NTNU – Trondheim
Norwegian University of
Science and Technology

Photography - Bridging the Gap between Needfinding and Visual Thinking

Andreas Wulvik

Mechanical Engineering
Submission date: June 2015
Supervisor: Martin Steinert, IPM

Norwegian University of Science and Technology
Department of Engineering Design and Materials

Summary

Founding on the pre-master project "Photography – a new tool in needfinding" (Wulvik et al., 2015)[a] as starting point and based on theory from the fields of needfinding and visual thinking, an explorative experiment has been piloted, run and analysed. The goal of the experiment was to test the influence of focal length and depth of field on the verbal ideation output from designers tasked with improving situations shown in pictures. Output was measured by: number of insights, solutions, questions and the total sum of these. Focal length and depth of field were given two values each, pushing towards extremes in both directions. Focal length was given the values of 24mm (wide-angle) and 100mm (telephoto). Depth of field is mostly controlled by aperture, and apertures of f/2.8 and f/16 (f/11) were used to create small and large depth of field. Four different situations were photographed with all four combinations of the focal length and depth of field (24mm & f/2.8, 24mm & f/11, 100mm & f/2.8, 100mm & f/16), a total of 16 pictures to be used as stimuli in the experiment. To get as much data as possible from the participants, each was shown four pictures, one from each situation and paired with different variable combinations. 18 participants were a part of the final experiment, each looking at four pictures. A total of 72 data points were collected.

Data was collected through audio recording and eye tracking. Each participant was placed in front of a monitor with an eye tracker placed underneath and given a hand-held microphone. During the experiment, the participant was asked to think out loud so that their thoughts could be captured with audio. The eye tracker captured their gaze position. After the experiment, audio data was coded into insights, solutions and questions based on an emergent coding scheme. Even though eye-tracking data was recorded, it was not finally used for the results of this pilot experiment due to difficulties with interpreting gaze data. Eye tracking technology and knowledge is rapidly improving, so if any breakthroughs in analysis should occur in the near future, data is stored and available for analysis.

No statistically significant results were found when running independent samples t-tests for focal length and depth of field. However, though not statistically significant, an indication towards the effect of increased output of solutions when using a longer focal length was found with a calculated significance value 0.894. This potential difference is corresponds with initial hypothesis H1: *Longer focal lengths provide more output because more equal attention is given to visual elements*.

Based on results from and identified error sources in this pilot experiment, areas of improvement for a second experiment have been proposed and are currently put into place.

[a] Pre-master project can be found in Appendix C

Sammendrag

Basert på prosjektoppgaven "Photography – a new tool in needfinding" (Wulvik, Balters, & Steinert, 2015) [b] har et utforskende eksperiment basert på teori fra needfinding og visual thinking blitt prøvd ut, gjennomført og analysert. Målet med eksperimentet var å undersøke effekten av brennvidde og dybdeskarphet på verbale resultater fra designere som fikk i oppgave å forbedre situasjoner vist i bilder. Resultater ble målt i: antall oppdagelser (insights), løsninger, spørsmål og summen av disse. Brennvidde og dybdeskarphet ble gitt to ekstremverdier hver. Brennvidde ble delt i 24mm (vidvinkel) og 100mm (telefoto). Dybdeskarphet er hovedsakelig styrt av blenderåpning, og blender f/2.8 og f16 (f/11) ble brukt for å skape henholdsvis liten og stor dybdeskarphet. Fire forskjellige situasjoner ble tatt bilde av med alle fire kombinasjoner av brennvidde og dybdeskarphet (24mm & f/2.8, 24mm & f/11, 100mm & f/2.8, 100mm & f/16), totalt 16 bilder til bruk i eksperimentet. For å samle inn mest mulig data fra deltagerne i eksperimentet ble hver vist fire bilder, et fra hver situasjon med forskjellige kombinasjoner av variablene. 18 deltagere var med i det endelige eksperimentet. 72 datapunkter ble samlet inn totalt.

Data ble samlet inn gjennom lydopptak og eye-tracking. Hver deltager ble plassert foran en skjerm med en eye-tracker plassert på undersiden og gitt en håndholdt mikrofon. Mens eksperimentet pågikk fikk deltagerne beskjed om å tenke høyt så tankene deres kunne fanges opp på lydopptaket. Eye-trackeren tok opp hvor de så på skjermen. Etter eksperimentet var gjennomført ble data fra lydopptaket delt opp i oppdagelser (insights), løsninger og spørsmål basert på et erfaringsbasert kodeskjema (emergent coding scheme). Eye-tracking data ble tatt opp i eksperimentet, men ble ikke brukt i resultatet av dette piloteksperimentet på grunn av vanskeligheter med å tolke data. Teknologi og kunnskap innen eye-tracking blir stadig bedre, og om det kommer noen store gjennombrudd i nær fremtid kan data hentes frem og analyseres.

Ingen statistisk signifikante resultater ble funnet etter å ha gjennomført en t-test på brennvidde og dybdeskarphet. Det ble funnet en indikasjon på statistisk forskjell på antall løsninger foreslått ved bruk av lengre brennvidder med signifikansverdi 0.894. Denne potensielle effekten korresponderer med en første hypotese H1: Lengre brennvidder fører til økte resultater siden oppmerksomhet blir mer jevnt fordelt mellom visuelle elementer.

Basert på resultater fra, og feilkilder i dette piloteksperimentet har forbedringsområder blitt foreslått for et andre eksperiment og blir nå tatt i bruk.

[b] Prosjektoppgave kan finnes i Appendix C

Acknowledgements
First of all I want to direct a big thanks to my supervisor, Professor Martin Steinert, for always being available to answer my questions despite an always triple-booked calendar. Thanks to Bruce Sampson from Department of Fine Arts for lending me his lenses and providing expertise in the field of photography. For providing some insights into the world of eye tracking, Moritz Mussgnug, Mila Vulchanova and Hendrik Eshuis have been a great help. I am also grateful to Stephanie Balters, Jørgen Andreas Bogen Erichsen, Carlo Kriesi, Kristoffer Slåttsveen, and last but not least, my dear Lise Simskar for their invaluable advice throughout the last semester.

Thank you all!

Preface
This is the master's thesis submitted spring 2015 by candidate Andreas Wulvik at Department of Product Development and Materials as the final part of the master's degree in Product Development and Production from NTNU. The topic is "Photography - Bridging the Gap between Needfinding and Visual Thinking", and will experimentally explore the effect of focal length and depth of field on the verbal ideation output from designers tasked with improving situations shown in pictures.

Table of Contents

List of Figures

List of Tables

Page intentionally left blank for formatting reasons

Page intentionally left blank for formatting reasons

1. Theory

1.1 Visual thinking

In his book “Visual Thinking” (1969), Rudolph Arnheim challenges the idea that perception and thinking are disconnected from each other. He also strongly argues against the assumption that language is a basis for perception and thinking. Robert McKim describes several representations of thoughts in his book “Experiences in Visual Thinking” (1972). He calls these different representations *vehicles*, and describes them as the “form” of thoughts. The different vehicles described in the book are: verbal language, non-verbal language e.g. mathematics, sensory imagery and feelings. Visual thinking concerns the representation of thoughts through the vehicle of sensory images. McKim writes that visual thinking is carried on by three kinds of visual imagery: What we *see*, what we *imagine*, and what we *draw*. These three kinds of imagery can be used independently, but true experts in visual thinking are able to use all three flexibly (McKim, 1972, p. 8–9). They have the ability to first *see* a problem from multiple angles, then *visualize* a number of potential solutions for what they see, and then *draw* their ideas in a few quick sketches to compare and evaluate later. The visual thinkers then move back and forth between these three modes until they solve the problem.

McKim argues the importance of visual thinking to support creativity through the example of perception (McKim, 1972, p. 25). Through perception we experience the world around us, and when thinking visually instead of verbally, we are able to perceive the raw material around us, opposed to labelling everything we see with language. By labelling everything we see around us, McKim expects stereotyped thinking as a result, constraining thinking to conventionalized meanings of words. The unlabelled nature of visual thinking should according to McKim have a salutary influence on thought processes, allowing us to explore ideas outside conventional labels.

1.2 Needfinding

Robert McKim, Rolf Faste and David Kelley from the Stanford joint product design program coined in the 80’ies the term “needfinding” as a method to study people in order to identify unmet and usually uncovered needs (Faste, 1987; Patnaik & Becker, 1999). These needs don’t necessarily have to be the end user’s. It can be equally important to find hitherto untargeted or unknown stakeholders and learn about their pains and needs. The true needs and “nuggets” a product or system can be built on are usually latent or hidden (Aldaz, Steinert, & Leifer, 2013). A favoured method for uncovering these needs is observation. Blessing and Chakrabarti (2009) writes that observation “can produce unadulterated, direct and potentially very rich descriptions of events and their context, because data is captured while the phenomena occur". Needfinding largely borrows from the in situ observation studies of ethnography and anthropology.

The limitation of observation is that the observer usually cannot capture and process all information in real-time. This makes retrospective data analysis central, as data can be viewed ex post (Patton, 2002). An appropriate medium has to be selected to capture the essence of the observation and supply data for the ex post analysis. In addition to allowing the observer to go through data from the observation to support his or her own thought processes, captured data can also be shared with participants not partaking in the observation. This can be equally important, since entire design teams seldom have the time or resources to go in the field for observation. The medium selected should therefore possess the ability to directly communicate information from the field, without the observer having to explain everything in detail, but rather allow the team members to form their own ideas and insights from the material. Acknowledging the importance of visual thinking, being able to experience the world without labelling and stereotyped thinking, pictures were chosen as the preferred medium to convey information from observations, allowing others to *see* for themselves.

2 Background

Previous to this master thesis, I submitted a paper called "Photography – a new tool in needfinding" to ICED2015 (International Conference of Engineering Design) together with PhD candidate Stephanie Balters and Professor Martin Steinert as my pre-master project. The idea of using photography as a tool in needfinding sprung out from a trip to Bangalore, India in October 2014. The goal of the trip was to do needfinding in the realm of hygiene products for a German company, more specifically how laundry is done in private homes and Dhobi Ghats (professional laundry facilities run by the Dhobi caste) in India. As an enthusiastic hobby photographer I took pictures of what I saw, keeping what I knew about visual thinking in mind, to better communicate my findings upon return to Norway. Being able to show the pictures when explaining my findings upon my return was really helpful. The pictures provided an anchor point for my explanations, giving the other team members an understanding of the situation quite quickly. A total of 381 pictures were taken during the trip to India, where 49 of them were selected for further use. After going through all the pictures, the idea emerged that pictures seem to possess different qualities in terms of the information they provide. Some were overview shots, providing lots of chaotic information to take in, while others were focused on a single object, directing all focus there. As a result *snapshots, emphasised pictures* and *illustrated pictures* were proposed as three specific picture types to be used in the needfinding process. These were presented along with an illustration of how the process could look like when using pictures consciously in the submitted paper (Wulvik et al., 2015), and are also described later in this chapter.

2.1 Snapshot

The goal of the snapshot is to supply the viewer with a problem context. It is meant to function as a deep dive into the real world where the needs and problems actually are. This picture type is detail-rich and open for interpretation, not focusing on any single object. The snapshot is meant as stimuli for the designer(s) to spark open-ended discussions, avoiding steering the viewer in any specific direction. The snapshot supports the early phase of the needfinding process (Wulvik et al., 2015), opening up the problem space as much as possible to capture all potential needs before focusing on any specific area..

Figure 2.1: Snapshots

2.2 Emphasized picture

The emphasized picture directs the attention of the viewer towards specific areas in the situation shown. The goal is to steer the attention towards sections selected by either the photographer or the needfinder based on their relevance for the project at hand. Attention can be controlled by both making conscious choices while taking the picture, e.g. tight framing and only leaving the interesting part in focus to put emphasis on certain areas, and through post-processing, e.g. background can be blurred, and distracting areas can either be cropped out or retouched away. The emphasized picture is meant to support the process of narrowing down the problem space and develop potentially valuable need statements further.

Figure 2.2: Emphasized pictures

2.3 Illustrated picture

Illustrated pictures are used when insights have been uncovered and the design team has reached consensus on which needs they are going to address with their solutions. This picture type strives to freeze information, making it as unambiguous as possible in order to avoid misunderstandings in a later stage on what the design team decided.

The illustrated picture is made by using existing snapshots or emphasized pictures, adding text and illustrations on top of them to give explanation to the insights found (Siddharth & Roy, 2015). In order to make information more clear, it is possible to darken unimportant areas to improve readability and direct focus by removing distracting elements.

This type of picture can function as handover from the needfinding phase, creating an anchor point for solution generation.

Figure 2.3: Illustrated pictures

2.4 Needfinding process

Combining the three picture types, *snaphots, emphasized pixtures* and *illustrated pictures*, with the iterative user-centred design process (Leifer & Steinert, 2011) and wayfaring (Steinert & Leifer, 2012), the photography needfinding process (Wulvik et al., 2015) in Figure 2.4 was proposed in the pre-master thesis. Each consecutive picture type is meant to support the process of narrowing down the problem space by selecting and refining information displayed iteratively. For each iteration, the design team can use the current picture as a starting point to form a need statement, and then use the contextual information in the picture to test the feasibility and urgency of the need, incorporating Thomke's & Fujimoto's "Design – Build – Test" cycles (2000) much in the same way as Leifer & Steinert do in their iterative user-centred design process. Inside each of these iterations, Alexander's idea of divergent and convergent modes of thinking is included (1964). When using the pictures, the viewer(s) should first try to explore as many ideas as possible, opening up the problem space. After this is done, the viewer(s) can start narrowing down the potential ideas based on feasibility, impact and other selection factors. These iterations run until the final need statement has been formed. With each cycle of the needfinding process, ambiguity is reduced, eliminating room for interpretation. In addition to be a working tool for generating and selecting ideas, the pictures in the needfinding process can also serve as documentation along the way, and be used as a basis for conveying findings to people outside the team.

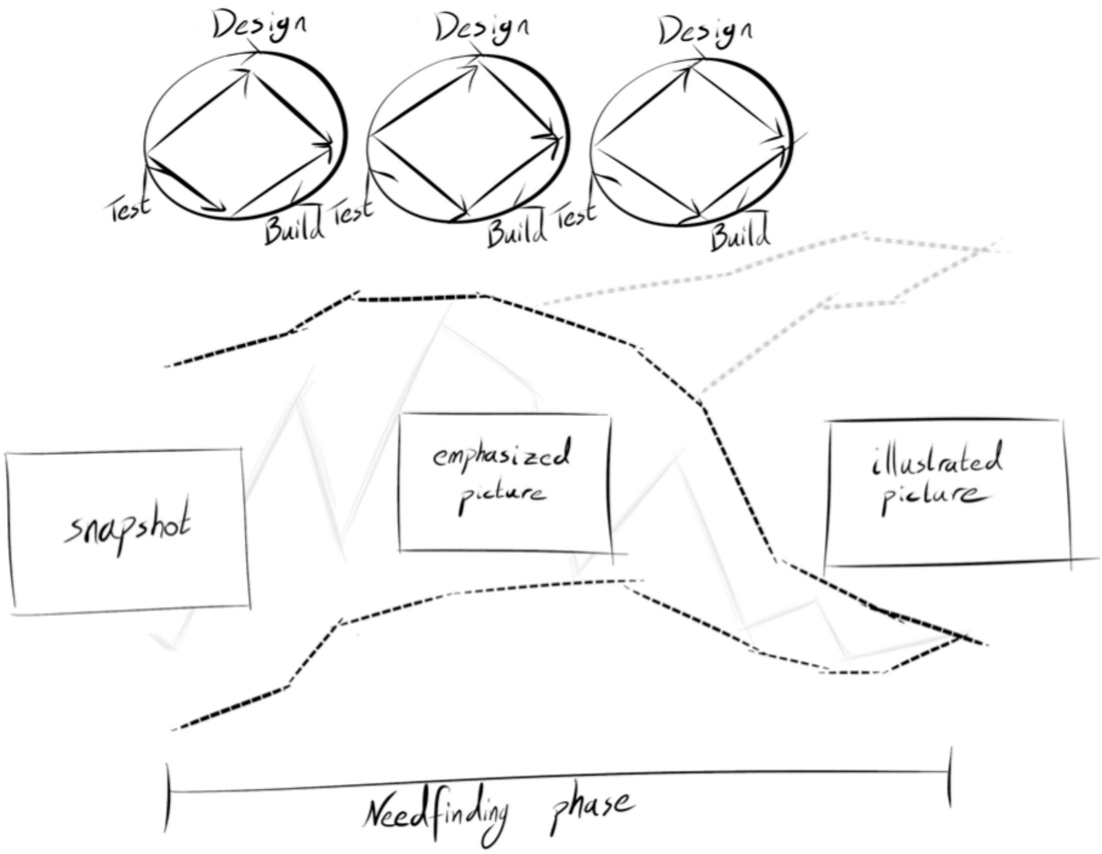


Figure 2.4: Photography needfinding process (Wulvik et al., 2015)

3 Experiment

Founding on the pre-master project, an explorative experiment to test the basic picture variables of *focal length* and *depth of field* is set up. The goal of this experiment is to gain some insight into the effects these variables have on the viewer's verbal ideation output in the needfinding phase of a project, and use this as a basis for developing the photography needfinding process further (Wulvik et al., 2015). In order to judge this effect, each participant's statements and gaze positions are selected as dependent variables. Statements will be captured in the form of audio recordings from each participant, and gaze position will be captured with eye tracking equipment. Recording and using audio to measure the effect of the experiment variables is based on the assumption that there is a one-to-one ratio between what participants think and what they say out loud. Before going more into depth on the experiment itself, the selected variables and their metrics are explained below.

3.1 Focal length

Focal length is a measure of the distance between image sensor and lens in order to focus sharply at infinity, i.e. objects at long distances (Jenkinson, 2011, p. 42). By changing the focal length, angle of view and distance in perspective are changed (Jenkinson, 2011). Lenses with long focal lengths have a narrow angle of view that can isolate subjects, and are called telephoto lenses. Short focal length lenses are called wide-angle lenses and have a wider angle of view that captures a larger portion of background, providing more context to the pictures.

I define distance in perspective as the perceived distance between objects along the depth axis in a picture. For short focal lengths, objects appear to be further away from each other. This is called *wide-angle stretching*. Longer focal lengths give the impression of objects being closer together than they really are. This is called *telephoto-compression*. These effects can be seen in Figure 3.1.

In the paper "Photography – a new tool in needfinding" (Wulvik et al., 2015), we hypothesise that changing focal length has a two-part influence on the viewer's output. First, by changing focal length, and therefore angle of view, the amount of background included in each picture varies. By including more background in the picture, the attention of the viewer keeps wandering and provides him or her with more insights. A narrower angle of view could then keep more focus on the object framed, concentrating all efforts on this. This could again lead to more specific and detailed output from the viewer. Second, the change in distance of perspective seems to move objects in the background back and forth respectively. Telephoto-compression brings the background seemingly closer to the foreground, making them appear more equal in size, and it was hypothesised that this leads to more equal treatment of foreground and background objects in terms of attention. The opposite effect is believed to be true for wide-angle-stretching; by creating a larger visual distance between the foreground and background, the objects in front appear larger, and should therefore receive more attention from the viewer. In terms of output, the compression should result in a more evenly spread spectrum of ideas, taking more of the picture into account. The stretching from using a wide angle should on the other hand underline the importance of the foreground object, and create ideas that are mainly focused around this.

The compression/stretching effect is believed to be the stronger of the two, and I state the first initial hypothesis as the following, H1: *Longer focal lengths provide more output because more equal attention is given to visual elements.*

Figure 3.1: Stretching/compression

3.2 Depth of field

Depth of field signifies how much of a picture that appears to be in focus (Pentland, 1987). In reality there is only an infinitely narrow plane in the picture that is in focus, but to the human eye, an area around this focal plane appears to be sharp. An object is defined as in focus when light rays reflected of the object converge on a single point on the image sensor of the camera. Light rays that are reflected from out of focus objects does not converge on one single point on the sensor, but rather hit it in a circle. This is called the *circle of confusion* (Figure 3.2), and is measured in millimetres. A point in space is considered sharp if the circle of confusion is so small that it is indistinguishable from a single point object in the final picture. The largest acceptable circle of confusion depends on several aspects. Print size, viewing distance and eyesight being the main aspects. Normally a circle of confusion of 0.025mm to 0.035mm is considered acceptable for full-frame image sensors.

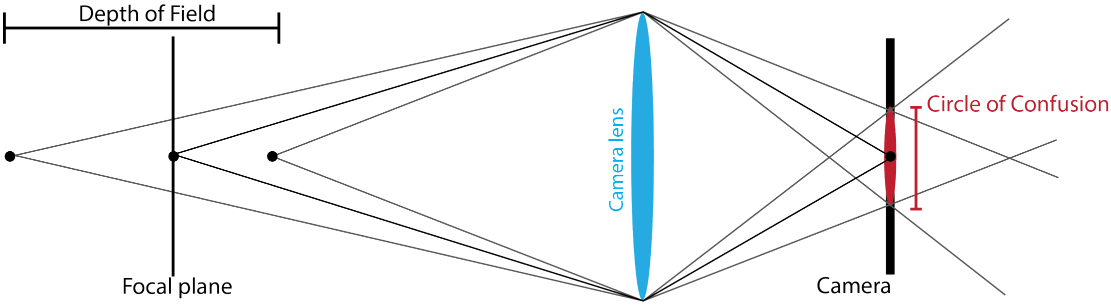


Figure 3.2: Illustration of DoF and circle of confusion

A *small* depth of field means that there is only a small area around the focal plane in the picture that is rendered sharply. Sharpness gradually decreases with distance from the focal plane as the size of the circles of confusion increase. The distance between the largest acceptable circles of confusion on each side of the focal plane, or two times the distance from the focal plane to the largest acceptable circle of confusion measures the depth of field.

A *large* depth of field means that a larger portion of the picture is rendered sharply, thus giving the impression that more of the image is in focus.

The main variable controlling depth of field in a picture is the effective aperture (Jenkinson, 2011, p. 39) of the lens used when taking the picture. Focal length also influences the depth of field, but on a smaller scale (Equation 2). The effective aperture can easiest be described as the relationship between the round opening letting in light to the image sensor and the focal length of the lens. This is a measurement of how much light is let in to the image sensor while camera shutter is open. Effective aperture is measured in f-stops. The f-stop is found by dividing the diameter of the light opening in the lens by its focal length (Equation 1). The denominator of the fraction gives the f-stop when the numerator is set to 1, so 25mm divided by 50mm is ½, resulting in an f-stop of 2.

$$\frac{1}{N} = \frac{d}{f} \tag{1}$$

Equation 1: f-stop (N) calculated by dividing light opening, d by focal length, f

The total depth of field can be calculated with Equation 2 shown below. It takes in the distance *u* from the image sensor to the focal plane, the f-stop *N*, circle of confusion *c* and focal length *f*. As shown in the table below, for constant magnification and circle of confusion, effective aperture is the main variable controlling depth of field.

$$u_{total} = \frac{2uNcf^2(u-f)}{f^4 - N^2c^2(u-f)^2} \tag{2}$$

Equation 2: Total DoF (Conrad, 2006, p. 8)

Depth of field	f/2.8	f/11 and f/16
24mm (1m focus distance)	0.22 m	1.28 m
100mm (4m focus distance)	0.22 m	1.05 m

Table 3.1: Calculation of depth of field for constant magnification

As can be seen from Table 3.1, a low f-stop number, or f-number (large light-opening relative to focal length) results in a small depth of field due to the wide spread of the light rays passing hitting the image sensor in relatively large circles of confusion (Figure 3.3). A high f-number (small light-opening relative to focal length) results in a large depth of field. When the light opening is small, light rays that pass through are more parallel, and thereby hit the image sensor in smaller circles of confusion.

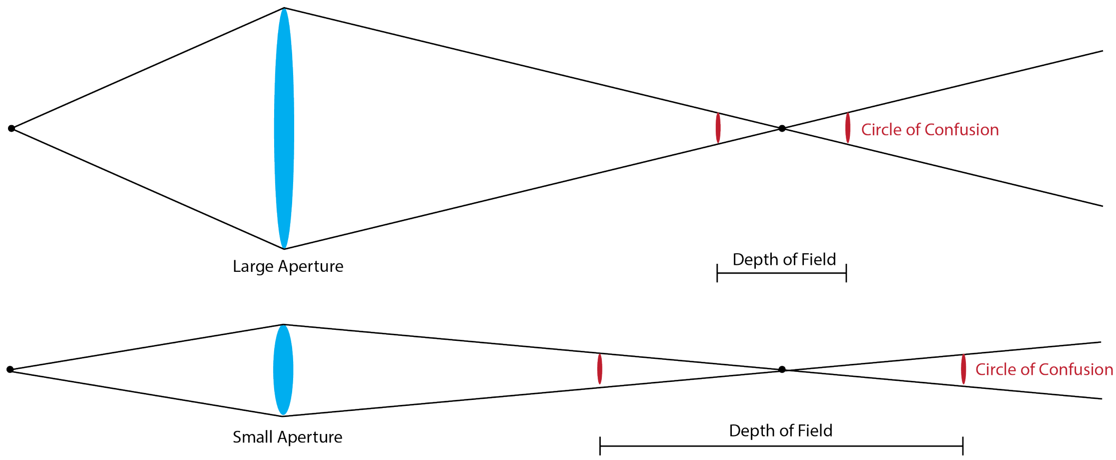


Figure 3.3: Influence of aperture on depth of field

As previously described in the paper by (Wulvik et al., 2015), depth of field is predicted to influence the amount of information the viewer extracts from a picture. The eye is attracted towards areas in sharp focus, and will therefore be emphasised by the viewer. Blurred areas are unpleasant to watch, making the viewers eye quickly move past them. This implicitly deems out-of-focus areas unimportant through the photographer's choice of not having them in focus. By controlling depth of field I hope to steer which areas in the picture, and thus the stimuli, the viewer is giving attention. From this, I hope to see a result in the type of output from the viewer when they are asked to ideate around each picture. Using a small depth of field, I hope to see output based around the small area in focus, converging on only a few areas of improvement. For larger depths of field, the wanted result is output of a broader nature. I hope to see ideas revolving around many different aspects in the picture, facilitating a divergent thinking mode, with a wide spectrum of ideas. Based on this, the second initial hypothesis is stated, H2: *A larger depth of field provides more output because of a larger area rendered sharp, and is thus given attention.*

3.3 Experiment metrics

When assigning values to focal length and depth of field i.e. aperture in this experiment, the visual effect and practicality of making test pictures must be balanced. For the experiment's sake, I would like to push each variable towards the extreme to gain as clear results as possible. For the practicality of making the test pictures, I would like the equipment to be as cheap and easy to use as possible. In the following chapters, I discuss the metrics to be used, taking both effect and practicality into consideration.

3.3.1 Selecting focal length

Values for focal length are set to 24mm and 100mm. 50mm is considered a "Normal" focal length on a full frame camera. This is because 50mm results in a perspective close to that of the human eye. 24mm was chosen as the extreme value because it is the widest focal length on full frame cameras that is close to distortion free. Wider focal lengths tend to create barrel-distortion approaching a fish-eye effect (Shah & Aggarwal, 1994). 100mm was chosen to balance the focal length of 24mm. The magnification between 24mm and 50mm is close to 2x, and similarly, the magnification between 50mm and 100mm is also 2x. Another reason for not using more extreme focal lengths is the feasibility of making test pictures of the same motive with both focal lengths i.e. more extreme values in focal length demands more space for the photographer to move back and forth when taking the test pictures.

3.3.2 Selecting aperture

Values for aperture, which controls depth of field, are set to f/2.8 for both 24mm and 100mm focal lengths and f/11 and f/16 for 24mm and 100mm respectively.
f/2.8 gives a fairly small depth of field, while it is at the same time possible to find lenses with this maximum aperture at fairly low cost, as it is maximum aperture that drives the cost of lenses. f/16 is set as the smallest aperture as smaller apertures start to render pictures unsharp due to diffraction (Rockwell, 2012). Because of different focal lengths, aperture has to be balanced to create the same depth of field. On the manual Nikon lenses used, f/11 and f/16 are consecutive apertures and gave the closest value for depth of field between focal lengths of 24mm and 100mm. The metrics can be seen in Figure 3.3 above.

3.4 Experimental setup

The two variables focal length and depth of field can be set up graphically as the two dimensions of this experiment (see Figure 3.4). Each quadrant describes a combination of two variables that will be represented by a picture.

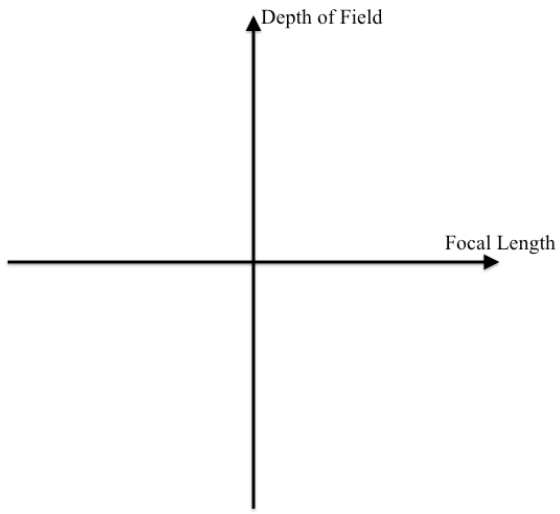


Figure 3.4: Focal Length vs. Depth of Field

The goal of the experiment is to investigate the impact of *visual* effects from changing focal length and depth of field, not changes in *information*. The same content should therefore be in all the pictures to keep the information constant. Scenarios for the test pictures should be selected to allow for maximal visual effect of experiment variables i.e. compression/stretching and large/small depth of field. To achieve this, the most important aspect is to have enough visual depth in the picture. This allows the effects of telephoto compression/wide-angle stretching and variable depth of field to be clearly visible. In order to activate test subjects while viewing each test picture, there should be an easily identifiable situation for the subject to quickly understand.

In preparation to the experiment, eight different picture situations with four pictures each were made, 32 pictures in total. Four of these situations were chosen for the experiment, a total of 16 pictures to use as stimuli. The four situations were selected based on their difference in content, striving to avoid any learning effects between them. Each participant is shown one picture from each of these situations, each with a different combination of the picture metrics i.e. 100mm & f/2.8, 100mm & f/16, 24mm & f/2.8 and 24mm & f/16. The four situations selected are: charging, coffee, flooding and bicycle. Each situation with corresponding four pictures can be seen below in this chapter.

In the first picture situation a person is trying to charge his VW eUp! with the quick-charger on campus (see Figure 3.5). He has some trouble with this as the car and the quick-charger use different standards for the charging plug.

Figure 3.5: Charging situation

In the coffee situation, a student is buying coffee with coins at the department vending machine while another student is waiting in line (see Figure 3.6).

Figure 3.6: Coffee situation

In the flooding situation, a group of fellow students were supposed to simulate building a wall of sandbags to protect against flooding (see Figure 3.7).

Figure 3.7: Flooding situation

For the bicycle situation a fellow student wore a hoodie while acting as he stole a bicycle (his own) by using the front wheel's quick-release mechanism (see Figure 3.8).

Figure 3.8: Bicycle situation

To get as much data as possible from each participant, one picture with a unique metric combination from each situation were grouped in a total of four picture sets labelled 1-4 (see Table 3.2). The experiment participants were shown these four different sets based on the order they did the experiment i.e. participant 1 got set 1, participant 6 got set 2 etc. Showing participants all situations and all metric combinations should contribute to reduce sources of error in output due to personal differences in participants and differences in situations.

Picture set	Charging	Coffee	Flooding	Bicycle
1	24mm & f/2.8	100mm & f/16	24mm & f/11	100mm & f/2.8
2	24mm & f/11	100mm & f/2.8	100mm & f/16	24mm & f/2.8
3	100mm & f/16	24mm & f/2.8	100mm & f/2.8	24mm & f/11
4	100mm & f/2.8	24mm & f/11	24mm & f/2.8	100mm & f/16

Table 3.2: Picture set metrics

A quiet room with the participant facing empty white walls was chosen as the location for the experiment. A 27" 4k monitor set to 1920x1080 pixels was placed on a high table with a barstool for the participants to sit on. The eye tracker used for the experiment was placed under the monitor on a small tripod that made it easy to set up for each new participant. The eye tracker was of the brand The Eye Tribe (The Eye

Tribe, 2015), which is a $99 budget tracker based on more high-end infrared technology. Behind the participant's left shoulder, a Nikon D5300 DSLR connected with a hand-held microphone is set up to capture the screen activity of the experiment for synchronisation purposes and to capture audio from each participant. The DSLR was framed on the monitor, so that the participants were not visible in the recording. Each participant was tasked to hold the microphone with one hand and to use the computer mouse to click trough the experiment with the other. The setup can be seen in Figure 3.9.

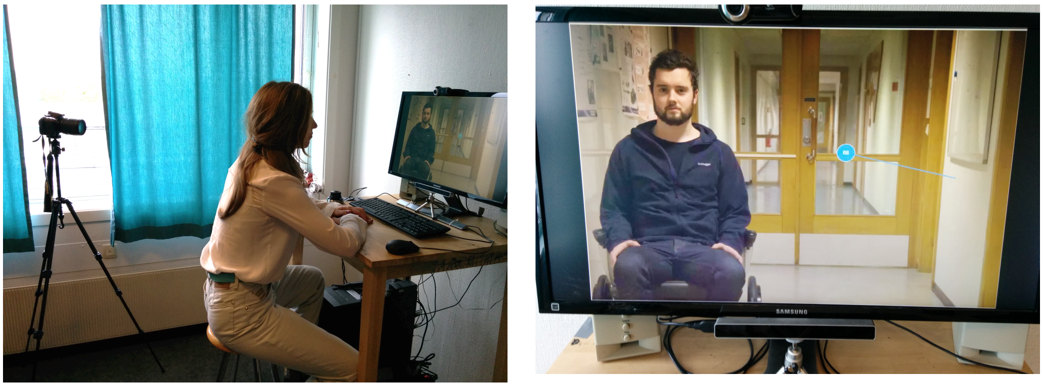


Figure 3.9: Physical experiment setup

The experiment was run through the eye trackers proprietary software platform, EyeProof (*EyeProof*, 2015), which is a cloud based service in beta version. Each experiment set was uploaded in separate "studies", with the four different pictures plus one picture that was common for all the sets. This picture, common to all the sets was shown first to all participants to make them comfortable and familiar with the task given before they were shown the real test pictures. This was done to reduce the risk of participants not understanding the task they were given, and thereby lose valuable data. In the studies defined in EyeProof, the four experiment pictures were set to appear in random order. The pictures uploaded to the test platform were 1348x900 pixels due to constraints in maximum file size from EyeProof. The task given to each participant before they were shown the pictures was the following:

In this experiment, you will be shown 5 pictures. Each picture will be displayed for 40 seconds. Your task is: "In how many ways would you improve the situation shown?" Immediately start thinking of how you can improve it. Please say out loud what you are thinking, as we will record audio. After the first picture, you may contact the experiment supervisor if you have any questions. You may at any point in time abort the experiment should you feel uncomfortable or have any other reason. There will be no repercussions for doing so. To start the experiment, please press "ok"

The time of 40 seconds was chosen after a few quick tests with members of the local research group. A time of 30 seconds seemed too short, as they didn't have time to

take in the information in the pictures. One the other hand, I didn't want to give the participants too long time either, as I wanted to see what the immediate reactions they had to the picture metrics were. The assumption is that if the time is too long, the participants will have time to look at the entire picture no matter what the effect of the test variables, and therefore any potential results would be obscured.
Before watching each picture, the participant is reminded of their task. A screen with the caption "In how many ways would you improve the situation shown? Press OK to *continue"* is shown, and they have to press an ok-button at the bottom of the screen to continue the experiment.
While each picture was shown on the screen, the participant was supposed to speak into the microphone and say everything that came to their mind. They were allowed to ask the experiment supervisor if they had any questions, or if something was unclear. This was done to avoid unusable data sets due to the participant not understanding the task at hand. The experiment supervisor's job was to keep the participants motivated, but not to tell them if they were doing it right or wrong. After the experiment has finished, the participants are given the following message:

Thank you for your participation! Please fill out the questionnaire provided by the experiment supervisor. We kindly ask you not to discuss the experiment with other people. This is in order to not influence them, should they participate in the experiment.

Prior to the experiment, all the participants had to sign a release form accepting that the data gathered is being used in research, and that the participant could at any time withdraw their consent to be a part of the experiment. After signing the release form, they were asked to answer the question *"How energetic do you feel? 1-low energy and 5-high energy"* before starting the experiment itself. After the experiment was finished, the participants were asked to fill out a questionnaire giving some background information about them (see Appendix B – Experiment questionnaire). This information could then later be used to look for patterns in the data collected.

4 Analysis

The experiment was run with 23 participants, all of them from Department of Product Development and Materials at NTNU. 21 participants were mechanical engineering students, one a PhD candidate and two professors at the department. Participant 11-15 had to be rejected due to missing audio recordings because the microphone was turned off. This left data from 18 participants, each shown four pictures. 72 data points were captured in total. The outcome of the experiment was two data streams – The camera recording and data captured through EyeProof. In order to combine these two data streams, Camtasia Studio (*Camtasia*, 2015) was chosen as appropriate software. The decision was based on usability, video editing capabilities and the possibility of doing screen recording.

The first step of combining data was to replay scanpaths in real-time in EyeProof, This was recorded with Camtasia in approximately 40-second clips. Scanpaths were chosen as representation of gaze data based on a conversation with Mortiz Mussgnug from ETH in Switzerland (Skype, 17.4.2015). Scanpaths (Chapter 4.2.3) shows the sequence of fixations/gazes, and makes it possible to look for connections between gaze position/time and audio output. Heat maps (Chapter 4.2.3) could have been used to represent the data, but the sequence of gazes is lost, and it is therefore difficult to compare audio output and gaze position. The recording had to be done manually, so the start and end points of each clip were not precisely at the beginning and end of the scanpath. This had to be trimmed later in Camtasia Studio.

Each participant was shown five pictures (including the "warm-up" picture), so five clips were made each, and with 18 participants, 90 video clips had to be processed.

When all scanpath clips had been recorded, each participant's clips were imported into Camtasia Studio along with the recording from the DSLR. The video and audio track were separated in the software after import. The first step was to cut the audio into clips matching each scanpath. If the participant kept talking after the 40-second duration of each picture, the audio clip was extended to cover their statements. All the audio in-between not contributing relevant data was removed to make the interpretation work easier. The video track was used to synchronise scanpaths and audio clips. The scanpath clips were on top of the video stream from the DSLR, so after the clips were in sync, the video was removed to save space in the final exported clip. Adding a 15-second caption of the experiment instruction given to each participant, and a 5-second task prompt before each picture put the finishing touches to the exported video. The timeline of the finished video clips can be seen below in Figure 4.1

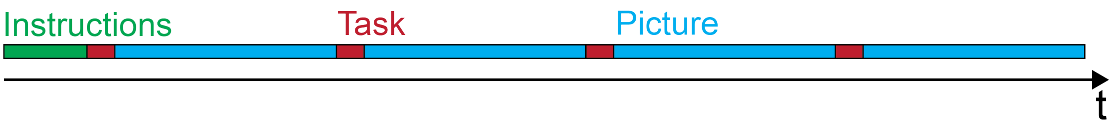


Figure 4.1: Video time-line

4.1 Breaking down audio into quantitative data

No pre-defined word protocol was used when processing the experiment data. An emerging coding scheme was instead chosen, as this was a pilot, designed to get an overview of the effects of focal length and depth of field. Listening through the audio data from the eight first participants, a first attempt at a coding scheme was proposed - counting all general statements. All observations or proposals were considered general statements, no matter if they are on topic or not. This was to look after any overall effect on output from the participant. The first eight participants were coded, and results were reviewed. The conclusion was that counting general statements were too vague, and not defined enough to make good arguments towards the effect of the picture variables. Three new categories for coding emerged after this discussion: Insights, solutions and questions. This coding scheme should present more robust ways of interpreting the audio source material, and also be more relevant for the realm of engineering design. *Insights* are statements of processed information, i.e. the participant makes a conclusion based on what he or she sees in the picture. One example could be the realisation of the bike in Figure 4.2 being stolen. *Solutions* are statements where the participant is making a suggestion on how to improve the situation in some way or other. Suggesting building more bike racks in the picture could be such a solution. All questions were put in the *question* category, no matter the subject. This was to make an attempt to connect with Ozgur Eris' (2003) research on how questions can drive the design process. Unfortunately there were too few questions recorded in this experiment to make such a connection. Having the new categories defined, the recordings were listened through again, coding material that fit one of the three categories along the way. The recordings were listened through twice to ensure nothing was overlooked.

Insights (needs) and solutions often come in pairs ("He needs to put on his shoes, so he should use a shoehorn"). I chose to count these separately, as one need can lead to multiple solutions, or one solution can cater to multiple needs.

Figure 4.2: Bicycle situation

4.2 Eye tracking

4.2.1 How does an eye tracker work?

Most modern eye tracking systems measure point-of-regard by the "corneal-reflection/pupil-centre" method (J. H. Goldberg & Wichansky, 2003). The eye tracker used for this is pilot experiment, The Eye Tribe, is a remote eye tracker i.e. mounted under the pc monitor as opposed to wearable trackers mounted on the participant's head. The tracker illuminates the participant's eyes with unobtrusive infrared light and captures reflections with an infrared camera. Light is reflected back, making the pupil appear as a bright disc ("known as the bright pupil effect" (Poole & Ball, 2006)). In addition to the light reflected back through the pupil, the corneal reflection appears as a small sharp glint (see Figure 4.3).

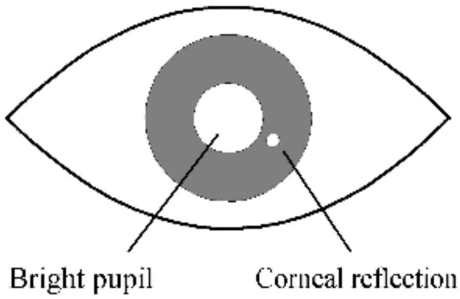


Figure 4.3: Corneal reflection and bright pupil as seen by the eye tracker (Poole & Ball, 2006)

After the tracker has identified the pupil and corneal reflection, a vector between them is measured, and the eye's point-of-regard can be found (see Figure 4.4).

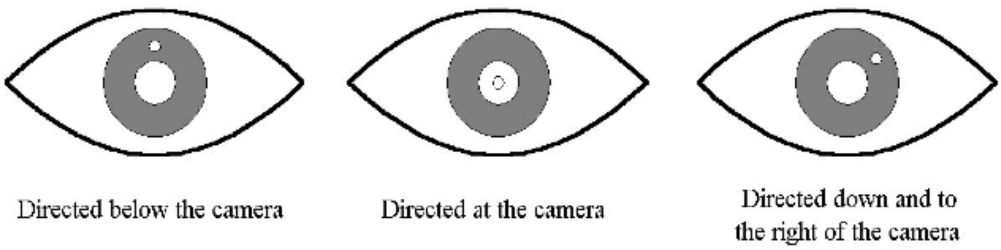


Figure 4.4: Corneal reflection position changing according to point-of-regard (Redline & Lankford, 2001)

Before gaze position can be recorded, the tracker must be calibrated. This is done by displaying a series of dots with known coordinates on the screen, and links them with the corresponding pupil/corneal reflection relationship. This is usually done between

9 and 13 times in a grid pattern to calibrate the tracker accurately over the entire screen (J. H. Goldberg & Wichansky, 2003).

4.2.2 Output data

Multiple metrics can be extracted from eye tracking experiments (Jacob & Karn, 2003; Poole & Ball, 2006). Three commonly used metrics are fixations, gazes and saccades.

4.2.2.1 Fixation

A fixation is a length of time where the angular position of the eye is relatively stationary, i.e. gaze position is constant (Jacob & Karn, 2003; Poole & Ball, 2006). This usually means movements less than two degrees of the eye for 100-200 milliseconds at low velocities. There is no one single way of identifying fixations, but rather multiple different algorithms in use today (Salvucci & Goldberg, 2000), although a minimum of 100 milliseconds is advised by Inhoff & Radach (1998). This means that it is difficult to compare eye-tracking experiments, since they can be based on different fixation identification algorithms.

The detection of a fixation (or rather end of a fixation) is also depending on the technical equipment used in eye tracking experiments. With increasing resolution and sample rates, it is possible to detect smaller saccades occurring between fixations (Jacob & Karn, 2003). In the end, the definition of fixations depends on the goal of each eye tracking experiment, and is set by the researcher.

Interpretation of fixations can vary a lot depending on context. When browsing a web-page, a large amount of fixations in a certain area can indicate interest, or it can be a sign that the area is complex and hard to understand (Jacob & Karn, 2003). When searching for information, a large number of fixations in an area can be a indication of difficulty to recognize the target item (Jacob & Karn, 2003). The duration of a fixation seems to be connected with processing time of the item viewed, and thus point towards that shorter fixations indicate ease of processing information. Shorter fixations therefore appear to be more meaningful for the viewer than longer fixations (Joseph H. Goldberg & Kotval, 1999).

4.2.2.2 Gaze

A gaze is the cumulative duration and average location of a series of fixations within a predefined area of interest (Jacob & Karn, 2003). The small saccadic movements in-between fixations is also counted towards the total gaze-duration. A fixation outside the predefined area defines the end of the gaze (Jacob & Karn, 2003). Number of gazes and gaze duration can be an indication of interest in various objects. Linking back to the discussion of the meaning of fixations, this has to be evaluated with the context in mind.

4.2.2.3 Saccade

Saccades are rapid eye-movements between fixations. Each saccade usually lasts 20-35 milliseconds (Poole & Ball, 2006). The saccadic movement is to fast to process information, so we do not get any insight towards what the eye is looking at. Even though the eye does not process information during saccades, recursive saccades can provide information about difficulty of processing information, i.e. saccades jumping back to previously looked-upon areas might be an indication that the viewer is having trouble processing what they see (Poole & Ball, 2006).

4.2.3 Interpretation of data

Modern eye tracking software is able to automatically convert raw gaze data into fixations and saccades. The software typically uses eye position or eye velocity to interpret the data (Jacob & Karn, 2003). As mentioned earlier, the minimum time threshold for defining a fixation may vary between different software, and can also be changed inside some as well. This makes it difficult to compare data collected across experiments, as what is defined as a fixation in one experiment could have been discarded in the other.

Two popular ways of visualising gaze data are *heatmaps* (Bojko, 2009) and *scanpaths* (Joseph H. Goldberg & Kotval, 1999). *Heatmaps* are graphical representations of fixation density, displayed as a colour overlay to the original picture (Figure 4.5a). A common colour coding is that colour temperature increases with density, i.e. red is high density and blue is low density. Having the fixation density overlaid the stimulus makes it easy to read. A heatmap shows the total fixations over a given time interval, but gives no indication of the sequence of and time spent on each fixation. It can be used to show areas of interest to the viewer, and to show data to others in a simple way. *Scanpaths* are defined as a chain of saccade-fixation-saccade sequences. Scanpaths are often displayed as multiple circles of various sizes (depending on fixation time), connected with lines representing saccades (Figure 4.5b). Fixations close together in both time and distance are often grouped together as gazes in these representations. Scanpaths have the advantage over heatmaps that they can show the sequence of fixations, allowing the researcher to gain some insight in what the participant is looking at while doing other things, and look for patterns in the gaze data.

Figure 4.5a & 8.3b: a) Heatmap and b) Scanpath

Even though the software can convert the raw gaze data into more usable variables, the researcher still has to interpret the output. Data can be analysed either *top-down* – based on cognitive theory or a design hypothesis, or *bottom-up* – observing data without any predefined theory linking eye movements to cognitive activity (Joseph H. Goldberg, Stimson, Lewenstein, Scott, & Wichansky, 2002). Following a top-down approach based on cognitive theory can for example interpret longer fixations as difficulty interpreting the particular objects meaning. A top-down design hypothesis could be that using a specific colour on a web page banner will increase the number of times it is looked at. A bottom-up approach has no predefined way of interpreting the gaze data, but rather looks for patterns and asks why they are looking there.

There are large differences in eye movements between participants doing identical tasks, making it very difficult to compare inter-participant data (J. H. Goldberg & Wichansky, 2003). In addition, it is very hard to understand if a fixation is a result of difficulty of interpretation, special interest, or just a “dead stare” while the participant is thinking of something else.

4.2.4 Analysis of experiment data

Participants were given the task of improving the situation shown, so that their eye movements could be attributed to cognitive activity (Just & Carpenter, 1976). As previously in this chapter, there are lots of difficulties connected with interpreting gaze data. Long fixations can be interpreted several ways – Difficulty of interpretation, special interest, or something as simple as a “dead stare” when the participant is resting his eyes somewhere while thinking of something completely different. Goldberg & Wichansky (2003) also writes that there are large differences in eye movements between participants, even when doing identical tasks. This makes comparison of scanpaths across participants very difficult.

The challenges of interpreting gaze data are so large that it is not considered worth the resources to spend time making sense of the data. Especially when inter-participant analysis has to be done across different pictures. Should there be some great insights into how gaze data collected can be interpreted in a valuable way, the data is captured and available for future analysis.

5 Results

All data gathered from the participants were entered into a spreadsheet to be interpreted in IBM SPSS Statistics Version 21.0. A partial example of the data can be seen below in Table 5.1.

ID	Sex	Age	Education	Occupation	Picture Set	Situation	Focal Length	Aperture	Insights	Solutions	Questions	Total
3	Female	24	Ind.eco/mech.eng	Student	3	Charging	100mm	Small	2	1	0	3
3	Female	24	Ind.eco/mech.eng	Student	3	Coffee	24mm	Large	2	3	0	5
3	Female	24	Ind.eco/mech.eng	Student	3	Flooding	100mm	Large	2	2	0	4
3	Female	24	Ind.eco/mech.eng	Student	3	Bicycle	24mm	Small	3	2	0	5
4	Male	24	mech.eng	Student	4	Charging	100mm	Large	0	1	0	1
4	Male	24	mech.eng	Student	4	Coffee	24mm	Small	0	3	0	3
4	Male	24	mech.eng	Student	4	Flooding	24mm	Large	1	2	0	3
4	Male	24	mech.eng	Student	4	Bicycle	100mm	Small	1	1	0	2
5	Male	54	PhD	Professor	1	Charging	24mm	Large	3	0	1	4
5	Male	54	PhD	Professor	1	Coffee	100mm	Small	0	1	0	1
5	Male	54	PhD	Professor	1	Flooding	24mm	Small	1	0	1	2
5	Male	54	PhD	Professor	1	Bicycle	100mm	Large	3	0	0	3

Table 5.1: Example of raw data from SPSS

5.1 Comparison of groups – Picture set

Before testing the effect of our experiment variables, focal length and depth of field, I wanted to investigate if there was any statistically significant difference in dependent variables between picture sets by using a one-way ANOVA (ANalysis Of VAriance) test. The ANOVA test is used to determine if there is any statistical difference between the means of two or more independent groups. It is not possible to determine which groups are statistically different from each other. This has to be done with Post Hoc tests, e.g. a Tukey test. The independent variable must be categorical, while the dependent variable(s) must be continuous. The different picture sets are set as the independent categorical variable, and number of insights, solutions, questions and the total sum of these are set as dependent variables.

When doing an ANOVA test, there are three assumptions that should be checked: First, the dependent variable should be approximately normal distributed for each of the independent variable groups. Second, there should be no significant outliers among the dependent variables in each group of independent variables, as these can have a large effect on the calculated mean and standard deviation of the group, which in turn can influence the result of the statistical tests. Third, the variance in each group of independent variables is equal (homogeneity of variance).

5.1.1 ANOVA Test

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Questions	1	16	.31	.479	.120	.06	.57	0	1
	2	16	.00	.000	.000	.00	.00	0	0
	3	20	.00	.000	.000	.00	.00	0	0
	4	20	.05	.224	.050	-.05	.15	0	1
	Total	72	.08	.278	.033	.02	.15	0	1

Table 5.2: Descriptives for dependent variable "questions" from the ANOVA test between picture sets

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Questions	Between Groups	1.113	3	.371	5.747	.001
	Within Groups	4.388	68	.065		
	Total	5.500	71			

Table 5.3: Results for dependent variable "questions" from the ANOVA test between picture sets

There were no statistically significant differences in the dependent variables insights, solutions and total between the different picture sets with a significance > 0.05, i.e. the null hypothesis of no statistical difference between test groups is kept due to a probability higher than 95% that the null hypothesis is true.

The significance given in the table is the probability of the null hypothesis (no statistical difference between group mean values) being true. The converse of this is that the probability of the actual hypothesis (statistical difference between group mean values) is 1 – significance value. For a significance of 0.05, the probability of the actual hypothesis being true is 95% (1 – 0.05 = 0.95). If the probability of the hypothesis is below 95% it is not considered significant.

The amount of questions was statistically significantly different between the different picture sets with a significance of 0.001, way below the threshold of significance of 0.05. To investigate this difference in amount of questions between the picture sets further, a Tukey Post Hoc test was conducted.

5.1.2 Tukey Post Hoc Tests

Multiple Comparisons

Tukey HSD

Dependent Variable	(I) Picture set	(J) Picture set	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Questions	1	2	.313*	.090	.005	.08	.55
		3	.313*	.085	.003	.09	.54
		4	.263*	.085	.015	.04	.49
	2	1	-.313*	.090	.005	-.55	-.08
		3	.000	.085	1.000	-.22	.22
		4	-.050	.085	.936	-.27	.17
	3	1	-.313*	.085	.003	-.54	-.09
		2	.000	.085	1.000	-.22	.22
		4	-.050	.080	.925	-.26	.16
	4	1	-.263*	.085	.015	-.49	-.04
		2	.050	.085	.936	-.17	.27
		3	.050	.080	.925	-.16	.26

*. The mean difference is significant at the 0.05 level.

Table 5.4: Tukey Post Hoc test for difference between picture sets with regard to dependent variable "questions"

From the Tukey Post Hoc test a statistically significant difference between picture set 1 & 2 (significance value = 0.005), 1 & 3 (sig. val. = 0.003) and 1 & 4 (sig. val. = 0.015) is found, all well below the threshold of 0.05 to reject the null hypothesis that there is no significant difference between the groups.

Going back to the raw data, a total of six questions in total recorded in the experiment, made by four participants. Five of these were from participants shown picture set 1, and two of the participants asked two questions each. Three questions were asked about the charging situation, two about flooding and one about the coffee situation. There is an equal amount of questions asked after seeing large and small depth of field pictures. The amount of questions asked sorted by focal lengths of 24mm and 100mm are four and two respectively. None of the questions asked were about the same topic. Based on this information, there is no obvious reason for the significant difference in questions asked based on picture sets. The difference seems to be mainly attributed to the individual participants' inclination to ask questions.

The negative result for statistically significant differences in output between picture sets is good for the experiment. This means that significant differences in output when testing for focal length and depth of field can most likely be attributed to those variables, and not to effects from picture sets.

5.2 Comparison of groups – Situation

An ANOVA test was also conducted to check for statistically significant differences in dependent variables between the different picture situations. This is done for the same reason as between picture sets – a negative result for the ANOVA (keeping the null hypothesis of no difference between means of test groups) strengthens the confidence in results for tests on focal length and depth of field being valid.

5.2.1 ANOVA Test

Descriptives									
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Total	Charging	18	2.72	1.227	.289	2.11	3.33	0	5
	Coffee	18	3.11	1.779	.419	2.23	4.00	1	7
	Flooding	18	3.33	1.455	.343	2.61	4.06	0	6
	Bicycle	18	4.06	1.662	.392	3.23	4.88	1	7
	Total	72	3.31	1.589	.187	2.93	3.68	0	7

Table 5.5: Descriptives for dependent variable "total" from the ANOVA test between picture situations

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Total	Between Groups	16.944	3	5.648	2.366	.079
	Within Groups	162.333	68	2.387		
	Total	179.278	71			

Table 5.6: Results for dependent variable "total" from the ANOVA test between picture situations

There were no statistically significant differences in the dependent variables insights, solutions and questions between the different situations with all significance values > 0.05, i.e. the null hypothesis of no statistical difference between test groups is kept due to a probability higher than 95% that the null hypothesis is true.

Although not statistically significant, there is an indication that there may be a difference in the total count of statements between the situations with significance 0.079. A Tukey Post Hoc test was conducted to investigate this further.

5.2.2 Tukey Post Hoc Tests

Multiple Comparisons

Tukey HSD

Dependent Variable	(I) Situation	(J) Situation	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Total	Charging	Coffee	-.389	.515	.874	-1.75	.97
		Flooding	-.611	.515	.637	-1.97	.75
		Bicycle	-1.333	.515	.056	-2.69	.02
	Coffee	Charging	.389	.515	.874	-.97	1.75
		Flooding	-.222	.515	.973	-1.58	1.13
		Bicycle	-.944	.515	.267	-2.30	.41
	Flooding	Charging	.611	.515	.637	-.75	1.97
		Coffee	.222	.515	.973	-1.13	1.58
		Bicycle	-.722	.515	.502	-2.08	.63
	Bicycle	Charging	1.333	.515	.056	-.02	2.69
		Coffee	.944	.515	.267	-.41	2.30
		Flooding	.722	.515	.502	-.63	2.08

Table 5.7: Tukey Post Hoc test for difference between picture situations with regard to dependent variable "total"

Assessing the Tukey Post Hoc test a significance level of p = 0.056 (94,4% probability that there is a statistical difference between test groups) was found, strongly indicating that there is a statistical difference between the charging and bicycle picture types. When comparing the charging and bicycle pictures (see Figure 5.1) two things come to mind. First, the bicycle situation is probably much more familiar for the participants than the charging situation, and therefore they have more ideas to contribute with. Second, one can argue that there is less content in the charging pictures compared to the bicycle pictures. This could also be a reason for why there are fewer contributions from the participants in total.

Figure 5.1: Charging and bicycle situation

Even though a strong indication towards a statistical difference between the charging and bicycle picture situations were found, there were no statistically significant results from the ANOVA test on differences between mean values of picture situations. Based on this, differences between test groups when testing focal length and depth of field can most likely be attributed to the experiment variables and not influences from picture situations and picture sets.

5.3 Comparison of experiment variables – Focal length

After establishing that there were no statistically significant differences in dependent variables between test groups in neither picture sets nor picture situations, the effect of focal length and depth of field can be tested. First, an independent-samples t-test was conducted to determine if there were differences in the dependent variables insights, solutions, questions and total count of statements between pictures shot with focal lengths of 24mm and 100mm. As with the ANOVA test, there are three assumptions that should be met when doing an independent samples t-test: First, the dependent variable should be approximately normal distributed for each of the independent variable groups. Second, there should be no significant outliers among the dependent variables in each group of independent variables, as these can have a large effect on the calculated mean and standard deviation of the group, which in turn can influence the result of the statistical tests. Third, the variance in each group of independent variables is equal (homogeneity of variance). Data are mean ± standard deviation, unless otherwise stated.

Group Statistics					
	Focal Length	N	Mean	Std. Deviation	Std. Error Mean
Solutions	24mm	36	1.50	1.082	.180
	100mm	36	1.92	1.079	.180

Table 5.8: Descriptives for dependent variable "solutions" for test on focal length

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Solutions	Equal variances assumed	.364	.548	-1.636	70	.106	-.417	.255	-.925	.091
	Equal variances not assumed			-1.636	69.999	.106	-.417	.255	-.925	.091

Table 5.9: Independent samples t-test for difference in means when testing for solutions based on focal length

No statistically significant differences were found between pictures with short and long focal lengths ($p > 0.05$), so the null hypothesis of no statistically significant difference between test groups is kept. Even though not statistically significant, there is an indication that the number of solutions proposed is influenced by the focal length used with a significance of 0.106 (89.4% probability that there is a statistical difference between test groups). If this is the case, pictures shot with a focal length of 100mm have more solutions proposed (1.92 ± 1.08) compared to pictures shot with 24mm (1.50 ± 1.08), a difference in mean values of 0.42 in favor of pictures shot with longer focal lengths.

The indication toward a higher output of solutions for pictures shot with a focal length of 100mm corresponds with H1: *Longer focal lengths provide more output because more equal attention is given to visual elements.*

5.4 Comparison of experiment variables - Aperture

An independent-samples t-test was conducted to determine if there were differences in the dependent variables insights, solutions, questions and total count of statements between pictures shot with small and large apertures i.e. large and small depth of field.

No statistically significant differences were found between pictures with small and large aperture for any of the dependent variables with significance > 0.05 i.e. less than 95% probability that there is a statistical difference between mean values of test groups. This result means that the null hypothesis of no significant difference between mean values of test groups is kept, and initial hypothesis H2: *A larger depth of field* *provide more output because of a larger area rendered sharp, and is thus given attention*, is rejected.

5.5 Comparison of experiment variables – inside picture situations – Aperture

In addition to explore the effect of focal length and depth of field across different picture situations, the local effects inside each of the situations were also tested. Similar to the test on focal length and depth of field across picture situations, an independent-samples t-test was conducted to determine if there were any differences in the dependent variables insights, solutions, questions and total count of statements between pictures shot with small and large apertures i.e. large and small depth of field inside each situation.

5.5.1 Situation = Charging

Group Statistics[a]					
	Aperture	N	Mean	Std. Deviation	Std. Error Mean
Questions	Large / Small DoF	9	.33	.500	.167
	Small / Large DoF	9	.00	.000	.000
a. Situation = Charging					

Table 5.10: Descriptives for dependent variable "questions" for test on depth of field in charging situation

Independent Samples Test[a]										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Questions	Equal variances assumed	64.000	.000	2.000	16	.063	.333	.167	-.020	.687
	Equal variances not assumed			2.000	8.000	.081	.333	.167	-.051	.718
a. Situation = Charging										

Table 5.11: Independent samples t-test for difference in means when testing for questions based on depth of field in the charging situation

No statistically significant differences between group mean values were found when conducting an independent samples t-test between pictures with small and large apertures of the charging situation with significance value > 0.05, i.e. less than 95% probability that there is a statistical difference between group mean values.

The independent samples t-test returned a significance value of 0.081, i.e. 91.9% probability that there is a statistical difference in mean values between the small and large aperture groups. Although not statistically significant, this is an indication towards a statistical difference between the groups. As previously discussed in chapter 5.1, there is no apparent reason for the frequency of questions being higher in one group or another, and based on the low amount of data, one should be careful with drawing any conclusions from this result. Due to this the null hypothesis of no statistically significant difference between test groups is kept, and H2: *A larger depth of field provide more output because of a larger area rendered sharp, and is thus given attention*, is rejected.

5.5.2 Situation = Coffee, Flooding and Bicycle

No statistically significant differences were found when conducting an independent samples t-test between pictures with small and large apertures inside the following three situations: coffee, flooding and bicycle with significance values > 0.05, i.e. less than 95% probability that there is a statistical difference between group mean values. Based on this, the null hypothesis of no statistically significant difference is kept and H2: *A larger depth of field provide more output because of a larger area rendered sharp, and is thus given attention*, is rejected.

5.6 Comparison of experiment variables – Split by situation – Focal Length

An independent-samples t-test was conducted to determine if there were any differences in the dependent variables insights, solutions, questions and total count of statements between pictures shot with focal lengths of 24mm and 100mm inside each of the picture situations.

5.6.1 Situation = Charging

Group Statistics[a]

	Focal Length	N	Mean	Std. Deviation	Std. Error Mean
Solutions	24mm	8	.88	.835	.295
	100mm	10	1.80	1.135	.359

a. Situation = Charging

Table 5.12: Descriptives for dependent variable "solutions" for test on focal length in charging situation

Independent Samples Test[a]										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Solutions	Equal variances assumed	.402	.535	-1.922	16	.073	-.925	.481	-1.945	.095
	Equal variances not assumed			-1.991	15.924		-.925	.465	-1.910	.060
a. Situation = Charging										

Table 5.13: Independent samples t-test for difference in means when testing for solutions based on focal length in the charging situation

No statistically significant differences were found when conducting the independent samples t-test between pictures with short and long focal lengths in the charging situation with significance values > 0.05, i.e. less than 95% probability that there is a statistical difference between group mean values.

There is an indication that there is a statistical difference between the means of each test group with regards to the number of solutions proposed with a significance level of 0.073, i.e. 92.7% probability that there is a statistical difference between the group mean values of pictures shot with focal lengths of 24mm and 100mm. If this is the case, pictures shot with a focal length of 100mm have more solutions proposed (1.80 ± 1.13) compared to pictures shot with 24mm (0.88 ± 1.83), a difference in mean values of 0.92 in favor of pictures shot with longer focal lengths.

This indication of a difference in number of solutions proposed between pictures shot with focal lengths of 24mm and 100mm in the charging situation corresponds with H1: *Longer focal lengths provide more output because more equal attention is given to visual elements.*

5.6.2 Situation = Coffee, Flooding and Bicycle

No statistically significant differences were found when conducting an independent samples t-test between pictures with short and long focal lengths inside the following three situations: coffee, flooding and bicycle with significance values > 0.05, i.e. less than 95% probability that there is a statistical difference between group mean values. Based on this, the null hypothesis of no statistically significant difference is kept, and H1: *Longer focal lengths provide more output because more equal attention is given to visual elements* is rejected.

Page intentionally left blank for formatting reasons

6 Conclusion and reflection

No statistically significant results on the effect of changing focal length and depth of field were found in this explorative experiment. There is some indication towards the effect of focal length, where the overall test on number of solutions proposed for 24mm vs. 100mm returned a significance of 0.894 and mean difference -0.417. This suggests that pictures shot with a 100mm lens elicits a slightly higher output of solutions than pictures shot with a 24mm lens. The same indicative result can be found in the test of number of solutions for focal length inside the charging situation. A significance of 0.927 was calculated with mean difference -0.925, also supporting the indication that a longer focal length elicits more solutions from the participants.

When considering the results, one should keep in mind potential error sources that could have influenced the results. One area that could have influenced the results is the content of the experiment pictures themselves. All the experiment pictures has been shot around Department of Product Development and Materials at NTNU, and are thus familiar to all the participants. This familiarity with the situations shown could contribute to even out the differences between test variable groups as the participants already have all information in their mind without really looking at the picture. One example of this is that several of the participants looking at the bicycle situation state that the parking spot shown is the only one under roof. This is not evident from the picture itself, and is thus source of error when assessing the effect of picture variables.

Another source of error in the pictures could be that there is not enough "depth" in the pictures to fully appreciate the effects of different focal lengths and depths of field. The only test picture with any real visual depth is the flooding scenario. There is open space behind all the subjects, and they are relatively far away from each other. The other three pictures have all interesting components arranged more or less in one plane, not showing any great effect of blurred versus sharp background from changes in depth of field or stretching/compression due to changes in focal length.

One other thing I noticed during the experiment was that participants didn't always understand what they were supposed to improve in the situation they saw – they didn't really see the problem. Insights from needfinding show that problems and user needs aren't necessarily obvious from the beginning, and one needs time to let impressions sink in before being able to identify them (Aldaz et al., 2013). For this experiment, I would have liked to have situations with clearer needs or problems for the participant to address due to the limited amount of time at their disposal.

Looking at the experimental setup itself in hindsight, it seems that the participants could have come up with more ideas if they had more time to let information sink in, as well as overcoming the initial obvious ideas. However, it is also possible that differences are evened out when the participants have enough time to look at the entire picture and overcome the effect of the picture variables tested. Increasing

amounts of data to be processed with longer times is also something to keep in mind, so small representative slices of data are preferred whenever possible.
Using an emergent coding scheme for analysis of audio data was suitable for this pilot experiment, as no obvious way of coding was in place beforehand. The iterative process of first looking at general statements and later decide to code for insights, solutions and questions made sure that the coded variables really were relevant. Categories are now in place and it is quite clear what to look for. A pre-defined word protocol should be in place for future experiment rounds. This should help remove subjective influences from the coder, and thereby make the dependent variables more robust for analysis.
Lastly, the statistical analyses were not corrected for outliers in the data sets and violations of the assumption that dependent variables were normally distributed. Outliers in data sets can have a large impact in calculated mean value, and will therefore strongly influence the significance level of tests comparing group means. For a later more refined experiment setup, the results should be corrected for any violations of assumptions to increase confidence in the results.

7 Further work

Even though no statistically significant results came out of this pilot experiment, there has been a great learning outcome both in terms of how to improve the experiment for the next round and learning new research tools for me personally. Because of all the insights towards how to do an experiment and potential influences on results identified, I believe a second experiment should be conducted with all this new information in mind.

7.1 Participant selection

Participants should come from different backgrounds, not just mechanical engineers from Department of Product Development and Materials. A broader selection might show differences in output based on participants' educational background, e.g. how do engineers compare to artists and people from humanities? Also, because of the rejected participants, the pilot was very unbalanced in terms of genders. Out of the 18 participants analysed, only three were female. In the next round, it would be preferable to have a more even gender distribution.

7.2 Experiment stimuli – Pictures

New pictures should be made for the second experiment. The new picture situations should be made in such a way that participants don't have any "background" knowledge about the situation outside what they can see in the picture and infer from that information. Detailed knowledge about the situation shown in the pictures might erase the differences of changing picture variables, and should be avoided as far as possible. This means that everyday situations could be problematic as stimuli, and more unusual situations are preferable for the next experiment iteration.

In order to obtain the different visual effects of changing focal length and depth of field, there has to be enough visual depth in the picture. Interesting content in the pictures should be found in at least two different planes, preferably three or more. This is to show the effect of blurred versus sharp areas of the picture and stretching versus compression between the planes when changing focal lengths. Lots of space to arrange pictures is important when considering this, and outdoor areas are preferable.

New pictures should show more obvious needs and problems for the participants to solve in order to activate their thoughts right away. This should be an entry point, but not necessarily the only aspect to ideate around. The challenge when making this picture is to not steer the participant into one specific solution, but rather giving them a focus point to anchor their divergent thought processes.

One concrete idea for new experiment pictures is to stand on a ship bridge with information panels in the foreground, the ship's deck in the middle ground, and a harbour in the background. This should also fulfil the requirements of having an unfamiliar situation for the participants.

7.3 Data capture

For the next iteration of the experiment a hands-free microphone should be used. There are two main reasons for this. A hands-free microphone can either be in the form of a headset, or a lapel microphone. Both have the advantage of being less intrusive than the large hand-held microphone used in the pilot experiment, hopefully making the participants more at ease. A new microphone would also pick up less ambient noise, and record a more high quality audio track. This should make it easier to analyse the audio data later.

Synchronising data streams manually as done in the first experiment is really time consuming. This entails having to export data from multiple sources and then combining and aligning them in separate software. A lot of work and time could be spared if screen activity, audio and gaze data were captured on one platform, e.g. iMotions Attention Tool and Tobii Studio (*iMotions - Attention tool*, 2015, *Tobii Studio*, 2015).

It could be interesting to record facial expressions with a webcam to gain some insights into participants' state of mind, e.g. frustration or confusion during the experiment. This type of analysis is included in the iMotions software suite, and already in use at this department.

7.4 Coding scheme

Before analysing data from the second experiment, a pre-defined coding scheme should be in place. The Language Acquisition and Language Processing Lab at Department of Linguistics and Literature, NTNU have expertise in this area, and will be a valuable partner when creating this scheme for coding audio data. If possible, someone else than the experiment supervisor should do the actual coding to avoid adding extra meaning to what participants state in the audio recordings.

7.5 Statistical analysis

Assuming that the second experiment runs smoothly and the proposed changes have the desired effect, the results should be run through the same statistical tests as in the pilot. Assumptions of no significant outliers, normally distributed dependent variables and homogeneity of variances should be checked before running the independent t-test and one-way ANOVA. If these are violated, the following tests should compensate for violations with appropriate methods.

8 References

Aldaz, G., Steinert, M., & Leifer, L. J. (2013). Instrumenting the user: Needfinding by combining data logging and immediate self-reporting. In *DS 75-7: Proceedings of the 19th International Conference on Engineering Design (ICED13), Design for Harmonies, Vol. 7: Human Behaviour in Design, Seoul, Korea, 19-22.08. 2013*.

Alexander, C. (1964). *Notes on the Synthesis of Form* (Vol. 5). Harvard University Press.

Arnheim, R. (1969). *Visual thinking*. Univ of California Press.

Blessing, L. T., & Chakrabarti, A. (2009). *DRM, a design research methodology*. Springer.

Bojko, A. A. (2009). Informative or misleading? Heatmaps deconstructed. In *Human-Computer Interaction. New Trends* (pp. 30–39). Springer.

Camtasia. (2015). (Version 8.5) [Windows]. Retrieved from https://www.techsmith.com/camtasia.html

Conrad, J. (2006). Depth of Field in Depth. Retrieved from http://www.largeformatphotography.info/articles/DoFinDepth.pdf

Eris, Ö. (2003). Asking generative design questions: a fundamental cognitive mechanism in design thinking. In *DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design, Stockholm*.

EyeProof. (2015). [Cloud]. Retrieved from eyeproof.net

Faste, R. A. (1987). *Perceiving needs*. SAE Technical Paper.

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, *24*(6), 631–645.

Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002). Eye tracking in web search tasks: design implications. In *Proceedings of the 2002 symposium on Eye tracking research & applications* (pp. 51–58). ACM.

Goldberg, J. H., & Wichansky, A. M. (2003). *Eye tracking in usability evaluation: A Practitioner's Guide. En: Hyona, J., Radach, R., Duebel, H (Eds.). The mind's eye: cognitive and applied aspects of eye movement research (pp. 573-605)*. Boston, North-Holland/Elesevier.

iMotions - Attention tool. (2015). (Version 5.4) [Windows]. iMotions. Retrieved from http://imotionsglobal.com/software/add-on-modules/attention-tool-remote-eye-tracking-module/

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes.

Jacob, R. J., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind*, *2*(3), 4.

Jenkinson, M. (2011). *The Portrait Photography Course: Principles, practice, and techniques The essential guide for photographers*. Peachpit Press.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4), 441–480.

Leifer, L. J., & Steinert, M. (2011). Dancing with ambiguity: Causality behavior, design thinking, and triple-loop-learning. *Information, Knowledge, Systems Management*, *10*(1), 151–173.

McKim, R. H. (1972). Experiences in visual thinking.

Patnaik, D., & Becker, R. (1999). Needfinding: the why and how of uncovering people's needs. *Design Management Journal (Former Series)*, *10*(2), 37–43.

Patton, M. (2002). *Qualitative Research and Evaluation Methods* (3rd Edition). London: Sage.

Pentland, A. P. (1987). A new sense for depth of field. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (4), 523–531.

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. *Encyclopedia of Human Computer Interaction*, *1*, 211–219.

Redline, C. D., & Lankford, C. P. (2001). Eye-movement analysis: a new tool for evaluating the design of visually administered instruments (paper and web). *Proceedings of the Survey Research Methods Section of the American Statistical Association, Available at Www. Amstat. org/Sections/Srms/Proceedings/y2001/Proceed/00248. Pdf*.

Rockwell, K. (2012). Diffraction. Retrieved June 7, 2015, from http://www.kenrockwell.com/tech/diffraction.htm

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications* (pp. 71–78). ACM.

Shah, S., & Aggarwal, J. K. (1994). A simple calibration procedure for fish-eye (high distortion) lens camera. In *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on* (pp. 3422–3427). IEEE.

Siddharth, & Roy, S. (2015). ID49: Research in Visual Ethnography Focusing on Markets of Kanpur (Vol. 1). Presented at the ICoRD'15 Research into Design Across Boundaries, CPDM, IISc, Bangalore: Springer.

Steinert, M., & Leifer, L. J. (2012). "Finding One"s Way': Re-Discovering a Hunter-Gatherer Model based on Wayfaring. *International Journal of Engineering Education*, *28*(2), 251.

The Eye Tribe. (2015, May 25). Retrieved May 25, 2015, from http://theeyetribe.com/

Thomke, S., & Fujimoto, T. (2000). The Effect of "Front-Loading" Problem-Solving on Product Development Performance. *Journal of Product Innovation Management*, *17*(2), 128–142.

Tobii Studio. (2015). (Version 3.3) [Windows]. Tobii. Retrieved from http://www.tobii.com/en/eye-tracking-research/global/products/software/tobii-studio-analysis-software/

Wulvik, A., Balters, S., & Steinert, M. (2015). Photography - A new tool in Needfinding. In *Photography - A new tool in Needfinding*. Milano, Italy.

Page intentionally left blank for formatting reasons

9 Appendix A – Statistical tests

9.1 Comparison of groups – Picture set

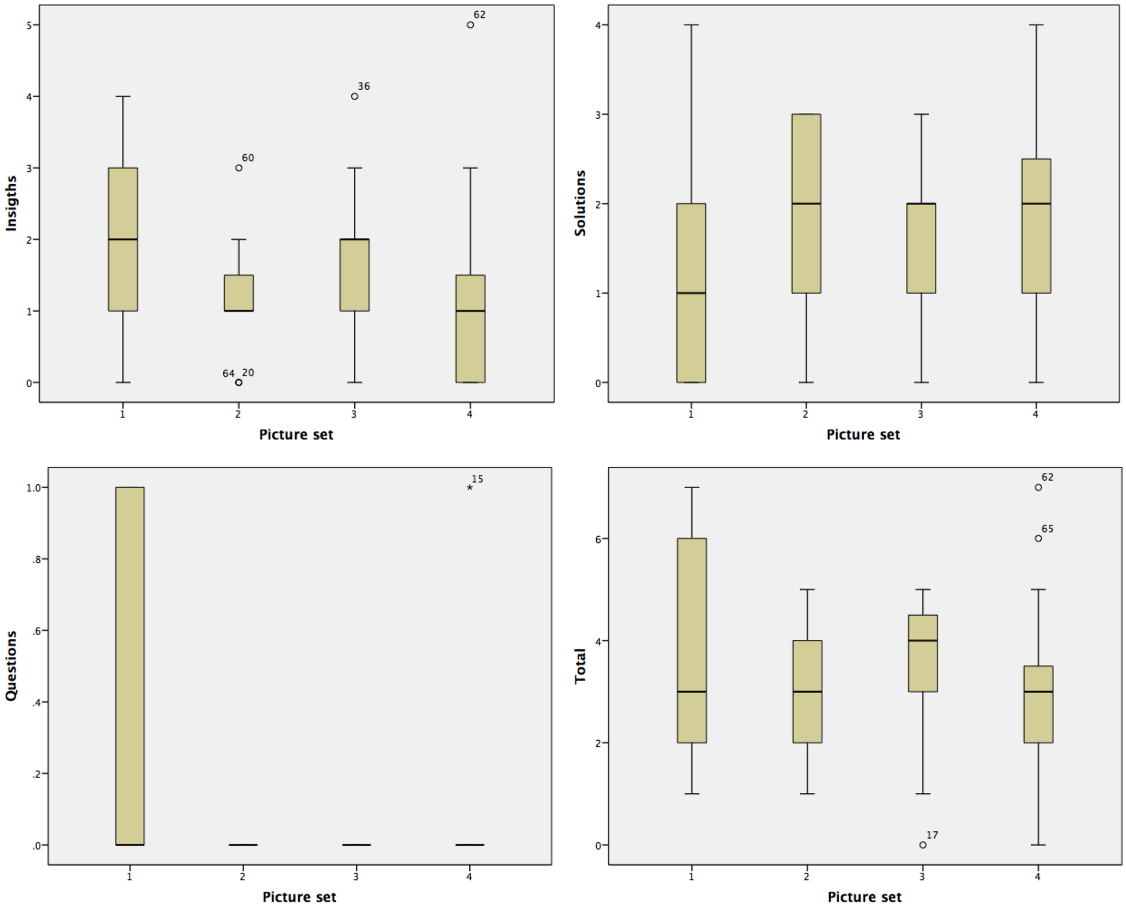
Case Processing Summary							
	Picture set	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	1	16	100.0%	0	0.0%	16	100.0%
	2	16	100.0%	0	0.0%	16	100.0%
	3	20	100.0%	0	0.0%	20	100.0%
	4	20	100.0%	0	0.0%	20	100.0%
Solutions	1	16	100.0%	0	0.0%	16	100.0%
	2	16	100.0%	0	0.0%	16	100.0%
	3	20	100.0%	0	0.0%	20	100.0%
	4	20	100.0%	0	0.0%	20	100.0%
Questions	1	16	100.0%	0	0.0%	16	100.0%
	2	16	100.0%	0	0.0%	16	100.0%
	3	20	100.0%	0	0.0%	20	100.0%
	4	20	100.0%	0	0.0%	20	100.0%
Total	1	16	100.0%	0	0.0%	16	100.0%
	2	16	100.0%	0	0.0%	16	100.0%
	3	20	100.0%	0	0.0%	20	100.0%
	4	20	100.0%	0	0.0%	20	100.0%

9.1.1 Test of normality

Tests of Normality[b,c]							
	Picture set	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	1	.189	16	.128	.908	16	.108
	2	.312	16	.000	.846	16	.012
	3	.260	20	.001	.901	20	.044
	4	.312	20	.000	.797	20	.001
Solutions	1	.188	16	.133	.871	16	.028
	2	.234	16	.020	.856	16	.017
	3	.316	20	.000	.844	20	.004
	4	.184	20	.076	.922	20	.108
Questions	1	.431	16	.000	.591	16	.000
	4	.538	20	.000	.236	20	.000
Total	1	.222	16	.034	.874	16	.031
	2	.167	16	.200*	.931	16	.253
	3	.266	20	.001	.841	20	.004
	4	.262	20	.001	.916	20	.084

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction
b. Questions is constant when Picture set = 2. It has been omitted.
c. Questions is constant when Picture set = 3. It has been omitted.

9.1.2 Boxplots to check for outliers



9.1.3 Test for homogeneity of variances

Test of Homogeneity of Variances				
	Levene Statistic	df1	df2	Sig.
Insigths	1.377	3	68	.257
Solutions	1.792	3	68	.157
Questions	37.361	3	68	.000
Total	3.090	3	68	.033

9.1.4 ANOVA Test

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Insigths	1	16	1.88	1.258	.315	1.20	2.55	0	4
	2	16	1.13	.806	.202	.70	1.55	0	3
	3	20	1.85	.988	.221	1.39	2.31	0	4
	4	20	1.20	1.281	.287	.60	1.80	0	5
	Total	72	1.51	1.138	.134	1.25	1.78	0	5
Solutions	1	16	1.44	1.365	.341	.71	2.16	0	4
	2	16	1.81	.981	.245	1.29	2.34	0	3
	3	20	1.75	.851	.190	1.35	2.15	0	3
	4	20	1.80	1.196	.268	1.24	2.36	0	4
	Total	72	1.71	1.093	.129	1.45	1.97	0	4
Questions	1	16	.31	.479	.120	.06	.57	0	1
	2	16	.00	.000	.000	.00	.00	0	0
	3	20	.00	.000	.000	.00	.00	0	0
	4	20	.05	.224	.050	-.05	.15	0	1
	Total	72	.08	.278	.033	.02	.15	0	1
Total	1	16	3.63	2.062	.515	2.53	4.72	1	7
	2	16	2.94	1.237	.309	2.28	3.60	1	5
	3	20	3.60	1.353	.303	2.97	4.23	0	5
	4	20	3.05	1.638	.366	2.28	3.82	0	7
	Total	72	3.31	1.589	.187	2.93	3.68	0	7

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Insigths	Between Groups	8.736	3	2.912	2.379	.077
	Within Groups	83.250	68	1.224		
	Total	91.986	71			
Solutions	Between Groups	1.550	3	.517	.422	.738
	Within Groups	83.325	68	1.225		
	Total	84.875	71			
Questions	Between Groups	1.113	3	.371	5.747	.0
	Within Groups	4.388	68	.065		
	Total	5.500	71			
Total	Between Groups	6.840	3	2.280	.899	.4
	Within Groups	172.438	68	2.536		
	Total	179.278	71			

9.1.5 Tukey Post Hoc Tests

Multiple Comparisons							
Dependent Variable	(I) Picture set	(J) Picture set	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Insigths	1	2	.750	.391	.231	-.28	1.78
		3	.025	.371	1.000	-.95	1.00
		4	.675	.371	.274	-.30	1.65
	2	1	-.750	.391	.231	-1.78	.28
		3	-.725	.371	.216	-1.70	.25
		4	-.075	.371	.997	-1.05	.90
	3	1	-.025	.371	1.000	-1.00	.95
		2	.725	.371	.216	-.25	1.70
		4	.650	.350	.256	-.27	1.57
	4	1	-.675	.371	.274	-1.65	.30
		2	.075	.371	.997	-.90	1.05
		3	-.650	.350	.256	-1.57	.27
Solutions	1	2	-.375	.391	.773	-1.41	.66
		3	-.313	.371	.834	-1.29	.67
		4	-.363	.371	.763	-1.34	.62
	2	1	.375	.391	.773	-.66	1.41
		3	.063	.371	.998	-.92	1.04
		4	.013	.371	1.000	-.97	.99
	3	1	.313	.371	.834	-.67	1.29
		2	-.063	.371	.998	-1.04	.92
		4	-.050	.350	.999	-.97	.87
	4	1	.363	.371	.763	-.62	1.34
		2	-.013	.371	1.000	-.99	.97
		3	.050	.350	.999	-.87	.97
Questions	1	2	.313*	.090	.005	.08	.55
		3	.313*	.085	.003	.09	.54
		4	.263*	.085	.015	.04	.49
	2	1	-.313*	.090	.005	-.55	-.08
		3	.000	.085	1.000	-.22	.22
		4	-.050	.085	.936	-.27	.17
	3	1	-.313*	.085	.003	-.54	-.09
		2	.000	.085	1.000	-.22	.22
		4	-.050	.080	.925	-.26	.16
	4	1	-.263*	.085	.015	-.49	-.04
		2	.050	.085	.936	-.17	.27
		3	.050	.080	.925	-.16	.26
Total	1	2	.688	.563	.616	-.80	2.17
		3	.025	.534	1.000	-1.38	1.43
		4	.575	.534	.705	-.83	1.98
	2	1	-.688	.563	.616	-2.17	.80
		3	-.663	.534	.604	-2.07	.74
		4	-.113	.534	.997	-1.52	1.29
	3	1	-.025	.534	1.000	-1.43	1.38
		2	.663	.534	.604	-.74	2.07
		4	.550	.504	.695	-.78	1.88
	4	1	-.575	.534	.705	-1.98	.83
		2	.113	.534	.997	-1.29	1.52
		3	-.550	.504	.695	-1.88	.78

*. The mean difference is significant at the 0.05 level.

9.2 Comparison of groups - Situation

Case Processing Summary							
	Situation	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	Charging	18	100.0%	0	0.0%	18	100.0%
	Coffee	18	100.0%	0	0.0%	18	100.0%
	Flooding	18	100.0%	0	0.0%	18	100.0%
	Bicycle	18	100.0%	0	0.0%	18	100.0%
Solutions	Charging	18	100.0%	0	0.0%	18	100.0%
	Coffee	18	100.0%	0	0.0%	18	100.0%
	Flooding	18	100.0%	0	0.0%	18	100.0%
	Bicycle	18	100.0%	0	0.0%	18	100.0%
Questions	Charging	18	100.0%	0	0.0%	18	100.0%
	Coffee	18	100.0%	0	0.0%	18	100.0%
	Flooding	18	100.0%	0	0.0%	18	100.0%
	Bicycle	18	100.0%	0	0.0%	18	100.0%
Total	Charging	18	100.0%	0	0.0%	18	100.0%
	Coffee	18	100.0%	0	0.0%	18	100.0%
	Flooding	18	100.0%	0	0.0%	18	100.0%
	Bicycle	18	100.0%	0	0.0%	18	100.0%

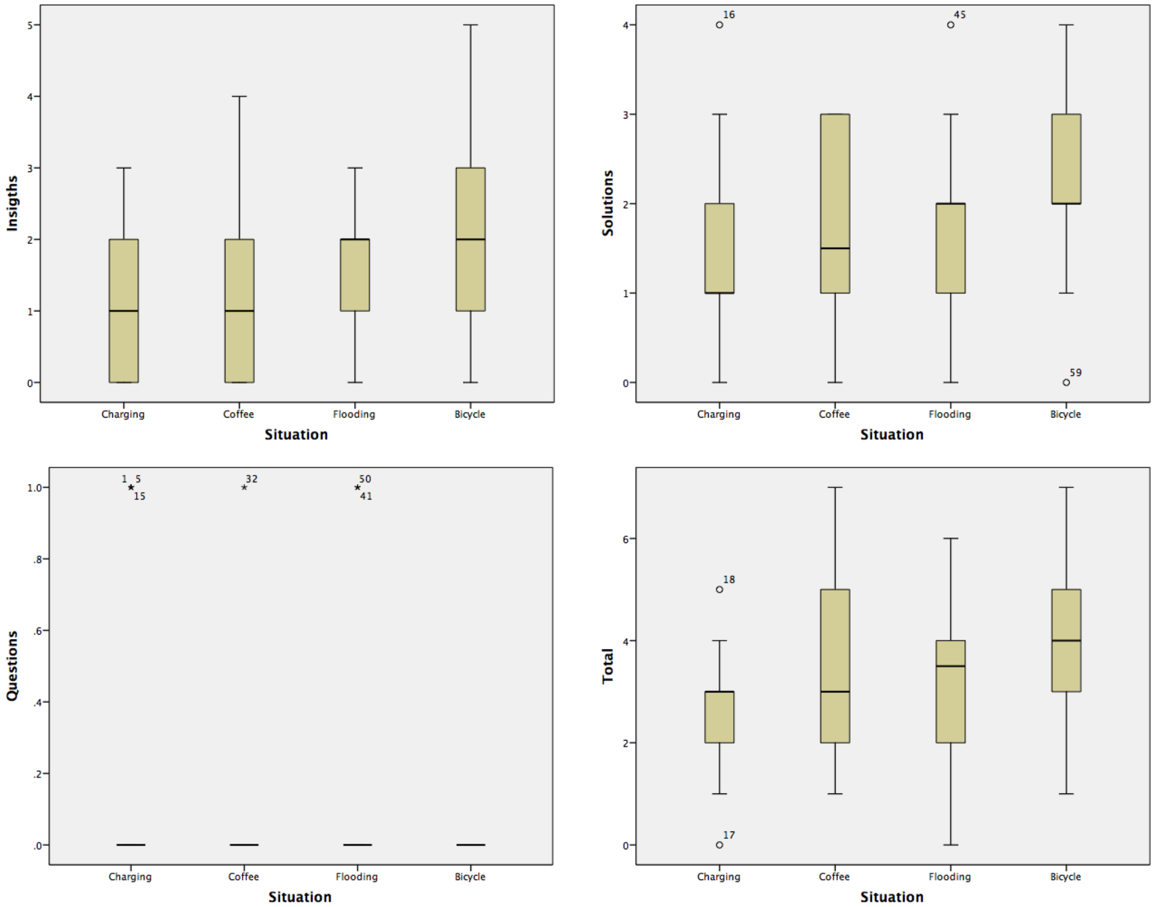
9.2.1 Test of normality

Tests of Normality[b]							
	Situation	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	Charging	.234	18	.010	.873	18	.020
	Coffee	.234	18	.010	.886	18	.033
	Flooding	.254	18	.003	.882	18	.028
	Bicycle	.202	18	.049	.925	18	.161
Solutions	Charging	.195	18	.070	.894	18	.045
	Coffee	.221	18	.020	.848	18	.008
	Flooding	.232	18	.012	.907	18	.078
	Bicycle	.265	18	.002	.902	18	.062
Questions	Charging	.501	18	.000	.457	18	.000
	Coffee	.538	18	.000	.253	18	.000
	Flooding	.523	18	.000	.373	18	.000
Total	Charging	.256	18	.003	.925	18	.159
	Coffee	.178	18	.135	.910	18	.087
	Flooding	.212	18	.031	.914	18	.101
	Bicycle	.182	18	.119	.952	18	.455

a. Lilliefors Significance Correction

b. Questions is constant when Situation = Bicycle. It has been omitted.

9.2.2 Boxplots to check for outliers



9.2.3 Test for homogeneity of variances

Test of Homogeneity of Variances				
	Levene Statistic	df1	df2	Sig.
Insigths	1.410	3	68	.247
Solutions	.640	3	68	.592
Questions	5.985	3	68	.001
Total	1.534	3	68	.214

9.2.4 ANOVA Test

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Insigths	Charging	18	1.17	.985	.232	.68	1.66	0	3
	Coffee	18	1.39	1.243	.293	.77	2.01	0	4
	Flooding	18	1.56	.856	.202	1.13	1.98	0	3
	Bicycle	18	1.94	1.349	.318	1.27	2.62	0	5
	Total	72	1.51	1.138	.134	1.25	1.78	0	5
Solutions	Charging	18	1.39	1.092	.257	.85	1.93	0	4
	Coffee	18	1.67	1.138	.268	1.10	2.23	0	3
	Flooding	18	1.67	1.085	.256	1.13	2.21	0	4
	Bicycle	18	2.11	1.023	.241	1.60	2.62	0	4
	Total	72	1.71	1.093	.129	1.45	1.97	0	4
Questions	Charging	18	.17	.383	.090	-.02	.36	0	1
	Coffee	18	.06	.236	.056	-.06	.17	0	1
	Flooding	18	.11	.323	.076	-.05	.27	0	1
	Bicycle	18	.00	.000	.000	.00	.00	0	0
	Total	72	.08	.278	.033	.02	.15	0	1
Total	Charging	18	2.72	1.227	.289	2.11	3.33	0	5
	Coffee	18	3.11	1.779	.419	2.23	4.00	1	7
	Flooding	18	3.33	1.455	.343	2.61	4.06	0	6
	Bicycle	18	4.06	1.662	.392	3.23	4.88	1	7
	Total	72	3.31	1.589	.187	2.93	3.68	0	7

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Insigths	Between Groups	5.819	3	1.940	1.531	.214
	Within Groups	86.167	68	1.267		
	Total	91.986	71			
Solutions	Between Groups	4.819	3	1.606	1.365	.261
	Within Groups	80.056	68	1.177		
	Total	84.875	71			
Questions	Between Groups	.278	3	.093	1.206	.314
	Within Groups	5.222	68	.077		
	Total	5.500	71			
Total	Between Groups	16.944	3	5.648	2.366	.079
	Within Groups	162.333	68	2.387		
	Total	179.278	71			

9.2.5 Tukey Post Hoc Tests

Multiple Comparisons

Tukey HSD

Dependent Variable	(I) Situation	(J) Situation	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Insigths	Charging	Coffee	-.222	.375	.934	-1.21	.77
		Flooding	-.389	.375	.729	-1.38	.60
		Bicycle	-.778	.375	.172	-1.77	.21
	Coffee	Charging	.222	.375	.934	-.77	1.21
		Flooding	-.167	.375	.971	-1.15	.82
		Bicycle	-.556	.375	.455	-1.54	.43
	Flooding	Charging	.389	.375	.729	-.60	1.38
		Coffee	.167	.375	.971	-.82	1.15
		Bicycle	-.389	.375	.729	-1.38	.60
	Bicycle	Charging	.778	.375	.172	-.21	1.77
		Coffee	.556	.375	.455	-.43	1.54
		Flooding	.389	.375	.729	-.60	1.38
Solutions	Charging	Coffee	-.278	.362	.869	-1.23	.67
		Flooding	-.278	.362	.869	-1.23	.67
		Bicycle	-.722	.362	.199	-1.67	.23
	Coffee	Charging	.278	.362	.869	-.67	1.23
		Flooding	.000	.362	1.000	-.95	.95
		Bicycle	-.444	.362	.611	-1.40	.51
	Flooding	Charging	.278	.362	.869	-.67	1.23
		Coffee	.000	.362	1.000	-.95	.95
		Bicycle	-.444	.362	.611	-1.40	.51
	Bicycle	Charging	.722	.362	.199	-.23	1.67
		Coffee	.444	.362	.611	-.51	1.40
		Flooding	.444	.362	.611	-.51	1.40
Questions	Charging	Coffee	.111	.092	.627	-.13	.35
		Flooding	.056	.092	.931	-.19	.30
		Bicycle	.167	.092	.280	-.08	.41
	Coffee	Charging	-.111	.092	.627	-.35	.13
		Flooding	-.056	.092	.931	-.30	.19
		Bicycle	.056	.092	.931	-.19	.30
	Flooding	Charging	-.056	.092	.931	-.30	.19
		Coffee	.056	.092	.931	-.19	.30
		Bicycle	.111	.092	.627	-.13	.35
	Bicycle	Charging	-.167	.092	.280	-.41	.08
		Coffee	-.056	.092	.931	-.30	.19
		Flooding	-.111	.092	.627	-.35	.13
Total	Charging	Coffee	-.389	.515	.874	-1.75	.97
		Flooding	-.611	.515	.637	-1.97	.75
		Bicycle	-1.333	.515	.056	-2.69	.02
	Coffee	Charging	.389	.515	.874	-.97	1.75
		Flooding	-.222	.515	.973	-1.58	1.13
		Bicycle	-.944	.515	.267	-2.30	.41
	Flooding	Charging	.611	.515	.637	-.75	1.97
		Coffee	.222	.515	.973	-1.13	1.58
		Bicycle	-.722	.515	.502	-2.08	.63
	Bicycle	Charging	1.333	.515	.056	-.02	2.69
		Coffee	.944	.515	.267	-.41	2.30
		Flooding	.722	.515	.502	-.63	2.08

9.3 Comparison of experiment variables – Focal length

Case Processing Summary

	Focal Length	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	24mm	36	100.0%	0	0.0%	36	100.0%
	100mm	36	100.0%	0	0.0%	36	100.0%
Solutions	24mm	36	100.0%	0	0.0%	36	100.0%
	100mm	36	100.0%	0	0.0%	36	100.0%
Questions	24mm	36	100.0%	0	0.0%	36	100.0%
	100mm	36	100.0%	0	0.0%	36	100.0%
Total	24mm	36	100.0%	0	0.0%	36	100.0%
	100mm	36	100.0%	0	0.0%	36	100.0%

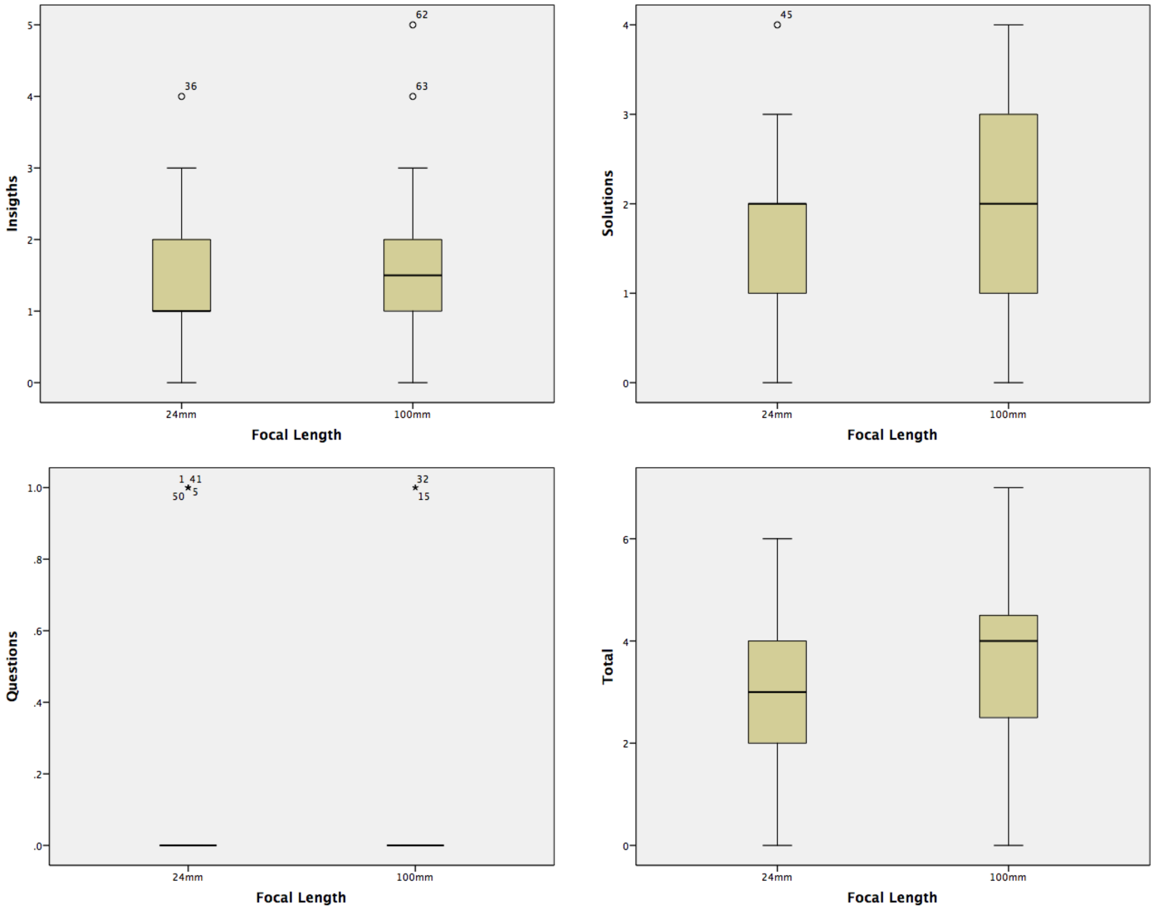
9.3.1 Test of normality

Tests of Normality

	Focal Length	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	24mm	.251	36	.000	.893	36	.002
	100mm	.180	36	.005	.907	36	.005
Solutions	24mm	.206	36	.001	.900	36	.003
	100mm	.191	36	.002	.917	36	.011
Questions	24mm	.525	36	.000	.366	36	.000
	100mm	.539	36	.000	.246	36	.000
Total	24mm	.210	36	.000	.940	36	.052
	100mm	.148	36	.045	.961	36	.232

a. Lilliefors Significance Correction

9.3.2 Boxplots to check for outliers



9.3.3 T-Test

Group Statistics					
	Focal Length	N	Mean	Std. Deviation	Std. Error Mean
Insigths	24mm	36	1.44	1.027	.171
	100mm	36	1.58	1.251	.208
Solutions	24mm	36	1.50	1.082	.180
	100mm	36	1.92	1.079	.180
Questions	24mm	36	.11	.319	.053
	100mm	36	.06	.232	.039
Total	24mm	36	3.06	1.433	.239
	100mm	36	3.56	1.715	.286

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Insigths	Equal variances assumed	1.419	.238	-.515	70	.608	-.139	.270	-.677	.399
	Equal variances not assumed			-.515	67.438	.608	-.139	.270	-.677	.399
Solutions	Equal variances assumed	.364	.548	-1.636	70	.106	-.417	.255	-.925	.091
	Equal variances not assumed			-1.636	69.999	.106	-.417	.255	-.925	.091
Questions	Equal variances assumed	2.965	.090	.845	70	.401	.056	.066	-.076	.187
	Equal variances not assumed			.845	64.002	.401	.056	.066	-.076	.187
Total	Equal variances assumed	1.492	.226	-1.342	70	.184	-.500	.372	-1.243	.243
	Equal variances not assumed			-1.342	67.865	.184	-.500	.372	-1.243	.243

9.4 Comparison of experiment variables - Aperture

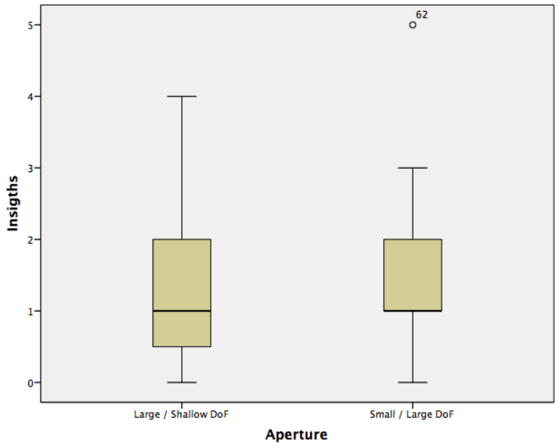
Case Processing Summary							
	Aperture	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	Large / Small DoF	36	100.0%	0	0.0%	36	100.0%
	Small / Large DoF	36	100.0%	0	0.0%	36	100.0%
Solutions	Large / Small DoF	36	100.0%	0	0.0%	36	100.0%
	Small / Large DoF	36	100.0%	0	0.0%	36	100.0%
Questions	Large / Small DoF	36	100.0%	0	0.0%	36	100.0%
	Small / Large DoF	36	100.0%	0	0.0%	36	100.0%
Total	Large / Small DoF	36	100.0%	0	0.0%	36	100.0%
	Small / Large DoF	36	100.0%	0	0.0%	36	100.0%

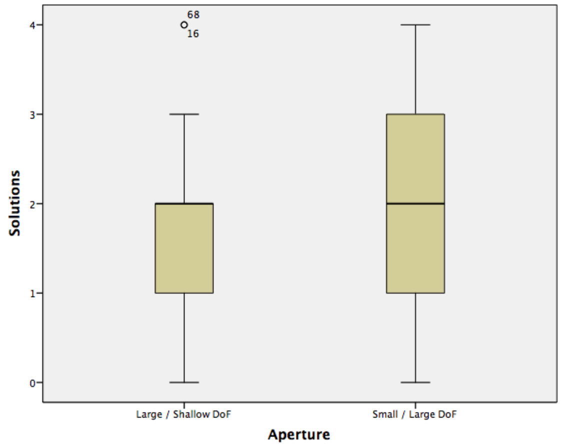
9.4.1 Test for normality

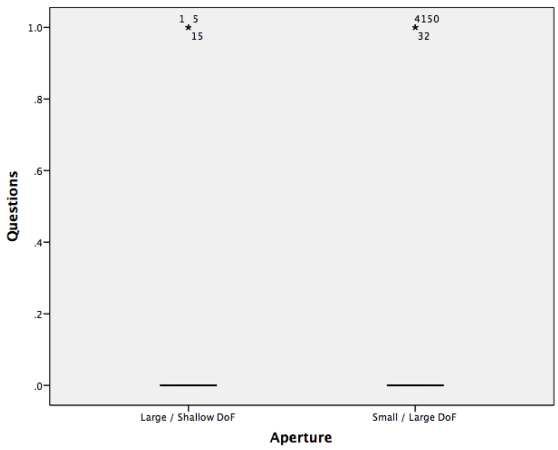
Tests of Normality							
	Aperture	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	Large / Small DoF	.196	36	.001	.894	36	.002
	Small / Large DoF	.234	36	.000	.889	36	.002
Solutions	Large / Small DoF	.215	36	.000	.902	36	.004
	Small / Large DoF	.172	36	.009	.917	36	.011
Questions	Large / Small DoF	.534	36	.000	.312	36	.000
	Small / Large DoF	.534	36	.000	.312	36	.000
Total	Large / Small DoF	.156	36	.027	.956	36	.162
	Small / Large DoF	.163	36	.017	.960	36	.211

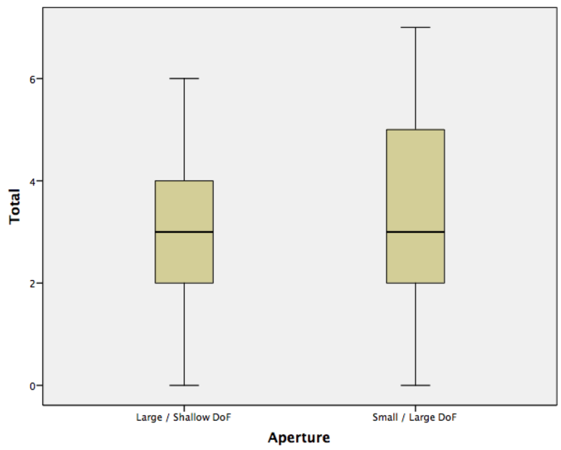
a. Lilliefors Significance Correction

9.4.2 Boxplots to check for outliers









9.4.3 T-Test

Group Statistics					
	Aperture	N	Mean	Std. Deviation	Std. Error Mean
Insigths	Large / Small DoF	36	1.42	1.156	.193
	Small / Large DoF	36	1.61	1.128	.188
Solutions	Large / Small DoF	36	1.64	1.073	.179
	Small / Large DoF	36	1.78	1.124	.187
Questions	Large / Small DoF	36	.08	.280	.047
	Small / Large DoF	36	.08	.280	.047
Total	Large / Small DoF	36	3.14	1.437	.240
	Small / Large DoF	36	3.47	1.732	.289

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Insigths	Equal variances assumed	.074	.786	-.722	70	.473	-.194	.269	-.731	.342
	Equal variances not assumed			-.722	69.960	.473	-.194	.269	-.731	.342
Solutions	Equal variances assumed	.171	.680	-.536	70	.593	-.139	.259	-.655	.378
	Equal variances not assumed			-.536	69.850	.594	-.139	.259	-.655	.378
Questions	Equal variances assumed	.000	1.000	.000	70	1.000	.000	.066	-.132	.132
	Equal variances not assumed			.000	70.000	1.000	.000	.066	-.132	.132
Total	Equal variances assumed	1.862	.177	-.889	70	.377	-.333	.375	-1.081	.415
	Equal variances not assumed			-.889	67.701	.377	-.333	.375	-1.082	.415

9.5 Comparison of experiment variables - Split by situation - Aperture

9.5.1 Situation = Charging

Case Processing Summary[a]							
	Aperture	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	Large / Small DoF	9	100.0%	0	0.0%	9	100.0%
	Small / Large DoF	9	100.0%	0	0.0%	9	100.0%
Solutions	Large / Small DoF	9	100.0%	0	0.0%	9	100.0%
	Small / Large DoF	9	100.0%	0	0.0%	9	100.0%
Questions	Large / Small DoF	9	100.0%	0	0.0%	9	100.0%
	Small / Large DoF	9	100.0%	0	0.0%	9	100.0%
Total	Large / Small DoF	9	100.0%	0	0.0%	9	100.0%
	Small / Large DoF	9	100.0%	0	0.0%	9	100.0%

a. Situation = Charging

9.5.1.1 Test for normality

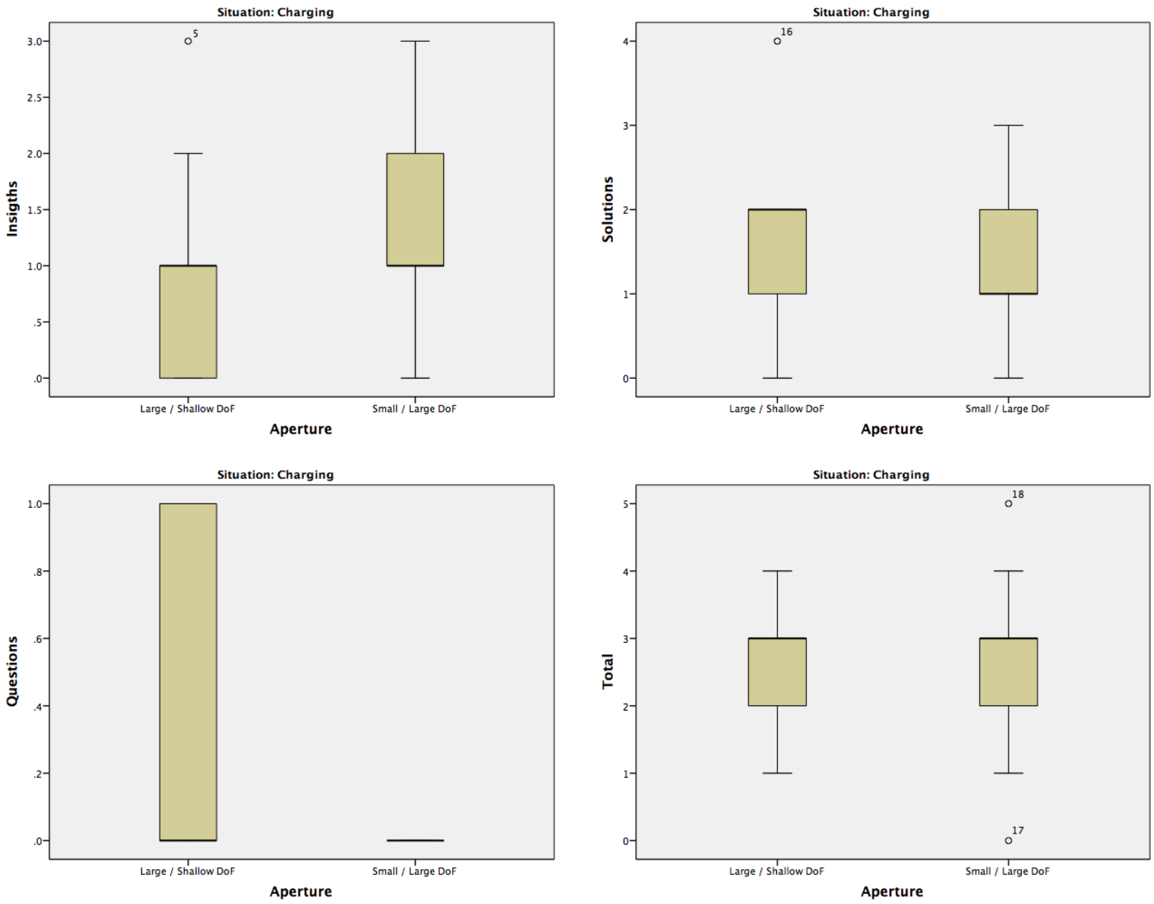
Tests of Normality[a,c]							
	Aperture	Kolmogorov-Smirnov[b]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	Large / Small DoF	.245	9	.127	.825	9	.039
	Small / Large DoF	.248	9	.116	.913	9	.338
Solutions	Large / Small DoF	.248	9	.116	.889	9	.195
	Small / Large DoF	.257	9	.088	.903	9	.273
Questions	Large / Small DoF	.414	9	.000	.617	9	.000
Total	Large / Small DoF	.257	9	.088	.903	9	.273
	Small / Large DoF	.255	9	.096	.940	9	.586

a. Situation = Charging

b. Lilliefors Significance Correction

c. Questions is constant when Aperture = Small / Large DoF. It has been omitted.

9.5.1.2 Boxplots to check for outliers



9.5.1.3 T-Test

Group Statistics[a]

	Aperture	N	Mean	Std. Deviation	Std. Error Mean
Insigths	Large / Small DoF	9	.89	1.054	.351
	Small / Large DoF	9	1.44	.882	.294
Solutions	Large / Small DoF	9	1.56	1.236	.412
	Small / Large DoF	9	1.22	.972	.324
Questions	Large / Small DoF	9	.33	.500	.167
	Small / Large DoF	9	.00	.000	.000
Total	Large / Small DoF	9	2.78	.972	.324
	Small / Large DoF	9	2.67	1.500	.500

a. Situation = Charging

Independent Samples Test[a]										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Insigths	Equal variances assumed	.081	.780	-1.213	16	.243	-.556	.458	-1.527	.416
	Equal variances not assumed			-1.213	15.517	.243	-.556	.458	-1.529	.418
Solutions	Equal variances assumed	.406	.533	.636	16	.534	.333	.524	-.778	1.444
	Equal variances not assumed			.636	15.156	.534	.333	.524	-.783	1.449
Questions	Equal variances assumed	64.000	.000	2.000	16	.063	.333	.167	-.020	.687
	Equal variances not assumed			2.000	8.000	.081	.333	.167	-.051	.718
Total	Equal variances assumed	1.039	.323	.187	16	.854	.111	.596	-1.152	1.374
	Equal variances not assumed			.187	13.710	.855	.111	.596	-1.169	1.391

a. Situation = Charging

9.5.2 Situation = Coffee

Case Processing Summary[a]							
	Aperture	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	Large / Small DoF	9	100.0%	0	0.0%	9	100.0%
	Small / Large DoF	9	100.0%	0	0.0%	9	100.0%
Solutions	Large / Small DoF	9	100.0%	0	0.0%	9	100.0%
	Small / Large DoF	9	100.0%	0	0.0%	9	100.0%
Questions	Large / Small DoF	9	100.0%	0	0.0%	9	100.0%
	Small / Large DoF	9	100.0%	0	0.0%	9	100.0%
Total	Large / Small DoF	9	100.0%	0	0.0%	9	100.0%
	Small / Large DoF	9	100.0%	0	0.0%	9	100.0%

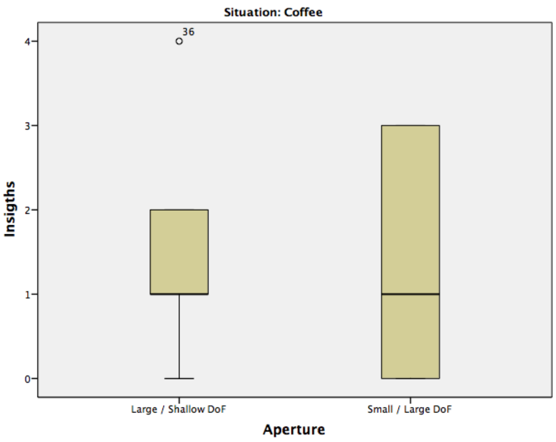
a. Situation = Coffee

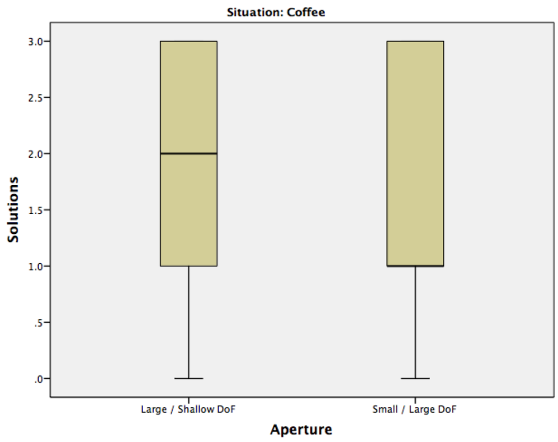
9.5.2.1 Test for normality

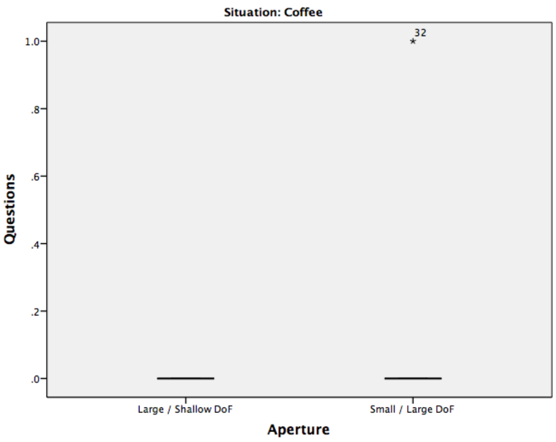
Tests of Normality[a,d]							
	Aperture	Kolmogorov-Smirnov[b]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	Large / Small DoF	.274	9	.050	.854	9	.083
	Small / Large DoF	.212	9	.200*	.826	9	.041
Solutions	Large / Small DoF	.206	9	.200*	.884	9	.172
	Small / Large DoF	.229	9	.191	.854	9	.082
Questions	Small / Large DoF	.519	9	.000	.390	9	.000
Total	Large / Small DoF	.201	9	.200*	.860	9	.096
	Small / Large DoF	.189	9	.200*	.916	9	.364

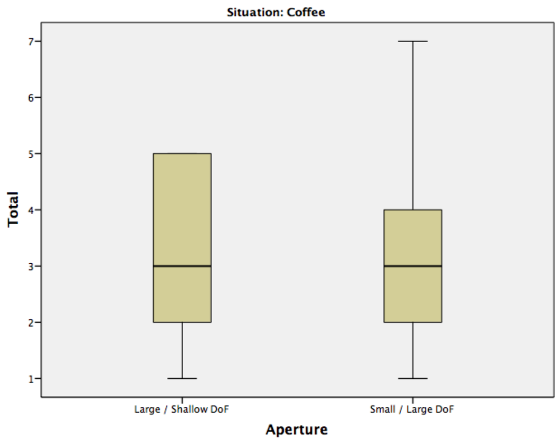
*. This is a lower bound of the true significance.
a. Situation = Coffee
b. Lilliefors Significance Correction
d. Questions is constant when Aperture = Large / Small DoF. It has been omitted.

9.5.2.2 Boxplots to check for outliers









9.5.2.3 T-Test

Group Statistics[a]					
	Aperture	N	Mean	Std. Deviation	Std. Error Mean
Insigths	Large / Small DoF	9	1.33	1.225	.408
	Small / Large DoF	9	1.44	1.333	.444
Solutions	Large / Small DoF	9	1.78	1.093	.364
	Small / Large DoF	9	1.56	1.236	.412
Questions	Large / Small DoF	9	.00	.000	.000
	Small / Large DoF	9	.11	.333	.111
Total	Large / Small DoF	9	3.11	1.691	.564
	Small / Large DoF	9	3.11	1.965	.655
a. Situation = Coffee					

Independent Samples Test[a]										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Insigths	Equal variances assumed	.760	.396	-.184	16	.856	-.111	.603	-1.390	1.168
	Equal variances not assumed			-.184	15.886	.856	-.111	.603	-1.391	1.169
Solutions	Equal variances assumed	.383	.544	.404	16	.692	.222	.550	-.944	1.388
	Equal variances not assumed			.404	15.764	.692	.222	.550	-.945	1.390
Questions	Equal variances assumed	5.224	.036	-1.000	16	.332	-.111	.111	-.347	.124
	Equal variances not assumed			-1.000	8.000	.347	-.111	.111	-.367	.145
Total	Equal variances assumed	.003	.957	.000	16	1.000	.000	.864	-1.832	1.832
	Equal variances not assumed			.000	15.654	1.000	.000	.864	-1.835	1.835
a. Situation = Coffee										

9.5.3 Situation = Flooding

Case Processing Summary[a]

	Aperture	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	Large / Small DoF	10	100.0%	0	0.0%	10	100.0%
	Small / Large DoF	8	100.0%	0	0.0%	8	100.0%
Solutions	Large / Small DoF	10	100.0%	0	0.0%	10	100.0%
	Small / Large DoF	8	100.0%	0	0.0%	8	100.0%
Questions	Large / Small DoF	10	100.0%	0	0.0%	10	100.0%
	Small / Large DoF	8	100.0%	0	0.0%	8	100.0%
Total	Large / Small DoF	10	100.0%	0	0.0%	10	100.0%
	Small / Large DoF	8	100.0%	0	0.0%	8	100.0%

a. Situation = Flooding

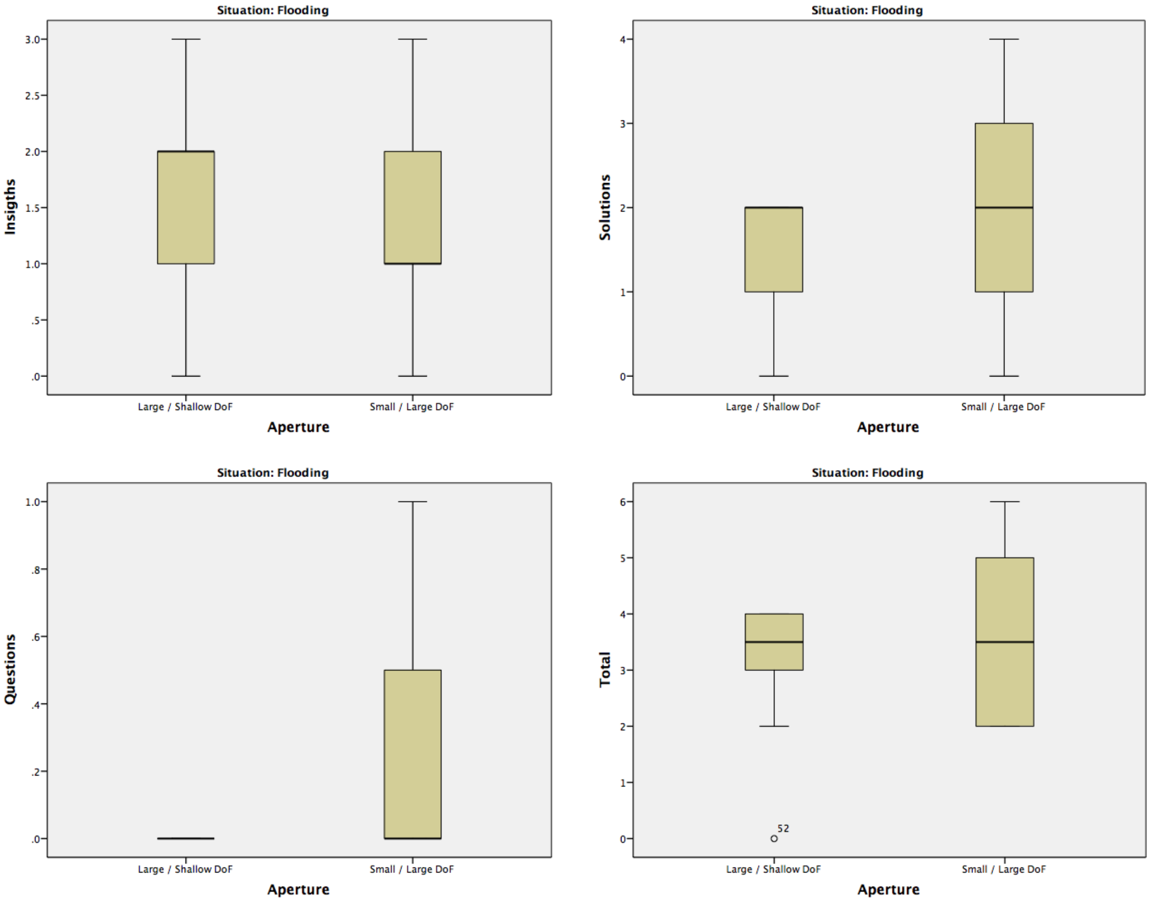
9.5.3.1 Test for normality

Tests of Normality[a,d]

	Aperture	Kolmogorov-Smirnov[b]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	Large / Small DoF	.342	10	.002	.841	10	.045
	Small / Large DoF	.284	8	.057	.906	8	.324
Solutions	Large / Small DoF	.362	10	.001	.717	10	.001
	Small / Large DoF	.152	8	.200*	.965	8	.857
Questions	Small / Large DoF	.455	8	.000	.566	8	.000
Total	Large / Small DoF	.269	10	.039	.744	10	.003
	Small / Large DoF	.208	8	.200*	.843	8	.080

*. This is a lower bound of the true significance.
a. Situation = Flooding
b. Lilliefors Significance Correction
d. Questions is constant when Aperture = Large / Small DoF. It has been omitted.

9.5.3.2 Boxplots to check for outliers



9.5.3.3 T-Test

Group Statistics[a]

	Aperture	N	Mean	Std. Deviation	Std. Error Mean
Insigths	Large / Small DoF	10	1.70	.823	.260
	Small / Large DoF	8	1.38	.916	.324
Solutions	Large / Small DoF	10	1.40	.843	.267
	Small / Large DoF	8	2.00	1.309	.463
Questions	Large / Small DoF	10	.00	.000	.000
	Small / Large DoF	8	.25	.463	.164
Total	Large / Small DoF	10	3.10	1.287	.407
	Small / Large DoF	8	3.63	1.685	.596

a. Situation = Flooding

Independent Samples Test[a]										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Insigths	Equal variances assumed	.173	.683	.792	16	.440	.325	.410	-.545	1.195
	Equal variances not assumed			.782	14.318	.447	.325	.416	-.564	1.214
Solutions	Equal variances assumed	1.069	.317	-1.180	16	.255	-.600	.509	-1.678	.478
	Equal variances not assumed			-1.123	11.437	.284	-.600	.534	-1.770	.570
Questions	Equal variances assumed	26.667	.000	-1.721	16	.104	-.250	.145	-.558	.058
	Equal variances not assumed			-1.528	7.000	.170	-.250	.164	-.637	.137
Total	Equal variances assumed	1.389	.256	-.751	16	.464	-.525	.699	-2.007	.957
	Equal variances not assumed			-.728	12.875	.480	-.525	.721	-2.085	1.035

a. Situation = Flooding

9.5.4 Situation = Bicycle

Case Processing Summary[a]							
	Aperture	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	Large / Small DoF	8	100.0%	0	0.0%	8	100.0%
	Small / Large DoF	10	100.0%	0	0.0%	10	100.0%
Solutions	Large / Small DoF	8	100.0%	0	0.0%	8	100.0%
	Small / Large DoF	10	100.0%	0	0.0%	10	100.0%
Questions	Large / Small DoF	8	100.0%	0	0.0%	8	100.0%
	Small / Large DoF	10	100.0%	0	0.0%	10	100.0%
Total	Large / Small DoF	8	100.0%	0	0.0%	8	100.0%
	Small / Large DoF	10	100.0%	0	0.0%	10	100.0%

a. Situation = Bicycle

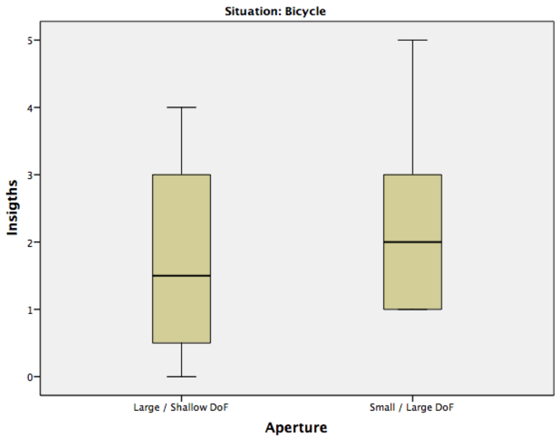
9.5.4.1 *Test for normality*

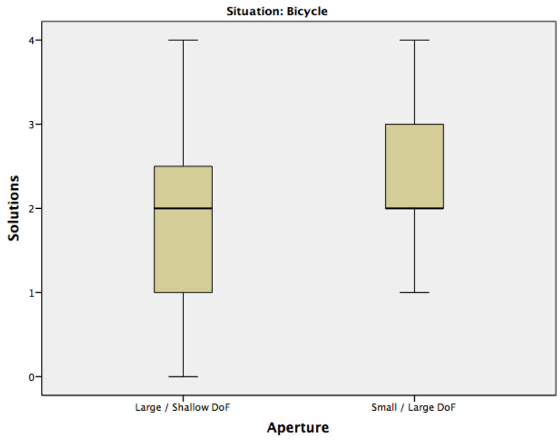
Tests of Normality[a,d,e]

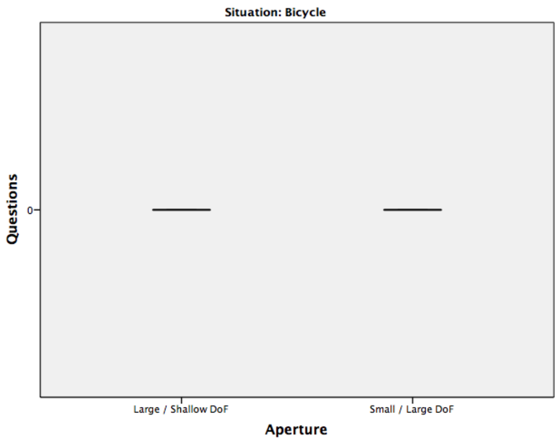
	Aperture	Kolmogorov-Smirnov[b]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	Large / Small DoF	.193	8	.200*	.920	8	.428
	Small / Large DoF	.231	10	.139	.824	10	.028
Solutions	Large / Small DoF	.210	8	.200*	.958	8	.792
	Small / Large DoF	.342	10	.002	.841	10	.045
Total	Large / Small DoF	.257	8	.127	.902	8	.301
	Small / Large DoF	.155	10	.200*	.969	10	.886

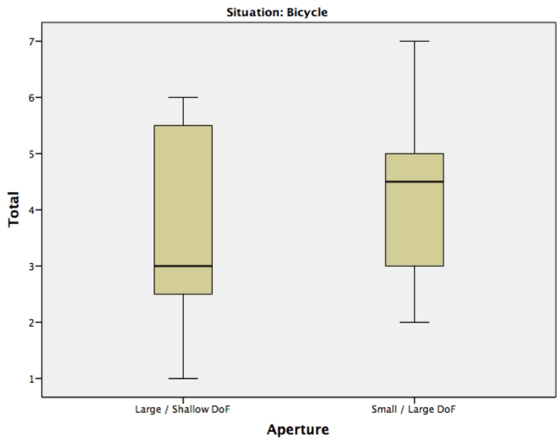
*. This is a lower bound of the true significance.
a. Situation = Bicycle
b. Lilliefors Significance Correction
d. Questions is constant when Aperture = Large / Small DoF. It has been omitted.
e. Questions is constant when Aperture = Small / Large DoF. It has been omitted.

9.5.4.2 *Boxplots to check for outliers*









9.5.4.3 T-Test

Group Statistics[a]

	Aperture	N	Mean	Std. Deviation	Std. Error Mean
Insigths	Large / Small DoF	8	1.75	1.488	.526
	Small / Large DoF	10	2.10	1.287	.407
Solutions	Large / Small DoF	8	1.88	1.246	.441
	Small / Large DoF	10	2.30	.823	.260
Questions	Large / Small DoF	8	.00	.000[b]	.000
	Small / Large DoF	10	.00	.000[b]	.000
Total	Large / Small DoF	8	3.63	1.847	.653
	Small / Large DoF	10	4.40	1.506	.476

a. Situation = Bicycle

b. t cannot be computed because the standard deviations of both groups are 0.

Independent Samples Test[a]

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Insigths	Equal variances assumed	.754	.398	-.535	16	.600	-.350	.654	-1.736	1.036
	Equal variances not assumed			-.526	13.986	.607	-.350	.665	-1.777	1.077
Solutions	Equal variances assumed	.888	.360	-.870	16	.397	-.425	.489	-1.461	.611
	Equal variances not assumed			-.830	11.636	.423	-.425	.512	-1.544	.694
Total	Equal variances assumed	.702	.414	-.982	16	.341	-.775	.789	-2.448	.898
	Equal variances not assumed			-.959	13.462	.354	-.775	.808	-2.515	.965

a. Situation = Bicycle

9.6 Comparison of experiment variables – Split by situation – Focal Length

9.6.1 Situation = Charging

Case Processing Summary[a]

	Focal Length	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	24mm	8	100.0%	0	0.0%	8	100.0%
	100mm	10	100.0%	0	0.0%	10	100.0%
Solutions	24mm	8	100.0%	0	0.0%	8	100.0%
	100mm	10	100.0%	0	0.0%	10	100.0%
Questions	24mm	8	100.0%	0	0.0%	8	100.0%
	100mm	10	100.0%	0	0.0%	10	100.0%
Total	24mm	8	100.0%	0	0.0%	8	100.0%
	100mm	10	100.0%	0	0.0%	10	100.0%

a. Situation = Charging

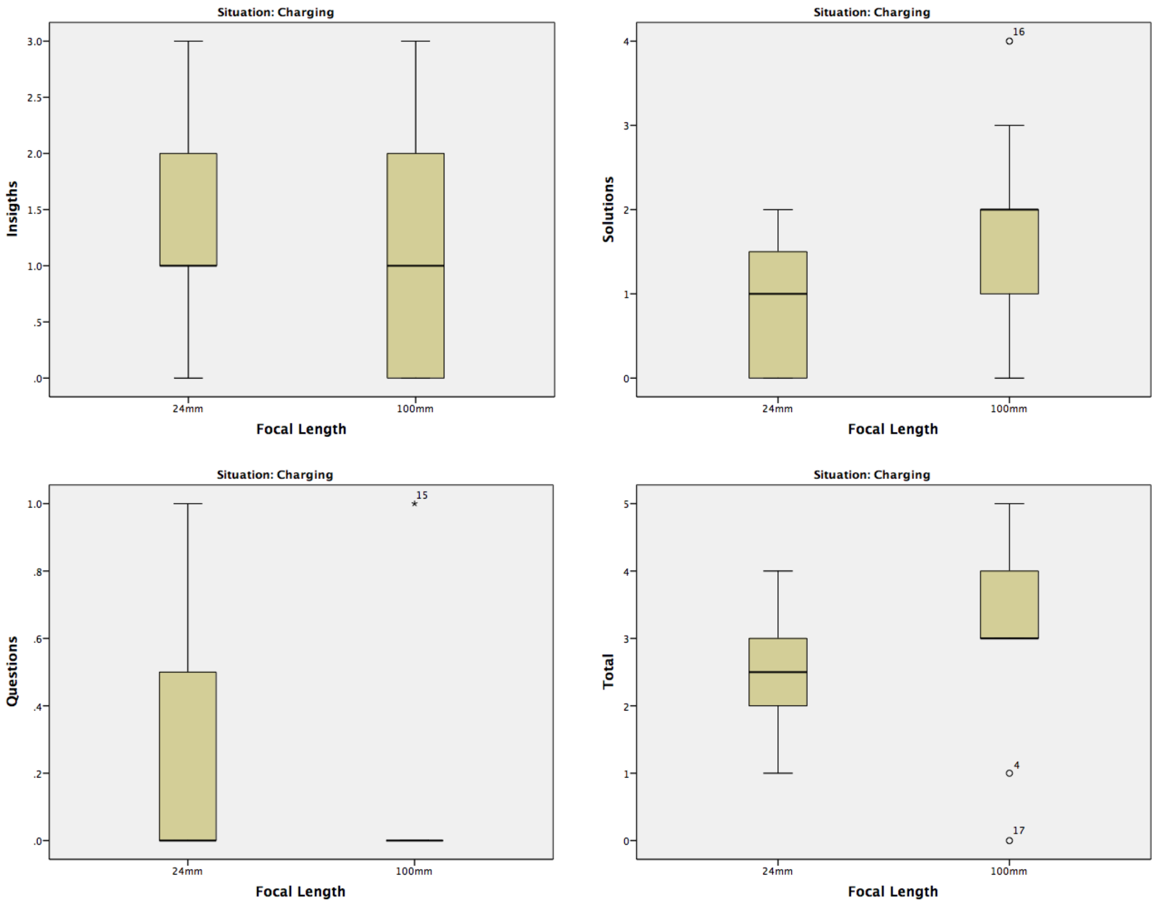
9.6.1.1 Test of normality

Tests of Normality[a]

	Focal Length	Kolmogorov-Smirnov[b]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	24mm	.284	8	.057	.906	8	.324
	100mm	.229	10	.148	.859	10	.074
Solutions	24mm	.228	8	.200*	.835	8	.067
	100mm	.230	10	.143	.933	10	.479
Questions	24mm	.455	8	.000	.566	8	.000
	100mm	.524	10	.000	.366	10	.000
Total	24mm	.205	8	.200*	.931	8	.522
	100mm	.328	10	.003	.876	10	.119

*. This is a lower bound of the true significance.
a. Situation = Charging
b. Lilliefors Significance Correction

9.6.1.2 Boxplots to check for outliers



9.6.1.3 T-Test

Group Statistics[a]					
	Focal Length	N	Mean	Std. Deviation	Std. Error Mean
Insigths	24mm	8	1.38	.916	.324
	100mm	10	1.00	1.054	.333
Solutions	24mm	8	.88	.835	.295
	100mm	10	1.80	1.135	.359
Questions	24mm	8	.25	.463	.164
	100mm	10	.10	.316	.100
Total	24mm	8	2.50	.926	.327
	100mm	10	2.90	1.449	.458

a. Situation = Charging

Independent Samples Test[a]										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Insigths	Equal variances assumed	.088	.771	.794	16	.439	.375	.472	-.627	1.377
	Equal variances not assumed			.807	15.851	.432	.375	.465	-.611	1.361
Solutions	Equal variances assumed	.402	.535	-1.922	16	.073	-.925	.481	-1.945	.095
	Equal variances not assumed			-1.991	15.924	.064	-.925	.465	-1.910	.060
Questions	Equal variances assumed	2.843	.111	.816	16	.426	.150	.184	-.239	.539
	Equal variances not assumed			.782	11.911	.449	.150	.192	-.268	.568
Total	Equal variances assumed	.280	.604	-.676	16	.509	-.400	.592	-1.654	.854
	Equal variances not assumed			-.710	15.379	.488	-.400	.563	-1.598	.798
a. Situation = Charging										

9.6.2 Situation = Coffee

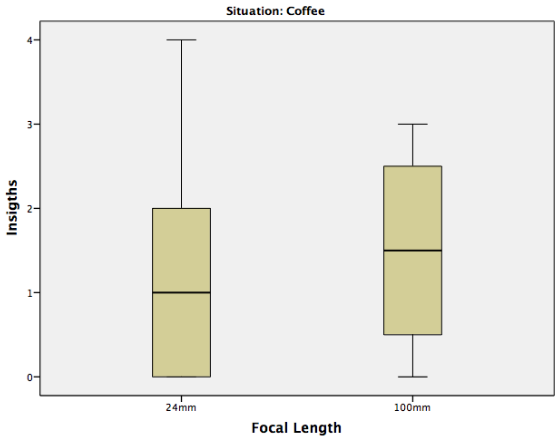
Case Processing Summary[a]							
	Focal Length	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	24mm	10	100.0%	0	0.0%	10	100.0%
	100mm	8	100.0%	0	0.0%	8	100.0%
Solutions	24mm	10	100.0%	0	0.0%	10	100.0%
	100mm	8	100.0%	0	0.0%	8	100.0%
Questions	24mm	10	100.0%	0	0.0%	10	100.0%
	100mm	8	100.0%	0	0.0%	8	100.0%
Total	24mm	10	100.0%	0	0.0%	10	100.0%
	100mm	8	100.0%	0	0.0%	8	100.0%
a. Situation = Coffee							

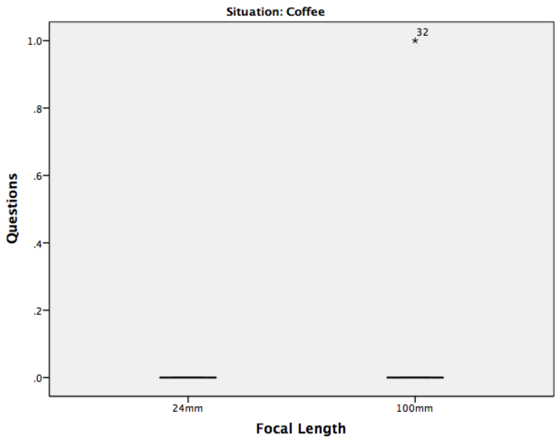
9.6.2.1 Test for normality

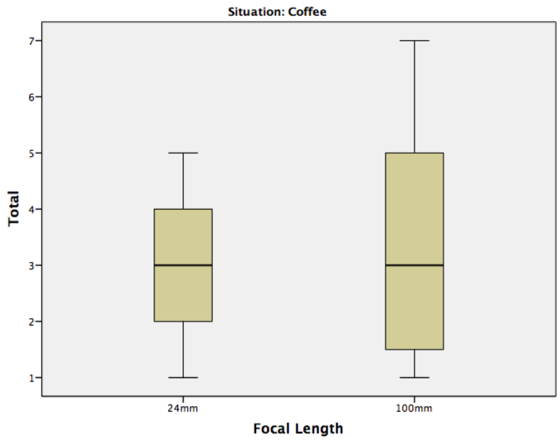
Tests of Normality[a,d]							
	Focal Length	Kolmogorov-Smirnov[b]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	24mm	.289	10	.018	.855	10	.067
	100mm	.162	8	.200*	.897	8	.274
Solutions	24mm	.195	10	.200*	.878	10	.124
	100mm	.240	8	.195	.858	8	.114
Questions	100mm	.513	8	.000	.418	8	.000
Total	24mm	.172	10	.200*	.917	10	.330
	100mm	.234	8	.200*	.904	8	.314

*. This is a lower bound of the true significance.
a. Situation = Coffee
b. Lilliefors Significance Correction
d. Questions is constant when Focal Length = 24mm. It has been omitted.

9.6.2.2 Boxplots to check for outliers







9.6.2.3 *T-Test*

Group Statistics[a]					
	Focal Length	N	Mean	Std. Deviation	Std. Error Mean
Insigths	24mm	10	1.30	1.337	.423
	100mm	8	1.50	1.195	.423
Solutions	24mm	10	1.60	1.174	.371
	100mm	8	1.75	1.165	.412
Questions	24mm	10	.00	.000	.000
	100mm	8	.13	.354	.125
Total	24mm	10	2.90	1.449	.458
	100mm	8	3.38	2.200	.778

a. Situation = Coffee

Independent Samples Test[a]										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Insigths	Equal variances assumed	.004	.952	-.330	16	.746	-.200	.606	-1.484	1.084
	Equal variances not assumed			-.335	15.753	.742	-.200	.598	-1.469	1.069
Solutions	Equal variances assumed	.000	1.000	-.270	16	.790	-.150	.555	-1.326	1.026
	Equal variances not assumed			-.271	15.193	.790	-.150	.554	-1.330	1.030
Questions	Equal variances assumed	6.914	.018	-1.127	16	.276	-.125	.111	-.360	.110
	Equal variances not assumed			-1.000	7.000	.351	-.125	.125	-.421	.171
Total	Equal variances assumed	3.348	.086	-.551	16	.589	-.475	.861	-2.301	1.351
	Equal variances not assumed			-.526	11.615	.609	-.475	.903	-2.449	1.499

a. Situation = Coffee

9.6.3 Situation = Flooding

Case Processing Summary[a]

	Focal Length	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	24mm	9	100.0%	0	0.0%	9	100.0%
	100mm	9	100.0%	0	0.0%	9	100.0%
Solutions	24mm	9	100.0%	0	0.0%	9	100.0%
	100mm	9	100.0%	0	0.0%	9	100.0%
Questions	24mm	9	100.0%	0	0.0%	9	100.0%
	100mm	9	100.0%	0	0.0%	9	100.0%
Total	24mm	9	100.0%	0	0.0%	9	100.0%
	100mm	9	100.0%	0	0.0%	9	100.0%

a. Situation = Flooding

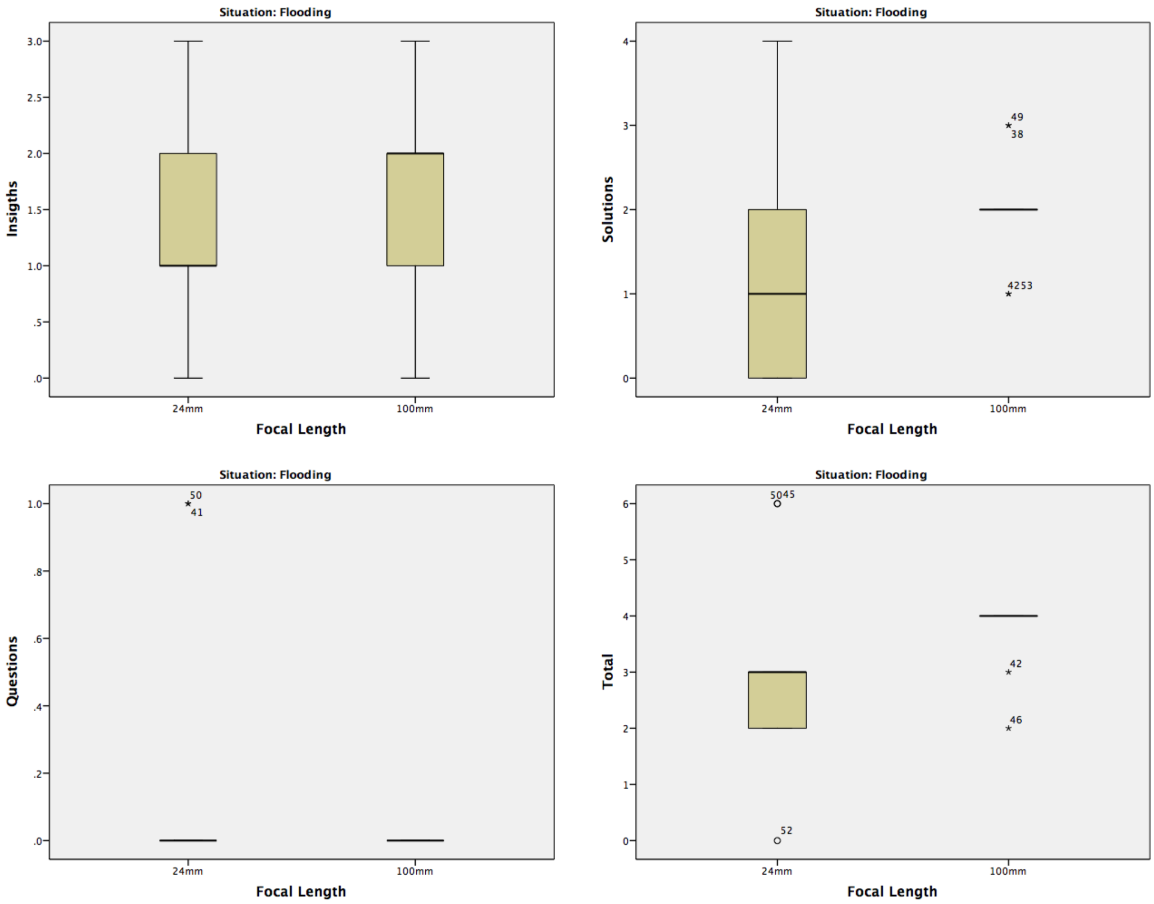
9.6.3.1 Test for normality

Tests of Normality[a,d]

	Focal Length	Kolmogorov-Smirnov[b]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	24mm	.248	9	.116	.913	9	.338
	100mm	.317	9	.010	.873	9	.132
Solutions	24mm	.196	9	.200*	.872	9	.130
	100mm	.278	9	.044	.833	9	.049
Questions	24mm	.471	9	.000	.536	9	.000
Total	24mm	.278	9	.044	.875	9	.138
	100mm	.459	9	.000	.564	9	.000

*. This is a lower bound of the true significance.
a. Situation = Flooding
b. Lilliefors Significance Correction
d. Questions is constant when Focal Length = 100mm. It has been omitted.

9.6.3.2 Boxplots to check for outliers



9.6.3.3 T-Test

Group Statistics[a]					
	Focal Length	N	Mean	Std. Deviation	Std. Error Mean
Insigths	24mm	9	1.44	.882	.294
	100mm	9	1.67	.866	.289
Solutions	24mm	9	1.33	1.323	.441
	100mm	9	2.00	.707	.236
Questions	24mm	9	.22	.441	.147
	100mm	9	.00	.000	.000
Total	24mm	9	3.00	1.936	.645
	100mm	9	3.67	.707	.236
a. Situation = Flooding					

Independent Samples Test[a]										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Insigths	Equal variances assumed	.049	.828	-.539	16	.597	-.222	.412	-1.096	.651
	Equal variances not assumed			-.539	15.995	.597	-.222	.412	-1.096	.651
Solutions	Equal variances assumed	3.864	.067	-1.333	16	.201	-.667	.500	-1.727	.393
	Equal variances not assumed			-1.333	12.226	.207	-.667	.500	-1.754	.421
Questions	Equal variances assumed	17.920	.001	1.512	16	.150	.222	.147	-.089	.534
	Equal variances not assumed			1.512	8.000	.169	.222	.147	-.117	.561
Total	Equal variances assumed	3.068	.099	-.970	16	.346	-.667	.687	-2.123	.790
	Equal variances not assumed			-.970	10.096	.355	-.667	.687	-2.196	.863
a. Situation = Flooding										

9.6.4 Situation = Bicycle

Case Processing Summary[a]							
	Focal Length	Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
Insigths	24mm	9	100.0%	0	0.0%	9	100.0%
	100mm	9	100.0%	0	0.0%	9	100.0%
Solutions	24mm	9	100.0%	0	0.0%	9	100.0%
	100mm	9	100.0%	0	0.0%	9	100.0%
Questions	24mm	9	100.0%	0	0.0%	9	100.0%
	100mm	9	100.0%	0	0.0%	9	100.0%
Total	24mm	9	100.0%	0	0.0%	9	100.0%
	100mm	9	100.0%	0	0.0%	9	100.0%
a. Situation = Bicycle							

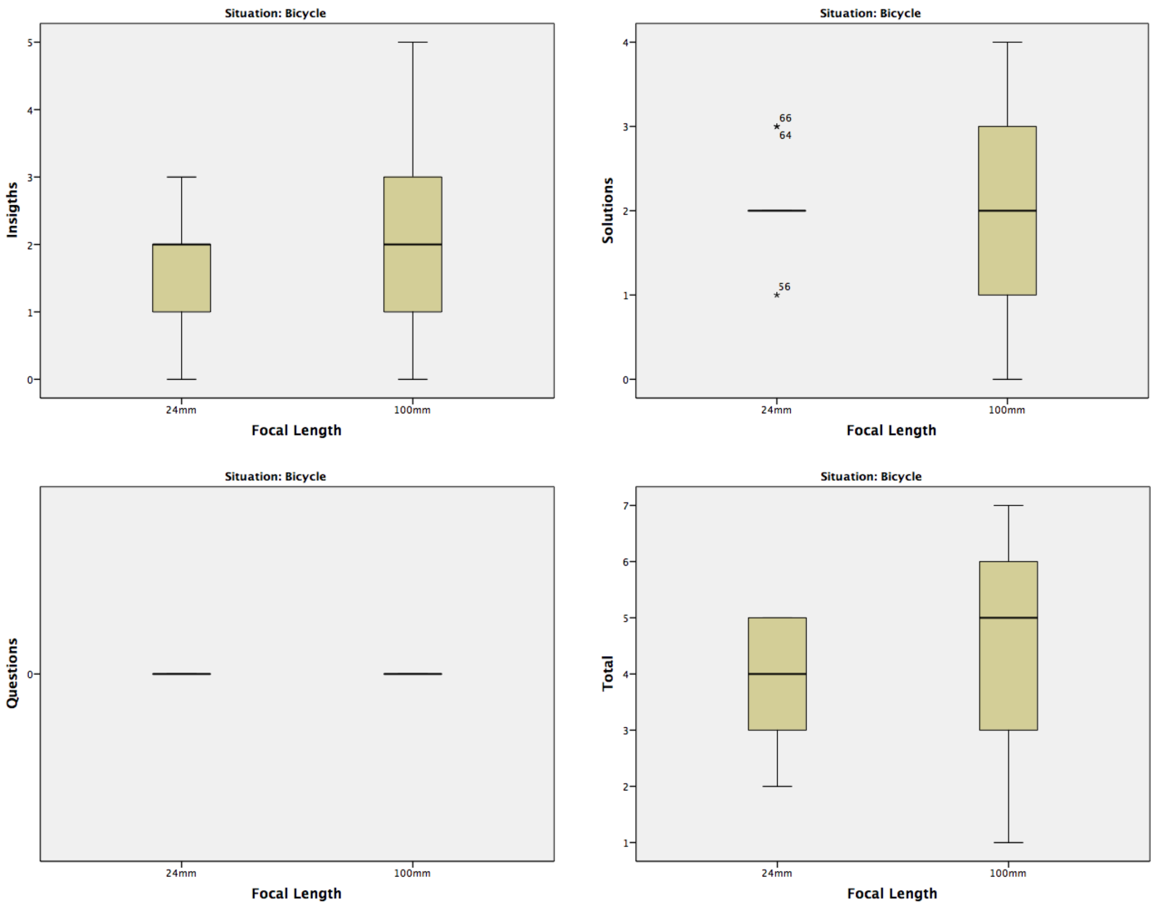
9.6.4.1 Test for normality

Tests of Normality[a,d,e]

	Focal Length	Kolmogorov-Smirnov[b]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Insigths	24mm	.192	9	.200*	.917	9	.364
	100mm	.216	9	.200*	.941	9	.588
Solutions	24mm	.351	9	.002	.781	9	.012
	100mm	.199	9	.200*	.931	9	.494
Total	24mm	.206	9	.200*	.884	9	.172
	100mm	.228	9	.194	.905	9	.285

*. This is a lower bound of the true significance.
a. Situation = Bicycle
b. Lilliefors Significance Correction
d. Questions is constant when Focal Length = 24mm. It has been omitted.
e. Questions is constant when Focal Length = 100mm. It has been omitted.

9.6.4.2 Boxplots to check for outliers



9.6.4.3 T-Test

Group Statistics[a]					
	Focal Length	N	Mean	Std. Deviation	Std. Error Mean
Insigths	24mm	9	1.67	1.000	.333
	100mm	9	2.22	1.641	.547
Solutions	24mm	9	2.11	.601	.200
	100mm	9	2.11	1.364	.455
Questions	24mm	9	.00	.000[b]	.000
	100mm	9	.00	.000[b]	.000
Total	24mm	9	3.78	1.093	.364
	100mm	9	4.33	2.121	.707

a. Situation = Bicycle

b. t cannot be computed because the standard deviations of both groups are 0.

Independent Samples Test[a]										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Insigths	Equal variances assumed	3.043	.100	-.867	16	.399	-.556	.641	-1.914	.803
	Equal variances not assumed			-.867	13.219	.401	-.556	.641	-1.937	.826
Solutions	Equal variances assumed	4.433	.051	.000	16	1.000	.000	.497	-1.053	1.053
	Equal variances not assumed			.000	10.992	1.000	.000	.497	-1.094	1.094
Total	Equal variances assumed	8.828	.009	-.698	16	.495	-.556	.795	-2.242	1.131
	Equal variances not assumed			-.698	11.967	.498	-.556	.795	-2.289	1.178

a. Situation = Bicycle

Page intentionally left blank for formatting reasons

10 Appendix B – Experiment questionnaire

Experiment questionnaire

Please answer the following questions in order to provide background information. The information will be used to look for trends in the data collected.

Participant ID: __________

Age: _______

Gender: M ☐ F ☐

Education:

____

Occupation:

____

Photo experience:

Beginner ☐ Intermediate ☐ Experienced ☐

Do you have any experience with needfinding? Yes ☐ No ☐

How energetic do you feel? 1-low energy and 5-high energy

____

Do you have any visual impairments? E.g. colour blind or reduced eyesight

____

Are you right- or left handed?
Right ☐ Left ☐

Page intentionally left blank for formatting reasons

11 Appendix C – Pre-master project

Introduction

Conveying information gained through needfinding to third parties always contain a strong element of bias. We propose to consciously use altered pictures to convey information and to initiate divergent or convergent iteration processes. Through this we propose the use of photography as a tool in needfinding.

Our intent is to support the needfinding in engineering design projects by generating complementary data and stimuli in the early phases of the iterative user centred engineering design process (Leifer and Steinert, 2011). User-centred design (Norman and Draper, 1986) has emerged since the 1950ies and is supported beyond design as core concept of understanding the deeper/hidden needs of the user (or any other stakeholder for that matter). In the 1980ies Robert McKim, Rolf Faste and David Kelley from the Stanford joint product design program coined the term "needfinding" and introduced a method to identify unmet and often untold needs (Patnaik and Becker, 1999). It has since become a corner stone in the general engineering design process (Eppinger and Ulrich, 1995). Its key method especially concerning the subliminal needs is observation. Generally the in situ observation studies on the tradition of ethnography and anthropology are favoured. However esp. retrospective data analysis is central as the observer is usually not able to catch all data in real time. The true needs and the "nuggets" a future product could be build upon are generally latent or hidden (Aldaz et al., 2013). Observation "can produce unadulterated, direct and potentially very rich descriptions of events and their context, because data is captured while the phenomena occur" (Blessing and Chakrabarti, 2009). However esp. retrospective data analysis is central, as the observer is usually not able to catch all data in real time (Patton, 2002). Steinert et al. (2012) include Alexander's idea of divergent and convergent thinking in their model, combining it with a series of design-build-test cycles as described by Thomke and Fujimoto (2000). As can be seen from Figure 1, they have made a model that shows how the process is gradually converging towards a final solution after running through several phases of diverging and converging iteration cycles.

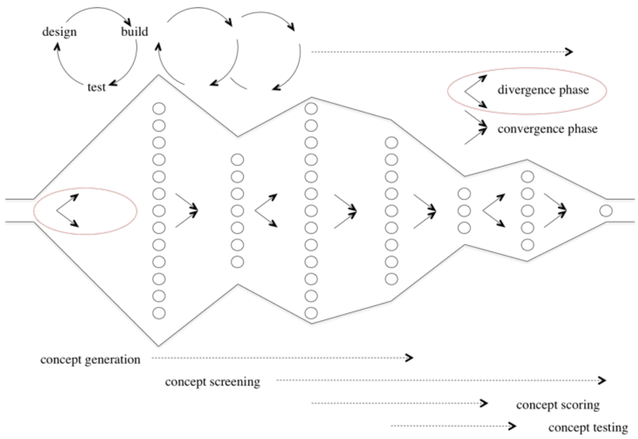


Figure 1: Divergent/convergent, iterative process. (Steinert et al., 2012)

Needfinding

The overall goal of needfinding in an engineering design project is creating a solid foundation to build a product or service on. Of central importance at the early stage of a new product or system design, is the clear understanding of first, who the actual user is and second, of which user needs have to be covered by the new design solution (Steinert and Leifer, 2012). Needfinding is the discipline of perceiving exactly these user needs (Faste, 1987). What is it that the user needs to do? It is important to be clear on what a need is. Faste (1987) describes a need as a perceived lack, something that is missing. A product built on a clearly defined need is much stronger than a product made because it was technologically feasible to do so (Faste, 1987). This is because a product anchored by a real need is something that the customer wants, and should therefore have a market.

As Steinert and Leifer (2012) describe with their Hunter-Gatherer model, the path of any development project is seldom a straight line, but rather a series of shorter sprints steering the direction of the project along the way. We adapt this way of thinking to the needfinding process, zooming in, and treating the final need statement as the goal of the hunt! Figure 2 depicts two different engineering design processes. Figure 2a shows the incremental, “classical” engineering design process. It is fixed target design, specifications and requirements based. With other words, the boundaries are defined in the beginning, and these shape the process. The engineering design team is moving towards the goal, adapting their solution to the specified requirements. Figure 2b is the radical engineering design process, adapting the constantly evolving boundaries described by Steinert and Leifer (2012) in their Hunter-Gatherer model. The figure shows there are multiple paths available, and depending on where the design team decides to go, the boundaries change accordingly.

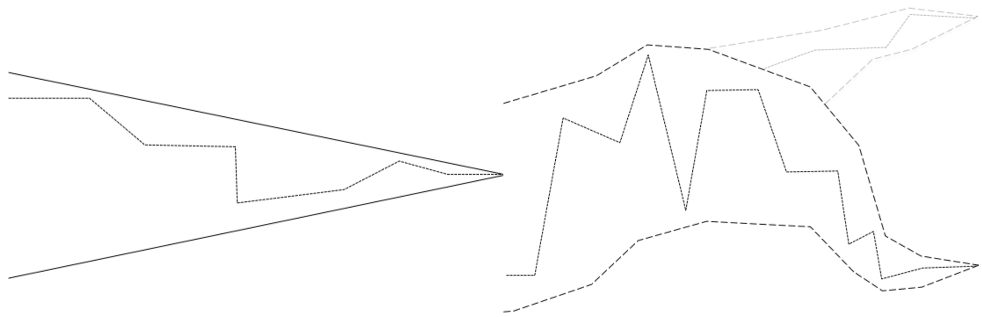


Figure 2a & b: a) Incremental engineering design process – adaptive in style. b) Radical engineering design process – innovative in style. (Steinert et al., 2012)

When doing needfinding, we have several tools at our disposal (Patnaik and Becker, 1999). Some of them enable us to experience needs and pains for ourselves, while some tools provide us with the experience of other people. We differentiate the tools into *unbiased*, what we experience directly, and *biased*, second-hand information. Among unbiased tools we find: observation, immersion and taking existing products apart. These tools are free from the interpretations and filtering of others. Biased tools are: video, storytelling, interviews, pictures and sketching to mention some. All of these have in common that they are a product of the needfinder's interpretation. Someone had to make a choice of what to communicate through their video, or what to tell when they were interviewed.
The biased needfinding tools allow us to control information provided, and through this we believe that it is possible to steer the attention of the designer and thereby have some control over the direction of the output. This is what we call priming or design fixation (Cardoso et al., 2009; Purcell and Gero, 1996).

Needfinding insight 1: By using biased needfinding tools, we are able to influence the needfinding process (by controlling attention)

There are mainly three aspects in the biased needfinding tools we can influence:

- The amount of information (how much)
- Selection information (what do we present)
- Emphasis on parts of the information (how it is communicated)

In this paper we will focus on the use of pictures in needfinding. We want to use pictures to firstly prime the design team with the content of the pictures, and secondly to capture results from team discussion in a visual manner through the team's selection and post-processing of pictures. Using photography as a research tool in an engineering design application, draws inspiration from the field of visual anthropology, learning how to observe and document findings (Collier, 1986). In his book, Collier (1986) describes the unbiased selection of information made by the camera. “The camera, by its optical character, has whole vision. No matter how select a unit we might wish to photograph, the camera faithfully records this specialized subject and also all other associated elements within focus and scope of its lens. This capacity makes the camera a valuable tool for the observer.” In addition to capturing the intended subject, there will always be additional contextual information in the picture frame, making the picture a much richer source of information than text. Text has the ability to only provide information of the intended subject, removing all superfluous information. The extra associated elements of the subject within focus and scope of the camera lens previously described by Collier are beneficial for needfinding, as the process is about creating an understanding for the situation where needs are to be found. This lets the viewer of the picture immerse himself in the situation depicted, based on the context supplied by background information in the picture frame. Collier pointing out the unbiased selection of information made by the camera is somewhat a paradox. Yes, the information captured within the focus and scope of the lens is to some extent unbiased, we will discuss later in the text how certain techniques can steer the focus of the viewer. On the other hand, a picture only captures an instantaneous moment in time, and only a small portion of the photographer’s surroundings. Overall the picture is a biased representation of reality, allowing us to control the perception of the viewer.

We know that pictures are widely used in needfinding exercises already. What has not been discussed previously is how different pictures may influence the thought process of the designer viewing them. Would it be possible to steer the needfinding process by using certain types of pictures?

Needfinding insight 2: By using certain types of pictures, it is possible to induce specific behaviours in the designer doing needfinding.

Photography

As a photographer, we have defined four main parameters that can be controlled without any major influence of the situation to be documented. These are depth of field, focal length, contrasts and framing. In this section, we give a brief introduction of each variable, and how it should influence the viewer.

Depth of field

Depth of field signifies how much of the picture will appear to be in sharp focus (Pentland, 1987). An object is in focus when light rays reflected of this object converge in a single point on the image sensor of the camera. Out-of-focus objects hit the image sensor in circles instead of converging to a point. These circles are termed "circles of confusion". When the circle of confusion becomes small enough to be indistinguishable from a point object in the final image, this point in space is considered to be in focus (Jenkinson, 2011, p. 38). A shallow depth of field means that there is only a small portion of the frame, close to the focal plane that will be in sharp focus. How much an object is blurred increases with the distance from the focal plane. A large depth of field results in a larger portion of the frame in focus. Depth of field is closely connected to the effective aperture of the camera. The aperture is the opening where light is let in to expose the film or sensor. The effective aperture of a lens is calculated by dividing the size of the lens opening d by the focal length f (Equation 1). Effective aperture is described by f-stops, so a 25mm lens opening divided by a focal length of 50mm results in the f-stop f/2 (Jenkinson, 2011). The lens f-number N is the denominator of the fraction describing the f-stop, *f/N*. The size of the aperture is inverse proportional to the depth of field i.e. a large aperture (big opening) results in a shallow depth of field. A small aperture (small opening) results in a larger depth of field. The reason for this is that rays of light that passes through a small aperture are more parallel and will result in smaller circles of confusion than when passing through a larger aperture.

$$f/N = \frac{d}{f}$$

Equation 1: Lens f-stop (Conrad, 2015, p. 5)

Total depth of field u_{total} can be calculated by Equation 2 from the focusing distance u, f-number N, focal length f and a predefined acceptable circle of confusion c, usually between 0.025mm and 0.035mm for full-frame image sensors. With constant magnification, focal length does not have any significant impact on the total depth of field. This means that we can assume depth of field is only dependent on effective aperture for constant magnification.

$$u_{total} = \frac{2uNcf^2(u-f)}{f^4 - N^2c^2(u-f)^2}$$

Equation 2: Total DoF (Conrad, 2015, p. 8)

We predict that depth of field will control the amount of information the viewer is able to extract from a picture. The eyes will be attracted towards objects in sharp focus, creating more emphasis on these. Out of focus objects are uncomfortable to watch, and is implicitly interpreted as unimportant by the viewer since the photographer has chosen to leave these objects out of focus. By using a large depth of field, we hope to spread the attention of the viewer throughout the picture, not giving emphasis to any specific region. A shallow depth of field will on the other hand steer the viewer's attention to the limited region in sharp focus.

Focal length

Focal length is a measure of the distance between the lens and the image sensor in the camera to focus sharply at infinity, i.e. objects far away (Jenkinson, 2011, p. 42). Different focal lengths give different effects in both angle of view and distance on perspective (Jenkinson, 2011). Lenses with short focal lengths are called wide-angle lenses, and lenses with a long focal length are called telephoto lenses. Wide-angle lenses can be used to include large portions of background, including the subject in a context, while telephoto lenses can be used to isolate subjects.

Changing focal length touches two aspects of how we can steer the viewer's attention. *First*, by changing focal length, we achieve different angles of view (Jenkinson, 2011). This influences the amount of background included in the picture, and thus the amount of contextual information captured. Due to this effect we predict that a shorter focal length (larger angle of view) will keep the attention of the viewer wandering, resulting in broader ideas and insights. Reversely, a longer focal length (narrower angle of view) will keep the viewer's attention more focused on the main subject. *Second*, the perspective of depth changes with focal length (Jenkinson, 2011). For shorter focal lengths, relative distances between objects in the depth-direction in the picture appear to increase. We call this *stretching*. Longer focal lengths create the opposite effect, called *compression*. This makes all objects appear to be closer to each other, and thus more similar in size. Using this, we predict that the stretching effect of the shorter focal lengths will put emphasis on the closest object in the picture frame, guiding the viewer's attention to hover around this object. Longer focal lengths will on the other hand compress the depth of the picture, making them appear more equal in importance.

Contrasts

Contrasts draw the attention of the human eye. This could be in the form of strong colours, or brightness of light (Ma and Zhang, 2003). Avoiding single points of contrast in a picture should keep the viewers attention moving around the frame to pick up as much information as possible. Using strong contrasts consciously will on the other hand help to steer the attention of the viewer towards certain areas of the picture.

Framing

By framing, we mean where we point the camera, how we orient the subject in respect to the background, and how tightly we crop pictures. This will affect both the amount of information included in the picture as well as which types of information that are presented together.

Framing the subject in front of an empty background should draw all the attention to the subject, while a detail-rich background will keep the attention of the viewer wandering around the picture. By changing the framing, the contextual information can be controlled, and thereby the interpretation of the viewer.

Needfinding insights from photography

Needfinding insight 3: By changing depth of field, focal length, contrast and framing, it is possible to control the output of the designer in terms of divergent/convergent thinking and guiding the direction of the project.

Picture types as needfinding tools

Based on needfinding insight 3, we propose three types of pictures using the discussed effects of the picture parameters to support design teams in the needfinding process. We call these characteristic picture types: Snapshots, emphasized pictures and illustrated pictures.

Snapshot

Snapshots (Figure 3) are used to start the problem reframing process, supplying unfiltered, detail rich pictures to spark open-ended discussion. The main strength of the snapshot is to supply a context, and is not supposed to steer the viewer in any specific direction. Snapshots are unedited photos of a given situation. Except for the inherent choice done by the photographer when choosing to take the picture as well as colour adjustment, the viewer is not steered to emphasize any part of the picture. These kinds of pictures should avoid using strong dramaturgical effects to draw attention to specific parts of the frame, but rather try to capture as much information as possible. In order to achieve these qualities, we propose to use a large depth of field, short focal lengths, framing with detail rich backgrounds and little use of contrasts.

Figure 3: Snapshots

Emphasized picture

The emphasized picture (Figure 4) directs the attention of the viewer towards certain parts of the frame, which are found relevant by the needfinder or the design team using different techniques, including both composition while taking the picture and post-processing. These techniques decrease the degrees of freedom by literally narrowing down the problem space in the picture frame. Emphasized pictures are of the same situation as the snapshots that have either been shot or retouched in such a way that focus is drawn to a specific part of the frame. This is done because the needfinding photographer has identified this as a point of interest, and wants to emphasize the importance of this detail to the viewer. This can either be done subtly with dramaturgical tools, or explicit with retouching. The photographer can use tools as shallow depth of field, longer focal lengths, clean backgrounds and contrasts to achieve steer the viewer's attention. To create emphasis in post-processing, the photographer can use crop, blur backgrounds and darken/lighten parts of the picture to increase contrast.

Figure 4: Emphasized pictures

Illustrated picture (Siddharth and Roy, 2015)

When insights start to crystallize and the design team reaches consensus on what needs they should address in their project, illustrated pictures (Figure 5) are used to freeze this information by striving to make this as unambiguous as possible. Everything considered unimportant by the photographer is removed by for example blurring and darkening. Important insights is then written and/or drawn on top to support and explain the content of the picture. Illustrated pictures should be used to explain a situation or an insight from the design team or the needfinding photographer. The goal is to ensure mutual understanding, by making information explicit. This can be seen as a further iteration of filtering from the emphasized picture, since the focus of the picture is analysed, and one specific bit of information is extracted. The illustrated picture can then be used as a basis for the resulting need statement that is handed over to the solution generation stage of the project.

Figure 5: Illustrated picture

Conclusion

We propose the conscious use of photography as a new tool in needfinding. By being in control of pictures taken and how they influence the viewer and design team, we hope to gain more insights toward the mechanics of the needfinding phase, and to be better equipped to increase the output quality, i.e. better the chances for the design team finding the “golden nugget”. Earlier in the text, three distinct types of pictures have been proposed to capture the various phases of the needfinding process (Figure 6): Snapshots, emphasized pictures and illustrated pictures. Each of these reducing opportunity for interpretation, focusing the attention of the viewer towards specific parts of the frame by obscuring or removing non-relevant information. The picture types have a dual purpose of priming the viewer with impressions, starting an iteration of narrowing down the final need statement, and to capture the outcome of each iteration in a visual manner to serve as either documentation of the process, or to prime the engineering design team in the next iteration cycle. We want to emphasize that the design team does not have to start the process with a snapshot and work their way through emphasized pictures and illustrated pictures. If the team already has sufficient knowledge of the problem, they can apply emphasized- or illustrated pictures directly in their process. We have already done some pilot testing of the tools, and they prove to be high value versus time spent based on the feedback from one of our company partners.

The next step for us will be to convert our needfinding insights into testable hypotheses. We are currently working on an experiment, exploring the effect of depth of field (see equation 1 & 2) and focal length combined with eye tracking to capture data. By the time of the conference we hope to be able to present some early data.

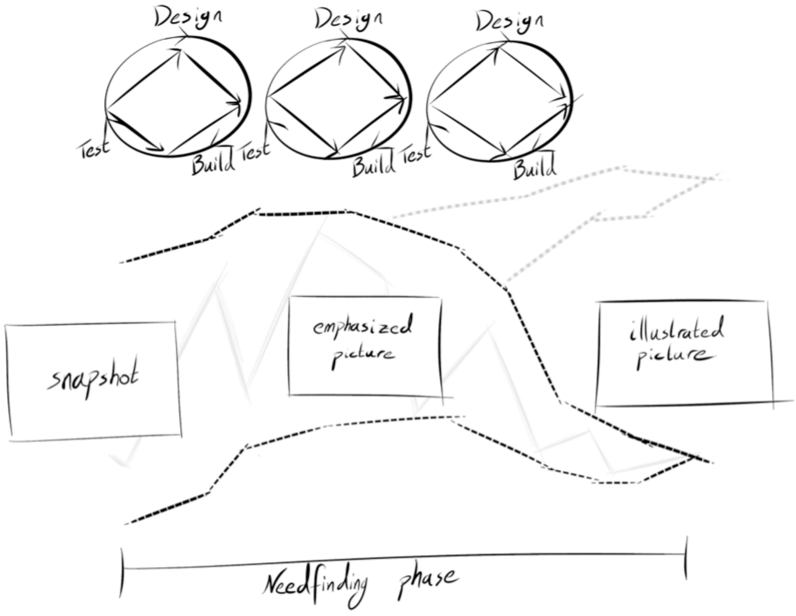


Figure 6: Photography needfinding process

Acknowledgements

We want to direct big thanks to Kristoffer Slåttsveen and Jørgen Andreas Bogen Erichsen for their valuable input in this project. A big thank you is also directed to Professor Amaresh Chakrabarti and his students for their hospitality when visiting Bangalore. This research is supported by the Research Council of Norway (RCN) through its user-driven research (BIA) funding scheme, project number 236739/O30 and by Henkel AG & Co. KGaA Duesseldorf; we would like to express our thanks for the continuing support.

References

Aldaz, G., Steinert, M., Leifer, L.J., 2013. Instrumenting the user: Needfinding by combining data logging and immediate self-reporting, in: DS 75-7: Proceedings of the 19th International Conference on Engineering Design (ICED13), Design for Harmonies, Vol. 7: Human Behaviour in Design, Seoul, Korea, 19-22.08. 2013.

Blessing, L.T., Chakrabarti, A., 2009. DRM, a design research methodology. Springer.

Cardoso, C., Badke-Schaub, P., Luz, A., 2009. Design Fixation on Non-Verbal Stimuli: The Influence of Simple vs. Rich Pictorial Information on Design Problem-Solving, in: ASME 2009 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, pp. 995–1002.

Collier, J., 1986. Visual anthropology: Photography as a research method. UNM Press.

Conrad, J., 2015. Depth of Field in Depth.

Eppinger, S.D., Ulrich, K., 1995. Product design and development. Prod. Des. Dev.

Faste, R.A., 1987. Perceiving needs. SAE Technical Paper.

Jenkinson, M., 2011. The Portrait Photography Course: Principles, practice, and techniques The essential guide for photographers. Peachpit Press.

Leifer, L.J., Steinert, M., 2011. Dancing with ambiguity: Causality behavior, design thinking, and triple-loop-learning. Inf. Knowl. Syst. Manag. 10, 151–173.

Ma, Y.-F., Zhang, H.-J., 2003. Contrast-based image attention analysis by using fuzzy growing, in: Proceedings of the Eleventh ACM International Conference on Multimedia. ACM, pp. 374–381.

Norman, D.A., Draper, S.W., 1986. User centered system design. New Perspect. Hum.-Comput. Interact. Erlbaum Assoc. Inc Hillsdale NJ.

Patnaik, D., Becker, R., 1999. Needfinding: the why and how of uncovering people's needs. Des. Manag. J. Former Ser. 10, 37–43.

Pentland, A.P., 1987. A new sense for depth of field. Pattern Anal. Mach. Intell. IEEE Trans. On 523–531.

Purcell, A.T., Gero, J.S., 1996. Design and other types of fixation. Des. Stud. 17, 363–383.

Siddharth, Roy, S., 2015. ID49: Research in Visual Ethnography Focusing on Markets of Kanpur. Presented at the ICoRD'15 Research into Design Across Boundaries, Springer, CPDM, IISc, Bangalore.

Steinert, M., Jablokow, K., Leifer, L.J., 2012. , in: EAGER: AnalyzeD - Analyzing Engineering Design Activities, Poster Presentation at the 2012 NSF Engineering Research and Innovation Conference, Sponsored by the National Science Foundation's Division of Civil, Mechanical and Manufacturing Innovation (CMMI), Boston, 9-12 July, USA.

Steinert, M., Leifer, L.J., 2012. "Finding One"s Way': Re-Discovering a Hunter-Gatherer Model based on Wayfaring. Int. J. Eng. Educ. 28, 251.

Thomke, S., Fujimoto, T., 2000. The Effect of "Front - Loading" Problem - Solving on Product Development Performance. J. Prod. Innov. Manag. 17, 128–142.

12 Appendix D - Problem text & Risk assessment

PDF with Problem Text and Risk assessment is added to next page

THE NORWEGIAN UNIVERSITY
OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF ENGINEERING DESIGN
AND MATERIALS

MASTER THESIS SPRING 2015
FOR
STUD.TECHN. Andreas Wulvik

Photography – Bridging the gap between needfinding and visual thinking

We propose the conscious use of photography as a new tool in needfinding. By being in control of photos taken and how they influence the viewer and design team, We hope to gain more insights toward the mechanics of the needfinding phase, and to be better equipped to increase the output quality, i.e. better the chances for the design team finding the "golden nugget". In the pre-master thesis (Fall 2014), three distinct types of photos have been proposed to capture the various phases of the needfinding process: Snapshots, emphasized pictures and illustrated pictures.

For the master thesis, we propose a confirmatory experiment on various dimensions influencing the effectiveness of photography as a tool in needfinding.

We hope to use the outcome of the master thesis as a repository for a future conference or journal paper.

Formal requirements:

Three weeks after start of the thesis work, an A3 sheet illustrating the work is to be handed in. A template for this presentation is available on the IPM's web site under the menu "Masteroppgave" (http://www.ntnu.no/ipm/masteroppgave). This sheet should be updated one week before the master's thesis is submitted.

Risk assessment of experimental activities shall always be performed. Experimental work defined in the problem description shall be planed and risk assessed up-front and within 3 weeks after receiving the problem text. Any specific experimental activities which are not properly covered by the general risk assessment shall be particularly assessed before performing the experimental work. Risk assessments should be signed by the supervisor and copies shall be included in the appendix of the thesis.

The thesis should include the signed problem text, and be written as a research report with summary both in English and Norwegian, conclusion, literature references, table of contents, etc. During preparation of the text, the candidate should make efforts to create a well arranged and well written report. To ease the evaluation of the thesis, it is important to cross-reference text, tables and figures. For evaluation of the work a thorough discussion of results is appreciated.

The thesis shall be submitted electronically via DAIM, NTNU's system for Digital Archiving and Submission of Master's theses.

The contact person is Martin Steinert

Torgeir Welo
Head of Division

Martin Steinert
Professor/Supervisor

NTNU
Norges teknisk-
naturvitenskapelige universitet
Institutt for produktutvikling
og materialer

ID	114	**Status**	**Dato**
Risikoområde	Risikovurdering: Helse, miljø og sikkerhet (HMS)	Opprettet	16.01.2015
Opprettet av	Andreas Wulvik	Vurdering startet	16.01.2015
Ansvarlig	Andreas Wulvik	Tiltak besluttet	
		Avsluttet	16.01.2015

Masteroppgave vår 2015

Mål / hensikt
Redusere risiko

Bakgrunn

Beskrivelse og avgrensninger
Risikovurderingen omfatter arbeid i forbindelse med skriving av masteroppgave

Forutsetninger, antakelser og forenklinger

Vedlegg

Referanser

Ansvarlig for risikovurderingen

Andreas Wulvik

Deltakere

Lesere

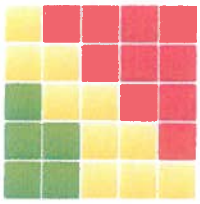
Enhet /-er risikovurderingen omfatter

- Institutt for produktutvikling og materialer

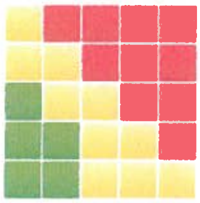
Følgende akseptkriterier er besluttet for risikoområdet Risikovurdering: Helse, miljø og sikkerhet (HMS):

Helse

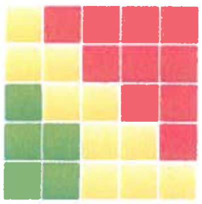
Ytre miljø



Materielle verdier



Omdømme



Følgende farer og uønskede hendelser er vurdert i denne risikovurderingen:

- **Stiv nakke og rygg**
    - Stiv nakke og rygg

Eksisterende og relevante tiltak:

Kaffepause for å strekke på beina

Tar regelmessige pauser for å røre litt på kroppen. Dette forhindrer stillesitting i lange perioder

Farekilde: **Stiv nakke og rygg**

Beskrivelse:

Stiv nakke og rygg pga stillesitting

Uønsket hendelse: ***Stiv nakke og rygg***

Beskrivelse:

Dårlig ergonomi kan føre til stiv nakke og rygg

Sannsynlighet: Sannsynlig (3)

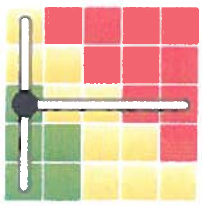
Begrunnelse: Fort gjort å ha dårlig ergonomi. Ikke spesialtilpassede kontorplasser for studenter.

Konsekvensvurderinger:

Aspekt: ***Helse***

Konsekvens: Liten (1)

Begrunnelse: Stiv nakke og rygg går fort over.



Følgende risikoreduserende tiltak er besluttet:

Revurdering av risiko etter besluttede tiltak:

Sluttresultat:

Farekilde: **Stiv nakke og rygg**

Uønsket hendelse: ***Stiv nakke og rygg***

Aspekt: ***Helse***

Risiko før tiltak: ● Risiko etter tiltak:

Endelig vurdering

Vedlegg