Master Thesis

Building a web-based experiment to capture and analyze cultural attraction

Oliveira Moreira Renan

2019

Eötvös Loránd University

Faculty of Education and Psychology

Master of Science in Computational and Cognitive Neuroscience

Building a web-based experiment to capture and analyze cultural attraction

Supervisor

Oliveira Moreira Renan Dr. Christophe Heintz

Cognitive Science at CEU

2019

**DECLARATION OF ORIGINALITY**

I, the undersigned Oliveira Moreira Renan, a student in the ELTE PPK Computational and Cognitive Neuroscience MSc in awareness of my criminal liability declare and confirm with my signature that the thesis titled Building a web-based experiment to capture and analyze cultural attraction is **my own independent intellectual work** and the use of the referenced, printed and electronic, literature occurred according to the general rules of copyright laws.

I understand that the case of a thesis plagiarism is:

* using a literal quote without quotation marks and without reference;
* using citation of content without referencing to it;
* using another person’s published thoughts as my own.

I, the undersigned state that I know the concept of plagiarism and I understand that in case of plagiarism my thesis will be rejected and in this case disciplinary action may be brought against me.

Budapest, 2019 ............................................

........................................................................

**CONFIRMATION OF THESIS CONSULTATION**

(a part of the thesis after filled out.)

Student’s name:

Student’s programme:

The thesis’s topic:

Supervisor’s name:

|  |  |  |
| --- | --- | --- |
| **Time of consultation** | **Topic of the consultation, notes, suggestions** | **Supervisor’s signature** |
| **19/02/2018**  **12:30** | First meeting, introduction to the theory of culture transmission, talk about the idea and possible project. |  |
| **27/04/2018**  **17:20** | Design of the project. |  |
| **09/08/2018**  **17:20** | First prototype of the website, refused, and planned a totally different way to collect the user input. |  |
| **18/10/2018**  **17:40** | Prototype of the website with the calibration scenario, talk about data storage, design of scenarios of factor of attraction. |  |
| **30/01/2019**  **14:00** | Presentation of the website ready with different kinds of scenarios (face, spiral, click here button). Discussion about improvement of them and data collection procedures. |  |
| **11/04/2019**  **17:45** | Final discussion about the data analysis, results and conclusions of the experiment. |  |

I allow the submission of the thesis.

Budapest, 2019.................................

........................................................................

supervisor’s signature

Acknowledgments

Living abroad is wonderful, the perks are numerous but at the same time it is challenging to deal with the loneness. Therefore, this is my opportunity to thank everyone that contributes to make my journey as special as it is. It will be overly personal and emotional, I’m aware it is erroneous, but forgive me.

First and foremost, I could not reach this point without my family, their support is immeasurable. I know it hurts being more than 10.000 km apart, but I really want to thank yours understanding when comes to my decision to live far away. My sister, you are the most reasonable and crafty person I will ever meet, my brother you are a sensitive and pally as much as our family needs, you two are the equilibrium of our household, I’m proud of being your brother. My father, the serene when matters, but the insane when comes to the aspects of enjoying life and racing. My mother, oh that is a complicated one, we are contrasting in so many aspects but at the same time I’m an exact replication of her, I want to thank you for loving me unconditionally. And to my grandfather and grandmother, thank you for teaching me what is love, always with a hug when I needed, yours place is the refuge and the paradise on earth, I will never be able to give back the amount of affection that I received, thank you so much.

I would like to express my very great appreciation to my supervisor Dr. Christophe Heintz for having the patience to guide and support me with great ideas. Deep gratitude towards Ildikó Király and every professor of the program. I would like also to extend my thanks to my classmates, Raissa, Dogu, Tiago, Salomé, these two years together were formidable because of your presence, special credit to Raissa for listening my customary whining. Szabina for being a great partner and friend. And Vini for being the longest friend, life drift us apart but we still brothers. Much obliged to everyone that was not mentioned but crossed paths in life with me.

Olivera Moreira Renan

Abstract

Cultural Attraction Theory has been topic of discussion in the cultural evolutionary field in the last 20 years, in this thesis I built a web platform with the intent to capture and document cultural attraction and enable other scientists to test their hypothesis. The experiment consists of a learn-and-copy game, where the user sees the position of a cross in the screen, and after an intermediary screen, it is asked to answer the position of the respective cross, on this context it is possible to insert stimuli in any of those screens which could work as an attractor.

Using this web platform, I implemented the experiment with three different stimuli (face shaped, spiral, button), collected the data and thereafter analyzed it with the aim of finding any indication of cultural attraction. Unfortunately, no relevant evidence of cultural attractors was found. This document contains a briefly introduction to the concepts of culture transmission and CAT.

Keywords: culture transmission, Cultural Attraction Theory, web platform, experiment, game, cross, position, attraction.

## Table of Contents

[Abstract I](#_Toc6261922)

[Table of Contents II](#_Toc6261923)

[List of figures, tables and equations IV](#_Toc6261924)

[Part I: Background 1](#_Toc6261925)

[Introduction 1](#_Toc6261926)

[Motivation of the research 2](#_Toc6261927)

[Research aim and objectives 2](#_Toc6261928)

[Research question 3](#_Toc6261929)

[Research approach 4](#_Toc6261930)

[Thesis structure 5](#_Toc6261931)

[Theoretical foundations 7](#_Toc6261932)

[Culture transmission 7](#_Toc6261933)

[Cultural Attraction Theory 10](#_Toc6261934)

[References 13](#_Toc6261935)

[Part II: Experiment / Empirical studies 15](#_Toc6261936)

[Design and Implementation 15](#_Toc6261937)

[Architecture 16](#_Toc6261938)

[Angular framework. 17](#_Toc6261939)

[Firebase hosting and database structure. 18](#_Toc6261940)

[Scenarios for testing 21](#_Toc6261941)

[Calibration. 22](#_Toc6261942)

[Face. 23](#_Toc6261943)

[Spiral. 24](#_Toc6261944)

[Button. 25](#_Toc6261945)

[Data collection 25](#_Toc6261946)

[Calibration phase 25](#_Toc6261947)

[Testing phase 26](#_Toc6261948)

[Attractor condition phase 27](#_Toc6261949)

[Data analysis 27](#_Toc6261950)

[Data extraction and cleaning 27](#_Toc6261951)

[Data calculations 30](#_Toc6261952)

[Results 34](#_Toc6261953)

[Transformation on the memory process. 35](#_Toc6261954)

[Documenting attraction. 42](#_Toc6261955)

[Part III: Conclusion 45](#_Toc6261956)

[Discussion and future work 45](#_Toc6261957)

[Research objectives: Summary of findings 45](#_Toc6261958)

[Limitations 45](#_Toc6261959)

[Future research 46](#_Toc6261960)

[References 48](#_Toc6261961)

[Appendix 49](#_Toc6261962)

List of figures, tables and equations

[Figure 1 - Thesis structure. 6](#_Toc6261793)

[Figure 2 - Data structure of Cloud Firestore 18](#_Toc6261794)

[Figure 3 - Data structure example from the experiment 19](#_Toc6261795)

[Figure 4 - Face shaped image used on the scenario. 23](#_Toc6261796)

[Figure 5 – Spiral used on the scenario 24](#_Toc6261797)

[Figure 6 - Click here button used on the scenarios. 25](#_Toc6261798)

[Figure 7 – Data containing noise. 29](#_Toc6261799)

[Figure 8 - Data after the process of cleaning. 30](#_Toc6261800)

[Figure 9 – Specific case of same image size with different screen heights. 32](#_Toc6261801)

[Figure 10 - Boxplot graph with the distance from the original position to the user position, threshold of 15% from the attractor. 36](#_Toc6261802)

[Figure 11 – Scatter graph with the distance from the original position to user position per interaction, threshold of 15% from the attractor 37](#_Toc6261803)

[Figure 12 - Scatter plot for the face scenario, using a threshold of 15%. 39](#_Toc6261804)

[Figure 13 - Scatter plot for the click here button on the top right scenario, using a threshold of 15%. 39](#_Toc6261805)

[Figure 14 - Distribution of the difference of DO and DU for the face scenario with 15% threshold. . 40](#_Toc6261806)

[Figure 15 - distribution of the difference of DO and DU for the click here top right scenario with 15% threshold. . 41](#_Toc6261807)

[Figure 16 - Distribution of the difference of DO and DU for all the scenarios of attraction with 15% threshold. 41](#_Toc6261808)

[Figure 17 - Enumeration of the quadrants in a matrix 14x14 that maps the screen coordinates. Each cell will contain the total amount of appearances. 43](#_Toc6261809)

[Figure 18 - Representation of screen/matrix for the calibration scenario. The highest value is highlighted in red. 44](#_Toc6261810)

[Figure 19 - Representation of screen/matrix for the button bottom left scenario. The highest value is highlighted in red. 44](#_Toc6261811)

[Figure 20 Appendix - Boxplot graph with the distance from the original position to the user position, threshold of 25% from the attractor. 49](#_Toc6261812)

[Figure 21 Appendix - Scatter graph with the distance from the original position to user position per interaction, threshold of 25% from the attractor. 49](#_Toc6261813)

[Figure 22 Appendix - Scatter plot for the click here button on the bottom left scenario, using a threshold of 15%. 50](#_Toc6261814)

[Figure 23 Appendix - Scatter plot for the spiral center scenario, using a threshold of 15%. 50](#_Toc6261815)

[Figure 24 Appendix - Scatter plot for the spiral left scenario, using a threshold of 15%. 51](#_Toc6261816)

[Figure 25 Appendix - Scatter plot for face scenario, using a threshold of 25%. 51](#_Toc6261817)

[Figure 26 Appendix - Scatter plot for click here top right scenario, using a threshold of 25%. 52](#_Toc6261818)

[Figure 27 Appendix - Scatter plot for click here bottom left scenario, using a threshold of 25%. 52](#_Toc6261819)

[Figure 28 Appendix - Scatter plot for spiral center scenario, using a threshold of 25%. 53](#_Toc6261820)

[Figure 29 Appendix - Scatter plot for spiral left scenario, using a threshold of 25%. 53](#_Toc6261821)

[Figure 30 Appendix - Distribution of the difference of DO and DU for the face scenario with 25% threshold. 54](#_Toc6261822)

[Figure 31 Appendix - Representation of screen/matrix for the button top right scenario. The highest value is highlighted in red. 54](#_Toc6261823)

[Figure 32 Appendix - Representation of screen/matrix for face scenario. The highest value is highlighted in red. 54](#_Toc6261824)

[Figure 33 Appendix - Representation of screen/matrix for the spiral top left scenario. The highest value is highlighted in red. 55](#_Toc6261825)

[Figure 34 Appendix - Representation of screen/matrix for the spiral center scenario. The highest value is highlighted in red. 55](#_Toc6261826)

[Table 1 - Detailed explanation of data variables 20](#_Toc5895158)

[Table 2 - Average distances and amount of closer/farther points 38](#_Toc5895159)

[Equation 1 – Distance between two points....................................................................... 30](#_Toc6257523)

[Equation 2 – Calculate attractor face................................................................................ 32](#_Toc6257524)

[Equation 3 – Calculate attractor spiral left..…. 33](#_Toc6257525)

[Equation 4 – Calculate attractor spiral center................................................................... 33](#_Toc6257526)

[Equation 5 – Calculate attractor button bottom left 33](#_Toc6257527)

[Equation 6 – Calculate attractor button top right 33](#_Toc6257528)

Part I: Background

# Introduction

The word ‘culture’ can draw multiple interpretations or even meanings for each of us, but some of the most accepted and well-marked definitions can be seen in the Merriam-Webster dictionary as: “the customary beliefs, social forms, and material traits of a racial, religious, or social group” or “the integrated pattern of human knowledge, belief, and behavior that depends upon the capacity for learning and transmitting knowledge to succeeding generations” (Merriam-Webster Dictionary, 2018). These two definitions together offer a great idea of what is culture and how it is vital in our life.

The second definition cannot be overlooked but still debated in the scientific community, more specifically when it refers to “...learning and transmitting knowledge to succeeding generations”. This part is the main topic of discussion of the theoretical approach in this thesis. Touching some topics as: How is human culture transmitted? What are the mechanisms that affect the process of transmission? Draw a parallel of culture transmission with genetic transmission, memetics, the cultural evolutionary theory, and more specifically one of those approaches within the cultural evolutionary theory called Cultural Attraction Theory. This latter approach to cultural evolution is based on the hypothesis that systematic transformations occurs during cultural transmission. Those systematic transformations that lead to the cultural success of a type of items (a story, a practice, an artefact, …) has been named ‘cultural attraction’.

With purpose of exploring cultural attraction, I created a web-based experiment which allowed multiple users around the world to perceive then copy the location of an item on a screen. This resulted on small amount of data that was analyzed with the goal to find patterns of transformation that could qualify as attraction. The whole creation of the experiment and the analysis of the results will be described in the empirical part of this thesis.

## Motivation of the research

My Information Technology background studies aligned with my working experience drove me to the construction of the website and the utilization of data science techniques to analyze data, together with my interests in social sciences, pushed me to develop the research on cultural evolution and more specifically Cultural Attraction Theory, which is currently discussed in the scientific community and has plenty of flourishing ideas and controversies. One of the controversies for example is considering memes as the unit of human cultural transmission, other example, regarding CAT, is the dilemma found on the article from (Claidière & Sperber, 2007) called “The role of attraction in cultural evolution” which defends a probabilistic view of cultural attraction that it was firstly introduced by (Sperber, 1996), against a deterministic interpretation of cultural attraction proposed by (Henrich & Boyd, 2002).

## Research aim and objectives

The aim of this research project is to create a web-based experiment that enables testing hypothesis related to CAT, with that purpose it was implemented several conditions of a learn-and-copy task, each condition included some possible features in the stimuli that could provoke attraction, i.e. ‘factors of attraction’. Besides the construction of the web-based experiment, it was collected data during 68 days with a total of 1296 data points, thereafter it will be performed data analysis testing the presence or existence of attractors on these scenarios.

The platform of the experiment was constructed with the aim to be available to other researchers and scientists interested on the topic. The code of the whole website structure is open source and available to be use by everyone, so the only required steps for the utilization would be the modification of the desired tested scenarios and deploying the website. Not only that, but the code of the data cleaning, calculation and analysis are also accessible. With this approach of adopting open source development, there is an aim of sparking a discussion about the topic and providing a platform where other researchers could test their own hypothesis and leverage on the website.

## Research question

The platform of the experiment and the data collected allows two possible lines of research, one of them is hypothesizing the position of the attractors and testing the validity of the assumptions, the second one is not assuming any attractor beforehand, tracking all the transformations and then documenting attraction.

Inside the line of research which assumes the position of the attractor, the focus of this research is to analyze the transformation of the mental representation (the ideational variant being the cross position) related to the short-term memory process of visualizing the cross position, storing the information and then retrieval of the correct position. On this approach, the relevant transformations are the ones connected with the instructions of the experiment, meaning answering the position of the cross the most accurate possible, and all the data which is an outlier will be considered noise and discarded. This is the main line of research of our experiment, and produces two questions:

* Can cultural attraction be produced with simple tasks? In particular, can we induce cultural attraction with a copying task based on visual memory?
* Based on our cognitive ability to identify and recognize faces (fusiform face area in human brain), does that generates cultural attractors in an image that resembles a human face?

Still inside the line of assuming the position of attractor, another approach would be considering all mistakes as part of the transmission process, this approach goes hand to hand with the cultural evolutionary framework in which the transformation are not perfect replications. It does not matter if the participant followed the main concept of answering correctly the position of the cross, all the transformations are valid, in the sense that the attractor could still play a role on influencing every possible output. This approach will not be studied in our experiment, but still valid to be mentioned.

The other line of research is not assuming the position of the attractors but accepting any transformation and then analyzing any patterns on the data which indicates attractors. For this approach it will be created a matrix that maps all the original position of the cross and the user feedback position of the cross and the research question would be:

* How can we observe and document attraction with a data set from a copying task?

## Research approach

(Verschuren & Doorewaard, 2010) describes five ways to conduct a research: a survey, an experiment, a case study, a grounded theory approach and desk research. While attraction is, according to CAT, a pervasive aspect in the formation of cultural phenomena, I have relied on the experimental approach for documenting attraction. The experimental approach has several advantages: first, it allows comparing different conditions for the same task, with different possible factors of attraction and therefore provide the means to test the validity of hypotheses about the existence of factors of attraction. Second, the experiment provides data that is promptly compliant with statistical analysis. This data provides a sandbox to test methods for documenting attraction and explore both lines of research explained before.

We decided to perform the experiment over the internet, in a way that would provide us a larger access to demographic and culturally different participants, this type of experiment is known as Internet-based experiment, Web experiment, online experiment, Web-based experiment, World Wide Web (WWW) experiment, and Internet experiment. (Reips, 2002). The structure of the Web-based experiment followed some of the standards and guidelines suggested on the paper of Reips.

This master thesis research can be separated in three stages of development: the first one was the exploratory literature studies about CAT and the planning and designing of the experiment. The second stage consisted on the coding and deployment of the website and the third and last one was the data collection, extraction, analysis and conclusions.

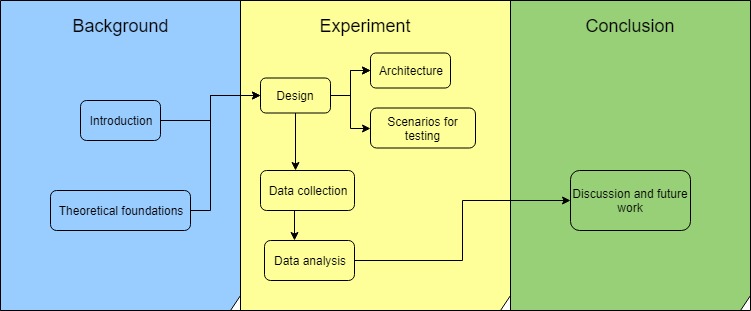
First stage: In-depth literature studies were fulfilled to enable the understanding of the research. Those studies together with discussions with Christophe Heintz made possible the elaboration of the design of the experiment. On this phase, it was also decided which programming language would fit the requirements and would be adopted in the project implementation, Angular 5, details on the decision will be explained on the section [Angular Framework](#_Angular_framework). It was also necessary technical studies regarding web programming, more specifically Typescript, programming language used by the Angular Framework. This phase on lasted approximately four months, from May of 2018 until August of 2018.

Second stage: The implementation, also known as coding of the website, used as the platform for the web-based experiments lasted for approximately six months from August of 2018 until January of 2019. This phase also contained the creation of the necessary connections to access the database (Cloud Firestore) and the deployment of the website on the Firebase platform, the details of this will be later explained on the section [Firebase hosting and database structure](#_Firebase_hosting_and).

Third stage: The last stage of the development of the research consisted in running the experiment, when the web-based experiment is online and accessible to the users, allowing the collection of the data. This collection period lasts for two months and 10 days, from 28th of January until 6th of April and subsequently, the data extraction from the database and analyses of the data are performed. In parallel with those tasks, this master thesis was being written with the purpose to provide the documentation of the research.

## Thesis structure

The structure of the thesis follows the instructions from the thesis manual from ELTE PPK Computational and Cognitive Neuroscience program, which recommends the separation of the structure into Theoretical and Empirical part as can be seen in the Figure 1:



*Figure 1* - Thesis structure.

The Part I, called Background provides a briefly introduction of the research developed, motivation, objectives and approaches utilized. Besides that, it also contains the Theoretical part which provides a base of understanding of the topic that will be researched and explored, eliciting definitions from different authors and researchers of the field and also the current discussions on the academic community.

The second part, called Experiment / Empirical Studies explains detailed all the phases of development of the experiment, the decision reasoning and the procedures utilized, from the design until the implementation and execution and lastly the data analysis.

The last part is the conclusion of the thesis, where it is discussed the results of the experiment, the conclusions reached, possible future works and limitations of the experiment.

# Theoretical foundations

## Culture transmission

Based on the definition of (Mesoudi & Whiten, 2008) and (Bisin & Verdier, 2008) culture transmission is the process in which information (knowledge, beliefs, preferences or norms of behavior) are passed from one individual to another one, within and across generations using social learning mechanisms as teaching, imitation or language.

It is possible to separate culture transmission in three different types: vertical, horizontal and oblique. Vertical transmission refers to the passage of traits from the parents to their offspring (top to down), and not necessarily has to be genetic related, for example it could be referent to transfer of knowledge from one parent to their adopted children. Other type is the horizontal transmission, which as the name suggests, in the hierarchy of generations, the transfer is from one side to another, this horizontally transmission consists of individuals on the same generation, as age peers, siblings, etc. The last type of transmission is called oblique transmission and covers the transmission from one individual of a specific generation to another of a younger one, excluding the parent-offspring relation, so for example, a teacher-student relation can be considered oblique transmission (Hershberger, 2014).

Culture transmission can be compared to genetic transmission, on the sense that both forms acquire traits or information from other individual (cultural traits/information versus genetic traits/information) and are intertwined in the process of evolution of our species, with the differences of the latter being more studied and with plenty of experiments conducted by population geneticists. While the former still has plenty of open discussions and far less experimental research (Mesoudi & Whiten, 2008).

Comparable to genetics, the approach to explain the cultural transmission principles and questions can also be viewed in an evolutionary approach, in the notion that culture itself evolves. Adopting this view, cultural transmission theories are called cultural evolutionary theories and follows the principles in accordance with the Darwinian ideas of *variation*, *differential fitness* and *inheritance*. According with (Mesoudi & Whiten, 2008) and (Mesoudi, Whiten, & Laland, 2004) those three principles can also be seen in culture transmission on the aspect that cultural traits vary across and within individuals and groups. Regarding differential fitness, some of the cultural traits will not be preserved and copied due to competition for expression, attention or memory space and some ideas are more attractive than others. And lastly, alike genetics, cultural traits are also inherited, but using the process of social learning.

There are evolutionary approaches to study culture evolution, one of them, known as Memetics, was originated from the term “memes” created by Richard Dawkins in the 70’s. It argues that memes are (cited in Sperber, 2000, pp. 163), “cultural replicators propagated through imitation, undergoing a process of selection, and standing to be selected not because they benefit their human carriers, but because they benefit themselves.” On this approach, the key word to be noticed is imitation, faithfully transmitted, meaning the information or the basic unit of culture is copied (replicated) from one person towards the other one. Thinking cultural transmission with the existence of “memes” allows scientists to use a Darwinian model of selection to explain the evolution of culture.

However, the explanation that Culture is composed by memes and the usage of memes on culture transmission have been contested in the scientific community for decades. Dan Sperber preaches that differently from genes, ideas are not transmitted perfectly identical from one brain to another. In his book (Sperber, Explaining Culture: A Naturalistic Approach, 1996) he explains why ideas are not replicated perfectly. Sperber characterizes mental representation as representation which exists inside its user. A mental representation can be a memory, a belief, an intention, etc. A public representation exists in the environment of the user. It can be a means of communications between individuals, such as a speech or a book.

Based on those two definitions, (Sperber, 1996) suggests that ideas are not transmitted faithfully between individuals, mostly because a mental representation is acquired by the act of observing others behavior or via linguistic communication. The mental representation in someone’s brain produces an output that can be observable by other individuals, named as public representation. This public representation is observed and captured by another person’s brain, which will need to infer the mental representation required to produce a similar public representation.

In other words, imagine a person X observes the behavior of person Y, and then X infers that Y have some belief Z and creates the “same” belief as Y. Using language, Y would tell X about Z, and X would start believing on Z also. In both cases, what matter is the transfer of the information Z. But as explained before the mental representation that X creates, may not be identical to the person’s Y mental representation. And a single public representation can also generate multiple mental representations in different individuals. On this whole process of perceiving, and forwarding information, there are systematic transformations on the mental representations.

This explanation explored by Sperber contrasts with the Memetics explanation, which defends that cultural transmission are based on the accurate replication of gene-like entities. Sperber thoughts are more connected with the view that ideas are “re-produced” (and not replicated), ideas are produced based on other ideas, and they share a deep connection, but they are not a perfect copy.

How this process of acquisition, storage and retrieval of information, also known as learning, is performed can be influenced by two types of cognitive mechanisms, inferential transformation and selective attention (Henrich, Boyd, & Richerson, 2008). Inferential transformation explains how mental representations are modified in ways to give an advantage to specific mental representations, what (Sperber, 1996) introduced the term “cultural attractors”. The other process that can influence is selective attention, which plays a role how individuals focus more on certain individuals to acquire their information, could be due to social status, a more prestigious person would attract more people to learn from their ideas.

Cultural Attraction Theory is based on the assumption that cultural phenomena result from attraction. Factors of attraction play a role in concentrating the population of cultural variants, thus making some cultural types more prevalent in the population of possible cultural variants. This happens when people produce cultural variants that are more often nearer to an attractor than further to it. Around these attractors, the imperfect replications would be if not identical, extremely close to it, this explain the existence of ideas which are stronger and spread to a population without literally copying. In (Henrich & Boyd, 2002), they provide a mathematical model and use the example of “Moon as person” and “moon as rock” as cognitive attractors, and that the distribution of the population will concentrate over that polarizing ideas.

Beside the Cultural Attraction Theory, which is the main focus of our study, (Lewens, 2018) explores four more lines of investigation which are the basis of research in Cultural Evolution: population thinking, cumulative culture, evolvability and cultural phylogenies. These four narratives, plus CAT can be associated in a way that they complement each other and provide an explanation for the cultural evolutionary framework without relying on the viewpoints of memetics.

Cultural transmission experiments are challenging mostly because culture is a phenomenon that occurs in large-scale and occurs through different generations. In the lab, it is impossible to simulate the same conditions, the groups are shorter and the time period smaller. (Miton & Charbonneau, 2018) describe some techniques used to emulate the features of real cultural populations and cumulative culture. They are: linear transmission chains (also called diffusion chains), replacement and closed-group.

In the linear chains, a first participant receives a stimulus and must remember it, the answer from this person will be the input for the next participant, this process is repeated multiple times and it is possible to see the transformation of answers. In the replacement method, the task consists of replacing the participants from the group and making them learn from the older members of the group. Briefly, the closed-group proposes that participants solve a task, and between the trials they can learn from each other. In our experiment we will not use any of those methods, explanation provided in the section [Design and Implementation of Part II: Experiment / Empirical studies](#_Design_and_Implementation).

## Cultural Attraction Theory

One of the models to explain the process of cultural evolution is Cultural Attraction Theory (CAT), which differs from other evolutionary approaches as it develops the idea of constructive convergence in culture transmission (Heintz, 2018). Constructive convergence refers to processes of cultural transmission that cause systematic transformation rather than faithful replication of cultural items across members of a population. However, these transformations are biased and favor the production of some cultural items, or ideational variants like (Buskell, 2017) calls. These transformation biases can be explained by the existence of factors of attraction, which stabilize the distribution of cultural items at a macro level, meaning entire populations and across generations (Claidière & Sperber, 2007).

(Sperber, 1996, pp. 111-112) presented this idea of attractors as “...abstract, statistical construct, like a mutation rate or a transformation probability...An attractor is not a material thing; it does not physically ‘attract’ anything”. Transformations can be biased in some direction, and at the end, creating a cluster around some point or points, those points are what is called cultural attractor. (Buskell, 2017) points that Cultural Attraction Theory have been used in different fields as anthropology, psychology, linguistic and cultural evolution.

One of the different interpretations of attraction in cultural evolution can be seen in the paper of (Claidière & Sperber, 2007) when he explains that (Henrich & Boyd, 2002) were the first to propose a formal model of cultural attraction, but their interpretation of cognitive attractors from Sperber is in what Claidière call “deterministic view”, differently from his interpretation of cognitive attractors which he calls “probabilistic view”. These interpretations affect how the cognitive attractors could influence cultural evolution.

The probabilistic view of attractors of (Claidière & Sperber, 2007) discuss about content-based attraction (inferential transformation) and distribution based transmission (selective attention) and concludes that the attraction force (probabilistic, which allows the ideational variants to get attracted either directions of the distribution) and the selective force are both in play at the same time, they will together influence and determine the outcome distribution of the population, depending on the peculiar properties and relative strength of both forces

What Claidière call deterministic view from Henrich and Boyd, is that the cognitive attractors will play a role influencing the start of the process, but after a while, only selective forces will determine the outcome. On the models proposed on (Henrich & Boyd, 2002) the outcome is determined by selection solely when the attraction is strong. On this approach, attraction would still be important but not a driving force for modeling the dynamics of cultural evolution. Claidière changes some parameters from the models of Henrich and Boyd and come to different conclusions, as attraction force being the determining force. At the end, Claidière explains that the view of attraction from Henrich and Boyd is deterministic because the cultural descent only moves in the direction of the attractor and he disagrees with that view.

# References

Bisin, A., & Verdier, T. (2008). Cultural Transmission. In S. N. Durlauf, & L. E. Blume, *The New Palgrave Dictionary of Economics* (pp. 1225-1229). London: Palgrave Macmillan. doi:10.1007/978-1-349-58802-2

Buskell, A. (2017). What are cultural attractors? *Biology & Philosophy*. doi:https://doi.org/10.1007/s10539-017-9570-6

Claidière, N., & Sperber, D. (2007). The role of attraction in cultural evolution. *Journal of Cognition and Culture 7*, 89-111. doi:https://doi.org/10.1163/156853707X171829

Heintz, C. (2018). Cultural Attraction Theory. 1-10. doi:10.1002/9781118924396.wbiea2311

Henrich, J., & Boyd, R. (2002). On Modeling Cognition and Culture: Why cultural evolution does not require replication of representations. *Journal of Cognition and Culture 2.2*, 87-112. doi:https://doi.org/10.1163/156853702320281836

Henrich, J., Boyd, R., & Richerson, P. J. (2008). Five misunderstandings about Cultural Evolution. *Human Nature 19, no. 2*, 119-137. doi:https://doi.org/10.1007/s12110-008-9037-1

Hershberger, S. L. (2014). Cultural Transmission. *Wiley StatsRef: Statistics Reference Online*. doi:https://doi.org/10.1002/9781118445112.stat06740

Lewens, T. (2018, 05 01). *Cultural Evolution*. Retrieved 04 10, 2019, from The Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/archives/sum2018/entries/evolution-cultural/

*Merriam-Webster Dictionary*. (2018, March 11). Retrieved 04 09, 2019, from Culture: https://www.merriam-webster.com/dictionary/culture

Mesoudi, A., & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509), 3489-3501. doi:https://doi.org/10.1098/rstb.2008.0129

Mesoudi, A., Whiten, A., & Laland, K. N. (2004). Perspective: is human cultural evolution Darwinian? Evidence reviewed from the perspective of the Origin of Species. *Evolution*, 1-11. doi:https://doi.org/10.1111/j.0014-3820.2004.tb01568.x

Miton, H., & Charbonneau, M. (2018). Cumulative culture in the laboratory: methodological and theoretical challenges. *Proceedings of the Royal Society B: Biological Sciences*. doi:https://doi.org/10.1098/rspb.2018.0677

Reips, U.-D. (2002). Standards for Internet-Based Experimenting. *Experimental Psychology, 49 (4)*, 243-256. doi:http://dx.doi.org/10.1026//1618-3169.49.4.243

Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach.* Oxford: Basil Blackwell. doi:https://doi.org/10.1093/mind/110.439.845

Sperber, D. (2000). An objection to the memetic approach to culture. *in Robert Aunger ed. Darwinizing Culture: The Status of Memetics as a Science. Oxford*, 163-173. doi:10.1093/acprof:oso/9780192632449.003.0008

Verschuren, P., & Doorewaard, H. (2010). *Designing a Research Project.* The Hague: Eleven International Publishing.

Part II: Experiment / Empirical studies

# Design and Implementation

The design and implementation of this project is a result of a collaborative work and ideas between the supervisor of the project and head of the Cognitive Science department of the Central Eastern University, Dr. Christophe Heintz and me, student graduating at the Master of Science program of Computational and Cognitive Neuroscience at the Eötvös Loránd University.

Due to the deeper and wider knowledge of Cultural Attraction Theory and the longer history on the field of Cognitive Science, the design of the experiment is a creation from Dr. Heintz. The implementation of the project, which contains development and deployment of the website are result of my expertise on the field of Information Technology, due to my bachelor’s degree in computer science and work experience on the aforementioned area.

The design of the experiment consists of a cross location game where the participant sees a diagonal cross and is asked to answer the position of it. At the beginning of the experiment, the cross will be demonstrated in a white background screen for one second, then in sequence it will be demonstrated a fill-out screen for half-second in a black background screen, this fill-out screen has the purpose of changing the visual fixation point of the user, subsequently the half-second, it appears another white background screen where is expected the participant to answer the position of the diagonal cross.

The concept behind this experiment consists on the understanding that during the cognitive process of visualization of the cross the player acquires information (the position of the cross), stores and then it transmits this information forward (the feedback position of the cross), however this process of perceiving/replying is believed to be imperfect and susceptible to external factors and transformations. The key concept and the goal of this experiment is to visualize if there is any possible scenario or factors, that would evidence the transformation of knowledge, in an individual level. Meaning that it would not happen an exact replication of the information, but a factor that plays a role on attracting these transformations to common points.

In the case of the designed experiment, the input and output are directed from/to the experimenter, because the goal of study is to analyze possible factors of attraction that are in play in the individual cognitive processes of acquisition, storage and retrieval of the information. But for a more elaborated research, it could be used a transmission chain to analyze these transformations in a larger scale, where a random walk would start from a seed (first input provided by the experimenter), and the output of the user would be an input of another user, following a chain until the last user.

On the following chapters, it will be detailed the architecture of the website, including the technical aspects of it, and the conditions tested, the reasoning behind each of them, and how they are related with the theory of factors of attraction. The code of the website will not be presented nor explained on this paper since is not the aim of this project approaching the nuances of computer programming, but rather test the Cultural Attraction Theory using a web platform which enables high availability and collection of data. That said, the code has comments in itself with the goal to facilitate the understanding of anyone interested on it and it is available on the personal GitHub repository of the author of this master’s thesis. That can be accessed at: https://github.com/RenanOm92/factorsAttractionFirebase.

## Architecture

The website was built using the Angular Framework and hosted on Google Firebase with the usage of their database. The website is available under the Uniform Resource Locator (URL): [https://factorsattraction.firebaseapp.com](https://factorsattraction.firebaseapp.com/).

The home screen of the experiment consists of a welcome page where the user receives information regarding the data storage and is asked to fill a text field with their e-mail. The next screen consists of the instructions of the experiment, explaining what the user can expect from next screens and guiding them how to participate correctly, this page contains a “Start” button. At the beginning of the experiment, it will be presented a black diagonal cross for just one second. Then automatically, it will be presented a black screen containing some different features, this screen lasts for half second or requires a click on a button to advance to the next stage. The last screen is a totally blank screen where the user should answer the position of the cross. After providing the position, the last screen is presented, where it shows the comparison between the original cross position and the user feedback position and offers the opportunity to the participant play again.

### Angular framework.

The project was developed using the framework Angular, which provides the possibility of creating applications compatible with cross platforms (web, mobile web, desktop) (Angular, 2011), fulfilling one of the aims of the design of the experiment which was having a website accessible by every type of user, at any moment, from anywhere. Generating this way, a spreader sample of whom would interact with the website, deeply connected with our goal that is to examine culture transmission, meaning that the broader the reach of the experiment, the more accurate would be the conclusions draw by it.

Angular is also an open source project led by Google and the community, which provides a great and stable environment for development, but at the same time still being updated with new features and being one of the most popular Frameworks for web development (Holderness, 2018). Another important advantage of Angular it is the two-way data binding feature, which enables the Model and the View being synchronized, allowing them to communicate between each other and any changes on the data would affect the visualization of it, used on the project for the calculation of the random position of the cross visualized in the start of each round.

With all these qualities listed above, Angular was chosen as the framework used for constructing the website of the experiment. The version utilized was Angular 5, which has released date on September 14th of 2016. (Angular changelog, 2016)

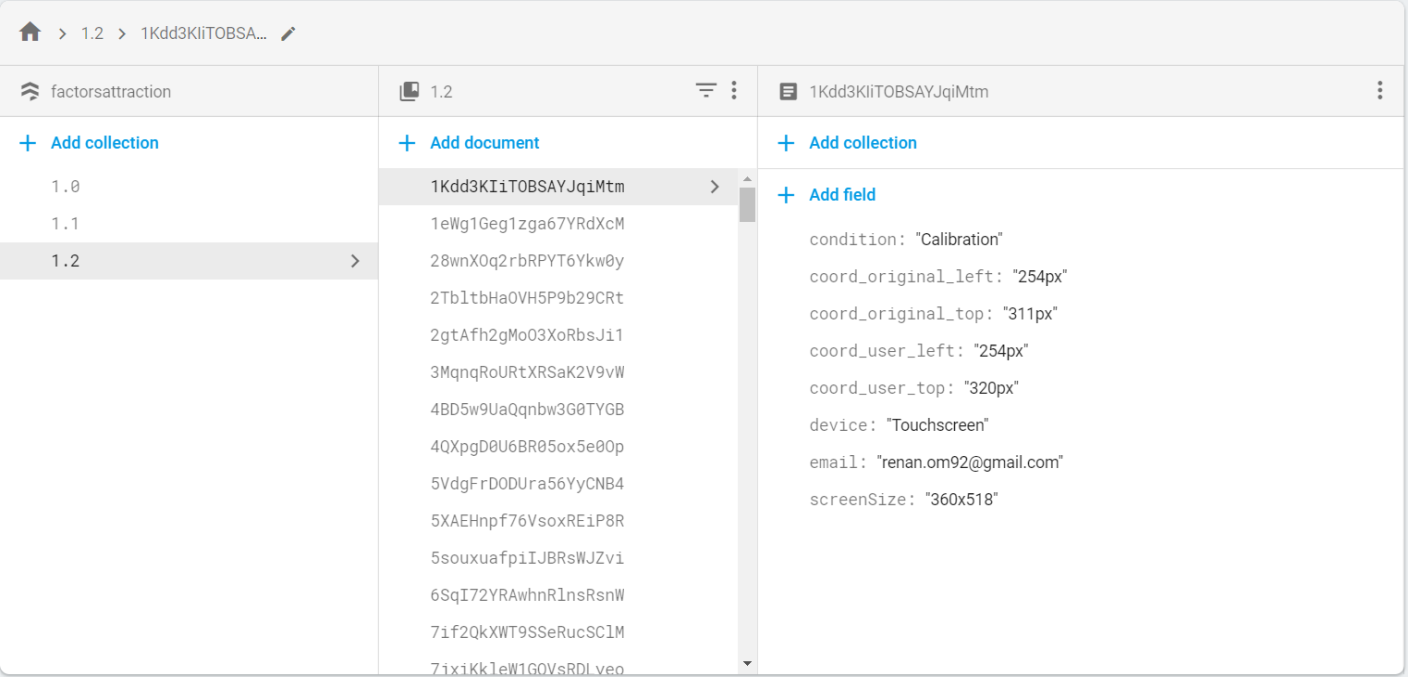
### Firebase hosting and database structure.

Firebase is a multifunctionality platform owned by Google which provides a large range of products, including hosting and database integration to web applications which were necessary for the deployment of this experiment (Documentation, 2016). Firebase was the chosen platform because it offers a free plan with the following specifications, which covers all the needs of the project: hosting the web application, 100 simultaneous connections, saving the data on a database with a limit of one *Gigabyte* per month, extracting the data from the database and custom domain. For the reasons listed above, the website was then hosted on Firebase with the domain name: factorsattraction.firebaseapp.com. And all the data collected from the experiment was saved on the Cloud Firestore Database accordingly with the data structure supported: Collection, Document, Data as shown on the Figure 2.



*Figure 2* - Data structure of Cloud Firestore. Source: (Cloud Firestore Data model, 2017). Retrieved from: <https://firebase.google.com/docs/firestore/data-model> Copyright 2019 by Google Developers

The experiment was released in different phases, as it will be explained on the section [Data collection](#_4._Collecting_data), and it each phase was saved as Collections using the semantic versioning approach (Best Practices When Versioning a Release, 2018). Each interaction of the user (also called as one round) it was saved as a Document with an unique identifier generated randomly by Firebase, and all the data referent to the round was saved inside the document. Each document on the final phase of the project had the following structure of data: *condition*, *coord\_original\_left*, *coord\_original\_top*, *coord\_user\_left*, *coord\_user\_top*, *device*, *email*, *screenSize*. The database structure can be seen on the Figure 3.



*Figure 3* - Data structure example from the experiment. Source: (Database, 2019). Retrieved from: [https://console.firebase.google.com/project/factorsattraction/database/firestore/data~2f](https://console.firebase.google.com/project/factorsattraction/database/firestore/data~2F) Copyright 2019 by Google Firebase

The meaning of each field value is better explained on the Table 1 below:

Table 1

*Detailed explanation of data variables*

|  |  |  |
| --- | --- | --- |
| Field | Description | Type of data and exemplification |
| Collection | Container of the document. Used for control of version. | Numbers on the format:  1.1  2.0 |
| Document | Unit of storage, represents one interaction of the user playing the experiment. It is an unique sequence of 20 characters composed of letters or numbers generated randomly by Firebase. | Sequence of 20 letters or digits:  1Kdd3KIiTOBSAYJqiMtm  4BD5w9UaQqnbw3G0TYGB  4QXpgD0U6BR05ox5e0Op |
| Condition | Represents the possible conditions in which the experiment can be played. | One of the following values:  Calibration  Face  SpiralLeft  SpiralCenter  ClickHereBottomLeft  ClickHereTopRight |
| coord\_original\_left | Represents the horizontal position on pixels which the cross was displayed at the beginning of the experiment. Where the 0 value is the left of the screen. | Number followed by pixel:  746px |
| coord\_original\_top | Represents the vertical position on pixels which the cross was displayed at the beginning of the experiment. Where the 0 value is the top of the screen. | Number followed by pixel:  520px |
| coord\_user\_left | Represents the horizontal position on pixels which the cross was answered by the user. Where the 0 value is the left of the screen. | Number followed by pixel:  777px |
| coord\_user\_top | Represents the vertical position on pixels which the cross was answered by the user.. Where the 0 value is the top of the screen. | Number followed by pixel:  12px |
| device | Contains which type of input device was used by the user. | One of the following values:  Touchscreen  Mouse |
| email | Contains the e-mail of the participant, used with the purposes of controlling the number of total interactions versus total of users. | Any possible value containing the e-mail format:  aaaaaa@aaaa.aaa |
| screenSize | Stores the value of the screen size of the device in which the participant used for starting the experiment. | Number x Number:  1366x626 |
|  |  |  |

## Scenarios for testing

When playing the game, the user can experience seven unique and different scenarios, they are: calibration, face, spiral centralized, spiral on the top right, spiral on the top left, “click here” button on the top right and “click here” button on the bottom left of the screen. Those scenarios will be detailed and explained in the following chapters. The scenario on which the user will play the game is randomly generated before the round starts (on the screen containing the instructions of the experiment in case of being the first round or at the end of the experiment for any rounds after that).

Even though the scenarios are randomly assigned, it was created a weighted value of appearance for each of them, with the purpose to have more appearance of a desired scenario. In the [Attractor condition phase](#_Attractor_condition_phase), which contains all the scenario and will be explained later, the weights are: 33.3% for face scenario, 33.3% for the spiral scenario, 22.2% for the “click here” button and 11.1% for the calibration scenario. The calibration scenario has a lower weight when compared to the other scenarios because in the initial phases of the experiment it was collected only calibration data, which caused the abundance of calibration data, making unnecessary expose this scenario with the same frequency in the final version of the experiment. The “click here” button has a slightly lower weight due to the fact that it has only two variances, in contrast with the spiral that have three.

### Calibration.

The calibration scenario was the first stage of the software development and contains the core of the experiment, and it also has the purpose of being a calibration for the following scenarios, meaning that the data collected with this scenario can be used as our control scenario to further comparisons, this is possible because the calibration scenario do not have any figure or image that could have a role of being a attractor.

The calibration scenario consists of a white background at the start page and feedback page with a black fill-out screen between them. The start page contains the original position of the cross and it is displayed for one second. Then automatically it is displayed the fill-out screen for half second and then the feedback page can be visualized until the user inputs some value for the cross.

The original position where the cross can be displayed at the beginning of the experiment is generated randomly but it follows some conditions and one parameter. The conditions that it must follows are: not being too close to the edges of the screen, nor being in the center. The cross cannot be in the center due to the appearance of the button “play again” on the center of the results page, which could lead the button to cover some results.

The parameter for calculation of the position of the cross is the screen size of the device of the user. Due to the reason that each user can have different devices (mobiles, personal computers, tablets) with different screen resolutions, it is necessary to capture the screen size on which the user is playing and then based on that, generate the original position of the cross.

Based on that two information, the original position of the cross is a random value in pixels between 5% of the screen size until 45% or 55% until 95% (for both height and width independently). For example, if the screen resolution (width × height) is 2000 × 1000 pixels (hereafter; px), the possible values for the width would be from 100px until 900px, and from 1.100px until 1.900px and for the height would be from 50px until 450px and from 550px until 950px.

### Face.

The face scenario has the same features of the calibration but with the addition of a background image in the start page and feedback page, this image, visible at Figure 4 has the intention of recalling the shape of a face with a missing eye. The purpose of this scenario is to try out the hypothesis that the missing eye could be a cognitive attractor. This hypothesis arises from the fact that humans have a specific area in the brain called Fusiform Face Area in which responds strongly to face stimuli (including shapes that resembles faces, as line). In the proposed figure, the right eye is missing on purpose, making the user feel the urge of the completion of the face shaped figure, generating a supposed cognitive attractor on the area of the right missing eye.



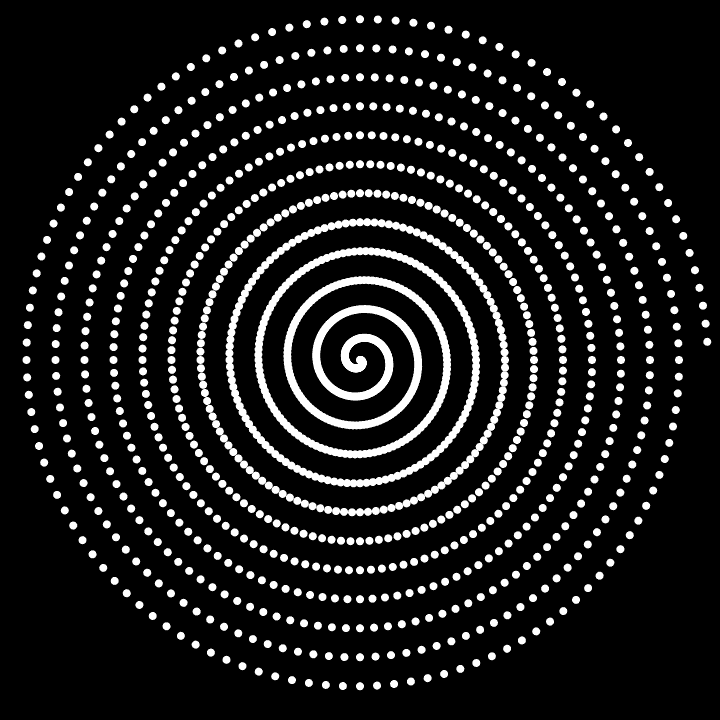
*Figure 4* - Face shaped image used on the scenario.

The face shaped image is always centralized on the screen and has the feature of missing the right eye. The size of the image is adjustable and relative to the size of screen, which means that the figure does not have a fixed value in pixels, so our analysis of the feedback of the cross need to be based on percentage of the screen size of each user, so in this way we are able to examine if the feedback of the user is located in an approximate position of the missing eye of the shaped face figure. With that said, the center of the missing eye position figure (a hypothesized attractor) is approximately on the coordinates 55% from the left and 27% from the top screen sizes.

Due to the position of the face and more specifically the missing eye being centered, I guided the generation of the random cross position to be more centered also. While the conditions of the cross position on the calibration scenario is being between 5% and 45% or 55% and 95%, in the scenario of the face, it was defined as between 30% and 49% or 51% and 70%. With these limits, the occurrence of the cross position on the center will be higher than in the calibration phase. At the moment of the design, it was believed that was worth generating values between 45% and 49% and 51% to 55%, meaning that the cross position could be further covered by the “Play again” button, affecting the usability and user experience quality but would increase the total of valuable samples for data analysis. The image size of the face shaped image will depend on the screen size and will be better explained in the section Data calculations.

### Spiral.

The spiral scenario like the previous one is an extension of the calibration, but it has the addition of showing a spiral on the fill-out screen for 0.5 seconds, this spiral, visible at Figure 5 has three different variances based on the location where it can appear (top-left, centralized or top-right) on the screen. This scenario has the purpose of to try out if the center of the spiral could be an attractor. The image measurements (width and height) of the spiral will always be half of the width screen.



*Figure 5* – Spiral used on the scenario. Source: (American tesol webinar halloween for young learners, 2013). Retrieved from: <https://americantesol.com/blogger/american-tesol-webinar-halloween-for-young-learners/>

### Button.

The button scenario is also based on the calibration, but it has the addition of a button on the fill-out screen which is necessary to be clicked to proceed with the experiment, this feature is different from all other scenarios, in which the fill-out screen is showed for 0.5 seconds then automatically the feedback screen is shown. The button on the Figure 6 has two different variances, based on the location (bottom-left and top-right) on the screen. The possible attractor tested on this scenario would be the center of button itself. The button has measures of 87x32 px.



*Figure 6* - Click here button used on the scenarios.

# Data collection

The website started gathering data on the day 28th of January of 2019 and is available until the present day, nevertheless at 06th of April of 2019 it was performed a snapshot of the database. This snapshot contains the data that will be used for analysis in this thesis. The web-based experiment was announced over the social media profile of the author of the thesis with the purpose to captivate users of different locations. A total of 1296 interactions were registered during the collecting period, in which 57 were performed by unique e-mail users, assuming that an unique e-mail is one person, results in a participant group size of 57 individuals, in which those participants could play in any of the factors of attraction hypothesized scenarios as also the calibration scenario.

## Calibration phase

The first deployment of the website, which followed the software development stages of the website itself, contained one and only the calibration scenario and it was published online on 28th of January of 2019. The data collected was saved into the Firsetore collection as version “1.0”. This first deployment did not collect the information regarding the device of the user, lasted for one day and collected 53 interactions.

On the 29th of January, there was deployed a new version, stored as version “1.1” which contained the device of the user, this version also only had the calibration scenario. The duration period of the “1.1” version was approximately four full days, lasted until 1st of February and collected 313 interactions.

The first two versions of the experiment collected 366 interactions of the calibration scenario, scenario which will be used as our control scenario and baseline for further comparisons. The calibration scenario will still be present in the next versions, but with the appearance rate reduced as explained in the section Scenarios for testing.

## Testing phase

On the day 2nd of February of 2019, the website was deployed in a testing phase, as known as version “1.2”, where introduced plenty of different scenarios with the purpose to test their usability and encounter possible misleading features, these usability tests were performed by myself and Christophe Heintz. Some scenarios (e.g. different face shaped image, white background fill-out) were not approved and will not be taken into consideration, meaning that they were not selected to the next phase and consequently will not be in the data analysis. The scenarios that were approved can be visualized on the section [Scenarios for testing](#_Scenarios_for_testing) and will be used on the data analysis. This phase lasted until 5th of February of 2019, a total amount of four full days and collected 49 interactions of the calibration scenario, 27 of the “click here” button top right scenario, 24 of the “click here” button bottom left, eight of the scenario spiral left, one of spiral center and seven of spiral right.

## Attractor condition phase

The attractor condition phase is the last collecting data phase of the experiment, and the one that contain all the scenarios designed at the beginning of the experiment, namely Calibration, Face Shaped, Spiral and Click Here Button, with all the possible variations. It started on 6th of February of 2019 and lasted until 06th of April of 2019, totalizing an amount of sixty full days. This phase collected 86 interactions of the calibration scenario, 307 of the face shaped scenario, 90 of the “click here” button top right, 76 of the “click here” button bottom left, 81 of the scenario spiral left, 94 of spiral center and 80 of spiral right.

During this phase, it was found a bug in the scenario Spiral on the top right of the screen (the spiral was not showed at the screen, consequently the fill-out screen was totally black and did not contain the hypothesized attractor), and all the data related to this specific scenario will not be taken into analysis.

# Data analysis

The data analysis can be separated in different sections, first and foremost, it was necessary to extract the data from the Cloud Firestore Database, where all the data was saved. After that, it was used the programming language Python to manipulate all the following activities: load the data, clean it, perform necessary calculations, utilize statistical measures and plot the results. The code and the data file are available at the git repository from the author (RenanOm92/factorsAttractionFirebase, 2019) under the folder “Data extraction and analysis”.

## Data extraction and cleaning

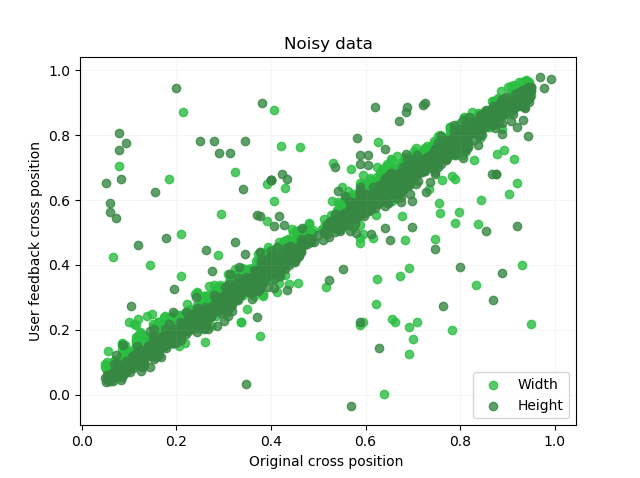
To perform the data extraction from Cloud Firestore, it was used the external package available at Node.js package manager (mostly known as npm) called “node-firestore-import-export”. Please visit their website (node-firestore-import-export, 2019) for a complete guidance and explanation of all functionalities from the package, on this thesis it will be only explained the basic tasks performed to extract data.

To install this package, is necessary to have Node.js installed and then execute on the Command Prompt the following command: *npm install node-firestore-import-export*. After the package being successfully installed, is necessary to get the credentials from the Cloud Firestore to be able to extract the data. On the Firebase Console, click on Project Settings (the gear icon button on top left of the page), navigate to Service Accounts and then click Generate New Private Key. There will be downloaded a JSON file with all the credentials necessary for the data extraction. With the package installed and with credentials downloaded, it is possible to perform the data extraction, to do so, the following command should be executed: *firestore-export --accountCredentials path/to/credentials/file.json –backupFile factorsAttraction.json*. It is necessary to adjust the path accordingly with the path of where it was downloaded the credentials file. As consequence, there will be created a JSON file called factorsAttraction.json containing all the data from the experiment.

Using python and the packages JSON and Pandas, the data was imported from the file factorsAttraction.json to a DataFrame, DataFrame is a 2-dimensional labeled data structure perfect for manipulating and performing the data analysis. For more information about how it was performed the action of transforming the JSON file into a DataFrame, check the file called dataReading.py on (RenanOm92/factorsAttractionFirebase, 2019). This file also contains the code performed in the data cleaning, data calculations and plotting the results.

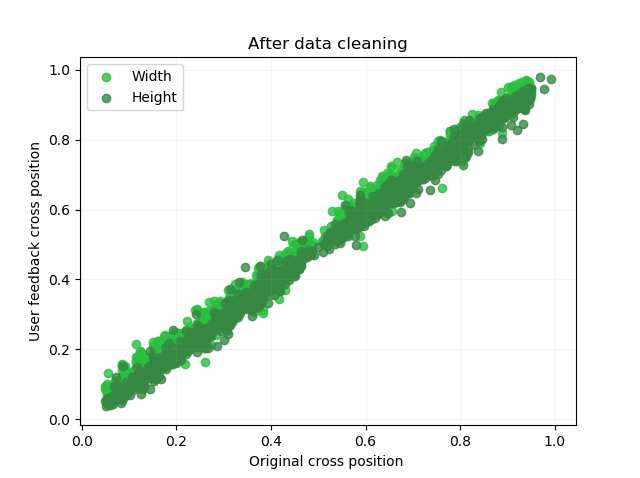
The data cleaning process consisted on two steps, the first one is transforming the data from the database in usable data, for example: the field “screenSize” was broken down in two columns: width and height, another example is the value left and top are transformed in X and Y of the cartesian coordinate system, and another one is the transformation of pixels in proportion of the screen size. The second step of the data cleaning consists on removing data from the DataFrame which will not be valuable to the analysis, this data is called noisy data. Noisy data on this experiment can be considered answers from the users where the feedback of the position is far away from the original position. This mistake could be a result of lack of attention from the user while playing, or even mistakes due to learning the game and another possibility is that those mistakes could be malicious answers with the purpose to damage the experiment.

On the Figure 7 is possible to visualize all the data collected on the experiment, each point represents the width or height values of an interaction, where the X-axis is the original position of the cross, and the Y-axis is the feedback from the user. The scale of the values is from 0 to 1, this value means the proportion of the position of the cross regarding the screen size. So, for example one screen resolution of 1000x1000px (width x height), where the original position of the cross is located at 500x200px, and the feedback of the user at 580x190px, the original position would be represented as 0.5x0.2, the feedback as 0.58x0.19. Then one point would be for the width (0.5,0.58) and another point would be for the height (0.2,0.19). If the feedback of the user is somewhat close to the original position, then when the point is plotted, it should be located on the diagonal, on the Figures 7 is possible to visualize the biggest distribution of points are located on the diagonal.



*Figure 7* – Data containing noise.

Based on this distribution, all the data that has more than 10% of discrepancy from the original width position to the user feedback width position will be discarded. The same rule is applied for the height values. After applying this rule, the data looked like the shown in the Figure 8. From a total of 1296 interactions, it remained a total of 1203 valid interactions.



*Figure 8* - Data after the process of cleaning.

## Data calculations

With the data cleaned, it is necessary to perform data calculations that will be required to analyze the results as: the distance between original and user feedback (hereafter DOU), position of the hypothesized factor of attraction, threshold radius of the influence of the attractor, etc. The first measurement that is calculated is the distance between the original position of the cross (a point in the cartesian coordinate system, where the width is the X-axis and height is the Y-axis) and the feedback of the user. All the coordinates here are also expressed in the scale 0 to 1 relative to the screen size. The formula (1) used to calculate the distance between two points is:

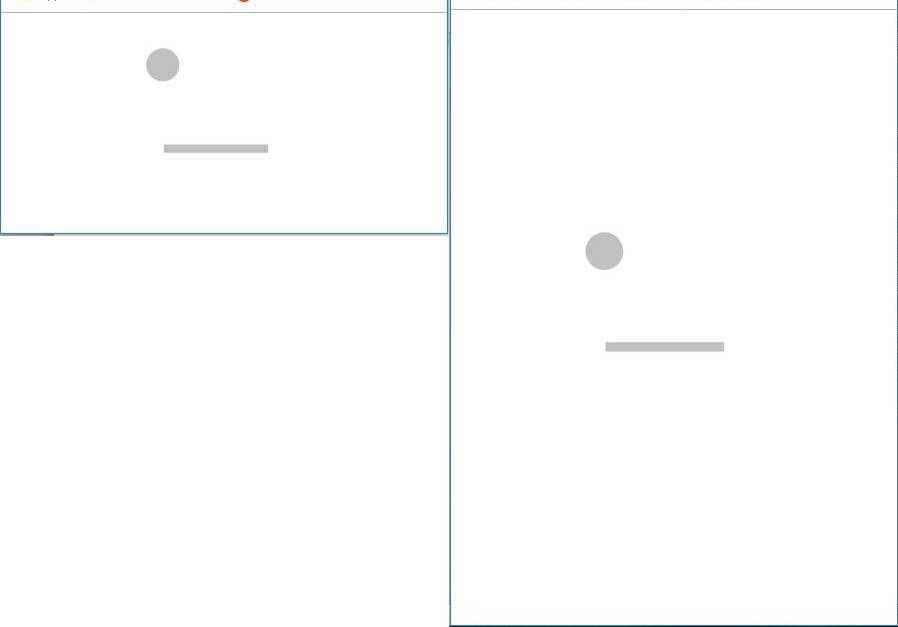
(x2 – x1)2 + (y2 – y1)2  (1)

The goal to calculate the distance between the original position and the user feedback is to validate if the calibration scenario has an average points distance lower than from the other scenarios supposed to contain attraction factors. In the calibration scenario, there is no variables to disrupt or effect the feedback of the user, so the average distance on this scenario should be considered the baseline error based on the lack of precision. In our hypothesized scenarios with factors of attraction, the tendency of the feedback of the user should get closer to the factor itself, resulting in an average longer distance from the original to the user position when compared to the average distance in the calibration. It was calculated the average distance from the points original position and user feedback of all scenarios separately, and the results will be presented in the next section.

Another important calculation is the location of the attractor in the cartesian coordinates system. This is a necessary step because some of the images used, as the face shaped image and spiral, have their dimension calculated relative to the screen size, meaning that the image size relies on the device from the user, the “click here” button image has a fixed size of 87x32 px. The face shaped image size depends on the aspect ratio of the screen size, the original image size has 1280 x 720 px, an 16:9 aspect ratio, or 1.777 ratio, and it will always keep this ratio (width versus height), so it is necessary to calculate if the screen size have a bigger aspect ratio (common on personal computers in our data), meaning that the image size will be based on the height. Or if the aspect ratio is smaller (frequent on cellphones/mobile devices in our experiment) and the image size will be based on the width screen.

The original spiral image has a dimension of 720 x 720 px, a ratio of 1 and will always have the image size based on the width screen, more precisely half of the width. So, if the screen resolution is 350 x 700 px, the image size will be 175 x 175 px. This rule applies for all spiral’s scenarios (left, center and right).

Not only the image size relies on the screen resolution as we have seen before, but the image disposition always depends on the screen resolution. The Figure 9 exemplifies the disposition of the face shaped image in two different screen resolutions with same width, which gives the same image size, but with different screen heights, affecting where the possible attractor would be. In the left image, the missing eye is closer to the top border of the screen, so the Y value is way higher than the Y value in the image on the right, where the missing eye is almost centered, having an Y value closer to 0.5 (50% of the height screen). Because of this behavior, it is necessary to calculate each interaction in an individual level and compare the original cross position and user cross position to the factor attraction point for each interaction in terms of the screen measurements (width and height).



*Figure 9* – Specific case of same image size with different screen heights.

The formula (2) for the face shaped image that gives the point of attraction of the missing eye relative to the screen size is defined as:

(2)

This formula (2) should be applied for both axis, X (width) and Y (height). Where the *screen size* refers to the width or height of the screen in pixels, the *image size* it is the width or height of the face shaped image in pixels, and the *location of the attractor in the image* refers to the missing eye in percentage in reference to the image, having a fixed value of 0.55 in the X axis and 0.73 in the Y axis. The result of this calculation is the relative position of the possible attractor (in a range from 0 to 1) in terms of the screen size. With this value is possible to compare if the user feedback cross position was closer or farther to the attractor compared to the original position.

For the spiral scenarios, the possible attractor to be analyzed is the center of the spiral, which is exactly at the center of the image for both height and width. As the spirals image always has half of the width screen size, the attractor on the X-axis for the left spiral has value 0.25 (meaning is at 25% of the screen width) and the center spiral has 0.5 (50% of the width size, exactly centered). For the Y-axis it is necessary to execute some calculations, for example the left spiral has a vertical alignment always at the top of the screen, so the formula (3) for calculating the attractor is:

(3)

For the center spiral the vertical alignment is centered, just like the horizontal alignment, so the formula (4) for the Y-axis is:

(4)

Both formulas give the attractor position in the Y-axis relative to the screen height in the scale from 0 to 1 to their respective scenarios.

For the “click here” button scenarios, the image size has a fixed width size of 87 pixels and height of 32 pixels. It will be taken in consideration the attractor as the center of the button for both axes, so the image size of 87x32 px should be multiplied by 0.5. The disposition of the left bottom button starts at 20% of the screen width and height in the cartesian coordinate system, and the right top button start from 80% from the cartesian coordinate system. Based on this information, the formula used to calculate the left bottom button for both axis separately, is:

(5)

The right top button values should also be calculated separately for the width and height and has the following formula:

(6)

The calculations explained so far produced the necessary information to proceed with the analysis: Relative position of the attractor for every interaction based on the scenario for the X and Y axis, and the average distance between the two points (original cross position and user feedback cross position) for every scenario. That information, aggregated with the relative position of the original cross position in the X and Y axis and the relative position of the user feedback cross position in the X and Y axis, enabled this study to analyze the possible impact of the hypothesized images in the users’ feedback.

With all the data calculated, the last step for the hypothesis of assuming the position of the attractors, is defining a threshold value where it should be investigated the radius of action of the attractor, meaning, how distant should be the cross position from the attractor so it could still suffer the attraction from it. On this question, a dilemma arises: taking in consideration a small radius of action, which explores less data but analyzes the points which will be more affected by the attraction factor or a larger radius of action that covers more data but at the same time may lose the effect of the attractor itself.

This dilemma can be solved with a wide and extensive data collection which generates plenty of data, enabling our study to concentrate only on the question of how far the attractor extends it radius of influence. To be considered inside the radius of attraction, the original position of the cross must be inside, or the user feedback has to be inside or both of them. Inside means a lower distance value than the threshold.

Unfortunately, this experiment had no funding and all the data collected was over social media from acquaintances of the author, this experiment collected 1296 unique interactions counting all the scenarios and calibration, 1203 were valuable after the data cleaning. From those 1203 interactions, 649 are of the five factors of attraction analyzed. Considering a radius of action of 15% of the total screen size from the attractor as the center (a diameter that covers 30% of the total size of the screen) it would leave only 163 unique interactions to all factors of attraction scenarios. Because of reduced number, it will be also analyzed a radius of action of 25% (diameter of 50%), meaning that half of screen size for height and width will be taken in consideration, on this case, there will be 342 unique interactions for all the factors of attraction scenarios.

## Results

The results are presented for two different lines of research: hypothesizing attractor position for the transformations regarding the memory process, and documenting transformation without assuming the position of the attractor.

### Transformation on the memory process.

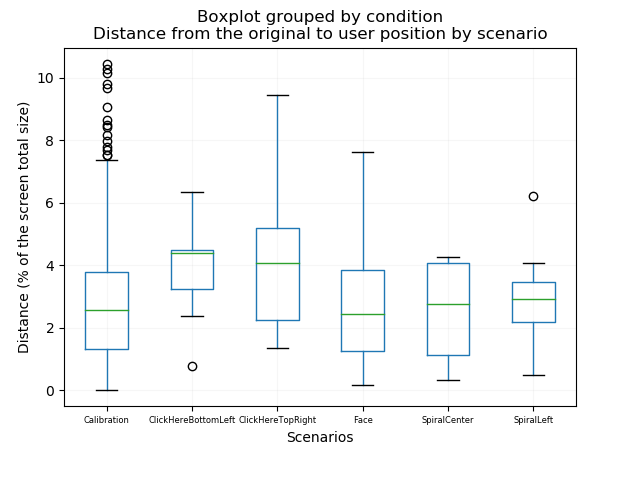
The results on this line of research can be investigated in two different approaches. The first one evaluates the distance between original and user feedback cross position. The second one compares the distance from the original position to the attractor with the distance from the user feedback to the attractor. All the data used on this section is cleaned (noisy was removed), and the threshold around the attractor are set on a radius of 15% or 25%.

#### Distance from the original position to user feedback.

On this approach, the null hypothesis consists on the assumption that all the scenarios would have the similar distance variance between the two points (original and user), this distance generated by the lack of precision of the user answers, common in all scenarios. The alternative hypothesis proposes the existence of an attractor in some images and scenarios, which influences the feedback from the user, this factor in conjunction with the normal error of precision, generates greater distances than in the calibration scenario.

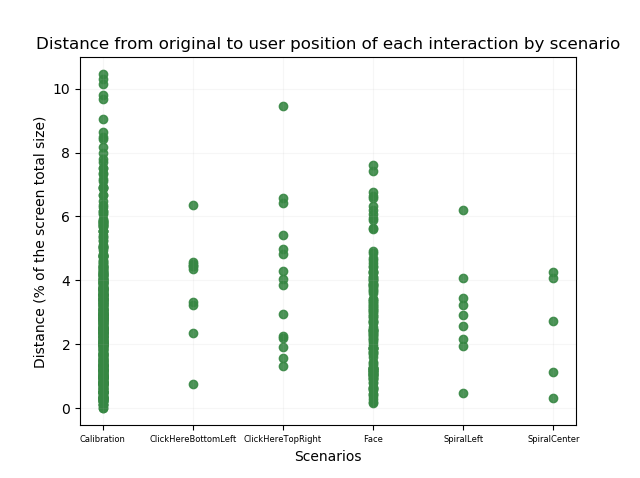
At the Boxplot graph on the Figure 10, using threshold of 15% from the attractor, it is possible to visualize the scenarios face and spirals have median value like the calibration scenario, likewise, the face, spiral and calibration scenarios share similarities with the boxes in the lower and upper quartile. Both scenarios containing the “click here” button presents somewhat a higher median value and quartiles. But looking at the Figure 11 where it elicits the amount of data, both “click here” scenarios are still covered by plenty of points of the calibration scenario, the dispersion of the calibration is much higher than the “click here” scenarios, so the difference seen in the box plot graph do not suggests any relevant variance, it is just result of lack of data, in my interpretation. For the verification of this perception, it was implemented an one-way ANOVA test regarding the 6 scenario/groups to validate if the differences between the group means are statistically significant.

The result of an one-way ANOVA from the data with threshold of 15% is [F(5, 629) = 2.07, p = 0.067, ηp2 = 0.016]. This result shows a low value for the eta-squared (1.6%), meaning that the model does not fit our data, and there is no relation between the differences of the distance related to the groups, this could be due to the lack of data and unequal size groups, not really having any relation between the distance and scenarios, or the non-existence of attractors. Based on this result, no relevant assumptions can be made to contradict the null hypothesis.



*Figure 10* - Boxplot graph with the distance from the original position to the user position, threshold of 15% from the attractor, presents a slightly difference in the median and box quartile for the scenarios with button click here.

In an attempt to utilize more data in the analysis, a threshold of 25% from the center of attraction towards the original or user points was taken into consideration. The box plot of this approach shares intrinsic similarities with the one on the Figure 10 from the 15% threshold and is available at the appendix as Figure 14. The one-way ANOVA result for threshold of 25% is [F(5, 808) = 2.35, p = 0.039, ηp2 = 0.014]. Also presenting an extremely low value for the eta-squared value which do not allows to draw any relevant conclusions regarding the distance from the original position to the user position across the scenarios explored.



*Figure 11* – Scatter graph with the distance from the original position to user position per interaction, threshold of 15% from the attractor, contradicts the assumption of the box plot graph, presenting the difference on the click here button scenarios can be just a matter of not enough data.

#### Distance from the attractor to the original and user position.

On this approach, it was calculated the distance from the attractor towards the original position of the cross (hereafter called DO) and likewise the distance from the attractor towards the user feedback position (hereafter called DU). With those two values, it was possible to compare them and see which one had the highest value, meaning that it would been farther from the attractor. Based on this comparison, it was possible to conclude for each interaction if the user feedback got closer or farther from the attractor when compared to the original position.

It was applied this technique for every attractor scenario and for both threshold cases (15% and 25% of the radius of influence of the attraction factor) and the results can be seen in the Table 2.

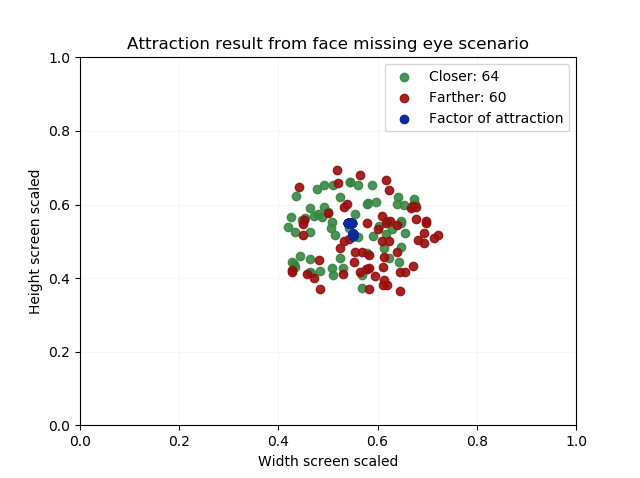
Table 2

*average distances and amount of closer/farther points.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario | Threshold | Average DO in % screen | Average DU in % screen | Total points got closer | Total points got farther |
| Click here bottom | 15% | 13.11% | 13.24% | 5 | 5 |
| left | 25% | 16.67% | 16.31% | 11 | 7 |
| Click here top right | 15% | 12.22% | 10.85% | 10 | 5 |
|  | 25% | 15.84% | 15.41% | 18 | 10 |
| Face | 15% | 11.30% | 11.10% | 64 | 60 |
|  | 25% | 15.80% | 15.72% | 126 | 132 |
| Spiral top left | 15% | 11.13% | 10.67% | 4 | 5 |
|  | 25% | 16.06% | 16.72% | 6 | 14 |
| Spiral center | 15% | 12.93% | 12.38% | 3 | 2 |
|  | 25% | 19.76% | 19.86% | 9 | 9 |

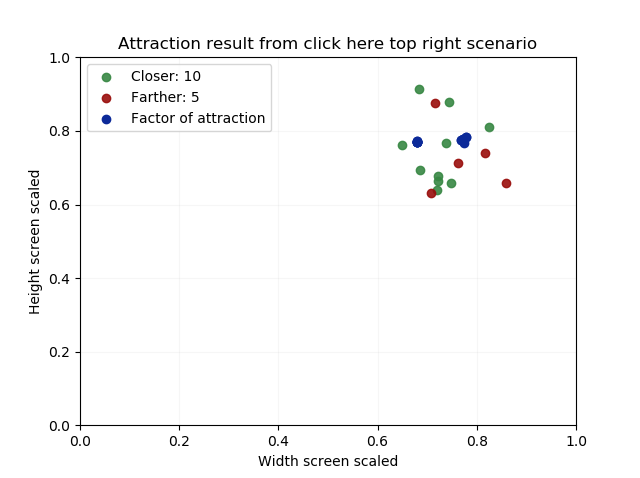
*Note*: provides the average distances from the attractor towards the original distance and user feedback, the amount of points which got closer or farther from the attractor for both thresholds (15% and 25%) for every scenario.

It was produced scatter plots to every scenario representing the user feedback points position, in those plots the points in green are the ones which got closer to the attractor and in red the ones that got farther. The blue ones are the attractor points, just to remind, the blue ones are not always on the same position, because it depends on the screen ratio, but the comparison regarding the distance was always done in an individual level per interaction. The Figure 12 represents the face scenario with threshold of 15% and shows 64 points which got closer to the attractor and 60 which got farther.



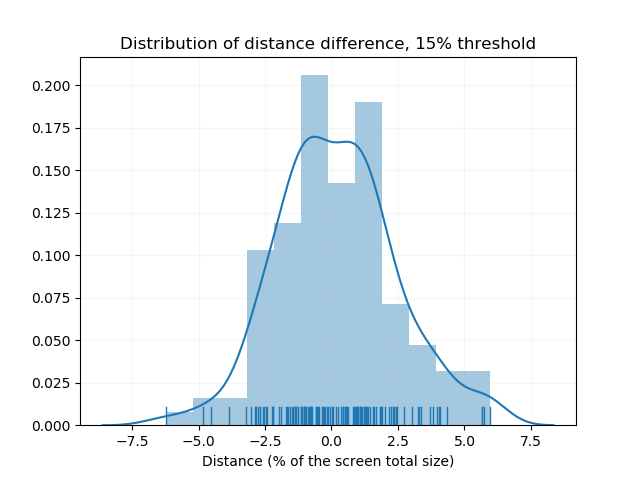
*Figure 12* - Scatter plot for the face scenario, using a threshold of 15%.

The scenario click here button on the top right shown on the Figure 13 is the one that produced the highest relation of attracted points across all the scenarios. The plots of the remaining scenarios for both thresholds can be seen on the appendix.



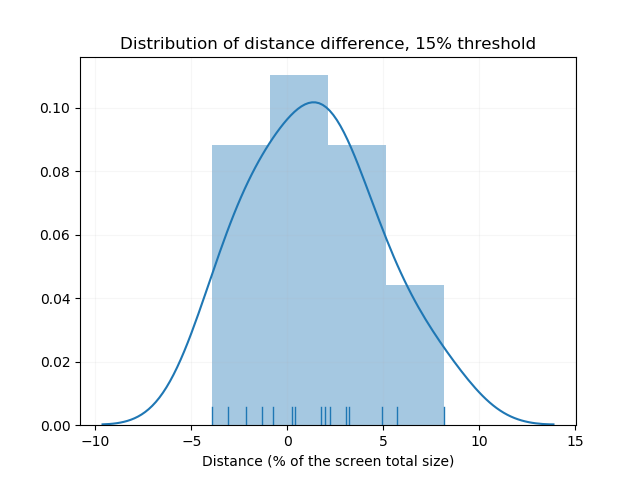
*Figure 13* - Scatter plot for the click here button on the top right scenario, using a threshold of 15%.

With the goal to analyze how the attractor influenced each scenario, it was plotted the distribution of the difference between DO and DU. The null hypothesis is that if there is no attractor in play, the head of distribution will be centered around the zero value. On the Figure 14 which demonstrates the distribution for the face scenario with threshold of 15% is possible to visualize the head around the zero.



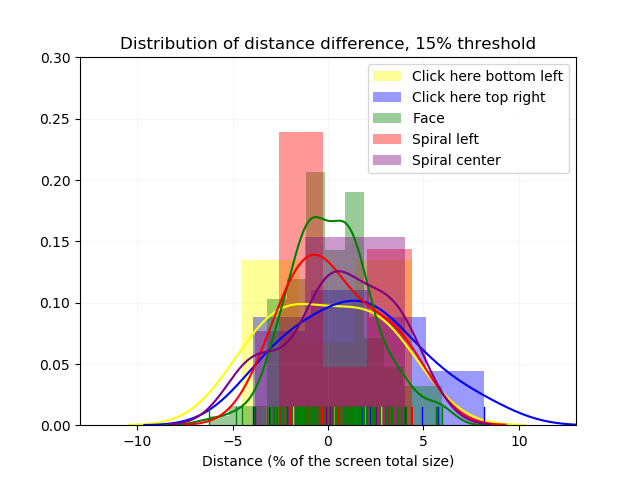
*Figure 14* - Distribution of the difference of DO and DU for the face scenario with 15% threshold.

It was also plotted the distribution for all the scenarios, but all of them have similar essence (head of distribution around the center), so with the purpose of not repeating information, it will not be demonstrated all of them one by one but just the click here top right button with 15% threshold, visible on the Figure 15, scenario which was mostly distinguished from others in regard of number of points which got closer/farther to the attractor.



*Figure 15* - distribution of the difference of DO and DU for the click here top right scenario with 15% threshold.

On the Figure 16, it was plotted all the five scenarios together to see how they contrast against each other, and as result all of them have the similarity of the head of the distribution around the zero value. The results for threshold of 25% are similar and it is available on the Appendix.



*Figure 16* - Distribution of the difference of DO and DU for all the scenarios of attraction with 15% threshold.

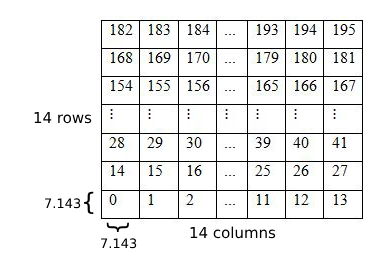
### Documenting attraction.

This line of research does not assume the position of the attractors, instead it accepts any transformation (the data cleaning part is not performed here, so it was used 1296 interactions) and then analyze any patterns per each scenario with the purpose to finding attractors that could be explained by the insertion of the figures. The data is divided per scenarios with 501 of the calibration scenario, 100 of click here button on the bottom left, 117 click here button on the top right, 307 of the face scenario, 82 of spiral left and 102 of spiral center. The spiral on the right was discarded by the same reason explained before.

For this approach it will be created a matrix that maps the screen coordinates. The screen will be split in squares, in which one of those squares will correspond to one cell of the matrix. All the data points referent to the original position of the cross will be assigned to one square, same for the user feedback position of the cross.

To map the screen size to a matrix, it is necessary to define what size the squares should be, with this purpose it was calculated the distance from the user original position to the user feedback position for each interaction and then the standard deviation of the totality of the data (1296), the standard deviation value found was ≈10.375. This value of the standard deviation will be considered the diagonal of a square, so it is necessary to find the size of the square, which will split the screen. When the diagonal has value of 10.375, the side of each square would be ≈7.33. Since the beginning, all the values are relative to the screen size (from 0 to 100%), then dividing 100 by 7.33 would provide 13.64 squares. With the purpose of having symmetrical squares, it was adopted the size of each square of 7.143 which will give 14 same sized squares for the width and height, see the Figure 17 for visualization of this matrix.

Having this matrix of 14x14 which covers the whole screen, it was calculated for each interaction where the original position of the cross would fit (matrix contains the total amount of appearances of the original position per square). It was created another matrix for the user feedback position of the cross. These matrices have 196 cells (14x14) and their enumeration can see in the Figure 17.



*Figure 17* - Enumeration of the quadrants in a matrix 14x14 that maps the screen coordinates. Each cell will contain the total amount of appearances.

With both matrices of 196 cells indicating the frequency of data points per square, it was created a new matrix of 196 x 196 which maps the original position of the cross (rows) to the user feedback (columns), in terms of square location. The next step was calculating the ratio for every row (each cell divided by the sum in the row), this action has the purpose to normalize the weight of original squares on the outcome. After that, it was calculated the sum of each column (which represents the user feedback). Each column out of the 196, represents one square of the screen, and the column with the biggest sum indicates the square that collected user feedback.

The data is presented in percentage, where the sum all off squares on the screen will totalize 100%. On this way it is possible to compare the results across the distinct scenarios. The Figure 18 contains the representation of the screen divided in squares for the scenario calibration, this matrix/squares contains the outcome of the calculations indicating the most ‘popular’ squares by the users.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0.61 | 0.2 | 0.43 | 0.3 | 0.3 | 0 | 0.46 | 0 | 0.43 | 0.3 | 0 | 0 |
| 0.3 | 0.3 | 1.33 | 1.25 | 0.41 | 1.59 | 0.3 | 0 | 0.77 | 0.3 | 0.61 | 0.67 | 0.79 | 0.61 |
| 0.61 | 0.2 | 0.51 | 0.43 | 0.61 | 0 | 0.15 | 0 | 0.61 | 0.61 | 0.37 | 0.53 | 0.59 | 1.02 |
| 0.12 | 0.67 | 1.38 | 0.24 | 0.47 | 1.02 | 0.72 | 0.61 | 0.92 | 0.51 | 0.51 | 0.97 | 0.7 | 0 |
| 0.12 | 0.72 | 0.50 | 2.42 | 0.12 | 0.72 | 0.61 | 0.43 | 0.3 | 0.72 | 0.56 | 1.02 | 0.56 | 0.72 |
| 0.61 | 0.61 | 0.56 | 0.78 | 0.56 | 0.56 | 0.77 | 0.3 | 0.51 | 0.82 | 0.82 | 0.63 | 0.34 | 0.72 |
| 0 | 0.15 | 1.18 | 0 | 0.61 | 0.82 | 0.72 | 0.3 | 0.61 | 1.04 | 0.87 | 0.15 | 0.15 | 0 |
| 0 | 0 | 0 | 0.61 | 0 | 0.12 | 0.1 | 0 | 0 | 0 | 0.68 | 0.77 | 0 | 0 |
| 0 | 1.11 | 0.55 | 1.04 | 0.17 | 0.17 | 0.43 | 0.61 | 0.98 | 0.43 | 0.34 | 1.16 | 0.56 | 0.61 |
| 0 | 0.3 | 1.01 | 0.3 | 1.34 | 1.17 | 0.86 | 0.41 | 0.61 | 0.49 | 0.53 | 0.68 | 0 | 0.36 |
| 0 | 0.58 | 0.56 | 0.84 | 0.52 | 0.86 | 0.82 | 0 | 0.82 | 0.41 | 0.2 | 1.69 | 1.33 | 0.3 |
| 0 | 0.92 | 0.61 | 10.8 | 0.7 | 0.49 | 0.3 | 0.3 | 0.72 | 1.13 | 0.2 | 0.69 | 0.15 | 0.15 |
| 0 | 1.44 | 0.92 | 0.5 | 0.58 | 0.44 | 0.08 | 0 | 0.72 | 0.98 | 0.10 | 1.02 | 0.15 | 1.23 |
| 0.2 | 0 | 0.1 | 0.7 | 0.12 | 0.7 | 0 | 0 | 0.3 | 0.74 | 0.3 | 0.92 | 0 | 0 |

*Figure 18* - Representation of screen/matrix for the calibration scenario. The highest value is highlighted in red.

The calibration scenario has no stimuli that was hypothesized as an attractor, so the scattered data over the whole screen is reasonable. The button on the bottom left scenario produced a matrix/squares as shown in the Figure 19, the expectancy was to have the highest values concentrated around the position of the button, but this was not fulfilled. All other scenarios did not show any indication of presence of attractor close to where the stimuli were place and their matrices are available at the appendix.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.49 | 0 |
| 0 | 1.49 | 0 | 0 | 0.74 | 0 | 1.49 | 0 | 0 | 0 | 0 | 0 | 0 | 2.98 |
| 0 | 0 | 0.74 | 0.74 | 0.74 | 0 | 0 | 0 | 0 | 1.49 | 0 | 0 | 0 | 0 |
| 0 | 1.49 | 1.49 | 0 | 1.49 | 0 | 0 | 0 | 0 | 0.74 | 2.23 | 1.49 | 0.99 | 0 |
| 0 | 0 | 0.24 | 0.24 | 0 | 0 | 1.49 | 1.49 | 2.23 | 0.74 | 0.74 | 0 | 0.49 | 0 |
| 0 | 0 | 0.49 | 1.74 | 1.49 | 4.47 | 0 | 0.74 | 1.49 | 1.11 | 0.74 | 0.74 | 1.49 | 2.98 |
| 0 | 0 | 1.74 | 1.49 | 0 | 0 | 0 | 2.23 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1.49 | 0 | 0 | 0 | 0 | 0 | 0 | 1.49 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0.74 | 0 | 0 | 1.49 | 0 | 0 | 0.74 | 2.98 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1.49 | 0 | 0 | 1.49 | 0.74 | 1.49 | 2.23 | 1.49 |
| 0 | 1.49 | 0.37 | 2.23 | 0 | 0.74 | 1.49 | 0 | 0 | 2.98 | 1.49 | 1.49 | 0 | 0 |
| 0 | 0 | 1.24 | 0 | 0.74 | 0 | 0 | 0 | 1.49 | 1.49 | 1.49 | 0 | 0 | 0 |
| 0 | 0.74 | 1.74 | 0.74 | 0.74 | 0 | 0 | 0 | 0 | 0 | 0 | 1.49 | 1.49 | 0 |
| 0 | 0.74 | 0 | 0 | 0 | 0 | 1.49 | 0 | 1.49 | 0 | 0 | 0 | 0 | 0 |

Figure 19 - Representation of screen/matrix for the button bottom left scenario. The highest value is highlighted in red.

Part III: Conclusion

# Discussion and future work

## Research objectives: Summary of findings

The main objective of this research was to build a web platform which provides a sandbox approach, where other scientists could create their own scenarios, modify the already existing ones and test their own ideas regarding cultural attraction. The web platform was successfully built, and it is available on the GitHub repository at (RenanOm92/factorsAttractionFirebase). The repository contains all the code of the website, the methods used for data analysis and the documentation, under an open source MIT License (The MIT License, n.d.) e, where there is no restriction to deal with the software and is it is free to be studied, used, modified, distributed, etc. The website with my experiments is deployed at Firebase and accessible under the address: [https://factorsattraction.firebaseapp.com](https://factorsattraction.firebaseapp.com/).

Based on the platform, we proposed an experiment containing different scenarios, which was performed the data collection and thereafter the data analysis. We analyzed the data in two different approaches: first, assuming the position of the attractors and filtering the data around those points, the second one considering the totality of the data on the screen. Both approaches did not produce any relevant indicators of the presence of attractors. Even though, our results did not fulfill our expectations, the construction of the platform provides the opportunity to the whole community to adjusts parameters accordingly with their own ideas and perhaps find different results.

## Limitations

The experiment had limitations which significantly impacted the whole development and results of the project. One of those was the lack of funding for collecting data, as result of this, the data collection was performed only by acquaintances of the author which is a considerable limited amount of people. Another factor that impacted the widespread propagation of the experiment was the absence of a data policy that protected and covered the participants of the experiment. These two facts together resulted in a small amount of data, especially after the data cleaning and data adaptions to the hypothesized factors of attraction.

Another issue was the lack of precision of the answers originated from participants that played using the touchscreen functionality (precision is the distance between original position of the cross versus the feedback from the user). Considering the 1151 interactions after data cleaning and which had the device data, 418 were from mouse and 733 from touchscreen. The overall precision was lower on touchscreen devices (average distance of 3.34% to all scenarios including calibration) compared to the interactions coming from mouse (average distance of 2.18%), making it even more difficult to detect possible factors of attraction. This represents a 1.16% difference in distance value relative to the screen size from touchscreen to mouse, a considerable value. Just to put in perspective the values of distance between calibration and all other scenarios (without filtering the threshold close to the factors of attraction) for mouse is 2.00% for calibration and all other scenarios is 2.30%. For touchscreen the calibration has 3.18% and all other scenarios is 3.43%. Which shows a gap way lower between scenarios than between types of devices.

A possible solution for the cellphone lack of precision is the introduction of a learning phase in which the user can get habituated to the experiment and possibly provide better results with more correctness. Another solution is performing this experiment only in devices which provides a mouse and excluding the data from touchscreen devices, as this experiment is supposed to analyze sensitive variations, the lack of precision can influence negatively the outcome.

## Future research

As future research it is possible to perform improvements on the web platform or modify the experiment with the aim to document and find attractors.

In terms of improving the quality of the web platform and engagement of the user while playing the game, it is possible to implement a tracking system which compares the answers of the user with his historical interactions or with the performance of other users, on this sense it provides the participant a more enjoyable feedback based on improvement/achievement which can stimulate the user to keep playing and providing more data. Another improvement on the website, would be the inclusion of a learning phase, which would guide the participant in the first rounds of the experiment, making the participant more comfortable and confident about the task to be performed, this improvement arose from observation of the behavior of the participants and also, from suggestion of the participants.

Regarding the experiment scenarios, in the future it would be indicated to use different stimuli figures, for example a face shaped image that is proven to activate the Fusiform Face Area, because the stimuli utilized in the experiment was created by me, without any testing regarding its effect. Overall, it would be suggested to proceed the research with the utilization of stimuli/images that are proven to have an impact on our cognitive process of visual perception and memory, because our images were selected without any prior testing.

# References

*American tesol webinar halloween for young learners*. (2013, 10 29). Retrieved 04 09, 2019, from American TESOL Institute: https://americantesol.com/blogger/american-tesol-webinar-halloween-for-young-learners/

*Angular*. (2011, 07 02). Retrieved 04 09, 2019, from Angular: https://angular.io/

*Angular changelog*. (2016, 09 13). Retrieved 04 09, 2019, from GitHub: https://github.com/angular/angular/blob/master/CHANGELOG.md

*Best Practices When Versioning a Release*. (2018, 01 23). Retrieved 04 09, 2019, from Codeship: https://blog.codeship.com/best-practices-when-versioning-a-release/

*Cloud Firestore Data model*. (2017, 10 03). Retrieved 04 09, 2019, from Firebase: https://firebase.google.com/docs/firestore/data-model

*Database*. (2019, 04 09). Retrieved 04 09, 2019, from Firebase: https://console.firebase.google.com/project/factorsattraction/database/firestore/data~2F

*Documentation*. (2016, 05 18). Retrieved 04 09, 2019, from Firebase: https://firebase.google.com/docs/

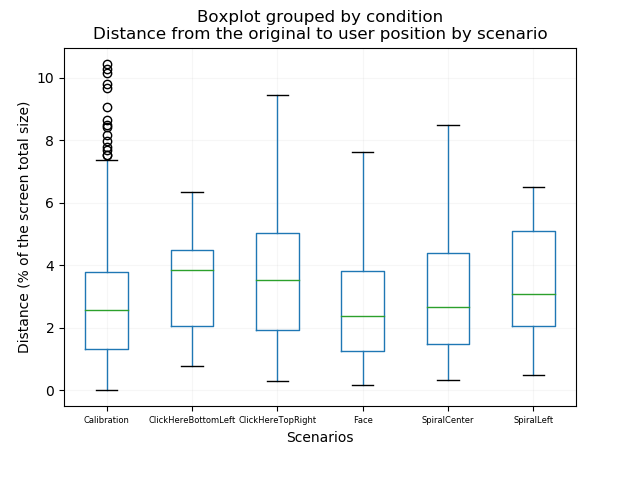
Holderness, A. (2018, 02 14). *Is the Angular Decline a Myth?* Retrieved 04 09, 2019, from Hackernoon: https://hackernoon.com/is-the-angular-decline-a-myth-e4cf563b72d6

*node-firestore-import-export*. (2019, 02 05). Retrieved 03 30, 2019, from npm: https://www.npmjs.com/package/node-firestore-import-export

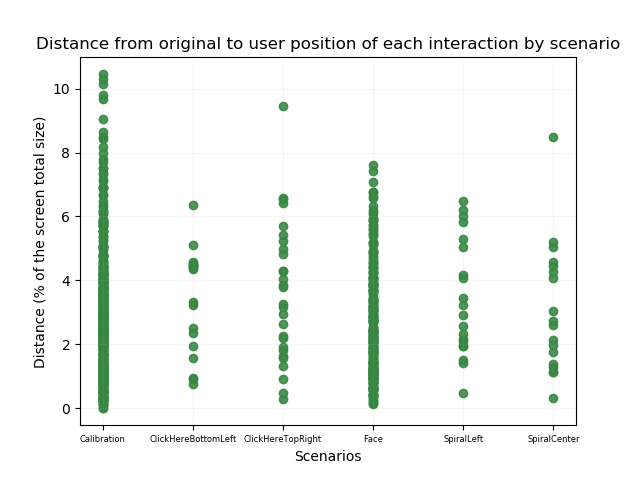
*RenanOm92/factorsAttractionFirebase*. (2019, 03 30). Retrieved 03 30, 2019, from GitHub: https://github.com/RenanOm92/factorsAttractionFirebase

*The MIT License*. (n.d.). Retrieved 04 16, 2019, from Open Source Initiative: https://opensource.org/licenses/MIT

Appendix



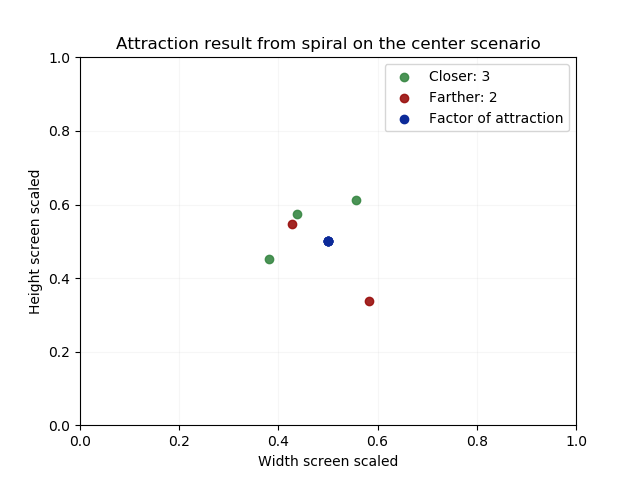
*Figure 20* Appendix - Boxplot graph with the distance from the original position to the user position, threshold of 25% from the attractor.



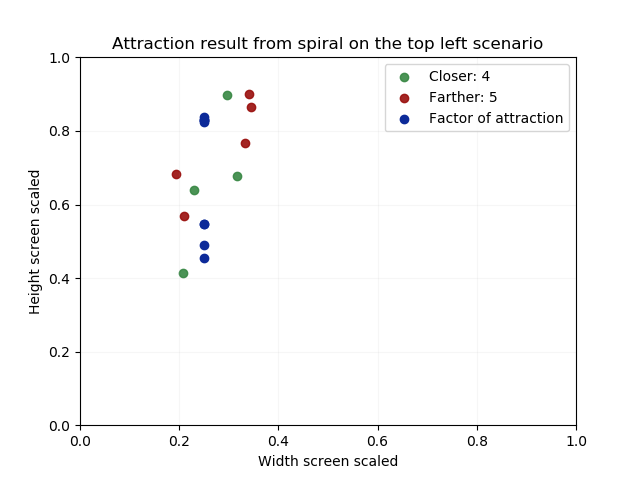
*Figure 21* Appendix - Scatter graph with the distance from the original position to user position per interaction, threshold of 25% from the attractor.



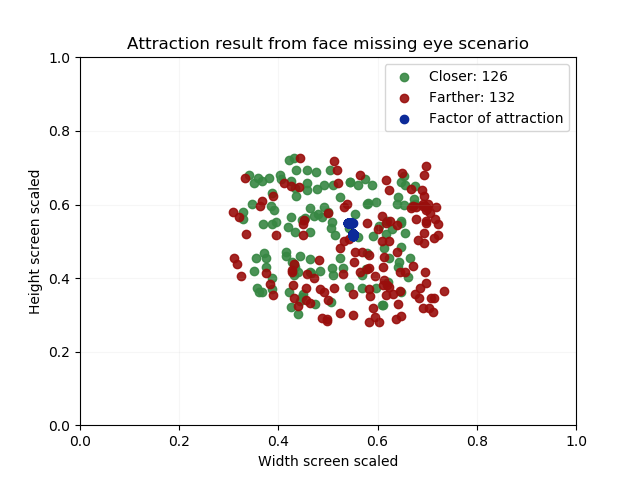
*Figure 22* Appendix - Scatter plot for the click here button on the bottom left scenario, using a threshold of 15%.



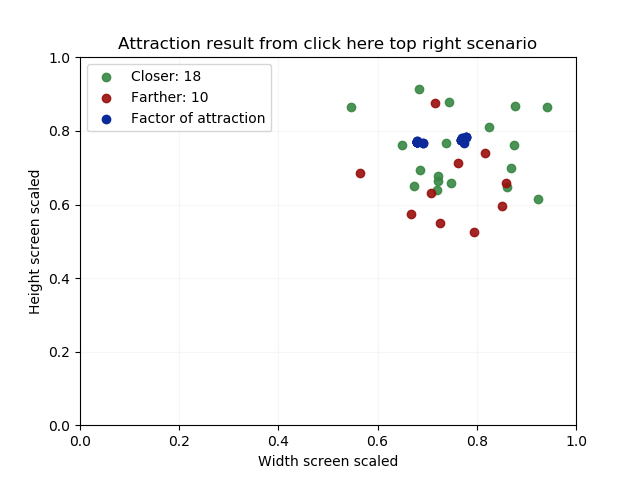
*Figure 23* Appendix - Scatter plot for the spiral center scenario, using a threshold of 15%.



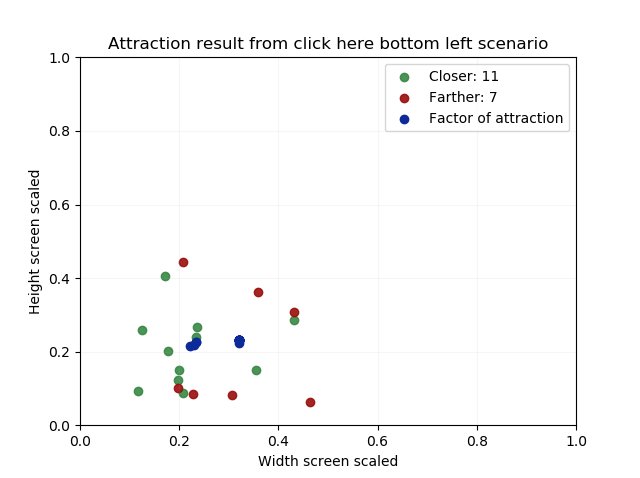
*Figure 24* Appendix - Scatter plot for the spiral left scenario, using a threshold of 15%.



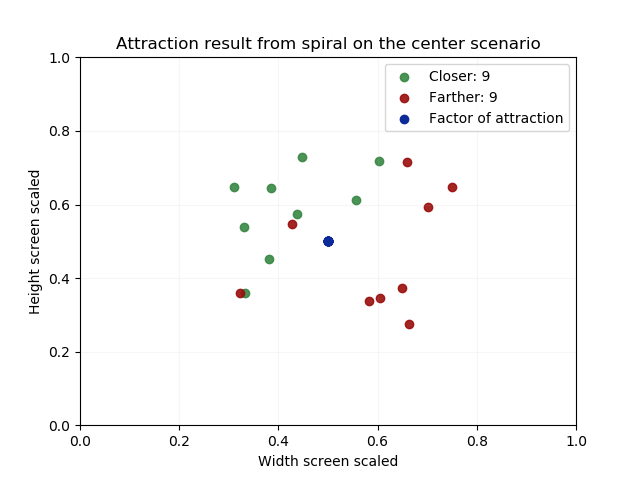
*Figure 25* Appendix - Scatter plot for face scenario, using a threshold of 25%.



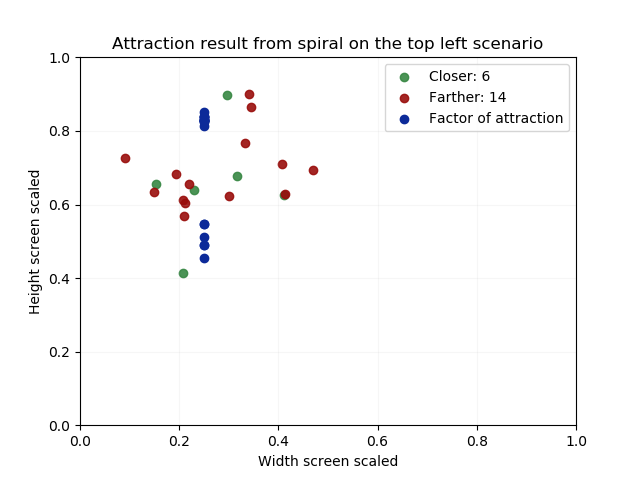
*Figure 26* Appendix - Scatter plot for click here top right scenario, using a threshold of 25%.



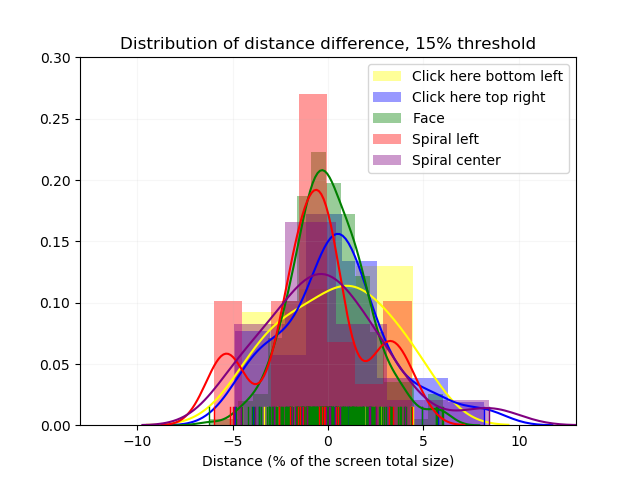
*Figure 27* Appendix - Scatter plot for click here bottom left scenario, using a threshold of 25%.



*Figure 28* Appendix - Scatter plot for spiral center scenario, using a threshold of 25%.



*Figure 29* Appendix - Scatter plot for spiral left scenario, using a threshold of 25%.



*Figure 30* Appendix *-* Distribution of the difference of DO and DU for the face scenario with 25% threshold.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1.16 | 0 | 0 | 1.16 | 0 | 1.16 | 2.32 | 0 | 1.16 | 2.32 |
| 0 | 1.16 | 0.58 | 2.9 | 0 | 1.16 | 0 | 0 | 0 | 0 | 0 | 1.16 | 0 | 0 |
| 0 | 1.74 | 0.58 | 0 | 0 | 0.96 | 0 | 0 | 0 | 2.32 | 1.74 | 0.58 | 0.58 | 0 |
| 1.16 | 0 | 1.16 | 1.16 | 1.16 | 0 | 0.19 | 3.48 | 0 | 2.13 | 2.71 | 0 | 2.32 | 0 |
| 0 | 1.74 | 0 | 0.58 | 0 | 0 | 0 | 0 | 0 | 2.71 | 0.58 | 0.58 | 0.58 | 0.58 |
| 0 | 0 | 0 | 0 | 2.32 | 0.58 | 1.16 | 0 | 0 | 0 | 0.58 | 0.58 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0.58 | 0 | 0 | 1.16 | 0 | 0 | 1.16 | 0 | 0 | 0 |
| 0 | 0 | 2.32 | 2.32 | 1.16 | 0 | 1.16 | 0 | 1.16 | 1.16 | 0 | 0.38 | 0 | 0 |
| 0 | 0 | 0 | 0 | 2.32 | 1.16 | 0 | 0 | 1.16 | 0 | 1.16 | 1.16 | 1.93 | 1.16 |
| 0 | 1.16 | 0 | 0 | 1.16 | 1.16 | 0 | 0 | 0.29 | 2.61 | 0 | 0 | 1.16 | 0.58 |
| 0 | 0 | .9 | 0.58 | 1.16 | 0 | 0 | 0 | 1.55 | 1.16 | 0.58 | 0 | 0 | 0.58 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.16 | 0.38 | 0.38 | 1.16 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1.16 | 0 | 1.16 | 0 | 0 | 1.16 | 1.16 | 0 |

*Figure 31* Appendix - Representation of screen/matrix for the button top right scenario. The highest value is highlighted in red.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0.25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0.25 | 0 | 1.28 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 1.88 | 1.56 | 0.23 | 0 | 1.56 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0.96 | 0.96 | 2.09 | 2.34 | 2.08 | 2.37 | 1.10 | 2.93 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0.32 | 3.67 | 1.70 | 2.96 | 1.49 | 3.05 | 2.46 | 0.11 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1.75 | 3.23 | 2.75 | 3.33 | 3.46 | 2.17 | 0.51 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1.28 | 1.98 | 1.79 | 1.72 | 2.78 | 2.44 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0.42 | 2.63 | 3.05 | 2.8 | 2.81 | 2.31 | 0.25 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 3.2 | 2.57 | 4.07 | 1.47 | 3.36 | 2.58 | 0.21 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0.36 | 0 | 0.87 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

*Figure 32* Appendix - Representation of screen/matrix for face scenario. The highest value is highlighted in red.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.81 | 0 |
| 0 | 0 | 0 | 3.63 | 1.81 | 0 | 0.9 | 0 | 0 | 0 | 0 | 1.81 | 0 | 0 |
| 0 | 0 | 0 | 0 | 3.63 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.81 |
| 0 | 2.42 | 1.81 | 0 | 1.81 | 0 | 0 | 0 | 0 | 2.72 | 1.81 | 1.81 | 0.9 | 0 |
| 0 | 0 | 2.72 | 0.45 | 0.9 | 0.9 | 0.9 | 0 | 0 | 2.72 | 0.9 | 1.81 | 0.9 | 0 |
| 0 | 0 | 1.81 | 0.45 | 0.9 | 1.21 | 0 | 0 | 0 | 0.6 | 1.21 | 1.81 | 0 | 0 |
| 0 | 0 | 1.81 | 0 | 1.81 | 0 | 0 | 0.6 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 3.63 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1.81 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6 | 0.9 | 1.81 |
| 0 | 0 | 0.6 | 1.81 | 0 | 1.81 | 0 | 1.81 | 0 | 0 | 0.9 | 1.81 | 1.81 | 0 |
| 0 | 0.9 | 0 | 0 | 0 | 0.9 | 0 | 0 | 0 | 1.81 | 1.81 | 0 | 1.81 | 0 |
| 0 | 0.9 | 1.81 | 0 | 1.81 | 0.9 | 0 | 0 | 0 | 1.81 | 1.81 | 0 | 0 | 0 |
| 0 | 1.81 | 0 | 3.63 | 0 | 1.21 | 0.6 | 0 | 0 | 0 | 0 | 1.81 | 1.36 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.45 | 0 |

*Figure 33* Appendix - Representation of screen/matrix for the spiral top left scenario. The highest value is highlighted in red.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.4 | 0 | 0 | 0 | 0 |
| 0.7 | 0.7 | 1.4 | 0 | 0 | 0 | 0 | 0 | 1.4 | 0 | 2.81 | 0 | 2.34 | 0 |
| 0 | 0 | 1.4 | 2.11 | 0 | 0 | 1.4 | 1.4 | 0 | 1.4 | 0 | 0 | 0.46 | 0 |
| 0 | 1.4 | 0.7 | 2.11 | 2.11 | 0 | 2.87 | 0 | 1.4 | 1.4 | 1.4 | 1.4 | 1.4 | 0 |
| 0 | 0 | 2.34 | 0 | 1.40 | 0.7 | 0 | 0 | 0 | 1.4 | 1.4 | 2.11 | 1.4 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0.7 | 1.4 | 0 | 0.7 | 0 | 0 | 0 | 1.4 |
| 0 | 0 | 0 | 0.46 | 1.4 | 1.4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 1.4 | 0 | 0 | 0.46 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1.4 | 0 | 0 | 1.4 | 0 | 0 | 0 | 0 | 0.46 | 0 | 2.81 | 0 | 0 |
| 0 | 0 | 0 | 3.52 | 0 | 0 | 0 | 0 | 1.4 | 0 | 0 | 0.7 | 1.4 | 0 |
| 0 | 0 | 0 | 0.7 | 0 | 1.4 | 0 | 0 | 0 | 2.34 | 1.64 | 1.40 | 0 | 0 |
| 0 | 0.7 | 0 | 3.52 | 0.7 | 0 | 0 | 0 | 0 | 0.46 | 0.46 | 1.87 | 0.93 | 0 |
| 0.70 | 0 | 0 | 0 | 1.4 | 1.4 | 0 | 1.4 | 2.11 | 0 | 0 | 0.46 | 1.4 | 0 |
| 0 | 0 | 0 | 0 | 2.81 | 0 | 0 | 0 | 1.4 | 1.4 | 0.7 | 0 | 0 | 0 |

*Figure* 34 Appendix - Representation of screen/matrix for the spiral center scenario. The highest value is highlighted in red.