A method to empirically investigate design manipulations

Erik Engervall
Abstract

A method to empirically investigate design manipulations

Erik Engervall

Marketplaces, similar to conventional e-commerce websites, rely on conversions as primary source of income. Conversions are visitors turned into customers, and a condition for this to happen is that visitors stay on the marketplace and interact with its content. A visitor that does not interact at all is considered a bounced visitor and is added to a metric called bounce rate. Plick requested a concrete method for executing a design experiment on their website where multiple design manipulations could be run in parallel and record data individually. During the thesis, said method was built and an experiment was executed with four unique design manipulations that were developed by combining two factors, compactness and vendor credibility. These factors were found through a theoretical analysis based on human visual functioning and emotional design. The experiment showed potential for both the theory and the method, but no statistically significant differences between the implemented design manipulations were found. The further use of this method is discussed as well as how to achieve a statistically significant result.
Acknowledgements

First and foremost I would like to thank everyone at Plick for their support and encouragement during this entire thesis work. Special thanks goes out to my supervisor Jimmy Heibert for his guidance and Christian Alex for always happily sharing his technical competence.

Next I would like to thank to my reviewer at Uppsala University, Mats Lind, for his endless patience and help with keeping me on the right track throughout the entire thesis work.

I would also like to thank my friends and family for motivating me and for helping me with proofreading and brainstorming.

Finally, I would like to thank all of my classmates from IT 12’ for having made these past 5 years an amazing and memorable experience.
## Contents

1 Keywords 10

2 Introduction 1

2.1 Background ........................................ 1

2.2 Plick - the marketplace .............................. 1

2.3 Commercial goals .................................. 2

3 Purpose 4

4 Part 1 - Theory 5

4.1 Web analytics ...................................... 5

4.1.1 Advantages with web analytics ................. 5

4.1.2 Disadvantages with web analytics ............... 6

4.1.3 Automatic data collection ....................... 6

4.1.3.1 Server log analysis .......................... 6

4.1.3.2 Page tagging .................................. 8

4.1.4 Google Analytics ................................ 9

4.1.4.1 Google Tag Manager .......................... 9

4.1.4.2 Important metrics in GA .................... 10

4.1.4.3 Technical difficulties ....................... 11

4.2 Human visual functioning ......................... 11

4.2.1 Visual thinking .................................. 11

4.2.2 Seeing the world ................................ 13

4.2.3 Perceiving the world ............................ 14

4.2.3.1 Process selection ............................ 15

4.2.4 Cognitive task ................................. 16

4.2.5 Visual query ...................................... 16

4.3 Affection in design ................................. 17

4.3.1 Emotional design ................................. 18

4.3.1.1 Visceral level ................................ 18

4.3.1.2 Behavioral level ............................ 19

4.3.1.3 Reflective level ............................ 20

4.3.1.4 Processes .................................. 21

4.4 Theoretical conclusion ............................ 21

5 Part 2 - Empirical investigation 23

5.1 Research questions ................................. 23
5.2 Limitations ............................................. 23
5.3 Introduction to testing ................................. 23
  5.3.1 The marketplace's current state of design ........ 23
  5.3.2 Time spent per page .............................. 25
  5.3.3 Perceptive emotion ................................ 26
  5.3.4 Design manipulations ............................. 28
    5.3.4.1 Design manipulation 1.1 ....................... 29
    5.3.4.2 Design manipulation 1.2 ....................... 30
    5.3.4.3 Design manipulation 2.1 ....................... 30
    5.3.4.4 Design manipulation 2.2 ....................... 32
5.4 Method ................................................. 39
  5.4.1 Data collection .................................... 39
    5.4.1.1 GTM configuration .............................. 39
    5.4.1.2 The Custom HTML Tag ......................... 42
5.5 Developing the design manipulations ................. 43
  5.5.0.1 HTML and JavaScript .......................... 44
    5.5.0.2 CSS and LESS ................................. 44
  5.5.1 Activating the experiment ....................... 45
  5.5.2 Interpreting data from GA ....................... 45
5.6 Results ................................................ 48
  5.6.1 Data from GA .................................... 48
  5.6.2 Chi-square test .................................. 48

6 Conclusions ............................................. 50

7 Future Work ............................................. 51
  7.1 Correlation between bounce rate and conversion rate .... 51
  7.2 Include more devices and platforms .................. 51
  7.3 More radical design manipulations ................... 51
  7.4 Effect size and statistical significance ............. 51
1 Keywords

Bounce rate, conversion, marketplace, peer-to-peer, e-commerce, web analysis, web design, perception, emotional design
2 Introduction

2.1 Background

The amount of internet users is steadily increasing every year[6] and it becomes increasingly important for developers and designers to understand how their users perceive their websites. For businesses whose sole product, and thus single source of income, is a web or mobile application, providing users with an experience that is superior to competitors might be the difference between success and failure, regardless of content and features. Even more so nowadays, considering the speed at which for example web applications can be developed and deployed. Compared to 5-10 years ago, developers do not need to be experts in programming to build relatively complicated applications covering the full stack (back- and frontend). The developer tool box has increased immensely and cloud development solutions such as Amazon Web Services[1] can provide scalable servers and databases at a click of a button. Also, modern development frameworks such as Angular.js, Django and Ruby on Rails further simplifies and speeds up the development process. Considering all these available tools, the real contest is not necessarily being first - but rather delivering the best user experience possible. It has been shown that roughly half of all projects’ codebases is devoted to constructing the user interface[12] which underlines the importance of properly built user interfaces.

2.2 Plick - the marketplace

Plick is a peer-to-peer e-commerce, henceforth referred to as a marketplace, web and mobile application for second hand clothing in Sweden. Conventional e-commerce websites offer fungible merchandise, normally at the expense of variation, whereas a marketplace can offer unique and varied items at a potential bargain. The marketplace’s traffic consists of visitors that can browse content and registered visitors, users, that can be both vendors and consumers or neither. A user becomes a vendor by posting an item on the marketplace. The item is posted by creating a so called ad, which is viewable and purchasable by others on the marketplace. Similarly, a user becomes a consumer by purchasing an item from a vendor. Whenever an item is sold, the item’s ad will be marked as such but remains on the marketplace for search engine optimization purposes. Users can also follow and be followed by other users.

Plick’s web traffic are primarily organic visitors, as seen in figure 1, from search engines such as Google and Bing. The queries which generate the marketplace’s organic traffic consists primarily of long-tail keywords. Long-tail keywords are search
queries that are longer and more specific than regular keywords. These kinds of queries are typical for visitors that already have a good idea of what they want and are closer to the point of purchase, meaning they are more likely to purchase something from one of the query results. Long-tail keywords are very cost efficient as the search competition gets smaller and smaller as the query gets longer. They are also highly organic in the sense that the users create them themselves when creating ads. Specializing in long-tail keywords rather than paying search engines to appear high on a wider spectra of results reduces the flat amount of visitors on the website. However, it does increase the likelihood that one of the long-tail visitors become a consumer due to the precision of the result [22].

2.3 Commercial goals

Marketplaces and conventional e-commerce websites often share rudimentary design components such as product images and related ads. They also share a high focus on conversion, which is the act of turning website visitors into consumers. Conversion is measured as the ratio of the number of paying customers to the total number of visitors, a ratio known as conversion rate.

In order for conversion to occur, some sort of interaction must first take place, either through purchase, content creation or continued browsing. No matter the interaction’s type, it affects the bounce rate. Bounce rate is the ratio of the number of visitors that leave without any interaction to the total number of visitors. The
ratio is presented as a percentage which indicates how well the website is performing in regards to encouraging visitors to browse or interact with content on the website. Bounce rate is highly relevant for marketplaces, and e-commerce websites in general, since interaction is a prerequisite for conversion.
3 Purpose

This thesis investigates ways to reduce bounce rate on Plick’s web marketplace. To achieve this, a theoretical analysis will be performed to identify potential factors that could influence bounce rate. When potential factors have been identified, they will be investigated empirically.
4 Part 1 - Theory

This chapter will focus on topics related to how web design can affect bounce rate. It will shed light on different web analysis methods, human visual functioning and emotional design. The long extent of this chapter is due to an identified lack of theoretical and scientific research in the subject of bounce rate and marketplaces. The chapter will outline the most important theoretical content followed by a hypothesis that will be put to the test in an experiment.

4.1 Web analytics

The term analytics refers to the discovery, interpretation and visualization of data. Analysts often study historical data in order to make more informed decisions in their work. Web analytics is a modern extension of analytics, the main difference being the datasets themselves and the method to acquire them. Web analytics is made possible through clever software programs that automatically collect visitor data[7]. The data can be structured and described in different ways, but for Google Analytics, the software program that will be used in this thesis, the measurements are called metrics. Metrics are a purely quantitative measurement and could for example be the total number of user sessions on a website[3]. The quantitative nature of metrics makes it a befitting measurement for benchmarking performance. Chapter 4.1.4 will cover how metrics can be used in the analytic process together with dimensions.

Web analytics’s area of application extends further than simply observing metrics. It is up to each individual analyst to interpret the data in a way that matches the platform. By comparing metrics before and after for example advertisement campaigns or major feature releases, analysts can assess the results and improve future campaigns or releases accordingly with much less difficulty.

4.1.1 Advantages with web analytics

Web analytics generates large, representative samples from authentic users. The data is unbiased in the sense that it is collected from users performing real tasks on the platform. There are no human intermediaries that can be biased during the recruitment process and influential during the task testing process.

Because web analytics automatically collect data, it is extremely cost efficient. It does not require immediate attention of staff members before or during data collection.
4.1.2 Disadvantages with web analytics

If not properly specified, web analytics can provide vague and not as qualitative data as other testing tools. It is difficult to get data on how visitors feel when interacting with the website or why they click as they click. Simply put, pin pointing the website’s less obvious weaknesses becomes ambiguous.

4.1.3 Automatic data collection

Automatic data collection plays a vital role in modern web analysis. A single visitor during a single visit can provide a lot of data, not only through direct interactions such as clicking, but also through more discrete data points such as being a new or recurring visitor. This sort of data would be extremely difficult, if not impossible, to gather manually.

There are primarily two methods when it comes to automatic data collection on websites, server log analysis as seen in chapter 4.1.3.1 and page tagging as seen in chapter 4.1.3.2[11].

4.1.3.1 Server log analysis

Server log analysis is a web analysis method that leverages the fact that all web servers are capable of saving incoming requests as entries in log files. The logging feature is typically enabled by default and the log files are stored directly on the web server. Entries are typically generated through client’s accessing the web server through a browser, but can also be the result of a search engine crawler paying a visit. The simplest form of request, and one of the most important in terms of web analysis, is a page request. A page request is sent from the client to the web server when visiting a particular page that resides on the web server. The format of the
log entries can differ depending on a number of factors such as operating system, web server software and even web server software version. Looking at the world’s most used web server software, Apache[20], the default format is called Common Log Format and has the following pattern:

%h %l %u %t "%r " %>s %b "%{Referer}i " "%{User-agent}i "

<table>
<thead>
<tr>
<th>Format string</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>%h</td>
<td>Remote hostname, typically the client’s IP address</td>
</tr>
<tr>
<td>%l</td>
<td>Remote logname (from identd if supplied)</td>
</tr>
<tr>
<td>%u</td>
<td>Remote user if authenticated request</td>
</tr>
<tr>
<td>%t</td>
<td>Time when request was received</td>
</tr>
<tr>
<td>%r</td>
<td>First line of the request, for example “GET / HTTP/1.0”</td>
</tr>
<tr>
<td>%&gt;s</td>
<td>Server status code response</td>
</tr>
<tr>
<td>%b</td>
<td>The response’s size (in bytes), including headers</td>
</tr>
<tr>
<td>%{Referer}</td>
<td>The page from which the request was initiated from</td>
</tr>
<tr>
<td>%{User-agent}</td>
<td>Client browser</td>
</tr>
</tbody>
</table>

Table 1: Server log formats

A practical example: 127.0.0.1 - erik [23/Feb/2017:13:37:00 -00] "GET /destination.gif HTTP/1.0" 200 2326 "http://engervall.com/" "Mozilla/4.08 [en] (Win98; I ;Nav)"

A user, Erik, visits http://engervall.com/destination.gif with http://engervall.com as origin. The request is sent on 23rd of February, is of size 2326 bytes and has success status code 200. The IP address is 127.0.0.1 and the user’s browser is Mozilla running on a Windows 98 operating system.

Popular websites will have web servers with immense log files consisting of tens of thousands of entries. In order to make sense of this large amount of practically unreadable data, third party software is used. The software will interpret the log files and visualize the data in a human readable way. The most compelling reasons to use server logs are, according to Sawmill’s executive white paper [15], listed below.

- No third party cookies are used, meaning that browsers that has cookies disabled by default will not be excluded from the results.
- Browsers with JavaScript disabled will also not be excluded as the method does not require JavaScript.
- In-depth tracking of all transaction with the server, not only page requests but also for example media and file requests.
- By keeping the data directly on the server, it is highly secure and provides extra confidentiality towards the users.
• The data is available from the moment the web server receives traffic.

There are some major disadvantages as well. As mentioned, the data is saved in server log files that just keeps on growing, which can cause memory issues quickly if a website is popular. The removal of old data would thus remove the last advantage in the list above. This will also add to the overall cost as the download of log files will require a lot of bandwidth. Additional costs is a machine to run the actual log analytics software on, as well as storing the log files on, and the cost of the log analytics software itself.

4.1.3.2 Page tagging

![Page tagging process](image)

Figure 3: *Page tagging process*

Page tagging is a web analysis method performed on the frontend. Code is typically injected on a website through a small piece of JavaScript code. The JavaScript code creates cookies in the clients’ browsers in order to keep track of their activity on the website. The activity is send as pure data points to a third party application that presents the data and allows for cross referencing of parameters, data visualization and much more.

A practical advantage of using third party tools such as Google Analytics, as seen in chapter 4.1.4, is that user information such as age, location and origin before arriving to the website usually is available as Google runs a large portion of web traffic through their search engine.

Page tagging requires JavaScript to run which can cause issues on certain feature phones or browsers with specific ad blocking extensions installed. Ad blocking browser extensions typically do not affect page tagging unless specifically configured.
4.1.4 Google Analytics

Google Analytics (GA) is a tool developed by Google that helps website owners to analyze their websites’ traffic. It is a free web analytics service that allows tracking and reporting of all traffic to the website that it is installed on. The service is implemented through a code snippet that is included on each page on the website. The snippet runs in the client browser and collects the client’s data which is sent to the associated GA account to be interpreted as reports.

4.1.4.1 Google Tag Manager

Google Tag Manager (GTM) is another tool developed by Google that exists in symbiosis with GA. GTM allows website owners to add something called Tags on their websites which gathers in-depth data from custom events such as button clicks, pageviews and other behaviors that might be of interest. The fact that Tags can be created through Google’s web application enables non developers to take part of the analytics process and opens up a world of possibilities. It also enables website owners to customize when and how certain data should be collected with fine grained control, without having to dig through code.

Getting started with GTM requires that one defines a container for one’s website, typically named after the domain name. GTM generates a code snippet associated with the container that has to be added on each page of the website, after which further interaction with the website’s code base is optional. GTM’s core concept revolves around Tags which are essentially HTML code that is injected into the visitor’s browser once GTM’s code snippet has executed. Tags exist to send activity data to third party applications such as GA. How and when tags are triggered is controlled by Triggers and Variables. Triggers are based on events such as pageviews and can be attached to Tags to specify when they should fire. Variables are used to help define Triggers and configure what data they send. There are many conditions, some examples are uniform resource locators (URLs), element classes or id and random number. A URL is simply a reference to a resource on the web. Finally, the dataLayer is an optional JavaScript object that helps manage the information that Tags gather from your website. The dataLayer is object code that can be used to pass events and variables to GTM, achieving high quality data.

Injecting code through GTM has the great advantage of simplicity in both code and architecture complexity. However, an obvious downside with manipulating design through JavaScript within a GTM tag is the sometimes visible flickering effect that appears. The flickering effect is when all the HTML, CSS and JavaScript has loaded but the GTM tag is still being fetched and has not started executing yet. When the GTM tag is completely fetched, the code will run and the design will be
updated, assuming of course that the visitor has been assigned a design manipulation.

4.1.4.2 Important metrics in GA

**Session** One of the most important metrics in Google Analytics is a session. A session is an aggregated group of interactions committed by a single visitor during a certain period of time on a website. The session lasts until the visitor either leaves the website or is inactive for, as default, 30 minutes. The aggregated group of interactions can of course be empty, meaning that a visitor entered the website but then instantly became idle. If the visitor left the website or was idle for more than 30 minutes, the session would be considered a bounced session.

**Pageview** A pageview is a metric that is recorded whenever a visitor visits a particular page on the website. The pageview will be associated to that particular page and will occur whenever the page is loaded, or even reloaded.

**Dimensions** A dimension is a way of looking at the data. For example, attributes such as Device and Browser can be used as dimensions. In a way, it is used to gain a clearer view of the audience and how they interact with the website. Important for a dimension is that the object it describes can have actual values. The dimension Device could have dimensions such as Device Name and Device Resolution whose values could be Android and 100x200 pixels. GA can also create custom dimensions if need be.

**Bounce rate** Bounce rate is a measurement of website visitors leaving a website without any interaction. Interaction is considered as the resulting effect that two or more objects’ actions have on one another. In terms of a website, an interaction could be clicking on a link which would lead to another page. In this thesis bounce rate is considered as the key performance indicator (KPI). A KPI is a term to describe activities that bring success to businesses. In terms of a marketplace, success means having a high conversion rate, which has a strong correlation to bounce rate. A bounce can be caused by a wide range of reasons from erroneous website behavior to user misinterpretation or the website not meeting users’ expectations.

The strongest advantage with a marketplace with a high percentage of organic visitors is the fact that each visitor has an obvious intent. What most companies spend a great deal of effort acquiring through payed means, search engine ads and sponsored content on other websites, is provided for free and on users’ initiatives nonetheless. Furthermore, knowing what ads the users’ land on, the marketplace can accurately generate related ads depending on category and other attributes. This
also gives an advantage in engagement encouragement over conventional e-commerce platforms as it is possible to know more specific information about the visitor right away.

An improved bounce rate can be achieved in various ways. Common methods are either calibrating the calculations or improving the user interface. Calibrating the calculations means excluding or including certain results into the statistics and can be helpful if for example it is believed that users with a short visit duration should not count and so forth. It is non-intrusive on the actual product and can only aid in improving data precision. The downside with this method is that visitors, which may be potential customers, are ignored due to specific behavior that may or may not actually be improving the statistics and helping in improving bounce rate. The preferred method, and the one which will be applied in this thesis, is to improve the user interface. In product development, time is often of the essence and logic dictates that the highest positive impact will be gained by trying to improve the bounce rate on the most common landing page. The landing page is the specific page that visitors arrive at after following a link on a search engine. Due to the marketplace’s high percentage of organic traffic from long-tail keywords, the most common landing page is an ad. Each ad has individual content but has the same design template, which will be perfect for design improvements.

4.1.4.3 Technical difficulties

Depending on the architecture of the website, certain problems can arise that can cause GA to report false data. If the website is using particular server side rendering methods and a link is clicked, it is not certain that the script that runs GA will realize that a new page has been loaded. To the client, the page appears like any other, but in reality only the necessary parts have been replace. The technique is used to relieve the loading burden from the frontend. The problem can easily be avoided through sending custom events to GA from GTM that a click has taken place. Through GTM additional information such as what link or button was clicked and where it leads can be extracted and sent to GA. The custom events from GTM can also be configured to act as either interactive events or non interactive events, meaning that one can register events that do not affect bounce rate.

4.2 Human visual functioning

4.2.1 Visual thinking

Human beings tend to overestimate their own capabilities and unconsciously perceive themselves as the greatest in many aspects. This includes our ability to apprehend
visual information from our surroundings caught by our eyes. While some people are more attentive than others, our ability to apprehend visual information is truly not as great as one might think. In fact, an experiment[16] conducted in 1998 by Daniel J. Simons and Daniel T. Levin put this inability to apprehend visual information to the test by experimenting on pedestrians. The experiment consisted of four actors that worked together to perform a trick on unsuspecting pedestrians. The pedestrians were approached by the first actor holding a map asking for directions. During the conversation two more actors appeared carrying a door behind which the final actor hid. The door had to split the first actor and the pedestrian so that the first and the final actor could switch places. The first actor would thus take the final actor’s place, hiding behind the door, and vice verse.

![Figure 4: The frames a-d are captured from a video of the pedestrian experiment. Frames a-c shows the substitution trick in action and frame d shows the two actors that performed the substitution. The image is from "Failure to detect changes to people during real-world interaction" [16], courtesy of Daniel J. Simons.](image)

Even though the final actor could differ in both subtle attributes such as hair color and more obvious such as clothing and even gender, more than 50 percent of pedestrians were unable to make the distinction and point out the substitution.

The experiment’s result is due to the tradeoffs that evolution has had to make with the human brain. The brain consumes high amounts of energy by default[17] and must thus be of a reasonable size. This limit in brain real estate introduces memory capacity limitations, it would not be practically feasible to store a copy of the surrounding visual world in memory at all times, so only the most critical
and currently relevant information is stored. It is much more efficient to adjust our narrow focus towards particular objects required for particular tasks. More on focus and tasks will be covered in following chapters.

4.2.2 Seeing the world

The process of seeing is not a trivial one and plays an obvious crucial role in our ability to perceive the world. The human eye is similar to a digital camera, or rather, a digital camera is similar to the human eye. The camera however functions well as a metaphor. Just like a digital camera, the eyes contain lenses that focuses an image of the world on the eyeball and has an array of light-sensitive cones that can read the three different colors that make up all colors that humans can perceive: red, green and blue. Another similarity is the ability to compress the image data, which the camera does in a digital fashion and the eye’s retina does by using its several layers to extract the significant and relevant information. However, the two differ greatly in their method of signaling the memory storage from the sensors, be it the brain or the digital memory storage in the camera.

The camera’s pixels are structured as a uniform grid whereas human vision is nonuniform. This entails that in order to see in perfect visual detail, the targeted object must be in the fovea, which is the absolute center of our field of vision. The further out towards the edges, the worse the vision gets. This can easily be empirically tested by considering any situation where a stranger stands next to you. The eyes can make out the figure of a person standing next to you, but can not provide any specific details such as who the person is.

Figure 5: Illustration of the visual field. Image from Visual Thinking for Design [21], courtesy of Colin Ware.
As pointed out in figure 5, the majority of our visual processing power is concentrated in the parafovea, which is only slightly larger than the fovea itself. The parafovea makes up less than 5% of the visual world which would explain the high frequency at which our eyes move, we are simply trying to put the visual world together piece by piece, much like a panorama photo which is a set of images pieced together in series to create a larger one.

The comprehension that we analyze but a fraction of the visual world has improved the way that designer go about their work. The realization that our visual world is provided on a need-to-know basis is something that will weigh heavily during the design development process in this thesis.

4.2.3 Perceiving the world

The process of sight is important for humans to understand the world, to perceive our surroundings. However, sight is not the same as perception, perception is much more than just seeing the world. Perception is described as gaining knowledge, becoming aware or identifying something using our senses. Because this thesis has a strong focus on websites, the most important sense is, naturally, sight. Hearing, tasting, smelling and touching are yet to become important senses when it comes to interactions with websites.

Perception correlates strongly with sight and, according to Colin Ware[21], is determined by one of two processes. It is either determined by a process called bottom-up or a process called top-down. The names correlates to the processes’ algorithms. Both algorithms consist of the same three steps, but in reversed order. The three steps, in the bottom-up process’s order, are features, patterns and objects. The top-down is thus in the reversed order, starting with objects and ending with features with patterns in between.

Features refers to the processing stage where the sensory input is received in the primary visual cortex of the brain, also known as visual area one (V1)[9], that is when something has been detected. Neurons in the cortex are tasked to detect features in parallel and the features detected can range from size information, colors and much more.

The features are then used to build complex patterns that are used to construct our visual field. Similar features are grouped and form patterns displayed in our visual field, for example three dots with the colors red, green and blue would be segmented into three patterns that make up our visual field.

Lastly, the objects are the result of careful pattern processing. The objects, or visual objects, consists of both visual and non-visual attributes and are interpreted as a set of linked features that may vary from person to person depending on dif-
ferent factors such as previous experience and physiological predispositions. The colored dots would be interpreted as objects and their features would, combined with each individual’s personal associations and previous experiences, create a sense of perception of the object. Furthermore, the features and associations will cause us to prepare an appropriate action, also based on previous experiences. For example, the green dot might be associated with positive feedback and will thus cause a preparation of an interactive action whilst the red dot might be associated with negative feedback and is likely to be avoided. Our perception of the object is a combination of the visual features, our previous experiences that has turned into associations and the prepared actions.

The objects are stored in visual working memory, similar to a short term computer cache, which can only hold up to about three objects at any one time for a short period of time. The visual working memory limit capacity is similar to the famous “seven plus or minus two” theory stated by George A. Miller in 1956[10] although they focus on different areas of memory. The seven plus or minus two theory refers to the amount of items that our short-term memory can handle in terms of retention and recollection of items and not the actual recognition as with visual working memory. This limitation of visual working memory explains the pedestrians’ failure to recognize the substitute actor in the experiment by Daniel J. Simons and Daniel T. Levin. The limitation further stresses the importance of not overwhelming users with information, either with superfluous content or extravagant animations as each of these will end up as objects in users’ limited visual working memory.

4.2.3.1 Process selection

Which process that is selected is highly situational and depends on the individual’s current motivation. Motivation is generally driven by a particular need, which is also highly situational and personal. The needs can be of primal nature such as dealing with hunger or ensuring safety, but can also be more modern concepts such as personal expression through social media. Whatever the motivation may be, it entails a cognitive task in order to fulfill the need. Exactly what a cognitive task is will be covered in the next chapter.

If, for example, you are motivated to find a fruit to deal with your declining blood sugar, you will have an object in mind and your cortex will attempt to pattern match preexisting knowledge and associations of colors, shapes and other attributes to correspond with the features provided through your eyes. This would be an example of the top-down process, the task is already set and the eyes are scanning to match features to an object. The bottom-up process is whenever an object is not specified and could be compared to aimlessly viewing the world, continuously...
translating features to objects through patterns. Taking a stroll down the street is a prime example of the bottom-up process at work.

4.2.4 Cognitive task

A cognitive task, sometimes referred to as a cognitive process, is similar to a regular task except for the prominent use of cognitive activity rather than physical. It still includes the common elements of a task such as goals, steps, skills and desired output. The core of a task is to achieve some sort of result, which seldom is achieved by a single action and is therefore divided into subtasks or, as John B. Carroll puts it, Elementary Cognitive Tasks (ECT)[2]. Carroll defines ECT as “An elementary cognitive task (ECT) is any one of a possibly very large number of tasks in which a person undertakes, or is assigned, a performance for which there is a specifiable class of ‘success’ or ‘correct’ outcomes or end states which are the be attained through a relatively small number of mental processes or operations, and whose successful outcomes depend on the instructions given to, or the sets or plans adopted by, the person.”. This quote has a disclaimer that reads that some authorities disapprove of the expression ECT as it can be argued that there is no such thing as an elementary task. However, Carroll explains that the expression merely points out that there exists tasks that are simpler than other tasks. Tasks in general are central in the understanding of usability. For instance, the International Standard Organization’s (ISO) defines usability as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use”[5]. Goals in this context can be understood as definitions of tasks to be performed.

A practical example of a cognitive task could be to shop for an item online. The primary cognitive task would entail a result of an item purchased. In order to complete the cognitive task, it would first have to be divided into ECT’s. Each ECT would produce a result, either a success or a failure, that would bring the user a step closer (or further away depending on the outcome) to completing the primary cognitive task. ECT’s such as navigating to a single product are essential for the primary cognitive task and should be simple to perform.

4.2.5 Visual query

ECT’s are great for breaking down cognitive tasks and understanding users’ train of thought. The ECT’s are however not automatically performed simply because they are defined. In order to perform an ECT we need to know the critical information required. How to obtain the critical information required varies from task to task, but more often than not involves our eyes and our ability to perceive. As mentioned
in previous chapters, our sight is highly concentrated to a small area called the fovea in which we clearly view the world. Where our fovea looks, where our focus lies, is determined through visual queries. A visual query, like the name suggests, is a query for visual information meant to aid the execution of a current ECT. Colin Ware explains the importance of visual queries in the design process as “Effective design should start with a visual task analysis, determine the set of visual queries to be supported by a design, and then use color, form, and space to efficiently serve those queries.”[21].

Visual queries is thus a big part of design and according to Colin the starting point of the design process. The queries are primarily for aiding the ECT, but can serve multiple purposes depending on the situation. Imagine if the current ECT is to determine if, for example, we want to buy the piece of clothing we are looking at. A successful ECT, in the case of having a cognitive task that is to complete a purchase of a piece of clothing, would mean that we approve of the piece of clothing. However, if we disapprove, it can be seen as a failure in the chains of ECT’s and the cognitive task is cancelled. In this moment, there should be other incentives for visual queries and new cognitive tasks. We can only query what is visible so it must be in view and it must catch our attention in such a way that it does not obstruct our original visual query, but helps us form a new cognitive task should the original fail. In the case of a marketplace, this artificial incentive could be something like related ads. This will be further examined and experimented on in chapter 5.

4.3 Affection in design

As explained in previous chapters, perception is the core of our ability to understand of the surrounding world and is thus also an important part of our decision making process. A common misconception of our decision making process is that it is solely based on rationality and logic. It is in fact, according to neuroscientist Antonio Damasio[13], much more based on the affective system. This assertion made by Damasio was the result of a study he conducted on patients with brain injuries that impaired their affective systems. The study revealed that due to the impaired emotional system, patients had an extremely hard time forming their own decisions in both important and everyday matters. It ranged from where they want to live, what products to purchase, what products to use and even what to eat. The lack of emotional response when confronted with a trivial choice leads to difficulties in making a decision, especially if the outcome is more or less the same. Booking a meeting on Monday or Tuesday, assuming both days are available, can become a most burdensome task since both alternatives has the same outcome and no logical or rational advantages. It is where an emotional response would help the decision
making process. This is even more relevant when looking at products and branding, more often than not when purchasing for example groceries we value the brand above a wide range of attributes such as pricing and quality. Even more so when it comes to clothing, where brands can have significant meaning to one’s personality and social expression.

Our decision making is highly dependant on our ability to perceive and feel affection, which would imply that perception and emotion has an unexplored connection. Emotional design can thus be derived to play an important role in the decision making process of marketplace visitors. How can a marketplace, or any website for that matter, engage visitors to become registered users? Further knowledge in the field of emotional design is required to form a hypothesis.

4.3.1 Emotional design

Design should be built to provoke emotion and to create a connection on a personal level with the user. The user should feel good when interacting with the product, because that good feeling helps the decision making process and brings the user one step closer to a decision.

Human emotion is extremely complicated due to the complexity of our brain structure. Yet, to this day it is not clear exactly how our emotions work and there is no real consensus between scientists on a definition. There are however theories of possible structures that help visualize how our emotions work. Donald Norman is well known for his work in the field of design and cognitive science and has together with professors Andrew Ortony and William Revelle from the Psychology Department at Northwestern University formed a theory[13] on how human emotion operates. The theory proposes that our emotions has three levels, the visceral level, the behavioral level and the reflective level. In short, the visceral level is biological instinct from birth, the behavioral level is developed during our upbringing and the reflective level is analytical and intellectual. These three levels each play different roles in our affective system and together enable us to feel emotion. Understanding how each level work can help designers understand which level to focus on for different products.

4.3.1.1 Visceral level

The visceral level is the most instinctive level that all humans alike share a similar copy of upon birth from our biological inheritance. Each person’s visceral level is not identical to everyone else’s, but the majority of people have a very similar starting point. A common feature is that the visceral level is highly responsive and provides a sort of first line of defense. It makes quick decisions whether something is good
or bad and communicates the result to the brain and the muscles so that action can be taken. If something is interpreted as bad, or even dangerous, our muscles tighten and prepare for physical action, such as fleeing from a dangerous predator. If something on the other hand is good, we tend to relax our muscles and brain since there is no need for alarm.

Lizards are a prime example of a reptile that acts only on the visceral level, meaning that lizards operate in a just-in-time manner. Whatever the lizard sees, the lizards instantly reacts to without taking past experiences into account or reflecting on the various choices available. The information available to the lizard is solely sensory information and the lizard can not learn from previous mistakes.

Humans has the enormous advantage over other animals that we are capable of using our reflective mind to overcome our biological heritage and build upon the visceral level. The journey of shaping our visceral level is a personal one that consists of learning through not only reflective thoughts but also through experiences that form our behavioral level.

In the context of website design, the visceral level would judge a website’s immediate appearance. The first few moments on the website would be decisive for the visceral response and if the visitor gets a bad feeling it is very likely that the visitor will leave without any interaction and, in other words, bounce.

4.3.1.2 Behavioral level

Our emotions towards things are highly affected by our past experiences. Our pre-programmed visceral response will always be important as it acts as the first line of defense and helps us to stay safe, but many situations are more complex than our most instinctive reaction. More importantly, as we learn more about something, our visceral response becomes less significant as the behavioral level gains president. For example, we are predisposed to be cautious around animals such as snakes and spiders because they are associated as potentially poisonous and lethal creatures. This caution can be reduced through enough interaction and experience, which is why it is not too uncommon that snakes and spiders are found as house pets, both harmless and poisonous ones. Interactions and experiences can on the other hand have the exact opposite effect and sometimes it does not take more than one bad experience with a spider for a person to develop arachnophobia, the fear of spiders. James Russel [14] talks about experiences being strongly associated with emotion. Russel explains how emotional meta-experiences, the resulting emotion associated with an experience, labels experiences as a particular emotion, for example fear or happiness. The idea is that all the combined emotional responses of an experience generates a coherent, singular emotion associated to that experience. For example,
winning a large sum of money from a poker game with your closest friends causes happiness from winning, but also anxiety and a sense of guilt for taking your friends’ money. The emotional meta-experience could then possibly be summed up as guilt and anxiety, even though you won the game and was happy for doing so. The experience from the poker game would be integrated into your general knowledge and the emotional association with poker games would be affected negatively as it would be covered with a feeling of guilt. It is worth remembering that not everyone’s visceral levels look exactly the same from the start, some people are inherently more afraid of spiders and some people have less empathy for fellow poker players.

One of the most important qualities of the behavioral level is adaptability. One can for example overcome a mild fear of height, spiders and other common fears through training. However, there are of course people with initially strong visceral responses to certain things, and will thus inherently have a much harder time overcoming certain fears. Our opinion of all things, including website design, is highly subjective and one user’s preference is another’s rejection. The opinion is also affected by the user’s emotions at that current time, which makes it all the more difficult. Simply put, a designer can not satisfy all users’ needs at all times. A designer can however make the best of an impossible situation and follow Marc Hassenzahl’s [4] words "Although a particular experience cannot be guaranteed, it can be made more likely by applying some of the already available knowledge (...) to design and by further, design-oriented research".

4.3.1.3 Reflective level

After the visceral and behavioral level is the reflective level, the highest level of them all. The reflective level does not control behavior nor have any direct access to the senses like the visceral level has. It instead relies on thoughts and attempts to affect the behavioral level. The reflective level is highly adaptable and differ greatly from person to person depending on a number of factors. The differences in how we reflect upon things can be due to cultural views, education, certain training or simply life experiences. All of these things combined is what makes us think differently about things and also makes us easily affected. A trivial example is why more expensive things are for the most part more desirable than cheaper, even though the quality is indistinguishable. The meaning of prestige and status is vague in a practical sense, as it is only a currency in our minds. A currency whose value has been set by the culture that we live in, a culture that constantly evolve and differ from one another. What is highly valued in one culture may be worthless in another.

The reflective level also casts, as the name suggests, reflection before, during and after usage of for example products. There is always an experience beyond
the first encounter with a product, a sort of remembrance that include the feel of the product and the values that we attached to it. The product can be anything from a physical object to an artistic expression such as a painting. Regardless of the product’s nature, the impressions gathered from interaction paired with the cultural values of each individual, we form an opinion that is likely to shape to probability of us reusing that particular product.

4.3.1.4 Processes

Similar to the chapter on perceiving the world, the affective levels also have bottom-up and top-down processes. The bottom-up process refers to an activity initiated at the visceral level and finishing at the reflective level. Top-down is, as the name suggests, in the reverse order. Also very similar is how the bottom-up process is driven by perception, and the top-down by reflective thought. Both processes rely on neurotransmitters whose primary functions is to control how neurons transmit neural impulses between nerve cells, either by enhancing or inhibiting them. The bottom-up process operates through our senses and what we can perceive from our environment and is initiated by the visceral level that passes quick and decisive judgement, instantly releasing neurotransmitters befitting of the response. From the top-down process the initiation is instead at the reflective level through thought, and will eventually travel down to the lower levels, at which the befitting neurotransmitters will be released in the same fashion as with the bottom-up process.

4.4 Theoretical conclusion

As seen in chapter 4.2 and 4.3, there are lots of similarities between the two different theoretical areas that can be used to help developers and designers in their work. The most important similarity being the different processes described by each author, pointed out in chapter 4.3.1.4. Ware helps us understand the processes of perception and human visual functioning and Norman helps us understand the processes of emotion and affection. Both authors describe similar flows from low to high level and vice versa and together provide a solid foundation for identifying two factors that will help create the hypotheses to reduce bounce rate.

- First factor: Vendor credibility
  - Emotional motivation
  - Recognizability

From an emotional standpoint it is important to appeal to the visitors’ emotions, they should feel safe and confident while using the service or else they will probably
not purchase anything. Safety is a visceral feeling and is decided immediately upon landing on the marketplace. The first impression should thus be friendly and should evoke trust for the marketplace. The marketplace should also design content and functionality in a recognizable manner in order to appeal to the visitors’ behavioral level, a level that is conditioned by the standards of modern web design. Deviating too far from modern web design standards can confuse the visitor which is likely to cause a loss of confidence for the marketplace. The reflective level is the most difficult to address as it is highly individual. The visitors’ impressions from the interactions with the marketplace will impact the probability that they revisit the marketplace. The primary interaction, in terms of conversion, is purchasing an item and since all purchases are initialized with a conversation between a user and a vendor, vendor credibility is the first of the two major factors. Vendor credibility can be achieved by improving the presentation of the vendor and by highlighting the vendor more confidently. Ideally, an increase in vendor credibility would result in an increase of users clicking the purchase button, thus decreasing bounce rate.

- Second factor: Compact content
  - Product presentation (design)
  - Association (related products)

Chapter 4.2 demonstrated the limitations of human visual functioning and perceptive abilities. Perception is to a large extent the result of detection, and detection can only be achieved if there is something to detect. In other words, content that is not immediately available for detection can not be perceived. No matter if the visitor’s perception is driven by the bottom-up process or the top-down process, interactive elements such as the purchase button and related ads should be detectable from the moment the visitor lands on an ad. As described in chapter 4.2.2, our visual world, or the marketplace in this case, is built on a need-to-know basis. Considering the nature of marketplaces, the option to purchase an item is highly relevant for the business and its visitors. This is why the second factor is compact and informative content. The idea of informative content is that the visitor should be met by a broad variety of interactive options right away, without feeling overwhelmed by them. The interactive elements’ locations and sizes should be optimized, both vertically and horizontally, and, if need be, moved or hidden to give room for more important elements.
5 Part 2 - Empirical investigation

5.1 Research questions

How is bounce rate affected by

- emotional factors such as vendor credibility, emotional motivation and recognizability and

- perceptual factors such as compact content, product presentation and associative content?

5.2 Limitations

The thesis will perform a single experiment which will investigate all four combinations of the two primary factors, vendor credibility and compactness. The experiment is limited to Plick’s peer-to-peer marketplace platform and will be active for 10 full days.

Only visitors accessing the marketplace through a mobile device are included in the experiment. This limitation is due to the complexity of responsive design and the time required to make design manipulations that contain the same functionality as the original design. More experimental designs are of higher priority than a single perfectly responsive one. Visitors that access the website through a non mobile device are unaffected of the changes and receive the original version.

5.3 Introduction to testing

5.3.1 The marketplace’s current state of design

The majority of the marketplace’s visitors access the marketplace from a mobile device. The most common operating system, as seen in figure 7, is iOS, which, together with Android, makes up for more than 62.5% of all visitors. The most common resolution is that of an iPhone 6, which has a screen size of 4.7 inches in diameter and a resolution of 375x667 pixels. After inspection, it also appeared that the iPhone 6 is a somewhat representable mean of all different mobile sizes and will thus be used when developing the design manipulations. The current state of the marketplace’s design, henceforth referred to as the original design, is shown in figure 6, where the visitor’s immediate landing view is marked by the horizontal purple lines. The figure is captured from an iPhone 6 and displays the most common landing page on the marketplace, an individual ad in its entirety. The ad is quite
long, but remember that visitors can only see a portion of it at a time, how much depends on the screen size.

Figure 6: Original design of a single ad as of mid April 2017. The ad is divided into two columns for readability. The purple lines indicate the immediate view presented to visitors that land on the ad.

Figure 7 and 8 provides sufficiently accurate data for us to confidently say that an iPhone 6 is the most representative device on the marketplace and thus the one to
5.3.2 Time spent per page

Looking at the statistics from Google Analytics, it is clear that as the average time spent per page decreases, the bounce rate increases. This indicates that when visitors can not find anything of interest or anything that is easily available for interaction, they bounce. An increase in average time spent per page would thus imply a decreased bounce rate. According to Google Analytics’s engagement statistics for mobile traffic during the same period as figure 10, \( \frac{5139}{7296} \approx 70.4\% \) of sessions are less
than 10 seconds long and $\frac{4765}{7296} \approx 65.3\%$ of sessions only visit one page. Calculating the mean bounce rate from figure 10 gives us the same percentage. These percentages are also true for longer periods of time and can thus be used as comparison with the experimental results.

5.3.3 Perceptive emotion

With the statistics from figure 10 in mind, it can be concluded that the new design manipulations for the experiment should aim to catch the visitor’s eye. Visitors that spend little to no time on the marketplace are very likely to bounce. The impulse of leaving can be the result of many factors. If not for design it could be content, the visitor’s current mood, a sudden thought regarding something completely irrelevant and much more. However, it has to be assumed that the visitors enter the
marketplace with an intention of purchasing the item displayed in the ad, and if
unsatisfied would not deny a second try, perhaps through clicking on a related ad
or committing a custom search. The design must speak to the visceral level so that
users overcome the first few seconds and start browsing the marketplace’s content.

Due to the physical limitations to the parafovea mentioned in chapter 4.2.3, the
marketplace’s intent has to be obvious, there should be no question on how an item
from an ad can be purchased and there should be multiple objects to interact with
right away as well as incentives to scroll further. Creating incentive for visitors to
scroll is tough according to Nielsen whose research from 2006[8] revealed that 23% of

---

Figure 9: *The immediate view when landing on an ad.*

Figure 10: *An image of bounce rate vs average time spent on page from Google Analytics. The figure displays data gathered during March.*
visitors scroll on their first visit and merely 16% the second time. This means that a staggering 77% of visitors do not scroll what so ever during their first visit, which underlines the importance of the content above the fold and the incentives created to motivate scrolling. The expression above the fold refers to the first view the visitor is greeted by when landing on an ad, before any sort of interaction. The design manipulations has to somehow encourage and, almost more importantly, reward scrolling. If a visitor is provided an incentive to scroll but is met by disappointment, that visitor is unlikely to scroll again on the same website. It is thus also important to provide the visitor with a rewarding experience, and an easily accessible one at that.

When it comes to websites, assuming there is no audio, we can not perceive what we can not detect. If we are unable to perceive something, interaction with said thing will be impossible. Furthermore, any sort of emotional response, at any of the three previously discussed levels, will also be impossible. The most important design improvement for a marketplace is thus to create an environment in which the visitor is able to perceive as much content as possible, without causing a cluttered design. In a marketplace, visitors are most likely to have one of two cognitive tasks. Either they are looking to purchase an item or they are browsing ads, either for fun or for exploring. No matter the cognitive task, the ECT’s required are practically identical except for one or two exceptions. For example, ECT’s such as finding an ad, navigating to that ad, reviewing that ad’s images and description, sharing the ad, deciding whether the ad is interesting or not and finally either contacting the vendor if interested in purchasing or look for other content. The final tasks can be achieved either through clicking on related ads, committing a manual search on the marketplace or leaving the marketplace all together to use a search engine. The first two alternatives will not generate a bounce as they are interactions with the marketplace, the last one however is a bounce, assuming that the ad in question was the landing page. Note that the ECT’s finding and navigating to an ad is not covered in this thesis, but are nonetheless ECT’s required for visitors’ cognitive tasks.

5.3.4 Design manipulations

The design manipulations will be the foundation of the experiment. As discussed in 5.3.3, the content above the fold must be optimized and the incentive to scroll must be improved so that visitors reach the related ads. Keeping the fold in mind, the most important aspect of the design manipulations is thus compactness. Providing visitors with more content above the fold will not only lower the barrier for a purchase but will also create an incentive to scroll further or interact with the content available,
content that was previously hidden below the fold.

The second most important aspect, an aspect highly suited for marketplaces, of the design manipulations for this experiment is vendor credibility. Considering the fact that it is a peer-to-peer marketplace that requires interaction through chat prior purchase, the vendors’ credibility is of utmost importance as it may be an unconscious obstacle for some visitors. Unconscious obstacle refers to visitors not placing their trust in the vendor due to the presentation of the vendor, in other words, the design that visualizes the ad’s vendor.

These two aspects can be combined in various ways and can provide a total of four possible design manipulations as shown in table 2.

Table 2: Possible design manipulations

<table>
<thead>
<tr>
<th>Vendor credibility</th>
<th>Compactness</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>1.2</td>
</tr>
<tr>
<td>-</td>
<td>2.2</td>
</tr>
</tbody>
</table>

This will give a total of five different designs running in parallel, since the original design will be active as well. The design manipulations in table 2 are referred to as the numbers 1.1, 1.2, 2.1 and 2.2. As the table suggests, each number has different focus on the aspects mentioned previous in this chapter. For example, design manipulation 1.1 has low level compactness and a low focus on vendor credibility whereas design manipulation 1.2 as a high level compactness and still a low focus on vendor credibility. The design manipulations beginning with a 2 thus has a high focus on vendor credibility.

Lastly, a final design trick was implemented on all design manipulations. The instructions on how to create and sell ads was moved from each individual ad to the drop down menu. It might not seem like a big difference, but each browser has some sort of scroll progress indicator, and seeing that progress indicator move in a slow pace can be demoralizing for the visitor and cause a bad feeling.

5.3.4.1 Design manipulation 1.1

- **Compactness** -
- **Vendor credibility** -

Figure 11 shows the design manipulation 1.1 and figure 12 shows the immediate view upon landing on an ad with that particular design manipulation. Design manipulation 1.1 can be considered as the common starting point for all design manipulations and has a few specific differences from the original design. Firstly, as mentioned in
previous chapter, the instructions have been relocated from the ad to a separate page that can be reached through the drop down menu. Secondly, all superfluous paddings and margins have been removed and vertical content have, as much as possible, been transformed into horizontal content. For example, the interactive buttons for sharing through mail, Facebook or sending direct messages are placed to the left of the user profile rather than above and the ad’s extra images are placed to the right of the currently selected image rather than below. Furthermore, the purchase button’s text has been simplified from "LOGGA IN FÖR ATT KÖPA" (Login to purchase) to simply "KÖP" (Purchase) in order to encourage visitors to click it. The button will produce the same result, either a pop up asking the visitor to login if the visitor is not already or a chat box for messaging the vendor. The immediate view shown in figure 12 includes elements such as the vendor’s the purchase button, profile picture, meta information such as category, size, brand and likes, all of which are elements that were previously below the fold as seen in figure 9.

5.3.4.2 Design manipulation 1.2

- Compactness +
- Vendor credibility -

Figure 13 shows the design manipulation 1.2 and figure 14 shows the immediate view upon landing on an ad with that particular design manipulation. Design manipulation 1.2 was based on design manipulations 1.1 and has a high focus on compactness. The design manipulation 1.2 is exactly 400 pixels shorter in height which is equivalent to $\frac{400}{667} \approx 60\%$ of the height of an iPhone 6. The compactness is gained from further margin and padding reduction as well as removing the interactive buttons for sharing through mail, Facebook and sending direct messages. The direct message can still be sent as usual through pressing the purchase button marked "KÖP". The vendor’s profile picture has been moved next to the ad’s informative section and is thus horizontal content rather than vertical. The most important change however is the placement of the purchase button which resides just below the ad’s images, strongly encouraging visitors to press it.

5.3.4.3 Design manipulation 2.1

- Compactness -
- Vendor credibility +

Figure 15 shows the design manipulation 2.1 and figure 16 shows the immediate view upon landing on an ad with that particular design manipulation. Design manipulation 2.1 is not as concerned with compactness but instead highlights the vendor
Figure 11: Design manipulation 1.1. The purple lines indicate the immediate landing view.
5 PART 2 - EMPIRICAL INVESTIGATION

by placing the vendor’s profile image and associated information right below the ad’s images. The purchase button’s text is also extended with a personal touch by adding the vendor’s name and the ad’s price. This way it is abundantly clear from whom and for how much the item in the ad can be purchased, even more so when the vendor’s profile picture is closer to both the ad’s images and the purchase button itself.

5.3.4.4 Design manipulation 2.2

- Compactness +
- Vendor credibility +

Figure 17 shows the design manipulation 2.2 and figure 18 shows the immediate view upon landing on an ad with that particular design manipulation. Design manipulation 2.2 is the final design manipulation and has equally high focus on compactness and the vendor. The price has been placed in the bottom left corner inside the currently selected ad image and the purchase button is just below the ad’s images containing the relevant information regarding the vendor. Relevant information in this case is name and location. The vendor’s profile picture shares, similarly with
design manipulation 1.2, horizontal space with the ad’s informative section in order to reduce the overall height. This design manipulation is merely 2316 pixels in height, and looking at figure 18 reveals not only the vendor, the purchase button and all necessary information but also the upper half of the first row of related ads.
Figure 14: *The immediate view upon landing on and ad with design manipulation 1.2.*
Figure 15: *Design manipulation 2.1. The purple lines indicate the immediate landing view.*
Figure 16: *The immediate view upon landing on and ad with design manipulation 2.1.*
Figure 17: Design manipulation 2.2. The purple lines indicate the immediate landing view.
5 PART 2 - EMPIRICAL INVESTIGATION

Figure 18: The immediate view upon landing on and ad with design manipulation 2.2.
5.4 Method

This section will describe the method which consist of the necessary preparations and procedures to recreate the experiment step by step.

5.4.1 Data collection

Data collection is an important part of the thesis’s experiment. If the data collection is inaccurate and unspecific, the result is likely to be as well. There are lots of ways one can collect data from web traffic as discussed in chapter 4.1, but few alternatives are free and even fewer can provide data that is both detailed and in depth. GA, as described in chapter 4.1.4, allows for data gathering and data visualization. GA is equipped with the necessary tools to perform design experiments with parallel design manipulations but it requires that the design manipulations are kept on different URLs. Four new design manipulations would thus require four new URLs for each and every page, URLs that the marketplace’s visitors would use to share, for example, ads. Sharing an ad that is affected by a design manipulation, and thus has a different URL, would create a link to a page that only exists during the experiment. For example, a hypothetical ad with the original design would reside on the URL https://plick.se/ad_1, whereas the same ad with a design manipulation would reside on the URL https://plick.se/A/ad_1. If a visitor share the second URL on for example a social media, that URL will lead to an error page once the experiment is over, unless counter measures are implemented. Counter measures however, such as redirects, are not a viable long term plan if one plans to execute more experiments in the future. The URLs could also cause issues with search engine indexing unless one explicitly tells the search engine crawlers to ignore the experimental pages.

This is why GA by itself is not sufficiently advanced to perform an experiment with multiple design manipulations running concurrently whilst gathering data from each design manipulation independently.

5.4.1.1 GTM configuration

It is important to configure GTM correctly so that the data sent to GA is correct. The configuration, inspired by a web article [19], can be seen in list 5.4.1.1.

1. Create a GA and GTM account
2. Create a Container, preferably named after your domain.
3. Enable Random Number
PART 2 - EMPIRICAL INVESTIGATION

4. Create Trigger

- select Triggers from the side menu
- click on NEW
- enter a name for the trigger, the experiment’s name is **AB - Custom**
- click on Trigger Configuration
- set Trigger Type to Custom Event and enter a name, the experiment’s name is **abTest**
- set This trigger fires on to All Custom Events

5. Inject code on the website by creating a Custom HTML Tag

- select Tags from the side menu
- click on NEW
- enter a name for the tag, the experiment’s name is **AB - Custom HTML tag**
- click on Tag Configuration, select Custom HTML
- enter custom HTML code, do not forget surrounding <script> tags if it is JavaScript or <link> tag if it is CSS
- click on Triggering and select an appropriate trigger, the experiment’s trigger is All Pages

6. Create Data Layer Variable to extract information from the dataLayer

- select Variables from the side menu
- click on NEW in User-Defined Variables
- enter a name for the variable, the experiment’s name is **AB - dlv - eventAction**
- click on Variable Configuration
- select Data Layer Variable
- enter a Data Layer Variable Name that matches your dataLayer arguments provided through the Custom HTML tag just above, the experiment’s name is **eventAction**
7. Send data to GA by creating an Event Tag

- select Tags from the side menu
- click on NEW
- enter a name for the tag, the experiment’s name is AB - Google Analytics - Event
- click on Tag Configuration, select Universal Analytics
- select Event from Track Type
- select True from Non-Interaction Hit
- enter the Tracking ID from your GA account
- enter a Category, the experiment’s Category is named AB Test
- enter an Action, the experiment’s Action is named {{AB - dlv - even-tAction}} which is our newly created Data Layer Variable
- click on Triggering and select an appropriate trigger, the experiment’s trigger is AB - Custom

Figure 19: A visualization of the GTM configuration.

In short, the Custom HTML Tag AB - Custom HTML Tag is triggered on any pageview and loads the JavaScript code. The JavaScript code sends an
event through the dataLayer which is picked up by the Event Tag AB - Google Analytics - Event which has the trigger AB - Custom which triggers on all events. The Event Tag’s action is the Data Layer Variable AB - dlv - eventAction which allows GTM to receive data from the dataLayer as long as the push method, explained in chapter 5.4.1.2, includes the property eventAction. A visualization of GTM’s structure can be seen in figure 19.

5.4.1.2 The Custom HTML Tag

The Custom HTML Tag, in the experiment named AB - Custom HTML tag, is the key component in the GTM configuration. The tag contains JavaScript code that essentially performs the most critical parts of the experiment - identifying visitors’ device, dividing visitors between the design manipulations, making sure visitors get the same design each visit, executing JavaScript changes and sending activity data back to GA marked with which design manipulation the visitor visited. All this is achieved by less than 300 lines of JavaScript code.

- Determine if the visitor is accessing the marketplace through a mobile device by analyzing the screen’s dimensions. If the screen’s width is less than 767 pixels, the device is considered a mobile device and will participate in the experiment. There are other ways to determine the visitor’s device such as reading the browser information, but that method is not completely reliable and some visitors might not be accounted for.

- Which design manipulation that is assigned to the visitor is determined by the Random Number Variable previously enabled in chapter 5.4.1.1. The Random Number can be generated and accessed in our JavaScript code by assigning {{Random Number}} to a JavaScript-variable, which will then be different each time the script runs.

- When running experiments it is important that user experience is not affected negatively. A typical issue is design inconsistency, so when visitors move between pages on the website they should see the same design manipulation, or the original design, wherever they go. To keep track of which design manipulation each visitor is assigned to, the design manipulation’s number, explained in chapter 5.3.4, is saved in a cookie named PlickAB-cookie. The first thing to evaluate when a visitor accesses the site is whether or not that visitor is new or returning. This can easily be achieved by attempting to read the cookie named PlickAB-cookie. If there is no cookie by that name the visitor is determined as new and will be assigned a design manipulation based on the Random Number.
Once a design manipulate has been assigned to the visitor, a block of JavaScript code associated to that particular design manipulation will run. That JavaScript code will implement dynamic changes to the design by for example moving elements and changing certain texts. This part will be covered in more detail in chapter 5.5.0.1.

The final piece of code that is executed in the Custom HTML Tag is the dataLayer. The dataLayer sends activity data to GTM by using a method named push that takes an object as argument. In order for GTM to make any sense of the object, a few properties are absolutely necessary. Firstly, an event must be specified. This is the event that will eventually turn up in GA and is in this experiment named abTest. Secondly, a connection with our Data Layer Variable must be established so that GTM knows what to do with the data. The connection is established through the property eventAction that correlates to the Data Layer Variable Name previously defined in our Data Layer Variable from chapter 5.4.1.1 named AB - div - eventAction. The eventLabel property is superfluous in this particular experiment but could potentially be used to send additional data, its value is the same as the eventAction. As seen in the code below, eventAction is responsible for sending the visitor’s active design manipulation. Finally, the eventNonInteraction is set to true, so that the event is not interpreted as an interaction from the visitor but rather a data collection event.

dataLayer.push({ 'event':'abTest', 'eventCategory': 'ab-test', 'eventAction': prefix + ' - ' + changeID + ' - ' + variantID, 'eventLabel': prefix + ' - ' + changeID + ' - ' + variantID, 'eventNonInteraction': true });

5.5 Developing the design manipulations

Developing the design manipulations required skills in HyperText Markup Language (HTML), Cascading Style Sheets (CSS) and JavaScript. HTML describes websites and defines their content. It is known as the most fundamental building block of the modern web. CSS is a language that describe the appearance and presentation of a website written in HTML. JavaScript is a dynamic prototype-based language executed in the client’s browser.

Specific to Plick, knowledge of Git, a version control system that keeps track of file changes and is crucial for teams, and LESS, a dynamic style sheet language that can be compiled into CSS, was also used. The experiment could have been performed without both Git and LESS, but using the two techniques speeds up the development process significantly.
Worth mentioning is that GTM had a preview-mode that helps with local debugging. The preview-mode could be enabled in the header menu at their website.

5.5.0.1 HTML and JavaScript

The structured content on the website is outputted as HTML and can be moved, added, removed or edited in one of two ways. Either change the core code on the backend, which would make interaction with the backend an absolute requirement, or use JavaScript to manipulate the HTML on the frontend. The latter of the two does not require any interaction with the backend and can be achieved by extending our already implemented Custom HTML Tag’s JavaScript code, explained in chapter 5.4.1.2.

JavaScript is a powerful programming language and allows for creation and manipulation of elements on the web. By using JavaScript, the developer is freed from interaction with the backend and can thus apply changes much more rapidly and in a much more contained environment. If the code breaks in any way it is easily updated through GTM.

5.5.0.2 CSS and LESS

The experiment’s CSS was written in LESS and then compiled to CSS. Due to Plick’s development environment, writing plain CSS and including it in the Custom HTML Tag was a hassle during the development of the design manipulations as a lot of changes were made continuously, and it would have required a GTM container version bump each change. Instead, Plick’s local development environment was used and LESS became an option. This did however entail that the server had to be involved in the end, but note that it is not a requirement what so ever, it was merely convenient during the thesis’s experiment. Under different circumstances, in a different development environment or with a different time frame, using plain CSS inside the Custom HTML Tag would have made the entire experiment more containable and potential design errors easier, and especially faster, to debug and fix.

CSS uses class- and id-selectors to target and style the appearance of elements on a website. During the third critical step in the Custom HTML Tag, described in chapter 5.4.1.2, the JavaScript code also set the website’s body-id to the design manipulation’s number. The CSS for each design manipulation had selectors for corresponding body-id’s to ensure that the correct CSS was applied to the correct design manipulation. It was extra important that the CSS’s selector and the body-id were matching correctly because if not, the JavaScript code might have applied changes that were meant to fit together with different CSS code which would have
caused disorder.

5.5.1 Activating the experiment

The experiment was at this time only available locally in a development environment together with GTM’s preview-mode. In order to make the experiment public the new version from GTM, including all changes and added Tags, Triggers and Variables, had to be submitted. Due to the usage of LESS, the server also had to be updated. For Plick, this meant that the LESS code had to be pushed to a Git repository and then pushed to the Flynn’s remote repository. Flynn is a platform as a Service (PaaS) that provides an open source platform for running production ready applications. Flynn makes scaling servers, or rather instances as they are called, easy and comes with routing and traffic load balancing out of the box. Once the push to Flynn had gone through, the new CSS code was applied whenever the body-id fitted one of the selectors.

The experiment could then be considered live and began assigning design manipulations to visitors and collecting and sending data to GA.

5.5.2 Interpreting data from GA

The data sent to GA from GTM’s dataLayer was sent as events and was interpreted as such in GA. There was no prepared view in GA for analyzing the Event-data during a time period, even less so correlating it to bounce rate. The only solution was to create a custom report. The thesis’s experiment’s custom report can be replicated by following a few simple steps in GA, as seen in list 5.5.2.

1. Create new Custom Report

   • select CUSTOMISATION from the side menu
   • select Custom Reports
   • click on + New Custom Report
   • set an appropriate Title and Name, the experiment’s title and name are abTest Report and Report Tab
     • in Metric Groups, add Unique Events
     • in Dimension Drilldowns, add Event Action and Even Category
     • in Views, select the view that is related to the experiment

2. Setup the report to show essential data

   • click on + Add Segment just above the Report Tab
select All Users and Bounced Sessions, then Apply

- at the top of the grid below, click on Secondary dimension and select Event Category

- at the same level to the right, next to the search input, press advanced

- set first drop down menu to Include

- set second drop down menu to Event Category

- set the third drop down menu to Exactly matching

- enter the chosen category value from the Event Tag described in chapter 5.4.1.1 AB Test in the search bar to the right of the three drop down menus, in the experiment the category value is AB Test

- press apply

The Custom Report should resemble figure 20. Note that the figure’s statistics include Sessions and Avg. Session Duration, two metrics that are optional and not mandatory at all for the experiment.
Figure 20: *How the Custom Report should look like after proper setup.*
5.6 Results

The experiment was active during 10 days, from 2017-05-09 to 2017-05-18, and recorded a total of 5,892 sessions from users accessing the marketplace from mobile devices or devices of similar dimensions. The result can be considered impartial because the experiment was not announced in any way to the participants.

5.6.1 Data from GA

<table>
<thead>
<tr>
<th>Design manipulation number</th>
<th>0</th>
<th>1.1</th>
<th>1.2</th>
<th>2.1</th>
<th>2.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sessions</td>
<td>1229</td>
<td>1124</td>
<td>1187</td>
<td>1125</td>
<td>1227</td>
</tr>
<tr>
<td>Bounced Sessions</td>
<td>740</td>
<td>652</td>
<td>699</td>
<td>715</td>
<td>726</td>
</tr>
<tr>
<td>Non-bounced Sessions</td>
<td>489</td>
<td>472</td>
<td>488</td>
<td>410</td>
<td>501</td>
</tr>
<tr>
<td>Bounce Rate (BR)</td>
<td>60.212%</td>
<td>58.007%</td>
<td>58.888%</td>
<td>63.556%</td>
<td>59.169%</td>
</tr>
<tr>
<td>BR relative to original design</td>
<td>0%</td>
<td>-2.205%</td>
<td>-1.324%</td>
<td>3.344%</td>
<td>-1.043%</td>
</tr>
</tbody>
</table>

Table 3 reveals the experiment’s final results after being active for 10 consecutive days. The table looks at four different parameters, sessions, bounces sessions, bounce rate and bounce rate relative to original design. The sessions are, as previously described in chapter 4.1.4.2, aggregated interactions by a single visitor during a time span, and are primarily important when looking at bounced sessions. Bounced sessions are the measure of how many of the sessions that bounced, as described in chapter 4.1.4.2. The bounce rate in the table is derived from the bounced sessions divided by the sessions and the bounce rate relative to original design is each design manipulations’ difference difference in bounce rate compared to the original design’s bounce rate.

5.6.2 Chi-square test

In order to validate the result statistically, the chi-square test is used. The test determines the probability of an observed difference between distributions being the effect of pure chance or not[18].

<table>
<thead>
<tr>
<th>Table 4: Chi-square test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi square value</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Original design : Design manipulation 1.1</td>
</tr>
<tr>
<td>Original design : Design manipulation 1.2</td>
</tr>
<tr>
<td>Original design : Design manipulation 2.1</td>
</tr>
<tr>
<td>Original design : Design manipulation 2.2</td>
</tr>
</tbody>
</table>
Table 4 shows the result of the chi-square testing. A significant of 5% was chosen as a reasonable level to determine significance. Given this, the results indicate that none of the design manipulations had any statistically significant effect on bounce rate in comparison with the original design.
6 Conclusions

Creating impactful design is very difficult, each and every visitor is unique in their own way and will interpret designs differently. Designers can strive to design for the majority but can not design for everyone. The theory of emotional design and perception plays a great role in design experimentation. This thesis explores how knowledge of emotion and perception can be used as a foundation to create new experiences through design manipulations. Even though the particular design manipulations used in the experiment did not result in statistically significant results, the main ideas and factors used can still be relevant for future work. The differences in bounce rate between the different experimental groups were not large but could still prove, if they are indeed real effects, to be relevant for conversion rates. Therefore a longer experiment, leading to more users being exposed to the design manipulations, could be interesting to perform. Also, the design manipulations’ changes were likely too small to affect the bounce rate enough and might have yielded a clearer result had they been more radical. The theory and the method should thus not be judged too quickly from the results.

Web analysis through page tagging proved to be a vital tool for making this thesis’s experiment possible. With the thesis’s developed method, developers and designers can take part in parallel bounce rate analysis. The method is horizontally scalable, meaning that an arbitrary amount of design manipulations can be investigated in parallel. Considering that four design manipulations were built and deployed within the time frame of this thesis, a small business should have no difficulties using the method to perform at least the same amount of design manipulations on its own platform. The simplicity and cost efficiency of Google Analytics and Google Tag Manager outshines other options and are strongly encouraged. The Custom HTML Tag also allows novice developers to take part of experiments because it does not require expert knowledge in JavaScript or CSS.
7 Future Work

7.1 Correlation between bounce rate and conversion rate

The correlation between conversion rate and bounce rate on the web is a highly relevant subject to explore further. The two metrics feel instinctively correlated but lacks theoretical backing. By extending this thesis’s experiment with detailed conversion tracking, one could investigate the correlation between the two, something that would provide great insight for future design theory.

7.2 Include more devices and platforms

Due to the limitations in chapter 5.2, this thesis’s experiment only includes mobile devices or devices with similar dimensions. An obvious future addition would be to include desktop visitors as well, which can be achieved by adding desktop specific design to each design manipulation.

One could also explore the possibility of testing native applications and other platforms than marketplaces. There is no reason for why this thesis’s theory could not be applied on for example a blog or company website. The conversion would be relying on different metrics, but bounce rate and interaction is important on many platforms. The thesis’s method might not apply for native applications though because of the requirement of Google Analytics and Google Tag Manager.

7.3 More radical design manipulations

The probability of receiving clearer results in terms of bounce rate difference between design manipulations can be elevated by implementing more radical design in each design manipulation. Radical design manipulations in this case implies design alterations such as removing conventional elements, implementing special animations and so on. Such design manipulations would be prone to receive clearer if not unanimous results, but could also hurt the marketplace’s reputation if usability was challenged to a large extent. Basically, designers can not implement features or design ideas simply for the fun of it or to get a significant variation in their data, they have to think of the good of the platform as well.

7.4 Effect size and statistical significance

In order to possibly achieve a statistically significant result, either the amount of visitors would have to be increased or the design manipulations would have to be more radical, or both. The amount of visitors can be increased by either prolonging
the experiment’s period or conducting the experiment on a platform with more
visitors. Fewer design manipulations would also cause each design manipulation to
gain more individual visits.
REFERENCES

References


REFERENCES


# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Traffic origin distribution of the marketplace collected by Google Analytics during March 2017.</td>
</tr>
<tr>
<td>2</td>
<td>Server log storing process</td>
</tr>
<tr>
<td>3</td>
<td>Page tagging process</td>
</tr>
<tr>
<td>4</td>
<td>The frames a-d are captured from a video of the pedestrian experiment. Frames a-c shows the substitution trick in action and frame d shows the two actors that performed the substitution. The image is from &quot;Failure to detect changes to people during real-world interaction&quot; [16], courtesy of Daniel J. Simons.</td>
</tr>
<tr>
<td>5</td>
<td>Illustration of the visual field. Image from Visual Thinking for Design [21], courtesy of Colin Ware.</td>
</tr>
<tr>
<td>6</td>
<td>Original design of a single ad as of mid April 2017. The ad is divided into two columns for readability. The purple lines indicate the immediate view presented to visitors that land on the ad.</td>
</tr>
<tr>
<td>7</td>
<td>A list of popular devices from visitors during March gathered by Google Analytics.</td>
</tr>
<tr>
<td>8</td>
<td>A list of the most common screen dimensions from visitors during March gathered by Google Analytics.</td>
</tr>
<tr>
<td>9</td>
<td>The immediate view when landing on an ad.</td>
</tr>
<tr>
<td>10</td>
<td>An image of bounce rate vs average time spent on page from Google Analytics. The figure displays data gathered during March.</td>
</tr>
<tr>
<td>11</td>
<td>Design manipulation 1.1. The purple lines indicate the immediate landing view.</td>
</tr>
<tr>
<td>12</td>
<td>The immediate view upon landing on an ad with design manipulation 1.1.</td>
</tr>
<tr>
<td>13</td>
<td>Design manipulation 1.2. The purple lines indicate the immediate landing view.</td>
</tr>
<tr>
<td>14</td>
<td>The immediate view upon landing on an ad with design manipulation 1.2.</td>
</tr>
<tr>
<td>15</td>
<td>Design manipulation 2.1. The purple lines indicate the immediate landing view.</td>
</tr>
<tr>
<td>16</td>
<td>The immediate view upon landing on an ad with design manipulation 2.1.</td>
</tr>
<tr>
<td>17</td>
<td>Design manipulation 2.2. The purple lines indicate the immediate landing view.</td>
</tr>
<tr>
<td>18</td>
<td>The immediate view upon landing on an ad with design manipulation 2.2.</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>19</td>
<td>A visualization of the GTM configuration.</td>
</tr>
<tr>
<td>20</td>
<td>How the Custom Report should look like after proper setup.</td>
</tr>
</tbody>
</table>
## List of Tables

1. Server log formats .................................................. 7
2. Possible design manipulations .................................... 29
3. Results ................................................................. 48
4. Chi-square test ....................................................... 48