumers).

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

Address 7491 Trondheim
Org.nr. 974 767 880
E-mail: jonasnj@stud.ntnu.no
http://www.ivt.ntnu.no/ipm/
Visiting Address Richard Birkelandsvei 2b, Room P314, Gløshaugen
Phone + 47 90 60 77 50
Fax + 47 73 59 41 29

A6. The prototype is appealing.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A7. The prototype indicate that the makers are skilled.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A8. The prototype indicate that the makers possess skills and knowledge to make the finished product.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A9. Please estimate how much time the inventors need to make a product that can be sold to consumers. *Please give estimate in months*.

☐☐☐ Months *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

A10. What parts of the prototype do you feel is likely to change on the final product? *Please specify*.

IF NONE - MARK THIS BOX: ☐

A11. Is there any aspect of the prototype you think is likely to *not* work? *Please specify*.

IF NONE - MARK THIS BOX: ☐

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

A12. Do you possess knowledge about the technology used in the prototype? *Please specify*.

IF NONE - MARK THIS BOX: ☐

Part B. Market

The questions in this section are designed to collect information about your impression of the market aspects of the concept.

B1. Please estimate the sales price for a commercialized version of the concept in NOK.

NOK *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

B2. Would you buy the product at the price you estimated?

☐ Yes
☐ No

B3. Please estimate the size of the market the first year (i.e. how many will purchase the finished product the first year after the release).

I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX: ☐

B4. Please estimate the production costs for the final product in NOK.

NOK *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

B5. You believe the product will become a success.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

B6. What is your overall impression of the prototype?

Very bad ☐ ☐ ☐ ☐ ☐ *Very good*

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

B7. If you were asked, would you be willing to invest in the project?

☐ Yes

☐ No

B8. Did you have any knowledge about the concept before attending this survey? *Please specify*.

IF NONE - MARK THIS BOX: ☐

Jonas Neraal Jakobsen, M.Sc. Candidate

Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)

Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.

Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

2. Concept Specific Questionnaire Re-evaluation

This questionnaire is designed to collect your impression of the concept presented after you have seen the function of the prototype.

Instructions

Please fill out the form in accordance with previous instructions.

Part A. Technical

Please indicate your degree of agreement with the following statements, using the scale below.

Disagree	*Somewhat disagree*	*Neutral*	*Somewhat agree*	*Agree*
☐	☐	☐	☐	☐

A1. The inventors have identified and addressed the technical issues of the problem.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A2. The concept is likely to work as described above.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A3. The prototype is well made.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A4. You do **not** like the design of the prototype.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A5. The product is ready for commercialization (i.e. ready for selling to consumers).

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A6. The prototype is appealing.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A7. The prototype indicate that the makers are skilled.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

A8. The prototype indicate that the makers possess skills and knowledge to make the finished product.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A9. Please estimate how much time the inventors need to make a product that can be sold to consumers. *Please give estimate in months*.

☐☐☐ Months *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

A10. What parts of the prototype do you feel is likely to change on the final product? *Please specify*.

IF NONE - MARK THIS BOX: ☐

A11. Is there any aspect of the prototype you think is likely to *not* work? *Please specify*.

IF NONE - MARK THIS BOX: ☐

A12. Do you possess knowledge about the technology used in the prototype? *Please specify*.

IF NONE - MARK THIS BOX: ☐

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

Part B. Market

The questions in this section are designed to collect information about your impression of the market aspects of the concept.

B1. Please estimate the sales price for a commercialized version of the concept in NOK.

☐ *NOK* *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

B2. Would you buy the product at the price you estimated?

☐ Yes
☐ No

B3. Please estimate the size of the market the first year (i.e. how many will purchase the finished product the first year after the release).

☐ *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

B4. Please estimate the production costs for the final product in NOK.

☐ *NOK* *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

B5. You believe the product will become a success.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

B6. What is your overall impression of the prototype?

Very bad ☐ ☐ ☐ ☐ ☐ *Very good*

B7. If you were asked, would you be willing to invest in the project?

☐ Yes
☐ No

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

3. Concept Specific Questionnaire

This questionnaire is designed to collect your impression of the concept presented.

Case: Cinematic glasses

This is a prototype from a group making Cinematic glasses for smartphones. The glasses simulate a portable cinematic experience using lenses and casing.

Instructions

Please take a few minutes to make yourself familiar with the prototype. When you feel ready, please continue with the questionnaire you have been given.

Part A. Technical

The questions in this section are designed to collect information about your impression of the technical aspects of the prototype. Please indicate your degree of agreement with the following statements. The scale used is described below.

Disagree	*Somewhat disagree*	*Neutral*	*Somewhat agree*	*Agree*
☐	☐	☐	☐	☐

A1. The inventors have identified and addressed the technical issues of the problem.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A2. The concept is likely to work as described above.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A3. The prototype is well made.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A4. You do **not** like the design of the prototype.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A5. The product is ready for commercialization (i.e. ready for selling to consumers).

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A6. The prototype is appealing.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

A7. The prototype indicate that the makers are skilled.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A8. The prototype indicate that the makers possess skills and knowledge to make the finished product.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A9. Please estimate how much time the inventors need to make a product that can be sold to consumers. *Please give estimate in months*.

☐☐☐ Months *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

A10. What parts of the prototype do you feel is likely to change on the final product? *Please specify*.

IF NONE - MARK THIS BOX: ☐

A11. Is there any aspect of the prototype you think is likely to *not* work? *Please specify*.

IF NONE - MARK THIS BOX: ☐

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

A12. Do you possess knowledge about the technology used in the prototype? *Please specify*.

IF NONE - MARK THIS BOX: ☐

Part B. Market

The questions in this section are designed to collect information about your impression of the market aspects of the concept.

B1. Please estimate the sales price for a commercialized version of the concept in NOK.

NOK *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

B2. Would you buy the product at the price you estimated?

☐ Yes

☐ No

B3. Please estimate the size of the market the first year (i.e. how many will purchase the finished product the first year after the release).

I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX: ☐

B4. Please estimate the production costs for the final product in NOK.

NOK *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

B5. You believe the product will become a success.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

Jonas Neraal Jakobsen, M.Sc. Candidate

Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)

Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.

Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

B6. What is your overall impression of the prototype?

Very bad ☐ ☐ ☐ ☐ ☐ *Very good*

B7. If you were asked, would you be willing to invest in the project?

☐ Yes

☐ No

B8. Did you have any knowledge about the concept before attending this survey? *Please specify*.

IF NONE - MARK THIS BOX: ☐

Jonas Neraal Jakobsen, M.Sc. Candidate

Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)

Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.

Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

4. Concept Specific Questionnaire Re-evaluation

This questionnaire is designed to collect your impression of the concept presented after you have seen the function of the prototype.

Instructions

Please fill out the form in accordance with previous instructions.

Part A. Technical

Please indicate your degree of agreement with the following statements, using the scale below.

Disagree	*Somewhat disagree*	*Neutral*	*Somewhat agree*	*Agree*
☐	☐	☐	☐	☐

A1. The inventors have identified and addressed the technical issues of the problem.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A2. The concept is likely to work as described above.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A3. The prototype is well made.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A4. You do **not** like the design of the prototype.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A5. The product is ready for commercialization (i.e. ready for selling to consumers).

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A6. The prototype is appealing.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A7. The prototype indicate that the makers are skilled.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

A8. The prototype indicate that the makers possess skills and knowledge to make the finished product.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

A9. Please estimate how much time the inventors need to make a product that can be sold to consumers. *Please give estimate in months*.

☐☐☐ Months *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

A10. What parts of the prototype do you feel is likely to change on the final product? *Please specify*.

IF NONE - MARK THIS BOX: ☐

A11. Is there any aspect of the prototype you think is likely to *not* work? *Please specify*.

IF NONE - MARK THIS BOX: ☐

A12. Do you possess knowledge about the technology used in the prototype? *Please specify*.

IF NONE - MARK THIS BOX: ☐

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

Part B. Market

The questions in this section are designed to collect information about your impression of the market aspects of the concept.

B1. Please estimate the sales price for a commercialized version of the concept in NOK.

☐ *NOK* *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

B2. Would you buy the product at the price you estimated?

☐ Yes

☐ No

B3. Please estimate the size of the market the first year (i.e. how many will purchase the finished product the first year after the release).

☐ *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

B4. Please estimate the production costs for the final product in NOK.

☐ *NOK* *I DO NOT FEEL QUALIFIED TO ANSWER - MARK THIS BOX*: ☐

B5. You believe the product will become a success.

Disagree ☐ ☐ ☐ ☐ ☐ *Agree*

B6. What is your overall impression of the prototype?

Very bad ☐ ☐ ☐ ☐ ☐ *Very good*

B7. If you were asked, would you be willing to invest in the project?

☐ Yes

☐ No

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

5. Background Information Questionnaire

This questionnaire is designed to collect additional background information about you.

Part A: Professional Experience

The questions in this section are designed to collect information on your career and whether and how they have changed over time. Please give estimates wherever suitable.

A1. Do you work/have worked for an engineering company?

Answer: _____

A2a. Where do you currently work?

If you have no current employer: Skip to part B

Company: _____

A2b. For how long have you worked for your current employer?

Answer: _____

A2c. Are you assigned to a specific department by your current employer?

☐ Management
☐ Research and Development (R&D)
☐ Testing and Verification
Other: _____

A3. What is your position in the company?

Answer: _____

A4. Are you in a position where you make, or participate in, project decisions? *Please specify.*

Answer: _____

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

A5. Are you in a project management position?

Answer: _____

Part B. Education

The questions in this section are designed to collect information on your education.

B1. What is your current level of achieved education?

- ☐ High School
- ☐ College
- ☐ Bachelor's Degree
- ☐ Master's Degree
- ☐ Ph.D.

B2. When did you graduate?

Month ☐☐ Year ☐☐☐☐

B3a. Please record your primary area of specialization.

Primary Area
of Specialization: _____

B3b. Please record any additional areas of specialization you currently have.

IF NONE: MARK THIS BOX: ☐

1. Area of Specialization: _____
2. Area of Specialization: _____
3. Area of Specialization: _____

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

B4. Are you currently studying for a degree? If no, skip to part C. If yes, please specify:

- ☐ High School
- ☐ College
- ☐ Bachelor's Degree
- ☐ Master's Degree
- ☐ Ph.D.

B5. When do you plan to graduate?

Month ☐☐ Year ☐☐☐☐

B6a. Please record your primary area of specialization.

Primary Area
of Specialization: _____

B6b. Please record any additional areas of specialization you currently have.

IF NONE: MARK THIS BOX: ☐

1. Area of Specialization: _____

2. Area of Specialization: _____

3. Area of Specialization: _____

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

Part C: Demographic Information

The questions in this section are designed to collect some of your demographic information.

C1. Are you:

☐ Male
☐ Female

C2. In what year were you born?

Year of Birth: ☐☐☐☐

C3. What is your nationality (i.e. citizenship)?
Please specify if you have multiple citizenships.

Answer: _____

C4. What is your native language?

Answer: _____

C5. Please rate your English proficiency.

☐ None
☐ Basic
☐ Advanced
☐ Fluent

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen

Part D: Further Participation

D1. Are you willing to receive follow-up questions or surveys of this study via e-mail in the future? If yes, please write your e-mail address below.

E-mail address: _____

Part E: General Information

You have just participated in an experiment on concept evaluation based on different design parameters to get a view into your experience of the prototype to understand what affects your impression.

As priming you were given either high or low resolution prototypes, with or without key functions intact. The goal of this experiment is to provide qualitative data on the effects of the different parameters and how it affects people with or without a relevant knowledge background. This is to provide better understanding of the effects prototypes.

We wish to remind you to be confidential about the content of this experiment to provide non-biased conditions for every participant, as stated in the consent form. We hope you enjoyed participating, and thank you kindly for your commitment of time to this experiment!

Thank you for your time and participation!

Jonas Neraal Jakobsen, M.Sc. Candidate
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU)
Richard Birkelandsvei 2B, NO-7491 Trondheim, Norway.
Phone: +47 90 60 77 50, jonasnj@stud.ntnu.no, Skype: Jonas.neraal.jakobsen