Doctoral Dissertation
Doctoral Program in Computer and Control Engineering (34^{th}cycle)

Saliency Prediction in the Data Visualization Design Process

By

Luisa Fernanda Barrera León

******

Supervisor(s):
Prof. Fulvio Corno,

Doctoral Examination Committee:
Prof. Leonardo Flórez Valencia, Referee, Pontificia Universidad Javeriana, Colombia.
Prof. Dan Witzner Hansen, Referee, IT University of Copenhagen, Denmark.
Prof. Ana M. Bernardos, Universidad Politécnica de Madrid, Spain
Prof. Cristina Gena, Università di Torino, Italy
Prof. Andrea Sanna, Politecnico di Torino, Italy

Politecnico di Torino
2022

Declaration

I hereby declare that, the contents and organization of this dissertation constitute my own original work and do not compromise in any way the rights of third parties, including those relating to the security of personal data.

Luisa Fernanda Barrera León
2022

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino (ScuDo).

To Me

One woman was at the beginning, and another woman at the end.

Acknowledgements

First of all, I would like to thank the universe for allowing me to come to Turin and to be able to complete the last step of a long academic career. I thank my tutor Fulvio Corno, who, without knowing me, opened the doors of his research group and guided me (almost pushing me) during the whole process. I also thank Luigi de Russis, whose criticism helped me in every aspect.

The most special thanks to my family because nothing would be possible without them. To my mom, the light in my journey, who did not let me faint, who endured with me all the dramas and has always had more confidence in me than I had in myself. To my dad, who every day was looking out for me and encouraged me to continue. To my sisters and niece, especially Carolina, who put all their trust in me and ensured that I did not give up. To my friends, Cesar Diaz and Carlos Perez, who made my days at Poli fun and made me feel at home; also for teaching me so many things and for being with me every step of the way. To my other friends, Carolina Cardenas, Gustavo Ramirez, David Ribero, Andres Corredor, and Gabriela Zabaleta, thank you for believing in me and giving me your support, not only in the academic field but also during life's bad and good moments.

I want to thank Roberta Bardini, who was the first person to talk to me in the lab. She understood me, and was patient when I didn't know any Italian. Her beautiful personality bolstered my confidence in the "Italian world ." I want to thank Pedro Molina for all his love, reminding me to see the positive things in everything, and for looking after me in difficult moments. Finally, to Valentina Lopez, who gave me the final boost, and without whom literally this thesis would not have been written, thank you for your time and faith.

Abstract

Saliency models are computational algorithms that predict the degree of attention each visual element has for a human observer. Several studies have used these models to create patterns, standards, and object recognition algorithms. Nevertheless, according to our research, saliency has been an instrument that has always been implicitly exploited in design standards or commonly used patterns, mostly invisibly from the end-users (and graph creators) awareness.

An example of hidden saliency uses can be seen in Data Visualization (DataViz). DataViz applications are already coming with well-established patterns, such as color palettes, that have been extensively studied using saliency models. This means that DataViz designers implicitly use the benefits of the saliency model predictions when selecting a preset pattern. Nevertheless, they are unaware that those patterns have a specific visual attention impact on the final observer.

The main objective of this thesis is to bring the information provided by the saliency models closer to the graph designer in a DataViz design process. The idea is to make explicit the potential impact of each design decision in the graph design process and allow the graph designers to exploit this information in their work. To that end, the first part of the thesis is a more profound study of saliency prediction models theory, their currently practical uses in InfoVis, and the functionality and performance of the InfoVis saliency models. Due to the novelty of the InfoVis saliency prediction models, we performed experiments on how these saliency models behaved on specific statistical graphs to validate their performance.

Based on the performance experiments, we choose a saliency model to integrate it into the graph design process. To achieve this integration, we propose two approaches described in the second part of the thesis. In the first one (*Design Tool*), we explored a mechanism for assisting graph designers in attracting the observer's attention to specific relevant data they can choose, specify at design time. In the

second (*Measurement Tool*), we integrate the saliency model into a common DataViz application as a validation tool at the end of the design process. Therefore, the designer can visualize the visual attention implications of each design decision and iterate the visual elements in order to improve the result.

Six experts from academia and industry evaluated the developed approaches. In general, most results demonstrated that integrating saliency prediction models into the DataViz design process is a relevant and valuable technique. Notably, the experts mostly expressed that they had not seen this type of support in other tools, which means they have significant potential.

The presented thesis opens the possibility of linking two relevant areas of visualization, such as the study of salience and data visualization, not only to create new visualization techniques but also to bring the knowledge of both areas closer to the fewer experts. Finally, integrating these two areas, whose objectives intersect in decision-making support, could become a vital instrument to improve the final observer decision-making process.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms / Abbreviations

DataViz Data Visualization

InfoVis Information Visualization

Chapter 1

Introduction

1.1 Context

"Saliency detection is a principle mechanism to facilitate visual attention. A good visualization guides the observer's attention to the relevant aspects of the representation." by Jänicke and Chen [2].

Visual Saliency Prediction is a large field of study that combines cognitive-visual research with computer techniques. This field works on simulating how a human inspects an image and how this information is processed in the brain from a visual-cognitive point of view. That simulation is made by computer models called *saliency prediction models*. In particular, those models determine where a human observer will focus their attention. The saliency prediction models have been developed computationally by employing traditional algorithmic models and more recent deep learning technologies. Saliency prediction has been widely applied to object detection in images and videos, image re-targeting, virtual reality, augmented reality, and autonomous robots [3–5]. It is mainly utilized in laboratories for the development of new patterns, standards, and algorithms. Nevertheless, according to our findings, saliency prediction uses are hidden from end-user.

An example of that hidden saliency prediction uses can be seen in the Data Visualization (DataViz) process. Data visualization is the graphical representation of information and data, which aims to create visual artifacts that facilitate their analysis for a human observer. In the process of designing data visualizations, many authors start with the theory of preattentiveness (first stage of attention) [6–8], which is the foundation for Saliency Prediction.

In order to generate *effective* visualizations, these authors emphasize the significance of understanding the visual-cognitive elements of human attention. In DataViz means that the observer can quickly spot patterns that aid in decision-making. However, all graphical designers should have a thorough understanding of the aspects that must be considered in order to control attention in a representation correctly. As a result, many visualization design frameworks already include some notions in their foundation designs. For example, established color palettes have been thoroughly investigated using saliency prediction to determine their impact on viewer attention as well as each of the visualization techniques, which have also been validated with saliency methods to know their visual-cognitive impact.

Altogether, visual saliency prediction is an important area of knowledge for data visualization design, although it is hidden behind design patterns and cannot be understood by all graph designers. This is where our research begins, in the quest to make explicit the use of saliency prediction within the DataViz design process, making it more visible and valuable to the graph designer.

1.2 Motivation

> *"Design graphic representations of data by taking into account human sensory capabilities in such a way that important data elements and data patterns can be quickly perceived."* by Colin Ware [6]

DataViz is nowadays a fundamental part of decision-making in different contexts. DataViz's growth has made many types of profiles responsible for making graphs, ranging from designers, statisticians, and engineers to doctors. Some of the mentioned profiles do not necessarily have basic design knowledge, such as the visual impact implied by the usage of one or other color, or how the orientation of a graph can change the focus of data attention. The *graph designers* have to consider the features mentioned previously (e.g., color, orientation, spacial position) to obtain a graph that correctly represents their data set. Along with this, graph designers must adequately select the combination of those visual features to drive the observers' attention to the relevant data (see Fig. 1.1).

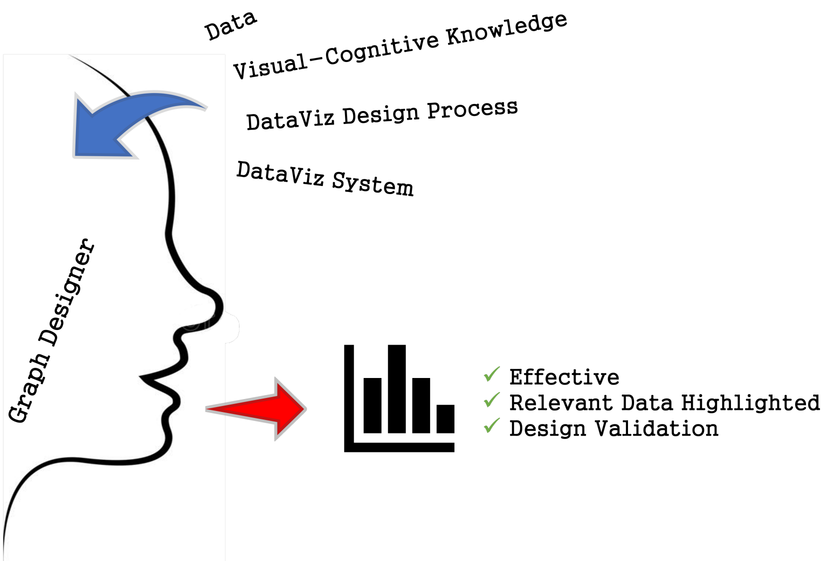


Fig. 1.1 Thesis Motivation

The design element selection process mentioned above is supported by frameworks such as Tableau, Power BI, or Microsoft Excel. Those frameworks offer pre-set layouts that guide the designer in selecting the visualization technique (e.g., bar charts, scatter plots, maps), the color palette, or the orientation. In most cases, those pre-established design elements have been previously studied, and their visual-cognitive impact is well known by experts [9]. The research literature concerning the techniques to generate and validate such design patterns employed by the visualization frameworks is voluminous. Unfortunately, much of this information is kept in highly specialized publications and is frequently written in a way that only the research scientist understand [10]. Access to this information could be very beneficial in supporting the graph designer in creating better visualizations in terms of making conscious design decisions based on a known visual impact of each of them.

Finally, it is also necessary to consider the relevance of the visual element's impact on the observer's cognitive process. According to Milutinovic *et al.* [11], in data visualization, the attractiveness effect (focus attention) impacts decision-making. The decision-maker can be influenced by the information presented, which can influence her mental image of the situation as well as her attributes, such as engagement and task knowledge. Milosavljevic *et al.* [12] found similar results, demonstrating that, when making quick decisions, visual saliency affects choices more than observer predilections. According to these studies, the observer's decisions highly depend on the attention areas in a graph, although they are not the only ones.

> *"Improving cognitive systems often means optimizing the search for data and making it easier to see important patterns."* by Colin Ware [6]

Altogether, it would be possible to support the designer in the design process by making explicit the visual impact that her design decisions could have on the observer through saliency prediction mechanisms. This thesis aims to integrate visual-cognitive concepts into the information visualization design process. We intend to bring those concepts to the graph designer's context and provide insight into how her design choices might affect the observer's perception. To achieve this, we established two main research aims:

1. Perform deeper research about the concept of salience prediction and its usages in Information Visualization (InfoVis).
2. Develop exploratory approaches to bridge the gap between the information provided by saliency prediction and the graph designer into a DataViz design process.

It is important to clarify that we make a difference between the term InfoVis and DataViz. InfoVis corresponds to all types of information representation (infographics, dashboards, statistical graphs). The term DataViz represents only the subgroup of representations corresponding to Statistical Graphs (bar charts, line charts, scatter plots, maps).

1.3 Main Thesis Contributions

The study of saliency prediction has been since 1980 [13] widely covered, but in areas that are often different from InfoVis. For this reason, a significant part of this thesis is focused on a broad study of saliency prediction (see Fig. 1.2). We started with saliency fundamentals research, then explored its use in the areas of Information Visualization (see Fig. 1.2.I), and finally, developed an in-depth analysis of the few models explicitly created for InfoVis images (see Fig. 1.2.II). Finally, in

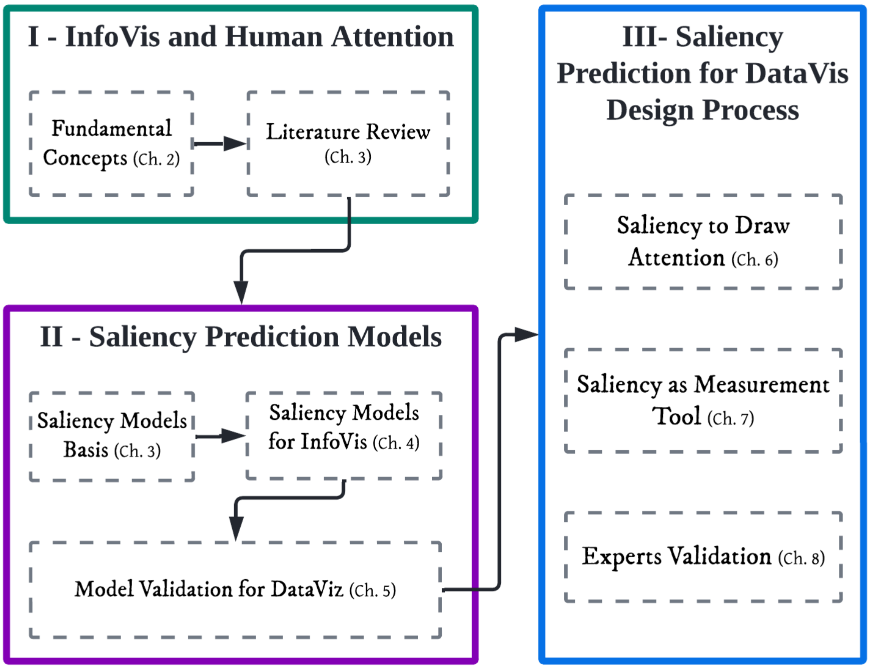


Fig. 1.2 Thesis contributions and thesis organization

order to achieve our goal of incorporating the saliency prediction benefits into the visualization design process, we developed two approaches, and they were validated by the academy and industry experts (see Fig. 1.2.III).

As results of this process, we generated four primary contributions:

Preattention Process in InfoVis Review. We conducted a literature review in order to understand more about how the visual attention process is employed pragmatically in InfoVis and if these uses are connected directly to the graph designer. However, as human visual attention is a broad knowledge area, we selected as the focal point the *Preattentive Process*. This process is the first stage in the human visual-cognitive mechanism and is the primary source of information to establish where the attention should be focused. Based on this process, the primary purpose of the literature review was to determine *how the preattentive visual process is used in information visualization*, with the focus of improving the observer's cognitive process in graph comprehension. Consequently, the literature review results revealed a classification of the preattentive process used for the InfoVis design process: as a design and measurement tool (see Chapter 3).

InfoVis Saliency Models Validation. Based on the results obtained from the literature review, we decided that it was important to carry out further research on saliency models and their development in the InfoVis area. Since this is a relatively new area, the oldest model was developed in 2017, and only two saliency models were found for DataViz. However, other classical models were considered during the investigation due to their previously validated performance with InfoVis images. We decided to reduce our scope to validate these models only with DataViz images (statistical graphs) because we wanted to validate the performance of these models, and doing so for all types of visualization would be more involved. These models were validated with eye-tracking collected data and three established metrics to evaluate their performance. Only one of the validated models showed a consistently high performance in the different validations conducted. This model was used as input for the development of two approaches in which saliency could be used in the two modalities found in the literature review (see Chapters 4 and 5).

Saliency Prediction as a Design Tool. Based on the studies mentioned above, we realized that we could use saliency prediction during different stages of graph design. The first option was to use it in the first design steps when the designer intends to make the data that she considers most relevant to stand out from the others. This can be done by helping the graph designer find the combination of visual elements (color, orientation) that will make the most relevant data more noticeable to the observer. In order to validate this approach, we developed a tool where the designer chooses the data to be highlighted, and the algorithm generates some visual elements combinations that achieve this goal. We utilized a saliency prediction model to validate where the attention is and select the one that has more saliency on the relevant data. This development has a Matlab and Python version (see Chapter 6).

Saliency Prediction as Measuring Tool. The second step in the graph design process where we notice saliency prediction can be included is at the end of the process. Saliency prediction can be employed as a measurement tool to validate how each design decision can impact the attention points. We integrated a saliency prediction model into a common data visualization framework to validate our approach. This development is entirely web-based, and the designer only has to add the data, select the graph of choice, make changes to the graph layout, and finally check if their design decisions would lead the observer to the area or data they desire.

1.4 Document Structure

Based on the research process presented in Fig. 1.2, the document is structured as follows:

Chapter 2: presents the fundamental definitions of Human Visual Attention and InfoVis. During the thesis, we found that some concepts related to the human vision process have been defined in different ways and addressed in various fields. That is why we consider it essential to clarify what concepts such as attention, preattentive, and salience mean for this research. Something similar happened with the concepts of InfoVis and DataViz, whose definitions are also described in this section. Finally, we present how both fields' concepts are connected from the visual-cognitive point of view.

Chapter 3: reports the process of literary analysis and its results. The literature review mainly aims at identifying how the preattentive visual process is used in information visualization, oriented to improve the observer's cognitive process in graph comprehension. This section presents the search methodology, the relevant findings, and a discussion about the development opportunities that emerged from the findings.

Chapter 4: introduces an overview of saliency prediction models, their development techniques, and the image dataset used for their creation and validation. Based on the literature review's findings, we established that it was essential to go deep into saliency prediction. In addition, this chapter aims to map the scenario of existing saliency prediction models for InfoVis images. During the study, we found that some authors established that classical saliency models had to be modified to the InfoVis images characteristics (e.g., color scale). For this reason, in this chapter, we present a detailed description of two saliency models created explicitly for InfoVis images together with an extensively used classical model.

Chapter 5: presents the results of the InfoVis saliency models validation. We perform a structured validation to select one of the models described in the previous chapter. To perform this validation, we collected eye-tracking data from the DataViz images data set created by us, and also we used some saliency standard metrics. In this chapter, we describe the validation process and the results that led us to choose one of the models and also reveal some insights into the behavior of saliency in clean graphs.

Chapter 6: We proposed an exploratory approach to systematically draw the observer's attention to relevant data by changing some visual elements. In this approach, we use the saliency prediction model (described in the previous chapter) to identify which combination of visual elements would most effectively highlight a piece of data chosen by the graphic designer. The purpose of this first development was to make convergence between the information given by a saliency model and the aim of emphasizing relevant data provided by the graph designer. This chapter describes the motivation, technique, and results of two development tools, one in Matlab and the other in Python.

Chapter 7: presents a second developed approach, in which the saliency prediction model is used as a measurement tool. We integrated saliency prediction into a traditional data visualization system to explore this second possible use of saliency as a measuring tool. This implementation seeks to allow the graph designer to estimate how each of her design decisions will affect the observer's attention. With this tool, it is possible to observe how each design decision made by the graphic designer can change the attention focus of the graph. Since it is implemented on a complete graph design framework, this tool works on all types of graphs. This chapter describes the motivation, the integration process, and the tool's results mentioned above.

Chapter 8: presents the protocol and results of a proof-of-concept validation. We conducted verifications with experts in data visualization from distinct viewpoints, one from academia and the other from industry, in order to validate the developments discussed in chapters 6 and 7. The primary goal was to confirm with Data Visualization experts that using Saliency Prediction in the DataViz design process was feasible.

Chapter 9: presents the thesis main conclusions.

Chapter 2

Key Concepts

This chapter aims at defining the main concepts in the areas relevant to this thesis: the Human Visualization Process and Information Visualization. The concepts presented in this section are described from a computer science perspective. Regarding Human Visualization, concepts were considered from the side of how they are simulated and modeled in the computational domain. Concerning InfoVis concepts, the concepts were defined from a Human-Computing Interaction (HCI) approach. During the research, we noticed that many authors had different interpretations of the concepts related to human vision and InfoVis, specifically about visual attention and the design process, respectively. Due to this interpretation's diversity, we decided to adopt the definitions established in this section for the whole thesis. Another factor considered in establishing the definitions was that the focus should always be on the observer's cognitive process impact. For instance, regarding the human vision concepts, we mainly focused on *attention*, the first step in the cognitive process. On the other hand, for InfoVis, we understand the cognitive process as the graph *reading and understanding process*.

2.1 Human Visual Attention

According to Ware [6], humans *"acquire more information through vision than through all of the other senses combined."*. In biological terms, about 20 billion brain neurons are dedicated to analyzing visual information. This process is a complex mechanism whose final aim is to find patterns, the crucial components of human cognitive activity [6]. This section describes those concepts related to the human attention process, a fundamental part of understanding InfoVis images.

Visual Attention. *Visual attention is the umbrella term used to denote the various mechanisms that help determine which regions of an image are selected for more detailed analysis* [14]. The term attention refers to the process that allows one to focus on some stimuli at the expense of others. According to [6], the entire attention process (in human vision) is being permanently "*tuned from top to bottom based on mental predictions and on what will be most useful to us*". This is the concept we focus on most in this thesis with respect to attention, *"what will be most useful"*.

Furthermore, in information visualization, attention is connected to the human cognitive process of how the observer understands a graph [15, 16], which means which information in the graph is most useful to the observer. Specifically, in cognition, *the visual working memory holds the visual objects of immediate attention* [6]. For this reason, attention is considered the earliest stage in the cognitive process. Related terms: visual perception.

Preattentive. According to [17], in the human visual process *a visual scene is analyzed at an early stage by specialized populations of receptors that respond selectively to such properties as orientation, color, spatial frequency, or movement and map these properties in different areas of the brain*. The term *"Preattentive"* includes all factors influencing this selection mechanism: the process (how it works) and the attributes (which visual elements influence the process).

Preattentive Attributes or Features. They are straightforward visual elements perceived without conscious attention. According to the definitions found in [18] and [19], *Preattentive Attributes* are classified in four groups: *Form*, which bundles line orientation, length, width, collinearity, size, curvature, spatial grouping, blur and numerousness; *Color*, including hue and intensity; *Motion* including flicker and direction; *Spacial position* made up of 2D position, (stereoscopic) depth, depth or convex/concave shape from shading. The authors also use the "visual element" concept in the literature to refer to preattentive attributes.

It must also be noted that not all preattentive attributes have the same impact on the preattentive process. Generally, the most marked impact is based on color, orientation, size, contrast, and motion or blinking, corresponding to the findings of neuropsychology [6]. On the other hand, in this thesis context, visual elements represent data and attract attention to important information.

Preattentive Process. Humans can simultaneously perceive a large number of visual attributes (e.g., color, orientation, shape) to direct their visual attention. Preattentive perception is done in parallel: each visual attribute is computed in parallel and then combined to select specific regions that are perceived without any conscious effort [20]. The preattentive process is fast (200 to 250 ms) and unconscious, in contrast to the attention process, which is done serially and is slower and conscious. In addition, this process decides what visual attributes are *offered up to our attention and easy to find in the next fixation* [21]

The preattentive process can be represented algorithmically by two types of models: *Bottom-up* (or *stimulus-driven*, or *global*) and *Top-down* (or *goal-directed*, or *local*). These models simulate the mechanism used to detect the salient visual subsets in the human vision system [8]. Currently, several computational models can emulate these preattentive models and combine them into a single model (see Attention definition) [6].

Saliency. The Saliency of an item —be it is an object, a person, a pixel, etc.— measures how easy it is visually identified and arises from its contrast and separation relative to other objects or the background [8]. Thus, saliency detection is considered a critical attentional mechanism that guides visual attention.

Formally, the *Perceptual Saliency* is the degree to which a target stimulus "pops out" in a set of stimuli [22]. Thus, for example, if the target stimulus differs by a single attribute (e.g., color) from the other objects, it is more salient; meanwhile, if it differs by a combination of attributes (e.g., color and form), it is less salient. Alternatively, the *Visual Saliency* may be defined as the nature or quality of a viewed object which gives it relevance or importance to the observer [23]. Related terms: visual attraction effect, focus attention.

Feature Map. contains the information of the features (preattentive attributes) extracted from every part of the visual field simultaneously. In the "Model of

Visual Information Processing," (see Fig. 2.1) proposed by [6], a set of feature maps is the result of the first stage. These feature maps retain implicit data about their spatial origins through their links back to visual stimuli. Then, these feature maps will be combined to perceive the whole object.

Saliency Map. *represents the conspicuity at every location in the visual field by a scalar quantity and to guide the selection of attended locations, based on the spatial distribution of saliency* [24]. A saliency map combines all *feature maps*. It is a visual representation that highlights the image regions on which the observer's gaze focuses the attention [25]. Commonly, it is represented with a warm color palette, in which red is in the area with more attention probability. The main objective of a saliency map is to evidence the degree of importance of each pixel in an image to the human visual system.

Heat Map. as saliency maps, is a visual representation of the human observer focus in an image. For this thesis, we called the heat map to the result of an eye-tracking experiment. This means it is an attentional map generated from the collected eye-tracker data.

Eye Tracking. *is a sensor technology that can detect a person's presence and follow what they are looking at in real-time* [26]. This technology records observers' eye movements (gazes) and transforms them into data streams. Those data streams include pupil position, fixation path, gaze vector for each eye, and gaze point.

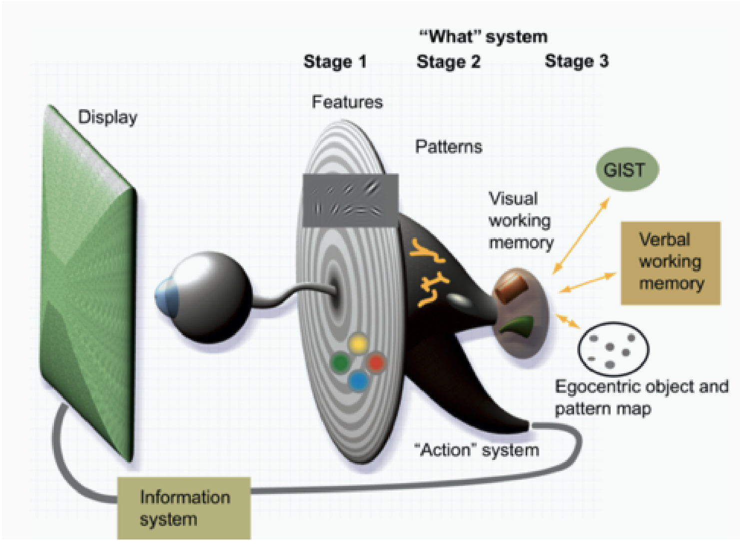


Fig. 2.1 "*A three-stage model of visual information processing.*" taken from Ware *et al.* [6]

2.2 Information Visualization

In the previous chapter, we covered the process of the human vision process, focused on the preattention stage, wherein concepts related to this early stage were described from a computing perspective. This section covers the concepts related to information visualization oriented to design and cognitive characteristics.

Information Visualization (InfoVis). *is the process of representing data in a visual and meaningful way so that a user can better understand it* [27]. Visualization provides an ability to comprehend huge amounts of data [19]. Also, Information Visualization integrates all types of data representation such as infographics, dashboards, word clouds, statistical charts (scatterplots, bar charts, or line charts), etc.

Data Visualization. (DataVis): is the graphical representation of information and data, *highlighting patterns and trends in data and helping the reader to achieve quick insights* [28]. In the bibliography, Data Visualization and Information Visualization meaning is nearly homonyms. For this thesis, we use Data Visualization to reference all statistical data visualization techniques such as bar charts, scatterplots, and line charts. We explicitly leave out of this group visualization techniques such as infographics or dashboards, that would be part of a broader concept (InfoVis for this research).

Graph Visual Elements. are all the graphical elements that compose a graph. For example, those elements could be textual (title, axes title, legends, axes labels) or forms (bars, points, lines, circles).

Relevant Data (information). is defined as a selected group of data that could be chosen by users or by an algorithm [29, 30], and represent data that is important to the object of study. The data relevance could be defined by an algorithm (e.g., filters) or by the graph designer. According to Ware in [6], *important data should be represented by graphical elements that are more visually distinct than those representing less important information.* In DataViz, in most cases, relevant data is highlighted with different preattentive attributes (e.g., color, position, size).

Highlighting. *The goal of highlighting is to make essential data points more visually prominent* [31]. Highlighting in InfoVis is used to guide the user in exploring the data, for instance, through visual elements to help the observer focus on relevant data or link data across multiple views [31]. According to Ware [6], highlighting is considered a computer-side operation that helps to find relevant patterns. Also, highlighting may be used to ensure the comprehension of charts with excessive information density [19].

Another important key concept for our research is the *visualization development stages* in InfoVis. According to Mazza *et al.* [7], the design process stages in InfoVis are:

Preprocessing and data transformations. In this stage, data is extracted from a data source and transformed into a structured format that a visualization application can use.

Visual mapping. In this stage, the graph designer defines which visual structures are used to map the data and their location in the display area. Three basic structures should be defined in this visual mapping stage: *spatial substrate*, the axes selection which defines the physical data representation space; *graphical elements* are everything visible that appears in the space (e.g., points, lines);

and *graphical properties* are the graphical attributes which the human vision is very sensitive (preattentive attributes).

Views creation. This is the last stage. The views are the final results of mapping data structures with visual attributes in this stage. Also, this stage has a refined process based on data design objectives (e.g., explore, communicate).

The design process described above is used for usual visualizations not related to the research of the new techniques developed or validations. The term "usual" is intended as a DataViz design for analyzing particular data, to find patterns, statistical behaviors, markers trends or important nodes, etc. For this thesis, when we mention *"early design stage"* we are referring to the *Visual Mapping stage* explained previously. It represents an initial design state because the graph designer has to make the design decisions, and the graph has not performed any validation.

2.3 Cognition, Human Visual Attention, and InfoVis

According to Rodrigues Jr. *et al.* [32], InfoVis *"provides faster and user-friendlier mechanisms for data analysis because the user draws on his/her comprehension immediately as graphical information comes up to his/her vision"*. This means the human vision system is strongly related to cognitive and memory processes that take action in data comprehension [33] and interpretation [14]. Rodrigues Jr. *et al.* [32] said that carrying out conscious management of the human vision system, mainly the preattentive process, should be the first stage towards the creation of a graph.

These human attention concepts are extensive. However, they are extremely useful for helping in the InfoVis design process by supporting graph creators to archive their data objectives and guiding the observer in its cognitive process. Design data visualizations by considering human perception capabilities lead to more accessible perceiving data elements and patterns for the observer [6].

According to [6], to design graphic representations of data, human sensory capabilities (visual perception and cognition) must be taken into account in such a way that relevant data elements and data patterns can be quickly perceived and processed (by visual working memory). According to Toker *et al.* [34], *"various cognitive abilities such as perceptual speed, verbal working memory, and visual* *working memory have been shown to impact user performance and/or user subjective experience with Infoviz task."* Also, Healy *et al*. [14] explain that the human visual system is strongly related to cognitive and memory processes that take action in data comprehension and interpretation.

In addition, the intended use of the concepts discussed in this chapter infers cognitive productivity. The measure of cognitive productivity *is the amount of valuable cognitive work done per unit of time* [6]. Improving cognitive productivity means that the observer can search for data and find relevant patterns in less time. In InfoVis, an eye-tracking device is common to get insights into observers' cognitive processes when reading the graph and to solve a definite task [35]. With this measuring technique, InfoVis creators can know which is the focus of attention of the graph and the performance in its comprehension.

Finally, it is essential to notice that the visual saliency impacts the InfoVis images comprehension and, therefore, decision making as well [11]. According to research conducted by Mulutinović *et al*. [11], the visual attraction effects (saliency) influence the decision-making in specific scenarios. In some cases, the decision-makers focus their attention on information that is arbitrary to their decision goals. The influence of saliency even over decision-maker preferences was demonstrated in [12].

In the next chapter, we present the results of a literature review to deepen how attention process knowledge practically influences the InfoVis design process.

Chapter 3

Literature Review

The previous chapter has presented some concepts related to human visual attention and its relationship with InfoVis. We noticed a voluminous theory in both areas, but we wanted to delve deeper and study how these theories have been integrated *practically*. For this research, *practically* means how human visual attention is used for the benefit of graph designers in InfoVis. Due to the extended concepts in the human attention process, we decided to focus our research on the *preattentive process*. As previously mentioned, the preattentive process is the first stage in the attention process and also impacts the InfoVis images comprehension [32, 6, 36, 11, 12].

3.1 Short Summary

We conducted a literature review to deepen how the visual attention process is used pragmatically in InfoVis and if these uses are somehow brought closer to the graph designer. The literature presented in this chapter mainly aimed at identifying possible responses to the following Research Question (RQ): *how the preattentive visual process is used in information visualization*, oriented to improve the observer's cognitive process in graph comprehension. Therefore, in selecting and analyzing the papers, rather than on the theory of how preattentive concepts impact the attention process, we focus on how these concepts have been used to improve data understanding (cognition) in information visualization. In other words, the purpose of our literature review was to provide an overview of how concepts related to the preattentive process are used *pragmatically* and implicitly in recent research.

In general, we found that the understanding of human attention process impact is often used *implicitly* in InfoVis design, in other words, based on analyses of attention models developed more than 20 years ago. In results section will discuss the *implicitly* concept in detail. Furthermore, experts in the field of design or human vision usually exploit the knowledge of how to correctly handle this observer's attention ability. Novices, who design graphics daily, seem not to be a significant target of current research. On the other hand, we identified a gap between activities done at design time and at the graph's validation time. The evaluation of graphs is done in laboratories and controlled environments and is highly time-consuming. Surprisingly, in our research, we did not find any tool that, at design time, may show the graph creator the impact that their design decisions might have on graph comprehension.

In the next sections, the process and results of the literature review, whose objective was explained above, will be presented. In addition, we also will discuss the findings and development opportunities that emerged from this research.

3.2 Review Methodology

The protocol we adopted for the literature review is a systematic review protocol consisting of three phases, detailed in the following sections: search strategy, selection, and classification.

3.2.1 Search Strategy

As a first phase of the literature review, we followed a *search strategy* that will lead us to find, in different sources of information, the most relevant articles to answer the main RQ. The search strategy was developed in three steps: the selection of search *sources*, the construction of the *search string*, and the definition of *filtering criteria*.

Concerning the selection of search ***search sources***, we based our decision on the main topic, "Information Visualization," and kept the preattentive concepts as a subtopic. This decision was based on the fact that we were interested in understanding, at the level of digital graph design, how the concepts related to the preattentive process were used pragmatically and not how the preattentive process operates. For this reason, the sources were selected because of their high relation with the topics of information visualization in computer science [37] since most conferences and journals relevant to the Research Question publish their papers in these sources. Therefore, four *sources* of information have been selected: ACM Digital library, IEEE Xplore, ScienceDirect, and Springer Link. On the other hand, although we were looking for the connection between InfoVis and the preattentive process, sources such as Pubmed were not considered since they approach the preattentive process from a more theoretical and neuroscientific approach, while we are looking for papers where these theories are applied in digital graph design.

For the construction of the ***search string***, we started using the keywords: *preattentive* and its spelling variation *pre-attentive*, and *information visualization* and its related term *data visualization*. As mentioned in section 2.2, InfoVis and DataViz have similar definitions and are commonly used without distinguishing one from the other. However, the search results were scarce, and for this reason, the search was expanded by adding the keyword *data highlighting*. In InfoVis, the 'highlighting' term denotes the focus data in a graph, a concept close to the preattentive process's output (see section 2.2). The results increased from 161 to 306 articles with this new keyword. The final *search string* therefore adopted in the literature review was: `(“pre-attentive” OR “preattentive” OR “data highlighting”) AND (“information visualization” OR “data visualization”)`. This *search string* was applied in the *Metadata* search field, that includes document title, author(s), publication title, abstract, and index terms. We are aware that many works where the preattentive process is exploited implicitly may have been excluded from the search; this is consistent with our goals of finding papers that *explicitly* tackle preattentive.

About the ***filtering criteria***, we concentrated the search on these knowledge areas: Computer Sciences, Human-Computer Interaction (HCI), Information or Data Visualization, Data Analysis, and User Interfaces. In addition to the concepts already discussed, we added the areas of HCI and User Interfaces. We selected these areas considering that human visual attention is extensively explored in HCI and User Interfaces areas. Both areas are related to how humans communicate with computers and how they perceive that interaction. In addition, InfoVis and DataViz are areas focused on the communication of data visually also associated with the HCI area. Another filter criterion was concerned with the period of the published research, and we selected the period between 2010 and 2021, inclusive. Historically, vision

Table 3.1 Search Strategy Summary

Step	**Result**
Sources	ACM Digital library, IEEE Xplore, ScienceDirect, and Springer Link
Search string	`‘(‘pre-attentive’’ OR ‘‘preattentive’’ OR ‘‘data highlighting’’) AND (‘‘information visualization’’ OR ‘‘data visualization’’)`
Filtering Criteria	*Knowledge areas*: Computer Sciences, Human-Computer Interaction (HCI), Information or Data Visualization, Data Analysis, and User Interfaces. *Period*: between 2010 and 2021.

attention has been developed since the 1980s and has been widely researched after that. However, this research aims to discover the current impact of the preattentive process in InfoVis's design stage, not its fundamentals, and for that, we focused on the last decade. For the last filter criteria, we considered only the articles in conference proceedings or journals as the document type.

Table 3.1 shows the search *sources*, the construction of the *search string*, and the definition of *filtering criteria*.

3.2.2 Selection

In this phase, we refined the articles list obtained in the search phase. The selection process had two steps. First, we ***refined*** the article list by removing those related to concepts outside the research scope. Then we ***selected*** the articles that met the research questions.

To ***refine*** the obtained articles list, we excluded the articles associated with the development of salience prediction algorithms, virtual reality, dynamic and 3D graphics, theoretical articles, vision problems (e.g. blurred vision or hypertropia), and surveys. Then, we manually searched those concepts in the article's title, keywords, and abstract.

The saliency prediction concept represents the set of models and algorithms that can predict the focal points of an image with varying degrees of accuracy.

In InfoVis, there are several saliency models [2, 38–40] that will be explained deeply in Chapter 4. We dismissed the saliency prediction algorithms concept because these articles present experiments and developments of attention prediction algorithms, but not how the saliency cognitively impacts the observer in the specific InfoVis domain. In addition, several papers currently summarize and evaluate those saliency algorithms [41–43]. Regarding the Virtual Reality, Dynamic and 3D articles, we excluded them because those were beyond the scope of the research. We are concentrating only on static and 2D images for the current study. About theoretical and survey articles, our research focuses primarily on the practical use of preattentive concepts in the area of InfoVis rather than on its theoretical aspects.

During the *selection* sub-step, we identified whether the use of *preattentive* concepts in the paper was actually oriented at improving the cognitive process in graph comprehension (RQ). Moreover, for this research, the concept of preattentive had to be a primary topic within the article; in fact, several articles used the word preattentive only to indicate human visualization's theoretical basis but did not focus their research on its exploitation. Therefore, in this selection process, some questions were established to choose those articles that would give us more explicit information about the use of *preattentive* in InfoVis; in fact, several reviewed articles were not explicit in how they used preattentive in their research. Therefore, the articles that responded to *at least one* of those four filtering questions (FQ) were selected to be analyzed. The filtering questions were:

- (FQ1) Are preattentive attributes used to improve the understanding of the graph?
- (FQ2) Are preattentive attributes used to draw the observer's attention to specific information?
- (FQ3) Are the presented article results about the cognitive influence of preattentive attributes?
- (FQ4) Are computational salience models used to measure the impact of the preattentive process?

To identify which articles answered our filtering questions, we performed a reading of each article's abstract, introduction, and conclusion. Besides, we looked

Table 3.2 Search Results

Step	IEEE	ACM	ScienceDirect	Springer Link	Total
Search Results	89	77	81	59	306
Refined Filter	73	59	68	48	248
Selected Articles	11	9	2	7	29

up the concepts used in our "search string" in the full text of the paper and analyzed the context in which they were used.

Table 3.2 presents the results of the selection process. The first row, "Search Results," shows the source whose results were used as inputs for the selected phase. The 'Refined Filter' numbers correspond to articles list refinement, discarding those related to concepts outside this research scope. Finally, based on the*filtering questions* presented above, 29 articles were selected to be analyzed. In Table 3.3, the FQ column shows the question numbers (one to four) which each article selected response. As can see, most of them respond to more than one question.

3.2.3 Classification

The results of the previous phases revealed that *preattentive use in the information visualization design to improve cognitive graph comprehension* (RQ) could be classified according to two main uses: (1) the preattentive attributes *as design components* and (2) the preattentive process as a *measuring tool*. In the first one, we assume that, in InfoVis design, the preattentive attributes are exploited to highlight relevant data or represent more data in a single graph. Moreover, for the second use, the preattentive computational models (prediction algorithms) are used as a method of graph evaluation, in general, to measure the cognitive processes of the observer. In section 3.3, this classification will be extensively discussed. At the end, 18 articles were classified as *as design components* and 11 as *measuring tool*.

Table 3.3 shows the list of the selected articles with their classification (Design Component or Measuring Tool) and author's names in chronological order. The "FQ" column represents the Filter Question (see Section 3.2.2) with which each selected article satisfies.

Table 3.3 Selected Articles

Title	Authors	Classification	FQ
Improving focus and context awareness in interactive visualization of time lines	Luz *et al*. [44]	Design Component	1, 3
Context-preserving visual links	Steinberger *et al*. [45]	Design Component	1, 2
Matse: the microarray time-series explorer	Craig *et al*. [46]	Design Component	1
Stacking-based visualization of trajectory attribute data	Tominski *et al*. [47]	Design Component	1
Leveraging cognitive principles to improve security visualization	Dunlop *et al*. [48]	Design Component	1, 2
Onset: a visualization technique for large-scale binary set data	Sadana *et al*. [49]	Design Component	1
Applying feature integration theory to glyph-based information visualization	Cai *et al*. [50]	Design Component	1
Comparing color and leader line highlighting strategies in coordinated view geovisualizations	Griffin *et al*. [51]	Design Component	1, 2, 4
Supporting supervisory control of safety-critical systems with psychologically well-founded infovis	Ostendorp *et al*. [52]	Design Component	1, 3
Using typography to expand the design space of data visualization	Brath *et al*. [53]	Design Component	2, 3
A space optimized scatter plot matrix visualization	Wang *et al*. [54]	Design Component	1
Cognitive benefits of a simple visual metrics architecture	King *et al*. [55]	Design Component	1, 2, 3
Font attributes enrich knowledge maps and information retrieval	Brath *et al*. [56]	Design Component	1
Keshif: rapid and expressive tabular data exploration for novices	Yalcn *et al*. [57]	Design Component	1
CorFish: Coordinating Emphasis Across Multiple Views Using Spatial Distortion	Richer *et al*. [58]	Design Component	1, 2
GeoBrick: exploration of spatiotemporal data	Park *et al*. [59]	Design Component	1, 2
Guidelines for cybersecurity visualization Design	Seong *et al*. [60]	Design Component	1, 3
Photographic High-Dynamic-Range Scalar Visualization	Zhou *et al*. [61]	Design Component	1, 2, 3,

Continued on next page

Table 3.3 – continued from previous page

Title	Authors	Class	FQ
Comparing averages in time series data	Correll *et al.* [62]	Measuring Tool	1, 2, 3
Does an eye tracker tell the truth about visualizations? findings investigating visualizations for decision making	Kim *et al.* [63]	Measuring Tool	2, 3, 4
Individual User Characteristics and Information Visualization: Connecting the Dots Through Eye Tracking	Toker *et al.* [64]	Measuring Tool	3
Highlighting interventions and user differences: informing adaptive information visualization support	Carenini *et al.* [65]	Measuring Tool	1, 2, 3
Eye tracking to understand user differences in visualization processing with highlighting interventions	Toker *et al.* [34]	Measuring Tool	1, 3, 4
Towards Facilitating User Skill Acquisition: Identifying Untrained Visualization Users Through Eye Tracking	Toker *et al.* [66]	Measuring Tool	1, 3, 4
Enhancing infographics based on symmetry saliency	Yasuda *et al.* [67]	Measuring Tool	2, 4
Pupillometry and Head Distance to the Screen to Predict Skill Acquisition During Information Visualization Tasks	Toker *et al.* [68]	Measuring Tool	1, 2, 3
Mitigating the Attraction Effect with Visualizations	Dimara *et al.* [69]	Measuring Tool	2, 3
Eye-tracking reveals how observation chart design features affect the detection of patient deterioration	Cornish *et al.* [70]	Measuring Tool	1, 3, 4
Incidental Visualizations: Pre-Attentive Primitive Visual Tasks	Moreira *et al.* [71]	Measuring Tool	1, 3, 4

3.3 Results Synopsis

The results of the search presented in section 3.2 showed that the preattention process was mainly exploited in InfoVis in two ways. The first one, **preattentive attributes as Design Components**, uses the knowledge about the capability of the preattentive attributes as part of the graph design components. This use implies manipulating these attributes to achieve different goals, in particular, instantly identifying relevant data or connecting data between graphs.

The second identified usage, **preattentive process as a Measuring Tool**, uses the preattentive process as a measurement technique, either using predictive models or instruments such as eye-trackers. Such measurement is generally used to assess attention on portions of a graph or determine the impact of preattentive attributes. Both categories are relevant to this survey's scope, as they aim to understand or improve the cognitive processes of the graph observer.

In this context, we present a synthesis of the selected articles grouped according to their classification.

3.3.1 Preattentive Attributes as a Design Component

We found several articles that use preattentive attributes as part of the design process, which is evident because color, shape, or size are fundamental design components in InfoVis. However, for our study, we interpret the use of these attributes as an explicit attempt, at design time, to steer and focus the observer's attention and help them in the cognitive process of graph interpretation. This means that we have been looking for studies in which the preattentive attributes were consciously used for their cognitive impact. We found that preattentive attributes have been used to highlight relevant information, link graphs, or represent several data sets in one graph. Besides, color, shape, and size are the most commonly used preattentive attributes.

During the classification phase (see section 3.2.3), we noticed that the preattentive attributes used in InfoVis are more than the commonly established ones (section 2.1). The authors also consider some design elements such as Glyphs as preattentive attributes due to their ability to draw attention to relevant information. In addition, we found in some articles that some attributes are also used as part of graph design methodologies. Based on these findings, within this Design Component category, we established three more subcategories to organize the different modalities employed by the preattentive attributes. These subcategories are: *Core Preattentive Attributes*, *Unusual preattentive attributes* and *Preattentive attributes into design methodologies*.

Core Preattentive Attributes

Regarding the use of *Core Preattentive Attributes* (e.g., color, shape, size), color is the most common preattentive attribute used to improve the comprehension of the graphs. The commonly used of color was demonstrated in different visualization techniques as linked graph [59, 45, 49, 54], dashboards [57], or to grouping related information on large scale time-series graphs [46]. More recently, a study by Zhou *et al.* [61] showed that changes in color characteristics such as the glare pre-attentively steer attention and focus the visualizer's attention onto high-value features. Similar results were obtained by Luz and Masoodian [44], they used blur to avoid attention on non-focal items to improve the observer's performance.

Although the color is the most used preattentive attribute, it is also used with other attributes to enhance the range of attention and give more information with a single representation in specialized InfoVis graphs. For instance, Tominski *et al.* [47] used color and spatial position as preattentive attributes to improve the trajectory graph (contextual data in a route). A similar study was presented by Dunlop *et al.* [48], they created a tool that combined shape, color, and size to identify and analyze threats in security data analysis. Furthermore, the research made by Cai *et al.* [50] also demonstrated how to optimize the visual search in a specific InfoVis graph combined with different preattentive attributes in a glyph. The last example found in combined preattentive attributes was the research made by Richer *et al.* [58]. They combined size, position, and shape to make selected data more visually prominent in a coordinated graph (visualization across the graphs).

Unusual preattentive attributes

Due to the many ways to represent data in InfoVis, preattentive attributes are not only used in specific charts (bars, boxplots, and others), but they can also be manipulated on text visualizations (e.g., Knowledge Maps). One example is the research presented by Brath and Banissi in [56]: they use font properties, such as bold and italic, to make some text visually preattentive and to add more information into textual displays. Also, they made a systematic exploration intending to include text (font) as a preattentive attribute in InfoVis [53].

Preattentive attributes into design methodologies

An interesting finding in the literature review was that some researchers had included preattentive attributes as a fundamental part of InfoVis design methodologies. For example, the study by Ostendorp *et al.* [52] showed that some preattentive attributes are fundamental design elements in the design of a safety-critical system's supervisory control dashboard. As part of the design process, the author selects the "most effective attribute" based on the effectiveness ranking of preattentive attributes considering the type of information. For instance, if the type of information is nominal, the possible visualization attributes are position, color hue, texture, connection, etc.

A similar approach was presented by King *et al.* [55], proposing a Visual Metrics Architecture to create dashboards including a pre-attentively property rank. For example, form and color are potent properties for preattentive detection. Another related research made by Seong *et al.* [60] showed a guideline for Cybersecurity information visualization. In this methodological guide, the authors encouraged designers to "*take advantage of the human visual system's ability to do preattentive processing by seeking to encode information pre-attentively visually*."

3.3.2 Preattentive Process as a Measuring Tool

As we already mentioned in section 2.1, human visual attention can be predicted, and the impact of each visual element can be evaluated using a computational model such as saliency maps. In addition, human attention can be measured with technical methods such as eye-tracking. Between both measurement methods, saliency prediction and eye-tracking, we found that eye-tracking is a common technique used in InfoVis to analyze perceptual and cognitive processes of visual tasks [72]. On the other hand, saliency maps are not a commonly used technique, but they can also predict InfoVis's visual perception. There are many saliency prediction algorithms [73], with a different approaches, most of them based on the Itti-Koch biological model [74] (see chapter 4). Regarding how these measurement methods are used in InfoVis, we found that they can be employed to establish design patterns, evaluate the observer's cognitive process, determine the influence degree in decision-making, and draw attention.

Several studies have investigated how preattentive attributes **influence decision-making** using preattentive process measurement methods. For example, Cornish *et al.* [70] research provided evidence of how the doctor's and nurse's ability to detect abnormal behaviors in patients' data increases with the use of color scales and marks. Similar research presented by Dimara *et al.* [69] showed evidence about how highlighted optimal choices could help decision-makers to focus on important information while ignoring distracting choices (decoy and distractor points). Another study made for Kim *et al.* [63] results demonstrated that different stimuli configurations, rather than the tasks themselves, could affect information-processing strategies when people make choices.

Concerning **establishing patterns** using practically preattentive measurement methods, Alberts *et al.* [9] research has provided evidence that the large numbers in dashboards concentrated the visual attention more than on the graphs. Also, they observed some context elements (e.g., human-like figures) get prime attention over the data. On the other hand, Moreira *et al.* [71] studies how simple visual tasks, such as finding points in a vertical or horizontal position, can be executed pre-attentively without a conscious process. The authors discover that the horizontal task is the most accurate and faster to execute pre-attentively, and color luminance is the worst.

Nevertheless, the most common usage is the use of the preattentive process measures methods to **evaluate the observer's cognitive process**. In a set of studies by Toker *et al.* [64, 34, 66, 68] and Carenini *et al.* [65], the authors present two main results about highlighting interventions: (1) highlighting interventions can improve visualization processing compared to receiving no interventions (task performance and usefulness); and (2) some specific visualization regions can cause low task performance in users with low values of specific cognitive measures.

The last finding in this preattentive usage classification is **drawing visual attention** exploding saliency methods. During our investigation, we found a few of these approaches. One is presented by Yasuda *et al.* [67], their approach detects a saliency region based on symmetry Gestalt principles. Then, they used that saliency to schematize visual images by approximating them as pairs of symmetric patterns extracted from object silhouettes. The results showed that this approach could draw more visual attention to the selected region of interest. Similar results were obtained by Jänicke and Chen in [2], showed that the use of saliency maps could be used to improve the search of patterns in complex diagrams (volume, tag clouds, and flow visualizations).

3.4 Relevant Findings

In general, we found that the preattentive process is currently used to highlight specific data, often in an 'implicit' way limited to specific data sets or visualization techniques. Also, the preattentive concepts are used to measure graph effectiveness. Besides, we noticed that in most cases, the research in preattentive on InfoVis *does not directly involve the graph designers*. The preattentive notion is still at the research level (patterns, standards, methodologies), but we did not find specific studies that seek to bring this knowledge to the graph designer. This designer's oblivion is one of the most relevant issues in the conducted research. Finally, we claim that future research should further develop how to integrate attention prediction models (evaluation) as an InfoVis design tool to support graph designer decisions, which will be one of the contributions of the thesis.

3.4.1 Highlighting and Data enhancement

Several articles demonstrated how preattentive attributes could be used to draw the observer's attention to specific or relevant data, using highlighting techniques [57, 70, 61]. Color, form, and orientation preattentive attributes categories are the most common highlighting techniques used to emphasize data, and at the same time, they optimize the visual search [50, 65, 34, 59, 63]. We also detected that highlighting might be applied according to the data type (e.g., identify threats [48]) or based on a design decision (e.g., Coordinate views [51, 45]).

In many investigations [63, 66, 64, 68, 67], the researchers recognize the importance of highlighting relevant points on the graph so that the observer can detect them more efficiently. The preattentive attribute handling also helps the observer make a more effective comparative data analysis. Other studies have used these preattentive measures to establish design patterns.

Preattentive attributes can also be used to enrich data. We identified two different techniques: adding more data and linking graphs. Adding more information in one

InfoVis graph can be done by integrating, commonly, color and shape preattentive attributes [59, 47, 56]. Integrating these attributes makes it possible to visualize different data in a single InfoVis graph. In the other technique, the preattentive attributes like hue, color, and marks are used to *link data*, which means showing the connection between graphs or displaying more details about the data [51, 45, 49, 46, 75].

3.4.2 Implicit and Unconscious design decisions

In most papers, using preattentive attributes as design elements is shown as something implicit, something that the researcher or graph designer already knows, or some data visualization systems handle by default. The implicit use of these attributes revealed an essential gap between the theory and the final users' intended use of this knowledge. Both expert researchers and data visualization systems are able to handle the impact of the preattentive attributes in the visual-cognitive process, but final users use them unconsciously. In most articles, it is clear that the studies are oriented to creating standards or understanding the attribute's attention behavior but do not help novice graph designers to understand the visual-cognitive impact of their design decisions.

One of the tough challenges for researchers in this domain is to bring the implications of the preattentive attribute's visual impact to the less knowledgeable in InfoVis design to make better and more conscious design decisions.

3.4.3 Graph designer oblivion

Although there is a significant theory about the cognitive impact of preattentive attribute manipulation, and as we showed above, preattentive also has more than one use in InfoVis, we noticed a gap between the theory and the needs of graph designers. For the purposes of this research, the *graph designer* are those who create InfoVis graphs daily, using any data visualization system. These graph designers come from different areas of expertise, such as professional designers or data analysis experts.

The results show that graph designers must have extensive knowledge about the preattentive process to understand how to use its attributes. For example, the data visualization systems have some visual elements, like color palettes, that have solid preattentive visual process fundamentals [76, 77]. However, in those systems, the visual-cognitive process is implicit, applying standard shapes or color scales without giving insights to the graph designer on how their use can affect the graph design objectives. Furthermore, using existing visual elements in data visualization systems does not provide insights for a novice user in many cases. For instance, a novice graph designer does not know how to handle visual elements to reduce the cognitive process's cost and maximize the observers' cognitive productivity.

3.4.4 Preattentive prediction and Graph Design Process

In InfoVis, the preattentive measurement techniques (eye-tracking and saliency maps) are utilized to get insights about an observer's cognitive processes when they read a graph. In this study, we identified that these techniques could be employed both to measure the cognitive impact of preattentive attributes and as a measuring tool to support graph redesign.

Within the studies, we observed many that used the preattentive concepts to evaluate their designs or new techniques to prove their possible effectiveness and establish design patterns [63, 66, 64, 68, 67]. However, only a few articles used the concepts to evaluate attention and improve the design of the graph, focusing on the overall design decisions for the whole graph instead of an individual attribute [67, 2].

Currently, some researchers are developing new methods and algorithms to generate attention prediction maps specifically for InfoVis images [39, 38, 40]. Those algorithms are the first approaches to predicting the visual and cognitive impact of preattentive attributes in the graph design stage. It is crucial to notice that these developments in saliency prediction can help reduce the gap between human vision attention knowledge and the graph design process.

3.5 Development Opportunities

Between the relevant findings exposed in section 3.4, we considered that the most significant opportunity is to develop mechanisms to close the gap between graph designers and human vision knowledge (observer's attention impact) to promote mindful design decisions.

Extensive literature exists on how each design decision, and selection of design elements such as color or orientation, can affect an observer's visual and cognitive processing. Given this vast amount of information, we consider it complex to bring all these concepts closer to every type of graph designer. Because each preattentive attribute can impact different levels of the graph design, and the impact also changes if a combination of them is made (see section 3.3.1). From this standpoint, bridging the gap between the graph designer and the impact of the preattentive attributes could be made by preattentive measuring methods (saliency maps and eye-tracking).

In section 3.4.4, we discussed how measuring the impact of preattentive attributes can be a promising tool for performing graph design improvements. Also, we pointed out that according to several studies, the focus of attention on an image affects the human cognitive process and improves the observer's understanding of the graph. Based on these findings, if a measurement attention method could be integrated into a visualization system, the graph designer could know the impact that each of her design decisions would have at the visual-cognitive level on the final observer.

However, in order to establish this bridge between saliency prediction and graph designers, it is necessary to go deeper into the topic of saliency prediction. First, it is essential to notice that in the literature, preattentive predicting methods are known as **Saliency Prediction**. Based on our investigation, only since 2017, some researchers have developed computational models to adapt salience prediction to InfoVis images. For this reason, we had to conduct a deeper study of human vision prediction in InfoVis, the models, and the algorithms that have been developed (Chapter 4).

The following two chapters will present more profound research about saliency prediction for InfoVis. Firstly, a short description of the selected saliency prediction models, and secondly, the description and results of an experiment to evaluate its functionality.

Chapter 4

Saliency Models in InfoVis

This Chapter aims at mapping the scenario of existing saliency prediction models approaches, how they work overall and which specific models exist for InfoVis.

4.1 Saliency Prediction Overview

For more than 40 years, researchers have been observing the behavior of human attention in natural images. These observations were carried out in the field of psychology since in the '80s, when some researchers applied them to the area of computation, creating computational models, so-called **Saliency Models**, to imitate human attention.

According to Yan *et al.* [78], currently, those saliency models can be classified in: **Classic Visual** Saliency Models and **Deep Learning Visual** Saliency Models. The first group considered the psychological and psychophysical basis, the models closer to the biological human vision process (See Fig. 4.1). And the second group, based on trainable data sets with images and their observers' data, generates deep learning models implemented in a fully automatic data-driven mode to extract the saliency regions. The following subsections briefly describe these models and the data sets used for their creation and validation.

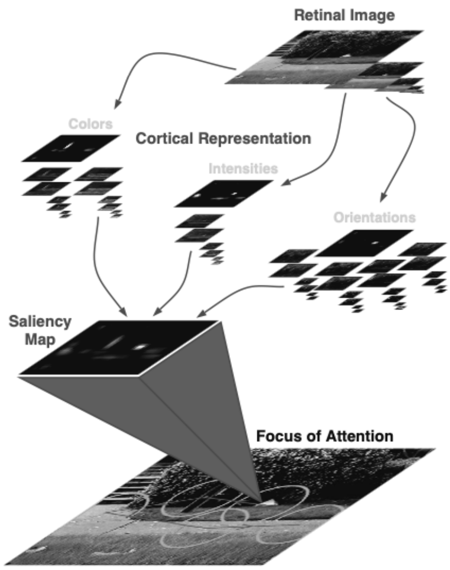


Fig. 4.1 "*Classical Saliency Model*" taken from [79]

4.1.1 Classical Saliency Models

Classical Saliency Models term refers to all those models developed based on psychological and neurobiological visual theories. Traditionally, based on the factors that drive attention, these Classical Models are divided into two types of models, namely **Bottom-Up** and **Top-down**. Bottom-up models represent the unconscious visual process (data-driven, task-agnostic model), and Top-down models are related to the visual-cognitive process (task-driven, task-specific model). In this section, each of these groups of models will be briefly explained.

Bottom-Up Model

This attentional mechanism is also called exogenous, automatic, reflexive, peripherally cued, or *stimulus-driven model* [5]. This model *measures how different an element is from its neighbors* [14] and tries to imitate the unconscious visual perception process. Bottom-up attention is fast, involuntary, and most likely feed-forward. An example of this is a red spot in a green field that could be a fruit or a predator [80]. This model uses low-level features such as color, texture, size, contrast, brightness, position, motion, orientation, and shape of objects that influence visual attention. Also, the influence factors come from solely from the visual scene [4].

In Bottom-up attention, the interest region must be different enough with respect to nearby features (e.g., a horizontal bar among several vertical bars) [5]. Although the information from each fixation influences our mental experience [14], the bottom-up process operates without prior knowledge about the image content [81]. Because this model is task-free, it has been more extensively developed. Task-free means that the observer does not have to solve a specific task over the image, for example, finding an object or the bigger bar in a chart [82].

According to Itti and Koch in *et al.* [83], several biological vision systems use a sequential computational technique to analyze complex visual scenes (images). Specific areas in the scene are chosen to be detailed depending on their behavioral significance or local image referent points. In primates, identifying elements and analyzing their spatial relationships typically requires fast and saccadic eye movements to concentrate the retina onto the object or subtle attention shifts. This process was described in *"Feature Integration Theory"* (FIT) proposed by Triesman and Gelade [13]. According to Borji *et al.* [5], the basis of most Classical Saliency Models is the FIT model.

The FIT model stated that *incoming visual information is first analyzed by early visual neurons, which are sensitive to elementary visual features of the stimulus (e.g., colors, orientations, etc.)* [82]. This analysis is performed in parallel over the whole visual field, using different spatial and temporal scales, and generating many cortical feature maps (see section 2.2). Each feature map represents the saliency proportion of a given elementary visual feature (preattentive attribute) in the visual field. Figure 4.1 shows an example of Treisman and Gelade's human attention process.

The "Feature Integration Theory" was extended by Koch and Ullman [84]. They proposed to create a single topographic and scalar saliency map where all feature maps are combined. Figure 4.2 shows the latest and completed computational model of "Feature Integration Theory" developed by Itti *et al.* [24]. The model will be presented in more detail in section 4.2.2.

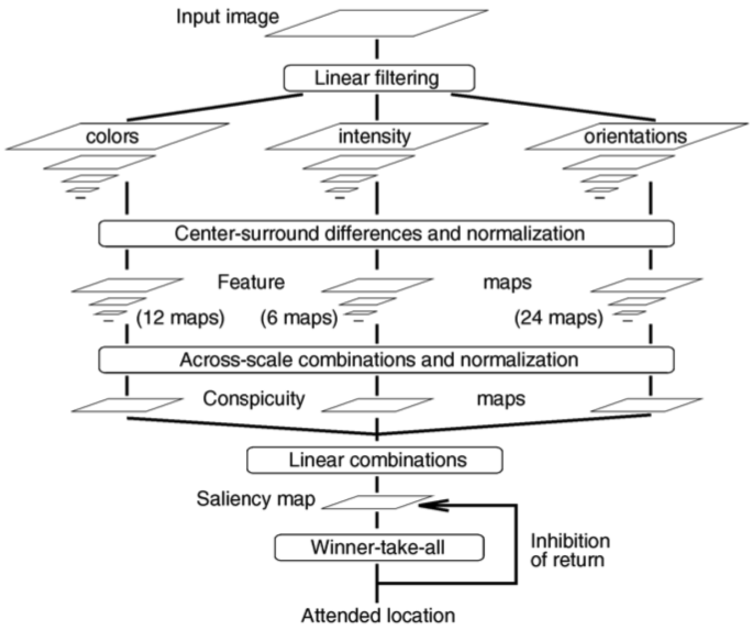


Fig. 4.2 Itti, Koch and Niebur computational approach of the “*"Feature Integration Theory"*” taken from [24]

The saliency model explained above is the most commonly used in the literature mainly because it is the one that comes closest to the biological and psychological process of human visual attention.

Top-Down Model

The Top-Down attention model, also called endogenous, voluntary, or centrally cued, is guided by cognitive factors such as tasks, prior knowledge, or expectations [4]. In contrast to the Bottom-Up, a cognitive task influences the visualization. For example, if you have to find a blue object, the regions of the image that have this color will attract your attention.

According to Healy *et al.* [14], the Top-Down approach *is a user-driven attempt to verify hypotheses or answer questions by "glancing" about an image, searching for the necessary visual information*. This model uses high-level features, and context-dependent features such as faces, humans, animals, vehicles, text, and others. The top-down preattentive model uses prior knowledge about the scene and/or about a given task [81].

At neural and psychophysical levels, bottom-up attention process responses in the first 125 milliseconds and top-down response approximately 100 milliseconds after [80]. This means that bottom-up attention could be rapidly executed because its effects are over simple properties, while the top-down attention model requires a cognitive process.

4.1.2 Deep Learning Saliency Models

The second and newest group of saliency models is the Deep Learning Saliency Models (DLSM). The DLSM *"are a single convolutional layer followed by a fully connected layer trained to predict fixations"* [85]. These models generalize the Itti classical visual model (see section 4.1.1) but add feature maps linear combination. Currently, there are more than 16 different static DLSMs whose performance is broadly studied in [5, 85, 78, 86]. Commonly, these deep learning models added more top-down information to predict saliency, for instance, the cognitive relevance of each generic object in the image.

The most significant difference between these models and the Classics mentioned above is the ability to extract higher-level features. For example, due to changes in image scale changes, text or face features in classical models can be overlooked as relevant objects.

According to recent studies [85, 78], The DLSM has been demonstrated that perform markedly better than classical saliency models based on hand-crafted features. In addition, these DLSM are pre-trained on large image datasets and with eye-tracking or mouse click data that provides more detailed information about vision behavior in different types of images.

4.1.3 Saliency Datasets

Both classical and deep saliency models, since their beginnings, have been validated and trained with *natural images*. Natural images include cartoons, art, satellite, life, landscapes, people, cities, line drawings, etc. The classification of these images can vary between the existing datasets [78], for instance, MIT300, SALICON, and CAT2000. Table 4.1 presents a summary of these datasets.

Table 4.1 Saliency Datasets Description

Name	Description	Reference	Method
MIT300	300 natural indoor and outdoor scenes	Eye-tracking.	[87]
CAT2000	4000 images from 20 different categories.	Eye-tracking	[88]
SALICON	10,000 images from the Microsoft COCO, 80 object categories	Mouse Clicks.	[89]
MASSVIS	Over 5000 static real-world data visualizations from 5 different categories.	Eye-tracking	[90]

Conventionally, saliency models have been validated using an eye-tracking technique (see section 2.1). The common databases as MIT300 [87] and CAT2000 [88] have collected fixation annotations from 39 and 24 observers, respectively. Other datasets such as SALICON collected the attention data from mouse clicks [89], with 60 observers for each image also using a crowdsourcing platform. From an InfoVis standpoint, one widely used is the MASSVIS database with eye fixations over 393 InfoVis images [90]. This database includes 2000 data visualization images from government reports, infographic blogs, news media websites, and scientific journals, of which exclusively 393 with eye-tracking data. MASSVIS collected eye-movement data from 33 observers over 393 images and at least 16 observers for each image.

Some databases, such as SALICON and MASVISS, provide additional metadata as contextual annotations, image categories, dimensionality, distinct colors (e.g., black and white), or recognizable objects (e.g., a human).

4.2 Saliency Models for InfoVis

According to Haass *et al.* [10], visual saliency prediction *"has been a valuable tool for studying how people process information in natural scenes"*. However, there is still a lack of work demonstrating the effectiveness of saliency prediction models with DataViz images. Therefore, this section will briefly describe the research carried out on saliency prediction in InfoVis until now. Additionally, according to some studies, three saliency prediction models that have the best performance with InfoVis images will be discussed.

4.2.1 Overview

Before going into detail about saliency prediction models in InfoVis, some of them only for DataViz, we will briefly describe some important features that characterize saliency models in InfoVis. Also, we will describe three selected saliency models for InfoVis.

Natural Images vs DataViz Images

One of the essential factors in the limited use of saliency prediction models in InfoVis is the difference between natural and InfoVis images. As explained in section 4.1.3, most saliency prediction models are designed and validated with images of natural scenes. However, those natural scenes have some visual characteristics that make the attention process different. According to studies performed by Matzen *et al.* [10] and Haass *et al.* [10] those image components are:

- **"Born Digital"**, unlike natural images, InfoVis images are created with digital tools, which makes it easier to isolate visual elements and even infer their cognitive relevance (top-down).
- **Small Objects**, in the classical saliency models, images are scaled to generate the feature maps (see section 4.1.1). In this process, small objects such as glyphs, numbers, text, and separate data points are susceptible to obscuring or smoothed.
- **Color Scale**, the researchers believe that human color perception changes accordingly to how the images have been created, which means that the image color space influences the process of the saliency prediction. For InfoVis images, some authors proposed that the Lab scale is more perceptually uniform than RGB for the saliency prediction.
- **White Space**, natural images tend to have objects everywhere that can focus the attention in contrast to the InfoVis images that normally have large and uniform color areas. This means that images in InfoVis are often in a unicolor background, which can generate more contrasting areas that may generate noise in the saliency generation.

- **Center attention**, classical saliency models use a center weighting approach (attention starts in the center of the image). However, saliency in photographs works well because the interesting objects are usually in the center, but in InfoVis, the relevant information emerges in any spatial location.

- **Text**, according to several studies [39, 40, 91, 38, 41, 42, 92, 10], text in InfoVis images receive a significant percentage of the attention. When a human reads a text, more time and eye fixations are required. Text in the InfoVis graph offers extra information about the context and details about the data. The text is an important visual element in the observer's data understanding. In contrast with natural images where text is rarely found, and thus it is not included as a feature attribute in classical saliency prediction models.

InfoVis Saliency Models

As explained above, classical saliency models have traditionally been used for natural images. Nevertheless, there are some studies where their efficiency has been evaluated on InfoVis images.

Polatsek *et al*. [41] attempted to study visual attention and saliency modeling in the context of task-based visual analysis (e.g., finding the lowest value on the graph). They investigated the effect of top-down factors on user visual attention by focusing on three low-level visual tasks: *retrieving the value of a specific data element*, *filtering data elements based on specific criteria*, and *locating an extreme attribute value within a dataset*. The authors collected eye-tracking data on the MASSVIS dataset [90] using those three visual tasks and compared the results to 12 saliency models to determine their accuracy in saliency prediction. However, only the DVS model (Matzen [38]) was explicitly designed for InfoVis images, while the others are based on natural images. As a result, according to their findings, observers pay special attention to data regions (i.e., data-containing components) throughout task-based visual analysis than during exploratory visual analysis.

Furthermore, the authors discovered that in particular tasks, such as *finding an extreme attribute*, the predicted attention points in statistical graphs are not always the most prominent and that more salient data points are not always faster to notice during task-based visual analysis [41]. Their research, however, has some limitations, such as the utilization of a diverse range of InfoVis images. The drawback is that the different factors, such as the various types of graphs used in the study, might skew the results.

Haass *et al*. [10] conducted a comparison analysis of three saliency models to see how they performed over a set of InfoVis images. On the MIT Saliency Benchmark website, the three saliency models are listed: Itti, Koch, and Nieber model [24] (henceforth called Itti-Koch model), Boolean Map-Based Saliency (BMS) [93], and Ensembles of Deep Networks (eDN) [94]. The authors selected MASSVIS images for the study to validate the previously listed models. According to the findings, the Itti-Koch model outperforms the other models. Based on the author's observations, the other models were created for natural scenes; however, the Itti-Koch model is based on a human visual processing system, which is a more neutral approach.

On the other hand, Livingston *et al*. [42] presented a study about Perceptual and Cognitive models used to predict saliency attention in statistical graphs (DataViz). They are classified as Perceptual models that are related to the perspective given by Treisman and Gelade [13]: Feature Integration Theory (see section 4.1.1). Regarding the Cognitive models, those predicting the graph saliency based on cognitive process depend on each visual element's perceptual effort. For example, Elzer *et al*. [95] defined perceptual effort as the time it takes to complete a task with the minimum number of fixations, and they assigned visual-cognitive weights on each graph element. The images on which the study was performed were those found in MASSVIS and new graphs created by them. Similar to the previous studies, among all the models evaluated, the perceptual models BMS and Itti and the cognitive model DVS were the ones that obtained the best results.

These three presented studies, Haass *et al*. [10], Polatsek *et al*. [41], and Livingston *et al*. [42], find that the **Itti-Koch model** is the best predictor of salience in InfoVis among the bottom-up models. Similarly, Polatsek *et al*. [41] and Livingston *et al*. [42] are both in agreement that the **DVS Matzen model** performs the best in saliency bottom-up and top-down prediction. Finally, all of the experiments revealed that text components are the primary center of attention in InfoVis images (e.g., tiles, axes names).

From the studies presented above, the saliency prediction in DataViz experiments has been mainly performed on saliency models designed for natural images. Few researchers have addressed the problem of creating models that precisely predict saliency for InfoVis. In our search we found three models expressly development

Table 4.2 InfoVis Saliency Models Overview

Feature	Itti-Koch	Bylinskii	Matzen
Development Year	2008 (last author update)	2017	2018
Baseline Conception	Biologically conceivable	Computational Approach	Biologically conceivable + Digital born images behaviour.
Model Classification	Classical Bottom-up (sec. 4.1.1)	Deep Learning (sec. 4.1.2)	Classical Bottom-up + Top-down (sec. 4.1.1)
Visual Features Channels	Color (RGB), Intensity and Orientation	FCN-32 standard parameters. Color (late fusion RGB-Depth)	Color (LAB color space), Intensity, Orientation and Text location prediction
Training (Root) Datasets(s)	Natural Images	MASSVIS	MASSVIS + Natural Images
Tested Dataset(s) (InfoVis)	MASSVIS	MASSVIS	MASSVIS + authors graphs
Development Strategy	Markov chains (GBVS algorithm [96])	Fully Convolutional Networks	Modified Itti-Koch + Text detection model
Development Language	Matlab	Python + Caffe DL framework	Matlab
Runtime (one graph image) *	*Simple*: 2.75s *Complex*: 2.48s	Simple: 51.64s Complex: 119.70s	Simple: 12.52s Complex: 68.95s

* In the feature "Runtime", *Simple* means with less visual elements (e.g., legend); and *Complex* with many visual elements.

for InfoVis: Matzen *et al.* [38], Bylinskii *et al.* [39] and Fosco *et al.* [40]. On the one hand, the **Matzen** model is an integration of bottom-up and top-down models. Moreover, the **Bylinskii** and **Fosco** algorithms, which were built using deep learning. Specifically, Fosco *et al.* [40] model integrates the Bylinskii model with a new deep-learning approach trained with a larger image dataset (called Imp1K).

As mentioned above, we found only two saliency models explicitly created for InfoVis images on which we will perform a more in-depth analysis: Matzen *et al.* [38], and Bylinskii *et al.* [39]. In addition, we considered the Itti-Koch model [74] for a deeper study based on the results of different authors presented above, demonstrating this model's high performance in DataViz images. Finally, it is essential to clarify that Fosco *et al.* [40], another InfoVis dedicated model, was excluded from the research for two reasons: first, it uses the Bylinskii model as a basis without modifications; second, it is trained with a broader spectrum of images, for example, posters and infographics. The above reasons make this model slightly distant from our prediction objective in statistical graphs.

Table 4.2 shows an overview of the mentioned models. Itti-Koch and Matzen models have similar bases because the Matzen model is based on Itti-Koch. However, the Matzen model was adapted for InfoVis images and their digitally born properties. On the other side, the Bylinskii model has a deep learning foundation, which means it is a trained model based on real eye-tracking data. Regarding Runtime values, we measured them in a Mac with the next characteristics: processor of 1.8 GHz Intel Core i5 dual-core; memory with 8 GB and 1600 MHz DDR3; and a graphics card Intel HD Graphics 6000 1536 MB.

The following sections describe how the three models work and their developed characteristics.

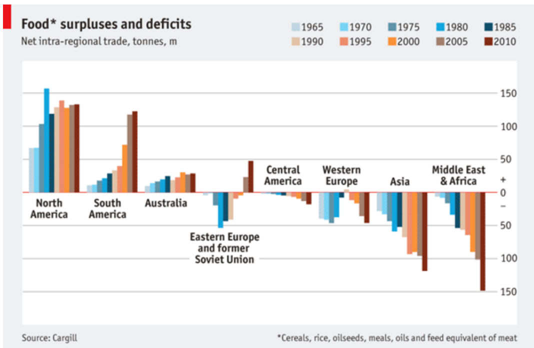
4.2.2 Itti-Koch (Classical Bottom-up Model)

The model proposed by Itti *et al.* [24] is biologically-inspired and the closest development to the Treisman and Gelade visual perception model (see section 4.1.1). Fig. 4.2 shows the complete process proposed by Itti-Koch. In addition, Table 4.3 presents an example for each feature channel, Color, Intensity, and Orientation, the resulting conspicuity maps (saliency map for feature channel), and the final Saliency map (base and blur image).

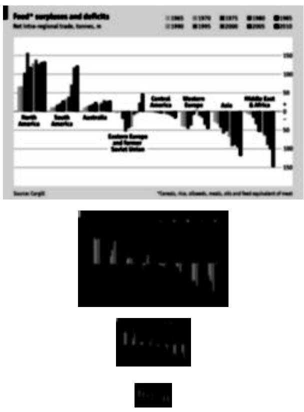
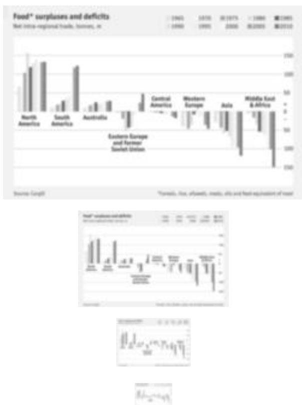
In this model, an input image is deconstructed into a series of multiscale neural "feature maps" that locate spatial discontinuities in the features of color, intensity, and orientation. Each feature map contains *non-linear spatially competitive dynamics*, which means that the activity of surrounding neurons influences the response of a neuron at a particular point in the map. The next target is determined by competition among neurons in this map, which results in a single winning region correlating to the most salient object. This model could be described in three main steps [86]:

Table 4.3 Itti-Koch Feature Maps examples.

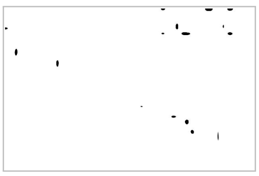
Input Image



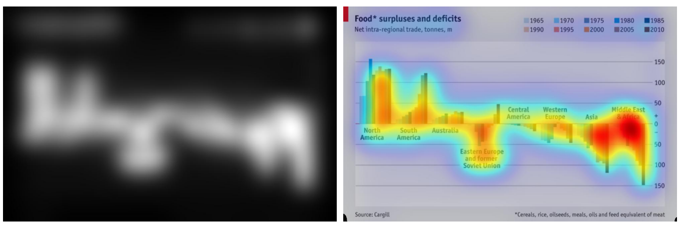
(A) Feature Maps

Color (red)	Intensity	Orientation (0^o)
		

(B) Conspicuity Maps

Color (for all RGB channels)	Intensity	Orientation (for all grades channels)
		

(C) Saliency Map (sum of Conspicuity Maps)



- **Extract Visual Features (Feature Maps)**. The basic visual feature used in most of the FIT base models are [82]: *Color*, *Intensity* and *Orientation*. However, the research presented by Wolfe and Horowitz [97] demonstrated that also *motion* and *size* are undoubtedly basic features. The saliency models based on FIT currently consider several visual features by adding impact weight to the whole attention process. This weight is defined by each saliency model based on different eye-tracking experiments [4].

  In this step, the locations that stand out from their surrounding are detected for each visual feature. This process is executed in parallel, as in the human vision process. In addition, from a biological perspective, the saliency models detect those stand-out locations on several scales by changing the image size ratio between the center and surrounding regions [24]. As a result of this step, a considerable number of feature maps are generated for each feature. Those feature maps are commonly represented as gray-scale images, in which the brightness of a pixel is proportional to its saliency.

  An example of this process is presented in Table 4.3.(A). The first column represents the feature maps resulting in the Color Red channel. Other feature maps can be generated for the rest of the RGB (red, green, and blue) scale. The second column shows the Intensity feature maps. Finally, the last one is Orientation in 0^o from 0^o, 45^o, 90^o, and 135^o that are evaluated in this channel. The number of maps generated is determined by each model. For instance, Itti-Koch's model algorithm (GBVS version [96]) generates 12 maps for *color*, 4 for *intensity* and 16 for *orientation* channels.

- **Compute Individual Feature Maps (Conspicuity Maps)**. In this step, the feature maps are summed up in independent saliency maps (for each feature), called conspicuity maps. According to Zhao and Koch [86], *this step uses biologically plausible filters such as Gabor or Difference of Gaussian filters or more sophisticated methods.* Some examples are the use of [86]: Bayesian statistics, discriminant center-surround hypothesis, entropy minimization algorithm to select fixations, stochastic models, and the Bernoulli mixture model.

  Firstly, over each feature map is applied a *weighting function* to determine the uniqueness of features. For instance, if the feature map has only a single bright region, this is a unique feature, and its weight is the highest. On the other side, if the feature map has several bright regions, its weight is lower. Then, each conspicuity map is normalized to prevent one of them has more feature maps than the others. The normalization phase could be made by simple normalized summation, linear combination with learned weights, and global non-linear normalization followed by summation or linear competition between saliency locations [98]. In the last phase, the sum-up was made using the methods mentioned above, using the resulting conspicuity maps. Table 4.3.(*B*) presents an example of the resulted conspicuity normalize map for each channel.

- **Integrate Feature Maps (Saliency Map)**: The last step is the creation of the saliency map. The resulting conspicuity maps are fused in a single attention topographic map. Based on the model presented in Fig. 4.2, a Linear Combination has been performed to sum up the conspicuity maps. According to Itti *et al.* [24], the linear combination compares *"the maximum activity in the entire map to the average overall activation measures how different the most active location is from the average. When this difference is large, the most active location stands out, and the map is strongly promoted"*. Then, as in the biological synaptic interactions, a winner-take-all is performed, which means that the most active locations remain, and the others are suppressed. Table 4.3.(*C*) shows the resulting Saliency map and its blur version visualization, which highlights better focal points.

Table 4.3 presents an example of how this saliency model works. First (A), the input image is discomposed based on three selected features (Color, Intensity, and Contrast). Then, a conspicuity map for each feature is generated, showing the regions where each feature "wins" the saliency (B). Finally, those conspicuity maps are combined, and the most relevant areas are selected and represented in a saliency map (C). As shown in Table 4.3, the resulting saliency map takes the most saliency locations from each feature and represents them in one map.

The resulted presented in Table 4.3 come from Graph-Based Visual Saliency (GBVS) algorithm [96]. The GBVS algorithm uses a Markovian approach to calculate its saliency maps [99]. The first step in this algorithm is to break up the input image into the three feature channels explained before. In this research, grayscale images have been taken as the input image in which intensity, size, and positional proximity among the faces are considered as the main parameters to estimate the saliency score.

Orientation is calculated by using the Gabor filter at 0, 45, 90, and 135 degrees. Intensity is the grey-scale version of the image calculated by eliminating the hue and saturation information while retaining the luminance with the following equation: $0.2989 * R + 0.5870 * G + 0.1140 * B$, where R, G, and B and the red, green and blue color channels respectively [99]. Colour is treated as two channels: Blue-Yellow (calculated $asabs(B - min(R,G))$, and Red-Green (calculated $abs(R - G)$). Salient regions are then located in each of these channels by computing:

$$d((i,j)||(p,q)) = log\frac{M(i,j)}{M(p,q)} \tag{4.1}$$

where $d((i,j)||(p,q))$ represents the connection between pixels, M(i, j) is the value of the pixel (i, j) in the feature map M (i.e., in the feature channels of color, intensity, or orientation). The next step is to connect each node (also known as a point or a vertex, which in this context represents a pixel) to every other node in each M, resulting in a completely connected graph. A Markov chain is then created. The system is memoryless and simply considers its present state and no prior state sequences when determining its next state [99]. Nodes are handled as states, and edge weights are treated as transition probabilities in the GBVS specified Markov chains. Higher values are obtained for nodes that differ more from those around them. This is because it is more likely to transition into subgraphs with lower similarity measures [99].

The saliency map's final result is obtained by linearly pooling each equilibrium distribution. The individual nodes in the GBVS algorithm behave similarly to the neurons in the visual cortex, communicating networked to identify regions of interest. Additionally, each node or region can do all calculations concurrently.

4.2.3 Bylinskii (Deep Learning Model)

Bylinskii *et al.* [39] developed two automated models that predict the *relative importance of different elements in data visualizations and graphic designs*. Due to the nature of the images, we will focus only on the model's description for DataViz images. The model is a neural network that was trained on several designs using human clicks and significance annotations. The authors investigated the model predictions in terms of ground truth importance and human eye movements using a novel dataset of crowdsourced relevance [100]. Researchers show how such relevance predictions may automatically be utilized for design retargeting and thumbnailing.

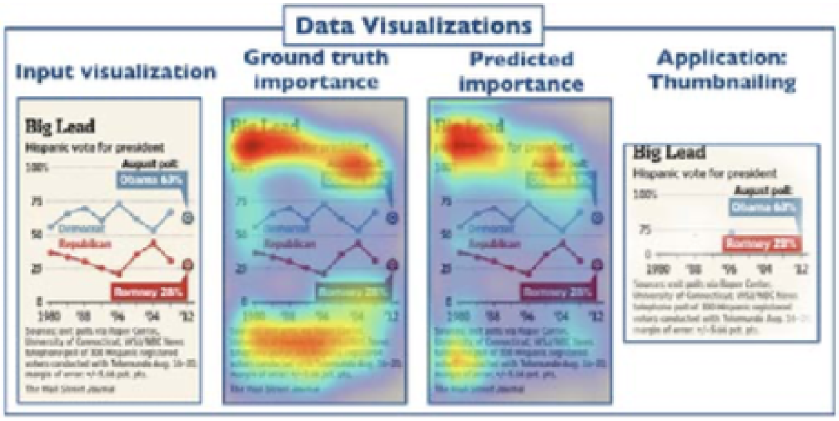


Fig. 4.3 *Bylinskii Deep Learning Saliency Prediction*. Image taken from [39]

The authors use the concept *"importance"* as a generic term to characterize the perceived relative weighting of design aspects. Image saliency is a form of *importance* that has been examined extensively. Traditional ideas of saliency, on the other hand, are based on bottom-up, pop-out effects. In contrast, Bylinskii *et al*. notion of importance is based on higher-level features such as the semantic categories of design components (e.g., title text, axis text, data points).

Bylinskii *et al*. model architecture is based on fully convolutional networks for semantic segmentation (FCNs), which are similar to some of the best-performing saliency models for natural images [101]. A directed acyclic graph of linear (e.g., convolution) and nonlinear (e.g., max pool, ReLU) operations over the pixel grid and a set of parameters for the operations are used to define FCNs. The network parameters are tuned using a loss function given a labeled training dataset. For this model, the authors used an FCN-32s for DataViz images. The Fig. 4.4 presents an example of how FCN32 deep learning works.

The model was parameterized according to the authors, as follows:

- The FCN-32s network was initialized with a base learning rate (lr) of $10e^{-5}$, scaled by a factor of 0.1 every 20K iterations. A larger learning rate makes the algorithm take giant steps down the slope, and it might jump across the minimum point, thereby missing it. Also, these parameters show that the

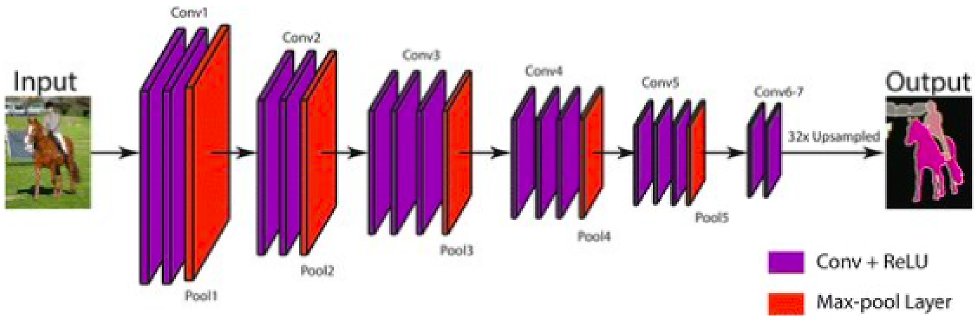


Fig. 4.4 FNC32 deep learning model used by Bylinskii saliency model. The image was taken from [1]

model must scale the image to move between epochs since it is an extensive deep learning network.

- With a momentum of 0.9 and a weight decay of 0.0005, a stochastic gradient descent solver was employed and run for 100K iterations. These parameters are used to modify the network's configuration after each training point in an effort to find the global minimum that minimizes the loss function. Regarding the momentum values, the model makes use of the standard value for momentum seen in many well-liked deep learning libraries.
- The model learning rate schedule was similar to the one used for semantic segmentation.

Regarding the data collection for the model, Bylinskii *et al.* used part of the MASSVIS dataset [90], and BubbleView interface [100]. In BubbleView, the observer sees a blurry image and therefore has to click on different segments of it to reveal small regions of the image, or bubbles, at full resolution. Initial investigations by Kim *et al.* [100] revealed a strong link between lab-based eye fixations and crowdsourced BubbleView click data. From the MASSVISS dataset, the authors took the eye movement data for testing their importance model predictions.

Some of the higher-level patterns in ground truth human annotations were captured by this model. For example, the model may learn to localize titles and appropriately weigh the relative importance of critical design components across a broad set of visualizations and designs. It is noteworthy to mention that this Bylinskii *et al.* model does not explicitly include aesthetics or design heuristics, instead focusing on simulating the behavior of the observer's attention on the content [40].

The Bylinskii model was not trained on systematic design variations, such as changes in font, text size, or element locations [39]. Nevertheless, its authors claim that the model *can correctly assign relative importance values to different design elements as they are moved around and resized.*

4.2.4 Matzen (Classical Bottom-up and Top-Down Model)

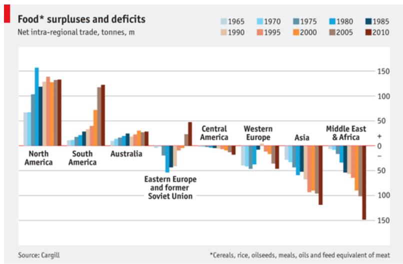
Top-down attention is extremely task- and situation-dependent in natural contexts, making it challenging to represent in any generic method. As a result, most extant saliency models exclusively include bottom-up attention. However, in DataViz, Matzen *et al.* [38] model integrates both classical Bottom-up and Top-Down approaches.

Matzen *et al.* in [91] presented a study where they demonstrated that high-level features such as *text* must be considered to perform a reliable saliency prediction model for DataViz. In addition, the authors suggest that the explanation for the poor performance of classical saliency models in InfoVis (see section 4.2.1) is that the spatial scales and visual features utilized by the classical saliency models are insufficient for DataViz images. The above statement is linked to the differences between a natural image and a DataViz image, described in section 4.2.1.

Based on these statements, high-level features, and DataViz image specifications, Matzen *et al.* [38] proposed the Data Visualization Saliency model (DVS). DVS model has two main components: a modify Itti-Koch bottom-up classical model (explained in section 4.2.2), and a *text recognition* algorithm as top-down high-level component.

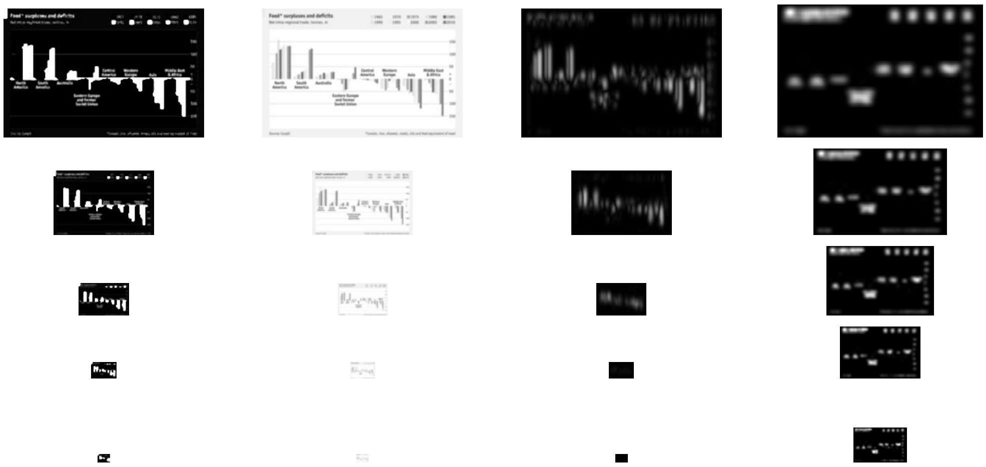
Table 4.4 Matzen's Feature, Conspicuity, Saliency and Text Maps examples

Input Image



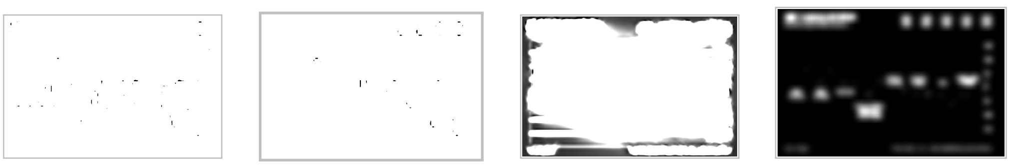
(A) Feature Maps

Color (CIO Lab)	**Intensity**	**Orientation (90^o)**	**Text**

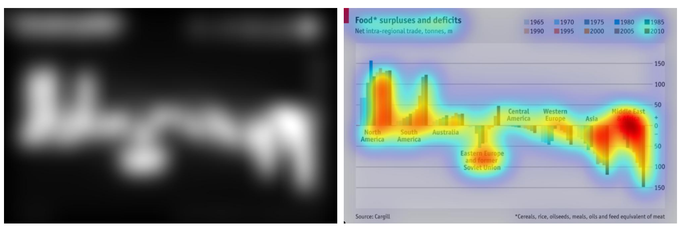


(B) Conspicuity Maps

Color (all CIO Lab channels)	**Intensity**	**Orientation** (all position grades)	**Text**



(C) Saliency Map
(sum of Conspicuity Maps)



The first DVS model component is a modified Itti-Koch model. Specifically, the DVS changed the channels used for *color features*. The original Itti-Koch model utilizes an essential color opponency representation based on RGB values. The DVS model adjusted the original technique by converting the representation of the input images into CIE Lab color space to resemble human visual perception better. Matzen *et al*. [38] claimed that the current saliency models often assigned low saliency values to bright and red regions, causing a more considerable difference with the human fixations map (e.g., eye-tracking data). In color theory, the CIE Lab scale reduces, as possible, the perceived difference between colors, bringing the color scale closer to an "organismal color perception" [102], which means that the CIE Lab model has the benefit of being perceptually consistent. As a result, values from feature maps computed over several color space channels can be meaningfully contrasted with one another [38].

The second DVS model component is a *Text Recognition* model. This integration was proposed because the images in classical Bottom-up models are resized in several scales, which might lead to the loss of small features. In particular, DataViz images have small objects that are susceptible to being smoothed in this process, including text elements. Furthermore, most of the saliency studies in InfoVis state that observers devote a great deal of attention to the text, which is not featured in existing saliency models. To resolve these problems, DVS integrated a *text saliency model* into the modified Itti model described above. Thus, the Matzen model has a Top-down approach because the observer considers the text a container of meaningful information (see Section 4.1.1).

Regarding the *text saliency model*, DVS employed a typical strategy in the text detection literature: extract the Maximally Stable Extremal Areas (MSER) that candidate text regions and then filter out non-text candidates using different text-diagnostic traits. Then a filter is performed to exclude non-text regions. Following the previous filtering, the remaining MSER areas were more likely to be letters or numbers. DVS generated three text-diagnostic edge features upon those remaining MSER areas to assess this probability using three text-diagnostic edge features: following a highly uniform background, capturing a specific topological characteristic of text, and the number of crossings between a vertical or horizontal scan line and the text edges is often an even number. In part A, Table 4.4 shows an example of the text detection model.

The integration of the two Matzen model main components, the modified Itti-Koch model and the text detection, is performed employing a *linear combination*. The formal equation is [38]:

$$s = (I + w * T)/(1 + w) \tag{4.2}$$

where I is the saliency result of the Itti-Koch modified model and T is the text saliency map. The parameter w defines the relative weight between I and T. The denominator, $(1 + w)$, provides a weighted average to preserve the saliency scaling from 0 to 1 intact. The w value was defined by performing a systematical manipulation of I and T and comparing the resulting saliency maps with MASSVIS eye-tracking data (see Table 4.1). The authors chose to use a weight of 2 because it was the value in which the performance metrics approached an asymptotic limit (continuous variable tending to 2).

Table 4.4 presents an example of the Matzen model process. In part (A), it is possible to see how visible are the graph details until the last image scale (with CIO Lab channels), compared with the classical Itti-Koch RGB color channels approach (see Table 4.3). In addition, in part (A), Matzen added the text detection model, making the graph text parts visible. Then, part (B) shows the summed maps with bases of the Itti-Koch model (conspicuity maps). Finally, as explained before, the four conspicuity maps are combined by linear combination.

4.3 Discussion

This section gives an overview of saliency models to contextualize how the computational prediction of human visual attention is performed. In addition, we further elaborated on current work on saliency prediction in InfoViz images.

One of the most important aspects was the difference between natural and InfoVis images. These differences were explicitly focused on features such as color scale, large empty spaces, and the importance of the text as an attention-getting element. In addition, we found that only two saliency models, one classical and one deep learning, have been built or adapted for InfoVis images: Bylinskii and Matzen. However, according to several authors, classical models such as the one developed by Itti-Koch also maintain acceptable performance.

Based on the study presented in this chapter, the behavior of these models left us with some concerns:

- Bylinskii and Matzen's models use the same dataset to train and calibrate their models, MASSVIS, which means the image categories in that dataset could limit them. On the other hand, the Itti-Koch model has been validated only with the same dataset, limiting the knowledge about its performance in different graphs.
- MASSVIS is a set of InfoVis images, from simple bar charts through scatterplots and maps to infographics. This image variety can be a disadvantage for the models because they are too generic, and these images often contain descriptive titles, annotations, logos, and text, noting the data source.
- In general, InfoVis images have other features such as orientation (vertical or horizontal) or the position of data-containing elements (e.g., bar in a bar chart), which could influence the focus points in an image.
- A common factor in all models described in this section is a tendency to devote most of the attention to the text, a relevant feature in the observer's cognitive process. However, we considered that an imbalance compared to the attention given to data-containing components.

Finally, the authors cited in this section highlight the importance and usefulness of these saliency models in the graph design process. However, in order to use any of these algorithms as a design support tool, we had to perform a validation of the performance of each of the models explained, considering the questions listed above. In the next chapter, we present the process and results of that analysis.

Chapter 5

InfoVis Saliency Prediction Models Validation

In the previous chapter, we described three saliency models developed or validated expressly for InfoVis: Itti-Koch model [74], which is based on a human visual processing system and uses a bottom-up approach; Bylinskii model [39], which is a neural model trained with real-world InfoVis images and has a top-down approach; and Matzen model [38], which combines a modified Itti-Koch (bottom-up) model with a text saliency model (top-down).

Although some authors have already performed validations on these models, we had four concerns about the models' behavior (see section 4.3):

1. These models have a remarkable performance because they can detect the image's text areas, which usually get most of the observer's attention. However, we want to verify whether there is a significant imbalance in the attention given to data-containing components (e.g., bars, lines).

2. According to our research, these models use the MASSVISS dataset (see section 4.2.1) to be trained and evaluated. However, using the same dataset to create and validate the models could make the predictions tailored without differentiating the characteristics of the visualization technique.

3. Related to the previous doubt, the MASSVIS dataset images often contain several context elements (e.g., logos, legends, annotations). Such context elements

generate additional information but do not represent the data, constraining the model's performance.

4. We identified that some representative InfoVis visual elements, such as orientation and position, are not considered in developing InfoVis saliency models.

The above questions suggested that the models should be evaluated in a different scenario to check possible biases due to the training effect. In addition, it is important to observe their behavior when varying some visual elements, particularly the text. To resolve the previous InfoVis saliency models behavior questions, we performed a set of validation experiments employing a ground truth collected data from an **eye-tracking device**. Fig. 5.1 shows the three performed experiments.

It is essential to clarify the fact that our experiments, as in the definition of the saliency models for InfoVis, use ***Exploratory Visual Analysis*** as a base task. As explained by Polatsek *et al.* [41], within the taxonomy of tasks related to data visualization, saliency models are developed within the exploratory analysis task. In this low-level task, the observer formulates a hypothesis about the data, which means that the observer will observe it *without having any specific task* set in advance, such as looking for the highest or the lowest value.

For the experiment, we plan to analyze and evaluate the selected models in two scenarios: **Saliency Models on MASSVIS dataset** and **Saliency Models and Clean Graph**. In the first experiment, *Saliency Models on MASSVIS*, we wanted to confirm the performance of these models with the frequently used dataset. In addition, we sought to analyze in more detail how the models behaved by predicting the graph saliency. That is, how saliency is predicted in visual elements such as bars or lines representing the data. Besides, we made a **Crop Analysis** to analyze the model's behavior when applied to data-containing elements.

For the second experiment **Saliency Models and Clean Graph**, we constructed a set of *Clean Graph* that are statistical graphs without context elements (e.g., logos, background images) and with text element variations. Next, we built saliency maps for each clean graph with each saliency model (Itti-Koch, Bylinskii, and Matzen). We also used an eye-tracking technique to acquire clean graphs of gaze data from 62 persons as ground truth. As the last step, we used three well-known saliency measures to compare the saliency maps obtained on each graph to the ground truth (CC, NSS, and AUC-Borji).

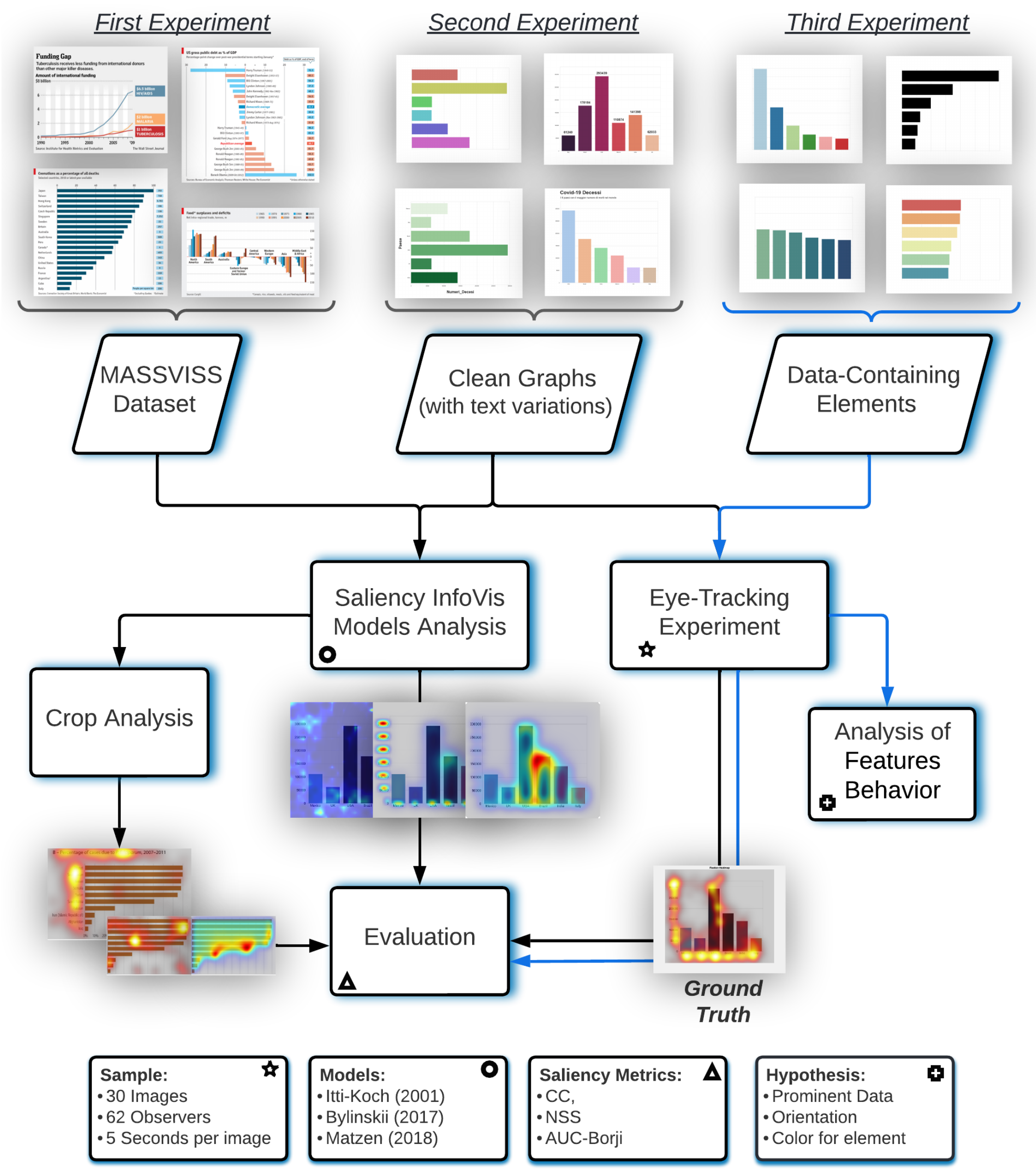


Fig. 5.1 Validation Experiments Process Overview

Because we found some shortcomings in the prediction of saliency in the second experiment, we conducted a third experiment (see Fig. 5.1): **Saliency Models and data-containing elements**. From this experiment, we could only extract some insights into the attention behavior in the data-containing elements (e.g., bars in a bar chart). The objective was to delve into the behavior of attention in the graph without textual elements. For this third experiment, we constructed a new set of clean graphs without textual elements and variations in color, orientation, and data-contained elements size. Experimental data were collected from 22 observers using the eye-tracking technique.

The following sections will present the experiments' inputs (graphs), valuation criterion (metrics), procedure, results, and insights. In addition, Appendix A has the link, and the folder description, to access the images and data resulting from the three experiments presented in this section.

5.1 Clean Graphs

As we stated previously, one of the shortcomings in the development and evaluation of the models was the exclusive use of the images provided by the MASSVIS dataset. For this reason, we produced a set of 30 statistical graph pictures named "*Clean Graphs*" (Fig. 5.2). These graphs were intended to avoid several design distractors that might cause saliency miscues in important visual components. Many of the images in the MASSVIS dataset have distracting feature elements such as logos, background images, captions explaining the data, double or combined graphs (such as bar charts and line charts), 3D graphs, and context shots (such as animal pictures). Furthermore, we could use the *Clean Graphs* to exhibit the same data set with alternative combinations of visual components. In this sense, we will know how accurate the saliency models were when some common visual elements were omitted or changed their placements.

The data visualization technique that we selected was **bar charts**. Bar charts were chosen as the principal graph because they are essential and widely used. Also, it is the most searched graph technique on Google and the most popular way to portray quantitative data, according to Visualization Universe [103], which analyses over 10,000 data points on data visualization-related queries. Furthermore, the bar graph is more adaptable than other graph techniques since it may modify some design aspects without altering their statistical nature. For instance, the bar positions can be reordered if the data is not tied to timelines.

We used the **Coronavirus (COVID-19) data-set** [104] to create the *Clean Graphs*. We showed the six nations with the highest casualties (USA, Brazil, India, Mexico, UK, and Italy) against the number of deaths at the time using data from COVID-19 Deaths Worldwide (293.439, 178.184, 141.398, 110.874, 62.033, and 61.240, respectively). This dataset was chosen because it produces a graph with a lot of variance among the subsets (i.e., each country).

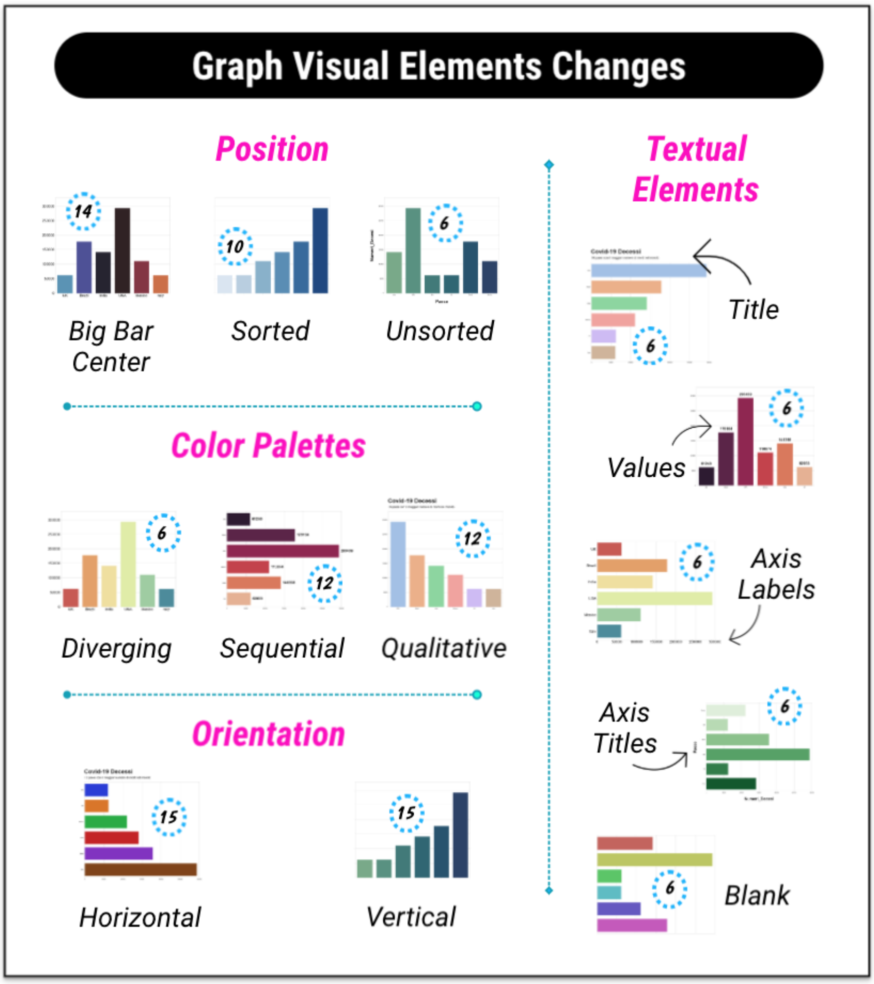


Fig. 5.2 Visual elements variations. The numbers in blue circles represent the amount of images that have that visual element characteristics.

In terms of visual elements variations, we varied whether or not the graphs featured the following textual elements: title, axis titles, axis labels, and data-containing elements labels. The *Baseline* was a white graph with no textual content. Then, as a common textual element, we added the *Axis Labels* since it is essential text information that every graph should have. Finally, we added the *Title*, *Values*, and *Axis Titles* to the remaining images independently. Figure 5.2 shows, on the right, an example for each variation made with and without textual elements.

We also included **three variables** that according to Itti *et al.* [24] are essential in determining saliency: **color**, **position**, and **orientation**. We employed three common color palettes (qualitative, divergent, and sequential) that varied in tonality (pastel and dark). We changed the color palettes to observe how the color degree affected the saliency forecast of the selected models. Since these models have a strong tendency to emphasize texts, they tend to overlook essential concepts of attention, such as the colors of an image. A sequential palette is used in 40% of the photos, while a category palette is used in another 40%. Finally, a divergent palette is used in 20% of the graphs since it is comparable to the sequential but has two colors as bounds. In Figure 5.9, *b* is an example of a sequential palette, *c* is a qualitative palette, and *f* is a divergent color.

Regarding the *position* variable, we shifted the data-containing elements to different locations (Big bar center, sorted or unsorted). Figure 5.2 shows an example of each position variation. It is clear that this cannot be done with all graphs, such as those depicting a time scale. However, we wanted to see if the saliency map differed depending on the size of the data-contained element. For instance, its position among the other bars determines the saliency of the bar representing the smaller data.

Lastly, we modified the graph's *orientation* from vertical to horizontal. This last variable was introduced since prior research revealed that there is a minor shift in saliency and reduced dispersion when the graph is rotated. Half of the images in the sample are vertical, and the other half are horizontal. Figure 5.2 shows an example of each orientation variation, horizontal and vertical.

It is important to note that by including these controlled variables, we are merely trying to check if the models are still accurate after all of these modifications or if they have no effect on the saliency prediction (see Fig. 5.2). For each *Visual Element Variation*, the clean graph can alter in numerous ways (*Features*).

Based on the variations described in Fig. 5.2, we sampled the 30 Clean Graphs utilized in the trials from a total of 120 potential variants. The number of graphs for each variation is shown in the blue circles in Fig. 5.2. For instance, 15 images have a horizontal orientation, six images have diverging color palettes, 12 images have sorted bar positions, and so on.

5.2 Saliency Metrics Validation

The MIT Saliency Benchmark project [105] is frequently used as a benchmark for evaluating saliency models. This project is an online source of saliency model performances and data sets. They score and summarize results for the most recent saliency models and maintain an up-to-date listing of other saliency data sets. The MIT Saliency Benchmark project uses measures that have been widely utilized to assess and compare salience algorithms to a set of known fixations. These metrics are classified as *location-based* (discrete fixation sites), *distribution-based* (continuous fixation map), and *Value-Based* Metrics, according to Bylinskii *et al*. [106] (average normalized saliency at fixated locations). In our research, we used three measures, one for each category.

A binary classifier assesses saliency maps in location-based metrics to see which pixels are fixated or not. As a consequence, the saliency map is used as a binary classifier to split the positive and negative point sets at different thresholds, and the area under the ROC (Receiver-Operating Characteristics) curve (AUC) is calculated [10]. The AUC statistic ranges from 0.0 to 1.0, with one being the most outstanding value and indicating that the saliency prediction is as near to human fixation as feasible. This statistic comes in three variants: AUC-Judd, AUC-Borji, and scrambled AUC (sAUC). The AUC measure, specifically the AUC-Borji, was chosen for this category because, according to Polatsek *et al*. [41], it is *"historically the most commonly-used metric for saliency evaluation"*.

Regarding *distribution-based* classification, these metrics look at the distribution of fixations rather than the "binary" fixation positions. These metrics are the Similarity Metric (SIM), Earth Mover's Distance (EMD), Pearson's Correlation Coefficient (CC), and Kullback-Leibler divergence (KL). The CC metric was used for this category. This metric *"measures the euclidean distance between the predicted saliency map and the normalized empirical saliency map"* [43]. Unlike other measures like KL, which penalizes misdetections heavily, the CC metric handles false positives and false negatives evenly [106]. Furthermore, the CC metric scale is between -1.0

and 1.0, suggesting that the connection between the saliency map and the empirical saliency map is described by a linear equation (heat map).

Only the Normalized Scanpath Saliency (NSS) metric is included in the last category, *Value-Based metric*. NSS metric "*first standardizes saliency values to have zero mean and unit standard deviation, then computes the average saliency value at human fixation locations*" [10]. As a result, a higher NSS indicates a better match between fixation locations and saliency predictions. The NSS score might be higher than one, depending on the distribution of fixations.

According to the literature, AUC, CC, and NSS measures are strongly recommended for saliency evaluation (e.g., [106, 41]). The CC measure allows us to assess how accurate the saliency map is in the heat map's most focused points in a balanced manner. On the other hand, the NSS measure helps evaluate relative significance in areas since it considers the distribution and order of fixations (scanpath). Finally, because the AUC metric evaluates the saliency of the entire image, it provides a more comprehensive assessment of the number of right points predicted by each model.

5.3 Experiments Methodology

5.3.1 MASSVIS validation

MASSVIS is a dataset composed of InfoVis images and consists of 393 images shown for 10 seconds each to 33 observers whose eye movements were tracked [90]. At least 16 observers viewed each image. Four examples of what MASSVIS contains are shown in Fig. 5.1 (*first experiment*). Of the 393 images present, we chose 110 with statistical graphs like bar charts or line charts. Because of their simplicity, we decided to make this filter and only measure with statistical graphs. This simplicity leaves us free of context images or logos, which can distract from the essentials of the graph. Additionally, we could also compare them with the behavior of the models in the case of clean graphs (see section 5.2).

In the validation using the MASSVIS images data set, each image was analyzed using two different approaches considering the initial saliency perceptions. The first approach was to make an analysis on the **complete image** with which we validated the effectiveness of the models shown in different studies. In the second, we make an **image crop**, looking only at the data-containing elements zone (the ones that actually represent the data).

For the experiment with the **complete image**, we took the statistical images extracted from MASVISS and generated the saliency maps with the three selected models. With these saliency map data, we could compare these models' behavior. To carry up this comparison (both graphically and numerically), we employed the saliency metrics described in section 5.2 by using the original Matlab code provided by MIT [105]. Considering that these saliency maps shown by all the models were consistent with the results shown by their researchers [10, 38, 42], we decided to perform a more detailed experiment on data zones (Crop validation approach).

The results of the previous experiment had not yet yielded any novel discoveries. Due to these results, we decided to perform a second one detailing the saliency more in-depth. The last experiment was to **crop the graph** on the data-contained zone and generated the saliency map with the three models. This experiment's main objective was to evaluate the accuracy of the saliency model over the graph without the title and other context elements. Based on studies on the saliency in InfoVis, it is clear that text elements within statistical images have a high impact [38–40]. As a result, some models use text detection algorithms to move the image's saliency to text elements.

However, as we stated before, our goal is to focus the saliency study on the representation of data that may be more cognitively meaningful to the observer. Specifically, for this experiment, we wanted to know how the three models perform compared to the observers' valid fixation points after cropping the image and excluding part of the text. We started the test with a sample of 11 images of MASSVIS. The images were cropped to exclude some of the text areas. A few images were also cut several times due to their composition and number of data elements. After the cuts had been made, the same analyses were repeated as the other two MASSVIS experiment approaches.

The data collected in this first scenario was consistent with prior evidence [38–40]. However, we observed that the models presented difficulties predicting saliency when removing the titles and others' attention-attracting elements. Therefore, we performed a new experiment with our eye-tracking data in the following scenario to obtain better evidence of the models' behavior in statistical images with and without attention-attracting elements.

5.3.2 Clean Graph Validation

Saliency prediction studies have shown that most attention is focused on textual elements, such as titles or axis names, as described in section 4.2. As a result, the saliency models have been modified to emphasize primarily textual elements. However, these models have also been trained and tested using real-world InfoVis images, so they have a lot of distractions. We aim to see how effectively current saliency models predict attention without those distractions and modify other design elements (textual elements, color, and location).

As a method to validate the saliency model's efficiency, we decided to perform an experiment using Eye-tracking. According to the studies presented in section 4.2.1, the saliency maps can be compared with the eye-tracking data through several metrics.

The first step was to collect the eye-tracking data from the 30 *clean graphs*. We had a total of 62 observers with the following characteristics: There were 57 men and five women; 51 aged between 14 and 18, 3 aged between 24 and 30, and 8 aged between 45 and 50. The *Tobii Pro Nano* model eye-tracker was employed in the investigation. This device is a USB-connected portable eye-tracker that records gaze data at 60 frames per second and is intended for fixation-based investigations. The information gathered from these images will be utilized to create the *Ground Truth*. Fig. 5.3 shows an eye-tracking process overview, the general data about the experiment variables, and the process developed for collecting the gaze data.

The experiment was carried out using the open OpenSesame [107] program. OpenSesame is block-based visual programming to conduct eye-tracking and psychological experiments. In appendix C, we report the program development overview, the main screens, and the process flow. The process developed in OpenSesame for the eye-tracking data collection followed the below steps (see Fig. 5.3):

1. Each observer was first placed in front of the computer to which the eye-tracker was connected.

2. The eye-tracker was then calibrated on the observer, adjusting the distance and height in relation to the screen. To make this calibration, the screen shows five points, intermittent, one for each corner and the center of the screen.

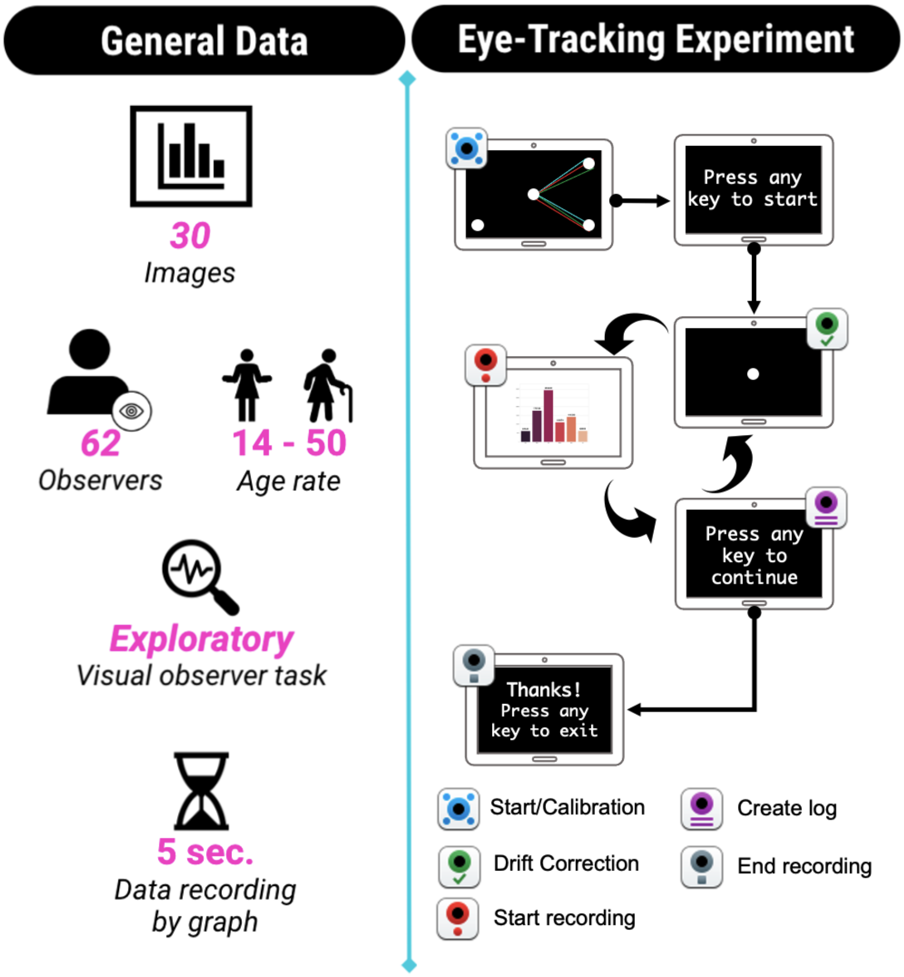


Fig. 5.3 Eye-Tracking main data and collection process for *Clean Graphs* experiment. The icons used in the process flow come from the OpenSesame program used to implement the process

3. The observer starts the experiment, pressing any key to start and then focalizing their attention in the center of the screen. At this point, the program executes a drift correction which establishes the Eye-tracker at the center of the screen (x=0, y=0).

4. The experiment began with each observer viewing each of the 30 graphs in the dataset for 5 seconds in random order, separated by a black screen. It should be noticed that between each image, a keyboard push was required, to allow spectators to move their gaze to the screen center. In this step, the program records the eye-tracker data and generates a text file containing the gaze point data required for each observer's analysis. Regarding the time selected for

eye-tracking data collection (5 seconds), is important to mention that we kept the parameters used by the model's authors [74, 39, 38] and state-of-the-art authors [41, 10, 42, 91, 106, 92].

5. As the last step, the data were cleaned to remove any irrelevant or unhelpful information before entering into the Matlab algorithm to generate the heat maps.

In parallel, to compare the Ground Truth and prediction models, saliency maps were generated on the same 30 *Clean Graphs* using the Itti-Koch, Bylinskii, and Matzen models. For the Itti-Koch model, as explained in section 4.2.2, the GBVS algorithm is commonly used and with the best perfomance [105]. For Bylinsky and Matzen models algorithms, we used the code provided by the authors (deep learning model and Matlab code, respectively).

Finally, we use the available saliency metrics algorithms (AUC-Borji, CC, and NSS) to compare the collected eye-tracking data and the generated saliency maps.

5.3.3 Data-Containing Elements

The last experiment was carried out as a deepening of the second experiment, considering some results such as the behavior of the salience in different graphs (see Fig. 5.1 *third experiment*).

Considering the relevance of the text in the attention in the graphs, we also wanted to deepen the behavior of the attention only on the elements that represent the data in the graph. This experiment is similar to the crop analysis performed in the first experiment (MASSVIS experiment). However, the crop analysis had a weakness in that the fixations points data had been collected in conjunction with all the textual elements, which meant that we were comparing with late attention to the data-containing elements. Additionally, as demonstrated by Polanski *et al.* [41], to generate a saliency model adapted to the InfoVis images characteristics, it is necessary to perform other analyses of the non-textual element's attention relevance.

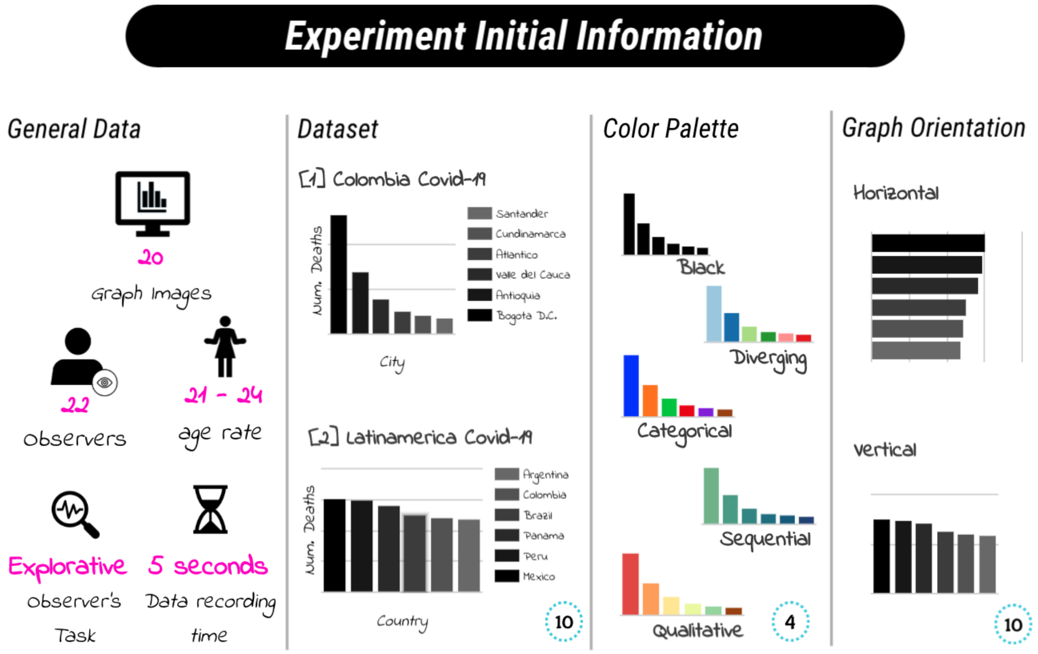


Fig. 5.4 Data-Contained Elements experiment general data. The numbers rounded by a blue circle represent the number of images for each visual element variation.

The process of this experiment is the same as performed in the second experiment with the clean graph. First, we built a second set of statistical images and carried up a gaze point collection with an eye-tracker. With the information given by the eye tracker, we analyzed the attention behavior in the data-containing elements and thus established some possible insights.

Figure 5.4 presents the initial data for the eye-tracking process. About the observers, we had 23 observers with the following characteristics: 13 males and 10 females, all aged between 21 and 24. Each observer looked at the images for five (5) seconds, continuing with the parameters used in the previous experiment. Also, the observer task was “explorative,” which means the observer only has to examine the graph.

About the graph visual elements variations (see Fig. 5.4). We eliminated all textual elements for this experiment in a new *Clean Graph* image set. We kept the variations in the color palettes, adding only a black palette in order to analyze if there was any change in the fixations in the absence of color variations. We also kept the change in orientation. Compared to the first clean graphs set, we did not consider the bars' position for this experiment. Since a more detailed experiment with more systematic variations is necessary to distinguish the impact.

On the other hand, we added the bar size variable using two different data sets. This variation would allow us to observe whether the difference between the heights of the bars generated a change in saliency. As is shown in Fig. 5.4, we use Colombia and Latinoamerica data about the number of deaths due to Covid-19. The Colombia data set has a pronounced difference in the number of deaths in its regions. The other dataset, Latinoamerica data, has more similar values, therefore, the bars have a less pronounced difference.

The experiment was executed as the previous one (see Fig. 5.3.*Eye-Tracking Experiment*). The only variation was in the transition between images because we eliminated the keyboard press as a transition. Instead, we added a visual control, the user had to look at a dot in the center of the screen, and when the program detected that the observer was looking at it, it moved to the following image. This control made the experiment go a little faster and did not exhaust the observer as much.

Overall, this experiment firstly demonstrated our hypothesis about the change in saliency when no text is present in the other important elements. Secondly, we observed several attention behaviors repeated between images, which can be taken as a baseline for improving InfoVis saliency models.

5.3.4 Experiments Limitations

Our study has shown some unknown aspects of visual attention behavior on data-containing elements in graphs. However, the study had some limitations. The first limitation is the age range of the observers. Approximately 87% of the observers were between the ages of 14 and 22. This age range is a reasonably young sample. However, like MASSVIS creators, we consider our observers to be categorized as novices because of their inexperience in understanding or reading graphs. Another experiment would have to be conducted with expert observers to verify if this influences in any way the attention on the graphs.

Second, we are aware of the sample in terms of the quantity and gender of the observers. Nevertheless, given the period of health emergency during these months, it was impossible to collect a set of observers with more uniform characteristics, even if we are not aware of any gender or age effects over saliency. However, we are unaware that these characteristics can affect the outcome of the experiments. Therefore, the variables of age, gender, and InfoVis experience are not considered for this study.

The third limitation is associated with the possible emotional reaction to the data presented. Although for both sets of graphs, we used updated pandemic data generated by COVID, only in the first set of images were the names of the countries and the numbers of deaths in each visible. However, all the observers were Italian, and even so, the bar representing Italy did not get much attention. The texts, in general, were more critical. In addition, the observers performed an exploratory task, indicating that there was no specific task that cognitively biased them toward specific data.

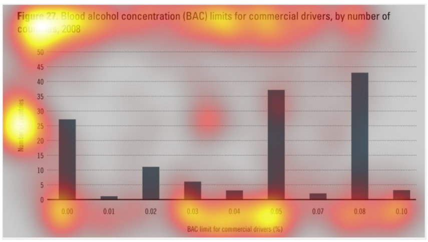
5.4 MASSVIS experiments results

The results of generating saliency maps with the three models on the MASVISS images were consistent with those presented by other researchers [93, 38, 42, 41, 10]. Table 5.1 shows each model's maximum, minimum, and average metrics values. As can be seen, on average, the Matzen model had the highest weight on the three metrics, followed by Bylinskii and then Itti-Koch.

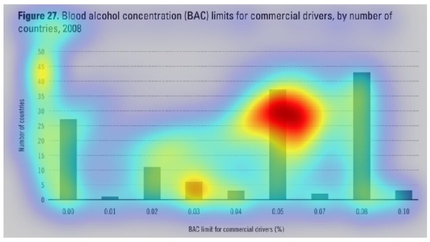
Table 5.1 Summary of metrics values in MASSVIS experiment

Metric	Value	Itti-Koch	Matzen	Bylinskii
AUC	Max	0.83	0.85	0.84
	Min	0.47	0.65	0.52
	Average	**0.69**	**0.76**	**0.71**
CC	Max	0.84	0.84	0.87
	Min	-0.2	0.24	-0.07
	Average	**0.41**	**0.63**	**0.6**
NSS	Max	1.53	1.71	1.57
	Min	-0.29	0.47	-0.09
	Average	**0.64**	**1.08**	**0.89**

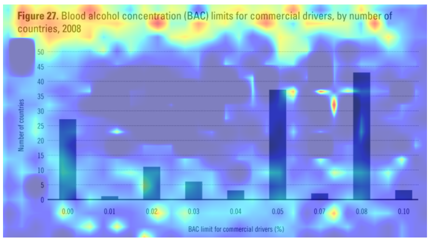
In addition to the metrics' results, the models' differences are visually distinguishable. Figure 5.5 shows an example of each model's saliency map in one of the selected statistical images. As we can see, the most accurate model in comparison with the observer's fixations (see Fig. 5.5a) is the Matzen model (see Fig. 5.5d).



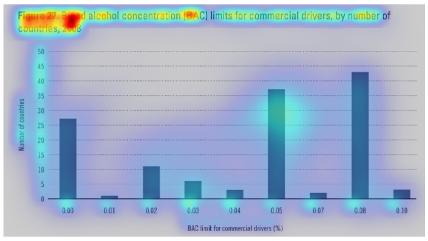
(a) Heatmap (eye-tracking data)



(b) Itti-Koch Saliency Map



(c) Bylinskii Saliency Map



(d) Matzen Saliency Map

Fig. 5.5 MASSVIS first experiment example saliency maps

Closely followed, the Bylinskii model in this image, at first glance, reveals a coincident saliency on the title and on the x-axis. However, it also reflects a wider distribution of attention rather than the Heatmap. Meanwhile, the Itti-Koch model shows the focus attention on the graph bars. Nevertheless, some images show different behavior. The result of each image can be seen in Appendix A.

Figure 5.6 summarizes the results of the three metrics used to evaluate the effectiveness of the models. In general, Bylinskii and Matzen models have the best scores in the three metrics, and the Itti-Koch model has a slightly lower percentage but is still near to them. The results also show a low behavior of the Bylinskii and Itti-Koch models, making them score below 0.0 (the lower limit of the three metrics). We took a sample of image 10% of lower values, and we observed some common characteristics:

- In the Itti-Koch model, the lower scores in the three metrics were in graphs with these characteristics: bar charts (all of the 10 lowers), the same color on each bar, small graph titles without legends, and also stacked graphs.
- About Matzen model, as Itti-Koch, has the worst scores in bar charts with the same bar colors, prominent data-containing elements area (many bars), and small or no titles. Other graphs among this group were line charts, most of them without titles and wide blank spaces.

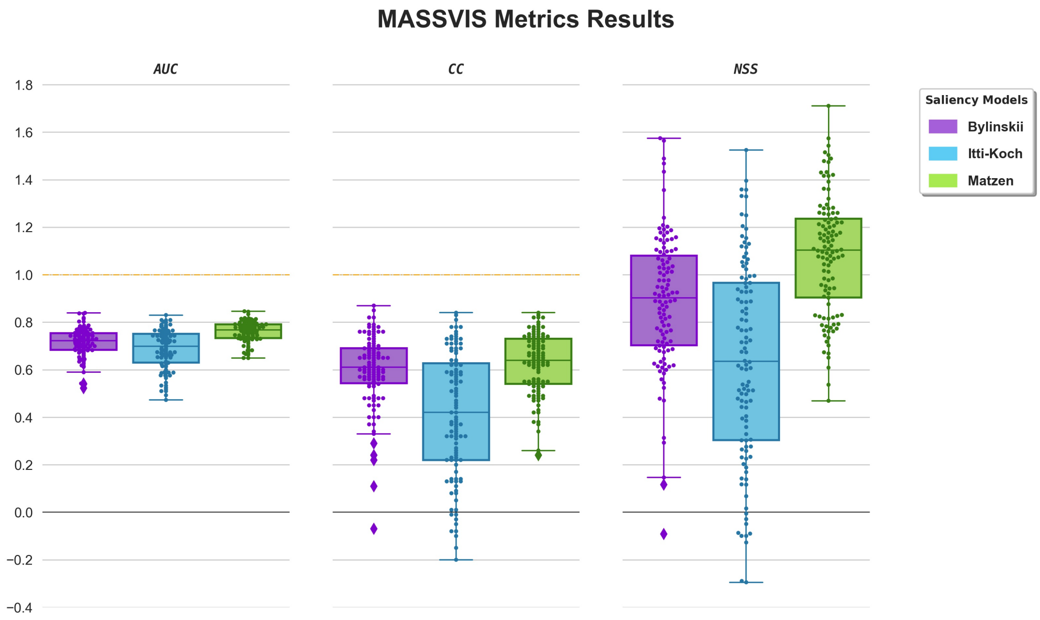
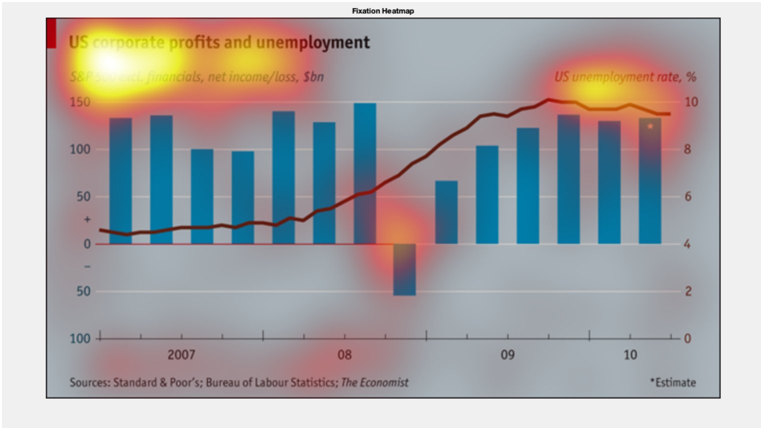


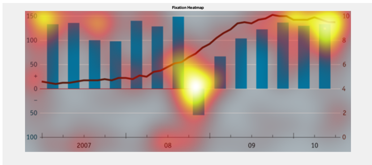
Fig. 5.6 All metrics results on MASVISS images. The orange line represents the AUC and CC metrics range limit.

- Regarding the Bylinskii model, unlike the others, almost all the graphs where it had a lower salience rating were untitled. In addition, most of them were line charts.
- If the graphs were cleaner, the model's performance was lower. By clean, we mean that they had in common white backgrounds, legends without much color or design, titles with little or no visible titles, and the same color for each graph element.

Analyzing these results, we have confirmation that Matzen, at least on these images, is the best algorithm of the three we are discussing. In addition, as already stated in the study description, the text elements have the highest salience among all the other elements in these statistical images. This occurs in the Matzen model and the fixations given by the MASSVIS data set. Although they scored low on the metrics, the other two models moved attention to the graph, which we consider a better measure of graph comprehension to some extent. Due to these results, and as explained in section 5.3.1, the next experiment (Crop Analysis) we performed was to detail these models' behavior only on the graph without context elements.

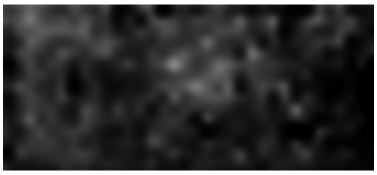


(a) Heatmap Complete Image

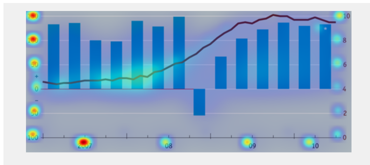


(b) Heatmap Crop Image

(c) Itti-Koch Saliency Map



(d) Bylinskii Saliency Map



(e) Matzen Saliency Map

Fig. 5.7 MASSVIS crop analysis saliency examples

For the Crop Analysis, to make the cuts on the images to get the saliency in the data-contained elements area, we had to choose an initial sample of 11 graphs. This initial sample was because the MASSVIS images are so diverse in viewing techniques, orientations, and sizes that it was complex to make automatic cuts of the images while maintaining the original fixation points data. Figure 5.7a is an example of a heatmap (observers fixations data) on the entire image, and Figure 5.7b is the cropped image heatmap. The results of this experiment are shown in Fig. 5.7. In this experiment, we made a cut-out excluding the original images' titles (see Figure 5.7b). Afterward, we took each cut and generated the saliency map in each model.

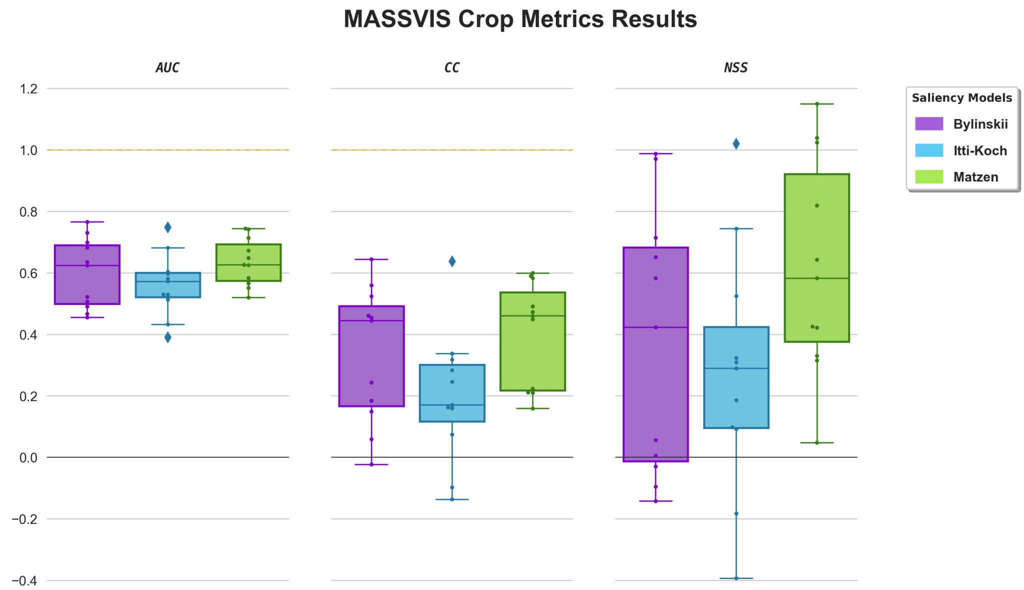


Fig. 5.8 MASSVIS crop analysis results

In general, the three analyzed models are getting into difficulties with saliency accuracy. Figure 5.8 presents the results from the metrics on the three models. Itti-Koch model decreased its performance by 9% for AUC and -23% for NSS. Although this drop in performance is significant, the model had a performance increase in crop images. It had a 4% improvement in the CC metric, which means more accurate in heat zone detection than the other models over the data-contained elements. On the other hand, the Matzen model had the most significant decrements, going from -24% in AUC to -66% in CC and 72% in NSS. These values are because the Matzen saliency model still prioritizes text areas, such as axis labels, which causes the saliency of data elements to be overshadowed.

Finally, concerning Bylinskii, it had drops of -13% in AUC, -32% in CC and -31 in NSS. These Bylinskii scores are not as bad as Martzen's, maybe because the model was trained with these same images, however, without a title, the model lost precision.

It is essential to understand whether the three models had an important decrease in their performance once the title was cropped. However, Matzen's model remains the best performer in all three metrics, followed by Itti-Koch. This leaves the question of what happens to these models when the graphs do not maintain their composition like those that are part of MASSVIS. Especially the second study on the cropped images allows us to hint at the potential difficulties these models can have if we want to detect the saliency in the data-contained elements.

The results showed that the data was consistent with our previous findings and did not diverge from the collected data. Additionally, we realized that this experiment had a limitation that could have affected not finding new salience patterns. This limitation was that the eye-tracking data given by MASSVIS were collected by showing the whole statistical image to the observers, generating distractors on the graph data-contained attention. This limitation will be addressed by the following experiments.

5.5 Clean Graph Results and Insights

This second experiment was conducted based on two fundamental premises:

- Bylinskii and Matzen models were developed using the same types of images as a basis, and their performance can be highly dependent on their structure. However, based on the previous experiment, the dependence of these models on text element detection limits their performance when the graphs change their structure.
- We do not know with certainty how the three models behave on other features like color, position, or orientation.

Based on these premises, we created the Clean Graphs with the variations explained in section 5.1. Then we compared the ground truth (eye-tracking data) with the saliency maps predicted by Itti-Koch, Matzen, and Bylinskii models. Fig. 5.9 presents an example of each Clean Graph variation, the ground truth fixation map (Eye Tracking column), and the saliency map resulting from each tested model. For instance, image (e) has only the axis titles, has a vertical orientation, and the bar position is unsorted. Bold numbers are the highest score.

Table 5.2 summarizes the models results for each graph variation. This table summarizes the results of the model's behavior with each type of variation. These results are the averages obtained in each of the metrics. The highest scores in each variation for each metric are highlighted in bold. At a glance, it can be seen that

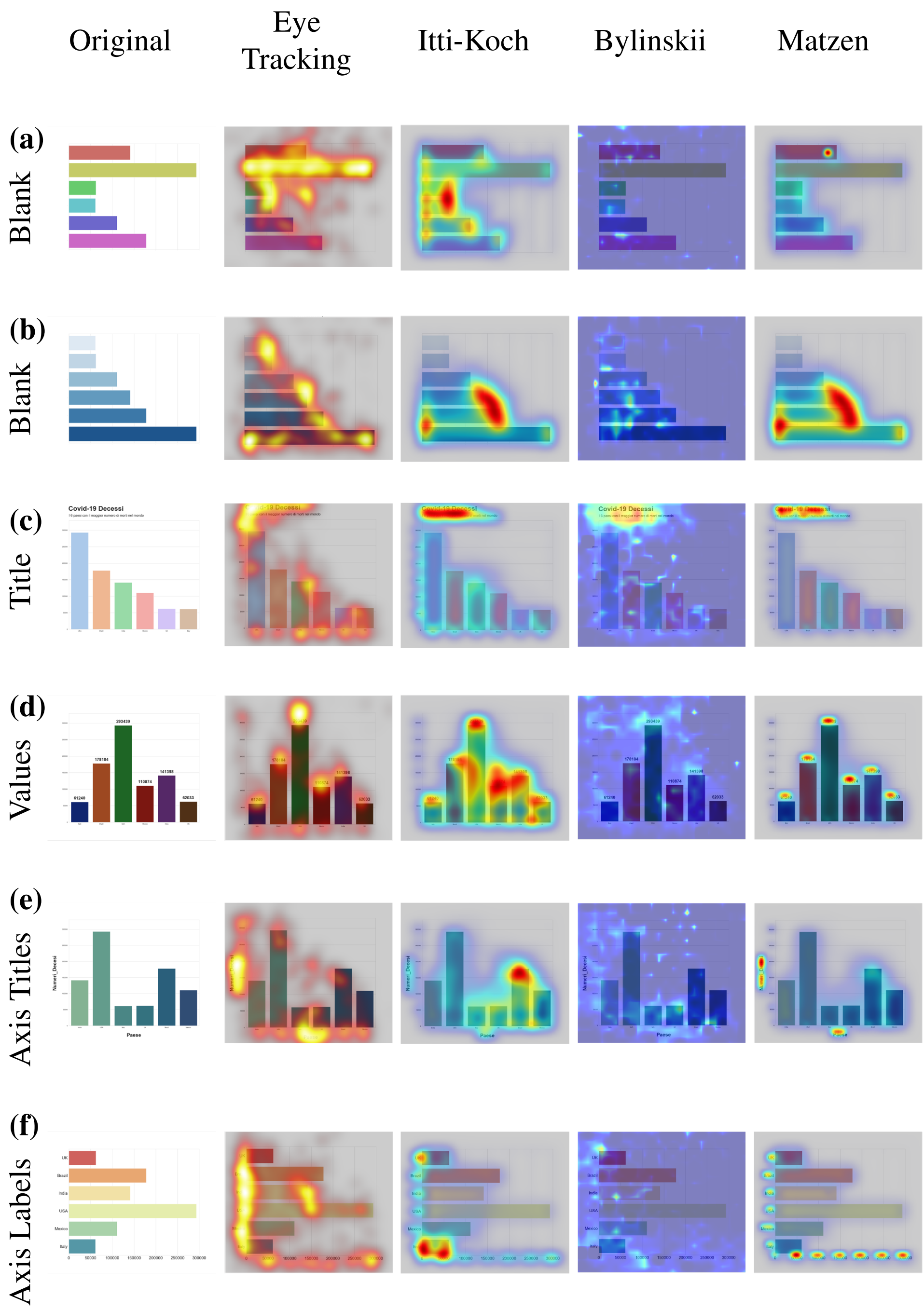


Fig. 5.9 Examples of *Clean Graphs Experimental Results*. The first column is the graph variation feature (textual element). The original graph images are displayed in the second column; each image represents a distinct version such as data **position** (sorted or unsorted), **color palette** (e.g., image *a* has a *Qualitative* palette), and **orientation** (vertical or horizontal). The third column is the ground truth. The saliency maps generated by each of the models are shown in the other columns: Itti-Koch, Bylinskii, and Matzen.

Table 5.2 Summary of metrics average in Clean Graph experiment for each text variation. Bold numbers are the highest score.

Metric	Variation	Itti-Koch	Matzen	Bylinskii
AUC	Blank	**0.76**	0.75	0.57
	Axis Labels	**0.69**	0.67	0.65
	Axis Title	0.67	0.67	**0.7**
	Bar Values	0.74	**0.75**	0.65
	Title	**0.73**	0.71	0.69
CC	Blank	0.57	**0.59**	0.2
	Axis Labels	0.41	**0.48**	0.43
	Axis Title	0.29	**0.53**	0.49
	Bar Values	0.39	**0.57**	0.28
	Title	0.55	**0.73**	0.58
NSS	Blank	1.04	**1.08**	0.37
	Axis Labels	0.66	**0.82**	0.65
	Axis Title	0.49	**0.87**	0.84
	Bar Values	0.89	**1.43**	0.56
	Title	0.84	**1.24**	1.06

Matzen is the one that obtains the best average in two of the three metrics. On the other hand, Itti-Koch obtains the best scores in only one metric. This leaves the Bylinskii model underperforming the other models.

In order to better understand these results, this section will analyze the values obtained for each of the variations organized into baseline (without text), textual, position, color, and orientation.

5.5.1 Baseline Clean Graphs

The Bylinskii model has the worst performance in all three metrics among the saliency models. For example, Bylinskii reports the lowest value of 0.02 (SD: 0.28) in the Baseline graphs for the NSS metric, whereas the maximum value for the same metric was 1.5 for the other models. The average for baseline graphs in the CC metric was 0.20 (SD: 0.16 SD), while the AUC metric was about 0.57. (SD: 0.06). The overall poor performance of the Bylinskii model indicates an apparent behavior: since it was created on images with numerous distracting elements, its effectiveness plummets when those aspects are removed. A glimpse of such behavior was seen in the previous experiment when we cropped the saliency area.

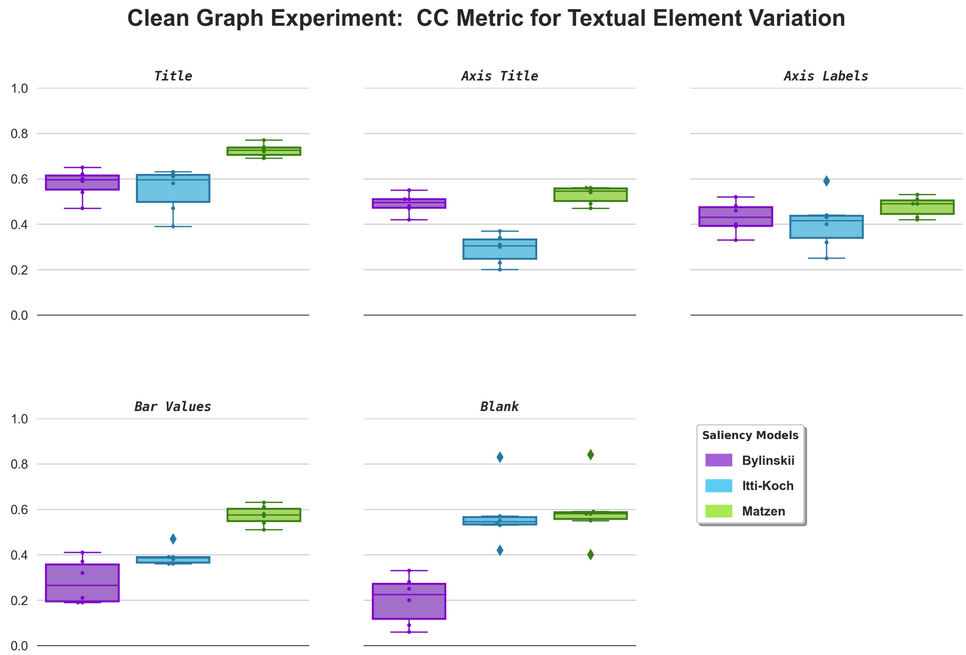


Fig. 5.10 CC metric results: baseline clean graph vs. textual elements

The Itti-Koch and Matzen models have remarkably similar overall behavior in all metrics. This behavior makes sense because the Matzen base method is a modified Itti-Koch algorithm. The distribution of their results in the CC measure (see Fig. 5.10) is relatively similar, with a maximum difference of 0.01 between their highest values and 0.07 between their lowest values. Furthermore, the Itti-Koch model has an accuracy of above 0.50 and 0.70 in the CC and AUC metrics, respectively. According to these results, the model performs quite well when no textual elements are in the graph.

The Matzen model outperforms the competition on three different metrics. The NSS achieves a value of 1.08, and it performs better than 0.50 in the other measures (0.58 CC and 0.75 AUC). These data indicate that when no text is available in the graph, Matzen's adaption of the Itti-Koch model only minimally improves model efficiency.

5.5.2 Textual Behaviour

The models report varied efficacy based on the metric they have evaluated, whether the graph contains or does not include particular textual elements. Consequently,

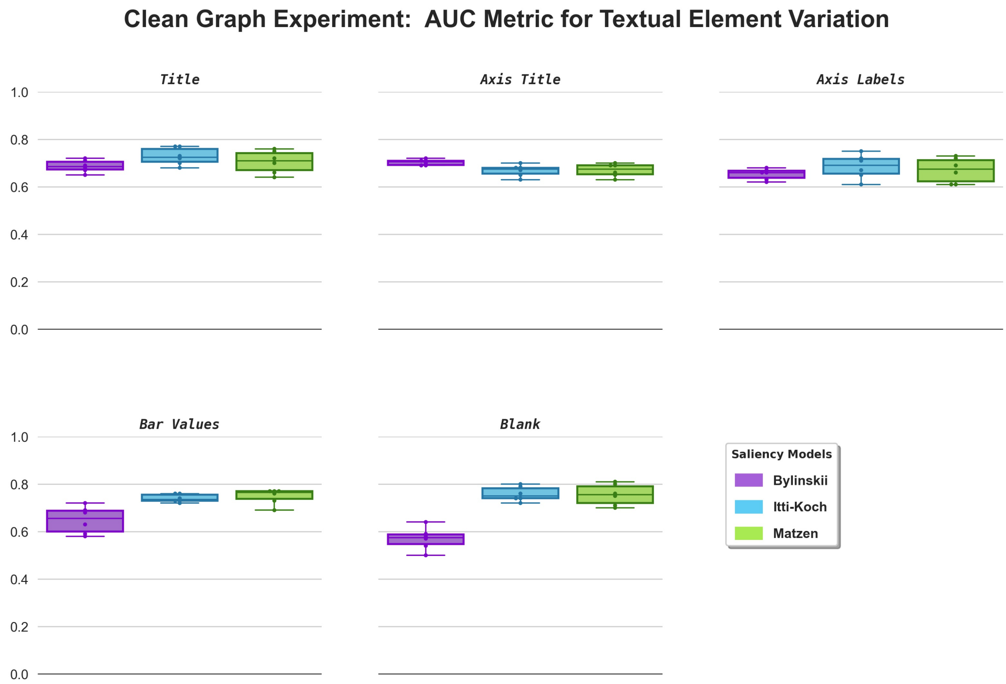


Fig. 5.11 AUC metric results: baseline clean graph vs. textual elements

unlike the baseline graphs, which exhibited identical tendencies, the score for each textual element shows significant disparities in the metrics results.

We used three saliency criteria to assess model correctness, as we did in the previous analysis: AUC, CC, and NSS (see Figures 5.11, 5.10, and 5.12, respectively). In general, the accuracy of the three models improves when the graph contains a *title* and *values above the bar*s (bar values). Given the nature of the basis images with which Matzen and Bylinskii models were implemented, this behavior was predicted. It's also clear that the Matzen model achieves the highest average values for most of the textual elements in all three metrics.

The AUC metric (Fig. 5.11), which considers all attention points when evaluating the models, reveals that the three models have accuracy with textual elements over 0.60. The Itti-Koch and Matzen models have the highest values when the graph contains values above bars. Both models focus on the text in Fig. 5.9.d, which illustrates this. Although Itti-Koch displays a wider saliency dispersion, it also highlights the AUC metric's relevant points. The Bylinskii model, on the other hand, gets the lowest score in this metric, even in graphs with values above bars (0.65). As seen in Fig. 5.9.c and f, the three models behave similarly for title and

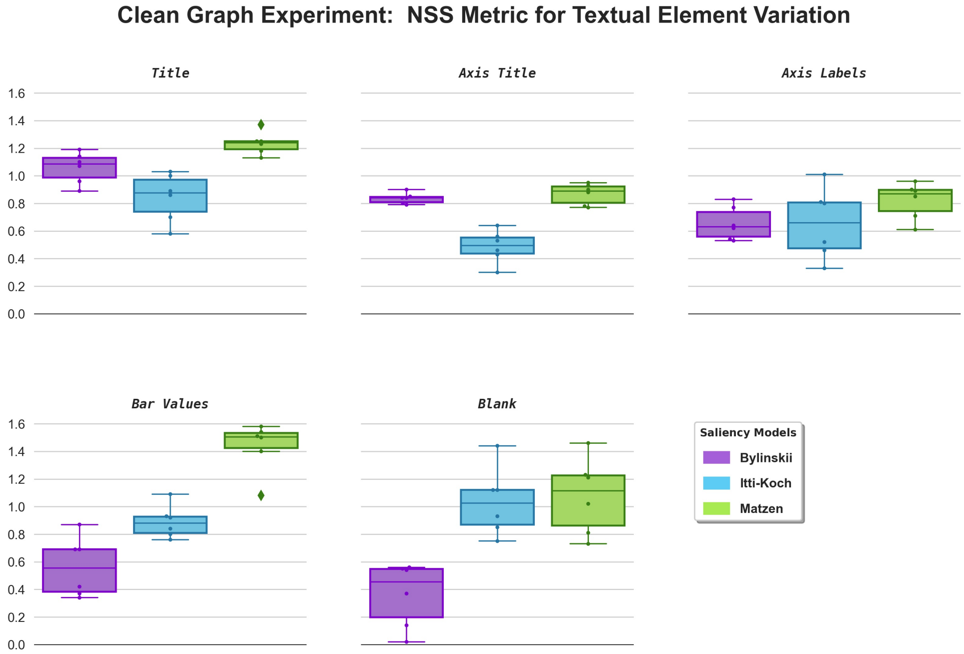


Fig. 5.12 NSS metric results: baseline clean graph vs. textual elements

axis label elements. According to the AUC metric, salience mainly depends on their recognition when textual elements are given. Because this metric measures all images, it is pretty soft with the saliency evaluation (see section 5.2). As a result, the other metrics highlight how the models have issues with the saliency of various textual elements.

The CSS metric, which compares the saliency map to the heatmap, demonstrates that the scores within each model vary significantly more (Fig. 5.10). Matzen's model distinguishes out in this metric since it received the greatest scores for all modifications in textual elements. These results suggest that the focus is still on the text, and those algorithms that include a text recognition component have an advantage. With titles, axes labels, and axes titles, this behavior may be illustrated in the Bylinskii model. Bylinskii, on the other hand, only performed 0.25 on average with the values above the bars (SD: 0.16). This impact might be because Bylinskii was trained on graphs with standard text placements (e.g., titles at the top of the graph), and it performs poorly when these are absent. In this metric, the Itti-Koch model, on the other hand, has the lowest accuracy among the models having axis title and axis label elements. As seen in Fig. 5.9.e and f. The Itti-Koch prediction lays a greater emphasis on the bars and less on the textual contents.

In general, the NSS measure (Fig. 5.12), which compares the saliency map directly with human fixation points, produced similar results to the CC metric. When saliency prediction was examined with the "values" element (see Fig. 5.9.d), the Matzen model performed significantly better than the others, in contrast to the CC metric. The NSS measure revealed the various scores, especially when the points are concentrated in certain image regions. This shift may be because the NSS metric assigns relative importance based on point dispersion. When these fixation points are extremely dense, and the saliency map reflects the same density, the NSS metric assigns a higher score. Fig. 5.9.d.illustrates this behavior. The saliency map generated by Matzen is very close to the ground truth, where the saliency is similarly heavily concentrated at each value on the bar.

Overall, the Bylinskii model's efficiency significantly varies when the graph's fundamental structure changes (e.g., without title). However, when the graph contains the title (see Fig. 5.10 and 5.12), we were wowed that the Itti-Koch model maintained efficiency of over 60% in CC and above 1.0 in NSS. Because the Itti-Koch structure does not distinguish between objects and text, our first hypothesis was that it would be the least efficient. However, when the other elements change, their performance improves. Conversely, using a text detection algorithm, the Matzen model demonstrated the expected behavior. However, it is noteworthy that it behaves similarly to the other models when the graph has only the axis labels. Visually, this result may be affected by the position of the graph (as in Fig 5.9.f respect to e).

5.5.3 Position Behavior

In terms of data-containing behavior, we wanted to see if the data position, in this case, the order of the bars, impacted the graph saliency. We divided the positions of data-containing elements into three groups: sorted, huge bar in the middle, and unsorted. In the *sorted* group, the graphs are displayed in a decreasing order by category (e.g., Fig. 5.9.b). In the *Big bar in the center* position, the bar indicating the largest data is placed in the center of the graph (e.g., Fig. 5.9.d). The last position group is *Unsorted*, where the graphs have an arbitrarily consistent position of the bars (e.g., Fig. 5.9.a). As a result of this analysis, we could observe that when the data is *sorted*, all three algorithms perform best, and when the largest bar is in the center, they perform worst. Fig. 5.13 and Table 5.3 show the results of each metric for each data position.

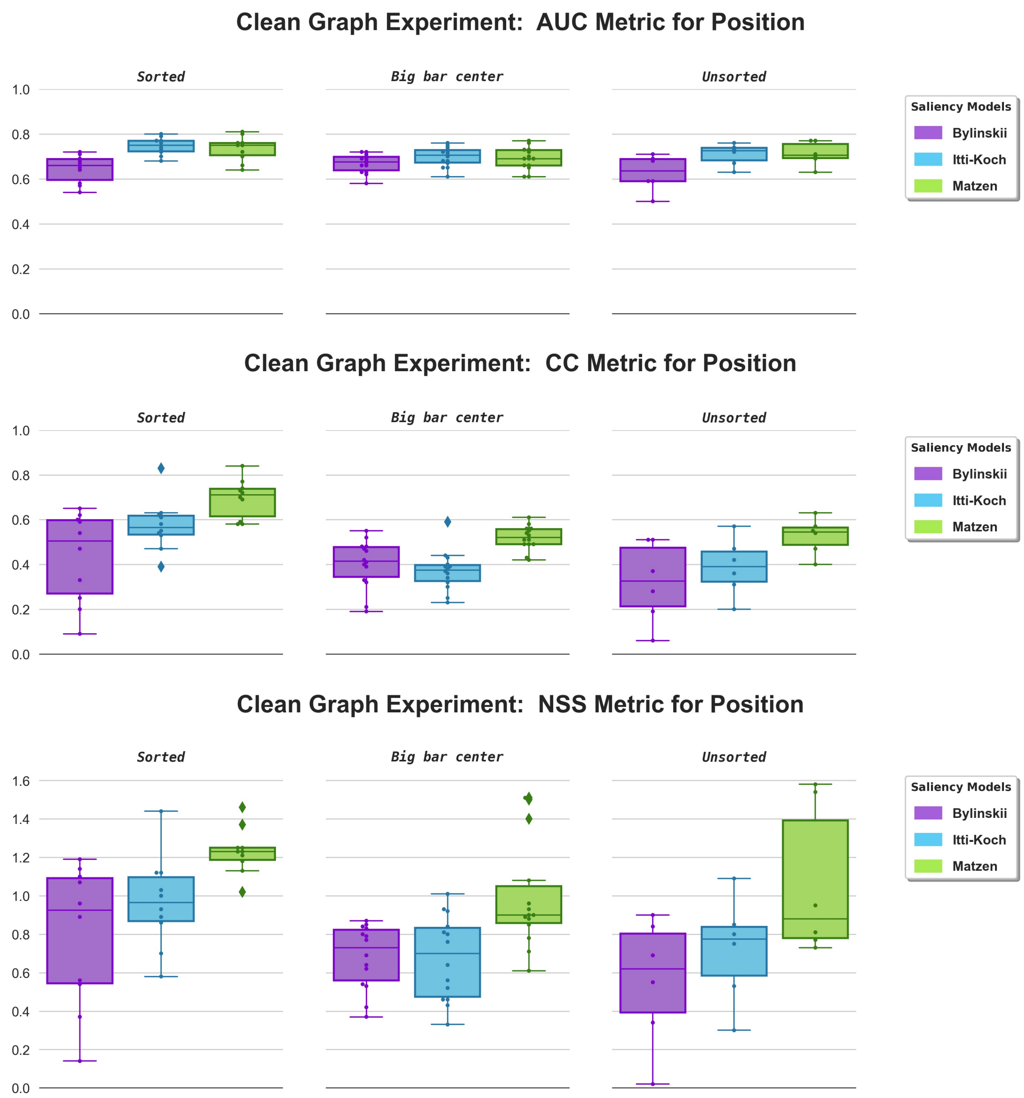


Fig. 5.13 Metrics results for each position category.

Table 5.3 Summary of metrics average in Clean Graph position variations.

Metric	Position	Itti-Koch	Matzen	Bylinskii
AUC	Big bar center	**0.7**	0.69	0.67
	Sorted	**0.75**	0.73	0.64
	Unsorted	**0.71**	**0.71**	0.63
CC	Big bar center	0.37	**0.52**	0.4
	Sorted	0.57	**0.69**	0.43
	Unsorted	0.39	**0.53**	0.32
NSS	Big bar center	0.68	**0.99**	0.68
	Sorted	0.97	**1.23**	0.8
	Unsorted	0.72	**1.06**	0.56

Based on the experimental data, when the huge bar is in the middle or is unsorted, the CC metric shows that saliency models lose at least 0.15 score points of accuracy with respect to the score obtained in the sorted position. Figure 5.9.a shows an example. According to eye-tracking research, the large bar attracts major attention. On the other hand, the three algorithms predict that intermediate bars will receive more attention.

In contrast, the NSS and CC metrics show that the large bar center position on the Matzen model had the greatest accuracy scores (see Fig. 5.13). However, the AUC metric shows that the Itti-Koch model has a slight performance gain over the other models. We presume this difference occurs because the text was placed above the bars in some graphs, such as the large bar center layout. The fixation distribution was limited on those kinds of clean graphs, as we mentioned in 5.5.1, and this narrow distribution influences the results provided by NSS and CC metrics [43].

5.5.4 Color Palette Behavior

Color is one of the most significant preattentive characteristics, and its influence has been researched extensively [13]. For this reason, in this second experiment, we employed diverse and predefined color palettes as part of our experiment to assess the colors' relevance to the saliency models under consideration. The selected color palettes were: Qualitative, Sequential, and Diverging.

The results show that the saliency models studied perform best with *sequential* palettes in terms of color palette division, CC, and AUC metrics (see Fig. 5.14). In addition, the performance of the *divergent* palettes degrades significantly in CC and NSS measures. Therefore, we anticipated that the colors chosen would substantially impact the saliency metrics. However, there was no discernible difference in the accuracy outcomes across the salience models.

Table 5.4 Summary of metrics average in Clean Graph color palette variations. The highest score for each model on each metric is in bold.

Metric	Color Palette	Itti-Koch	Matzen	Bylinskii
AUC	Divergente	0.69	0.67	0.65
	Qualitative	**0.73**	**0.72**	0.65
	Sequential	0.72	0.72	**0.66**
CC	Divergente	0.41	0.48	**0.43**
	Qualitative	**0.49**	**0.63**	0.41
	Sequential	0.41	0.58	0.36
NSS	Divergente	0.66	0.82	0.65
	Qualitative	**0.85**	**1.25**	**0.77**
	Sequential	0.77	1.06	0.64

From another point of view, as shown in Table 5.4, the behavior of the models remains similar to the previous variations. However, we could say that the classical models generally perform better when the color palette is qualitative (more difference between colors). In contrast to the classical ones, the deep learning model (Bylinskii) does not show any pattern of behavior since each metric has its highest values in different color palettes. However, these results may have a potential bias because not only does the color feature influences saliency, but also the graphs with qualitative color had the title, which makes the saliency prediction more accurate.

In comparison, the model's performance continues to be as in the other features. As a result, Matzen's saliency model continues to be the most effective, followed by the Itti-Koch model and Bylinskii.

5.5.5 Orientation Behavior

We made half of the clean graphs vertically and half horizontally (see Fig. 5.2). The accuracy of the three models in both locations is similar, as shown in Table 5.5. However, the three metrics' accuracy is consistently greater when the graph is oriented horizontally.

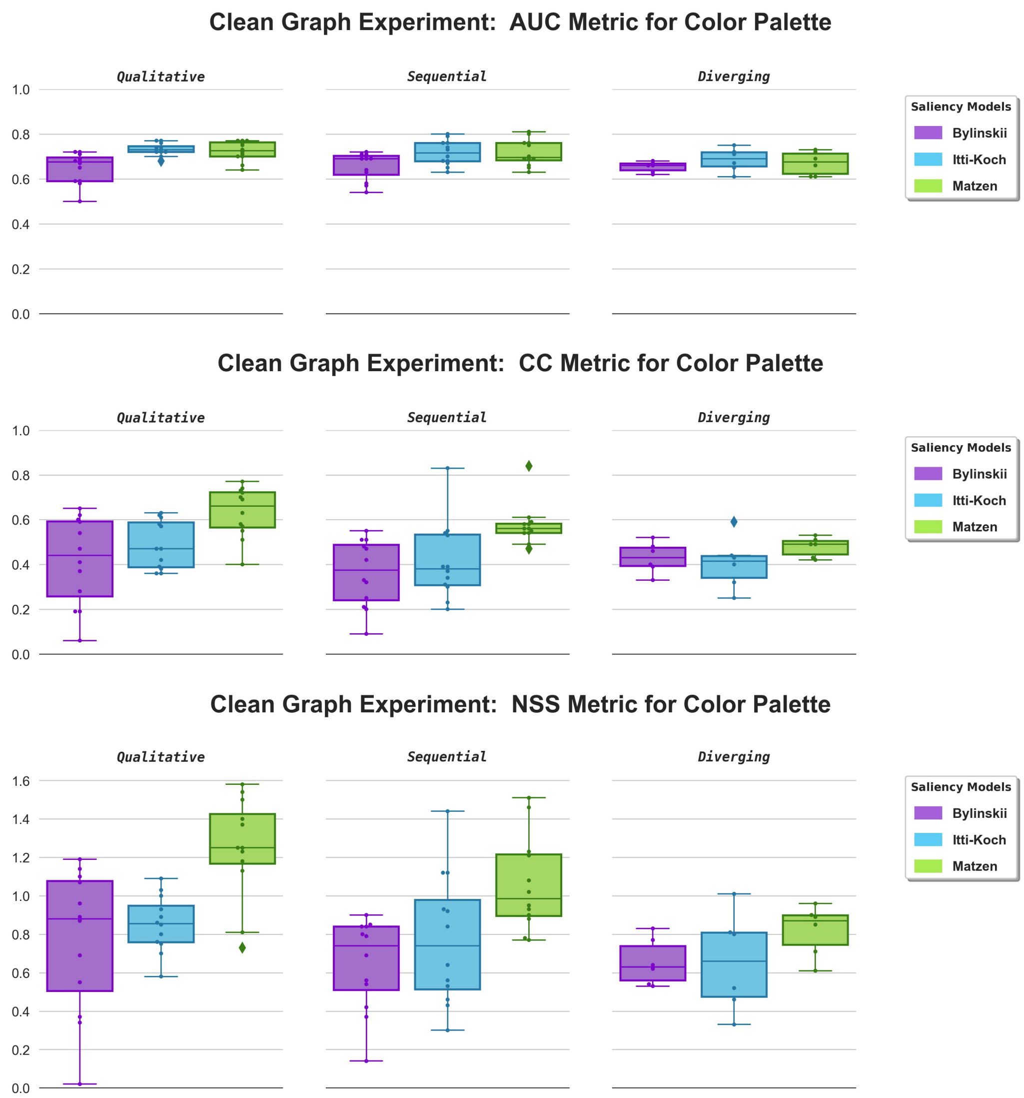


Fig. 5.14 All metrics results for each Color Palette variable

Table 5.5 Summary of metrics average in Clean Graph orientation variations.

Metric	Orientation	Itti-Koch	Matzen	Bylinskii
AUC	Horizontal	0.73	0.73	0.64
	Vertical	0.7	0.69	0.67
CC	Horizontal	0.48	0.61	0.37
	Vertical	0.4	0.55	0.42
NSS	Horizontal	0.85	1.14	0.65
	Vertical	0.72	1.03	0.74

Although the models' performance in the metrics results is consistent with the other variants, several elements stand out aesthetically. The saliency areas, for example, differ slightly depending on the graph orientation. These analyses, however, are outside the focus of this research and are part of a larger project. Overall, the Bylinskii model is the least successful, whereas Matzen is the most effective.

5.5.6 Comparison MASSVISS and Clean Graph experiments

Regarding the premises established at the beginning of this experiment, we could verify that, indeed, the models that have been generated or tested with the MASSVIS data set have a decrease in their performance.

Table 5.6 Comparison between MASSVISS and Clean Graph results.

Metric	Value	Itti-Koch Massvis	Itti-Koch CleanG	Matzen Massvis	Matzen CleanG	Bylinskii Massvis	Bylinskii CleanG
AUC	Max	0.83	0.8	0.85	0.81	0.84	0.72
	Min	0.47	0.61	0.65	0.61	0.52	0.5
	AVG	0.69	**0.72**	**0.76**	0.71	**0.71**	0.65
CC	Max	0.84	0.83	0.84	0.84	0.87	0.65
	Min	-0.2	0.2	0.24	0.4	-0.07	0.06
	AVG	0.41	**0.44**	**0.63**	0.58	**0.6**	0.4
NSS	Max	1.53	1.44	1.71	1.58	1.57	1.19
	Min	-0.29	0.3	0.47	0.61	-0.09	0.02
	AVG	0.64	**0.78**	1.08	**1.09**	**0.89**	0.69

Table 5.6 shows some metrics value references (maximum, minimum, and average) for each model in the two initial experiments. As can be seen in Table 5.6, the Itti-Koch model obtained better results in the experiment with clean graphs. On the other hand, the Matzen model has similar behavior in both experiments, although it

obtained higher scores in MASVISS, its average remains very similar between both experiments. The Bylinskii model showed a pattern of decreasing performance in all three metrics.

For more detail about the data obtained in this experiment, we created an interactive dashboard in Google Data Studio [1].

About the second premise, how the three models behave on other features (color, position, and orientation), we found some insights. The position and orientation of the data-contained elements apparently influence saliency behavior. Although the Clean Graph dataset was created thinking also about the variation of these characteristics, it was complex to know how much real influence it can still have on the prediction and, therefore, on the saliency performance. Further experiments on each element would have to be performed to get a deeper insight into the issue. In the following experiment, we investigate in more depth, on another clean graph dataset, how two of these elements can be important to complement the prediction of saliency in data-contained elements.

5.6 Data-Containing Elements Experiment

Based on the previous experiment results, we established another two premises: (*i*) saliency models behavior in graphs where textual elements are eliminated, and (*ii*) what could be those visual elements that influence the data-containing element's salience. It is essential to highlight that in the previous experiment, our clean graphs had at least the names of each point on the scale (axis labels) as a fundamental element of a graph. Our clean graphs eliminate any textual element in this experiment, aiming to deepen the saliency analysis only in the data area. Regarding the second premise, we constructed some hypotheses about saliency behavior that we solved with the collected data to understand more about saliency in data-contained elements. As in the previous experiment, we collected data with an eye tracker but on the second set of images (see Fig. 5.4).

[1] https://bit.ly/2sdExpe_Dashboard, Experiment Interactive Dashboard, last visited on October 14, 2022

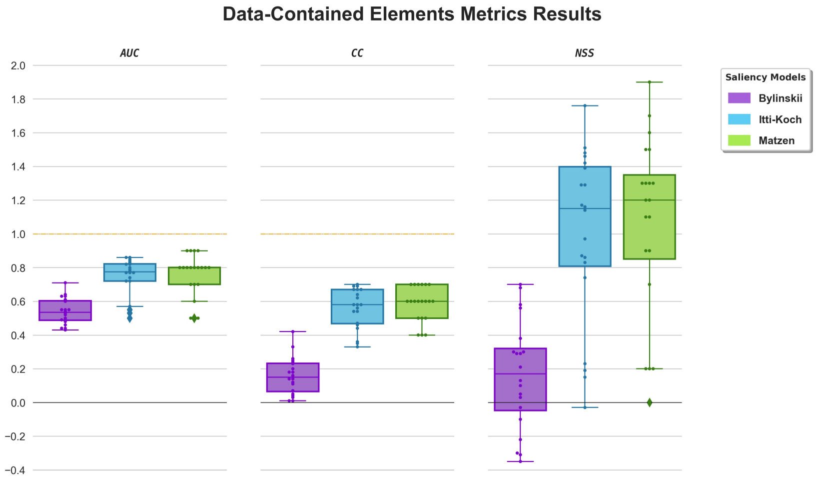


Fig. 5.15 Results of Data-Contained elements experiment for each Saliency Metric

5.6.1 Saliency Models Performance

This experiment rectified our hypothesis that some saliency models would drop their performance significantly if we completely removed the textual elements from the graph (see Table 5.7).

Table 5.7 Summary of metrics in Data-Contained Elements (**DataC**) vs. Clean Graph (**CleanG**) experiment

Metric	**Value**	**Itti-Koch CleanG**	**Itti-Koch DataC**	**Matzen CleanG**	**Matzen DataC**	**Bylinskii CleanG**	**Bylinskii DataC**
	Max	0.8	0.86	0.81	0.87	0.72	0.7
AUC	Min	0.61	0.5	0.61	0.49	0.5	-0.35
	Average	0.72	**0.75**	0.71	**0.74**	**0.65**	0.17
	Max	0.83	0.7	0.84	0.73	0.65	0.71
CC	Min	0.2	0.33	0.4	0.35	0.06	0.43
	Average	0.44	**0.56**	**0.58**	**0.58**	0.4	**0.54**
	Max	1.44	1.76	1.58	1.88	1.19	0.42
NSS	Min	0.3	-0.03	0.61	-0.04	0.02	0.01
	Average	0.78	**0.99**	**1.09**	1.05	**0.69**	0.16

The Bylinskii model was the worst-performing of the three models. Its values dropped by 320%, which is significant for performance. This decrease was especially evident in the NSS metric, which is the one that validates the saliency distribution. Furthermore, the metric that proves that this model is indeed strongly attached to a certain form of graphs is AUC. In this metric, which is the one that does not penalize erroneous predicted salience points, Bylinskii also had a significant decrease (from

0.65 to 0.17). This may mean that this model has difficulty predicting saliency solely on the visual elements representing the data.

The only model that showed improvement in this experiment was Itti-Koch (between 3.8% and 21.3% percentage of improvement). We assume that this is because by eliminating the text, the model would have no distractors outside of the three channels it handles (color, orientation, and intensity). Matzen model remains the highest scoring in all metrics (see Table 5.7 *Max values*). This behavior of the Matzen model was expected, since there is no text in the graph, the prediction is made by the Itti-Koch algorithm modified.

The results of each of the metrics on each of the elements that were varied in this experiment can be seen as following: Color Palette, Fig. 5.19, Orientation Fig. 5.20, and Data Set Fig. 5.21. Since the figures mentioned above are oversized, they can be found at the end of the chapter.

For more detail about the data obtained in this experiment, we created an interactive dashboard in Google Data Studio [2].

5.6.2 Features Behavior Hypothesis

Following the completion of all the studies, we were able to identify specific behaviors that we believed could improve the accuracy of the saliency prediction. These observed behaviors were categorized into three hypotheses, which were substantially answered using the clean graph data. The next sections detail each of these hypotheses and our response proposition.

The attention is mainly on the more prominent data-containing element

Our first hypothesis was about the height or size of the elements. We think that most of the attention would be on the largest bar, with any of the datasets used for the experiment. As shown in Fig. 5.16, Bar 1, which represents the largest data, is not the one with the highest average number of fixations. For the first dataset, where there was no significant difference between the bar sizes, it is Bar 2 that has the highest average number of fixations. In the second dataset, which has more equidistant bars,

[2] https://bit.ly/data-elements_results, Experiment Interactive Dashboard, last visited on October 14, 2022

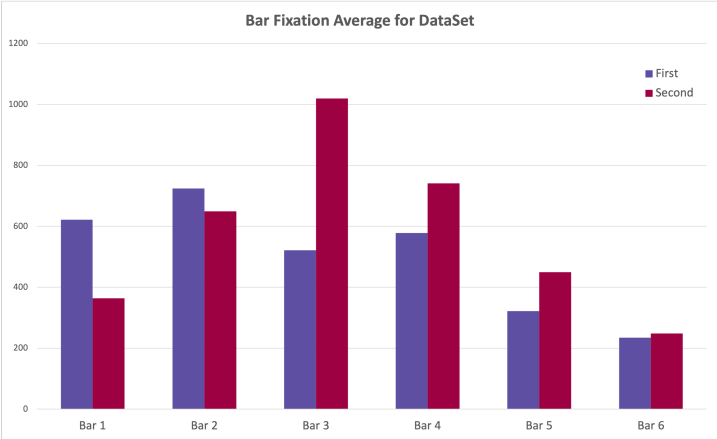


Fig. 5.16 Bar average fixations on the first and second datasets. The first dataset had a significant difference between the values of the subsets, the second one had more equidistant data. *Bar 1* represents the largest bar, and *Bar 6* is the smallest

the bar with the most significant average attention is the third (intermediate bar). Based on the averages of fixations, we can see that the attention behavior on the largest data-containing element varies according to its relationship with the others. Being the largest data-containing element does not give it the most attention on the graph.

The orientation influences the attention behavior

In Fig. 5.17, we presented the fixations average on each graph bar depending orientation (vertical or horizontal). This hypothesis is one of the most important we had since we noticed a change in the attentional behavior according to the orientation in the first experiment. As can be seen in Fig. 5.17, the fixation rate of each bar changes significantly with bar orientation. First, concerning the largest bar (B1), under the first dataset, the average number of fixations increases by about 50% when going from vertical to horizontal; in the second dataset, the increase is more significant, the fixations increase by 400% (on average). This aspect is crucial to developing a saliency prediction, as a graph can completely change the attention if it changes the orientation. Similar behavior can be seen in bars 2 and 5.

However, when the data-containing element becomes smaller, it gets more saliency when the graph is vertical. This performance is most visible in the second dataset, where the increase in attention on *Bar 3* and *Bar 5* is at least 100 fixations

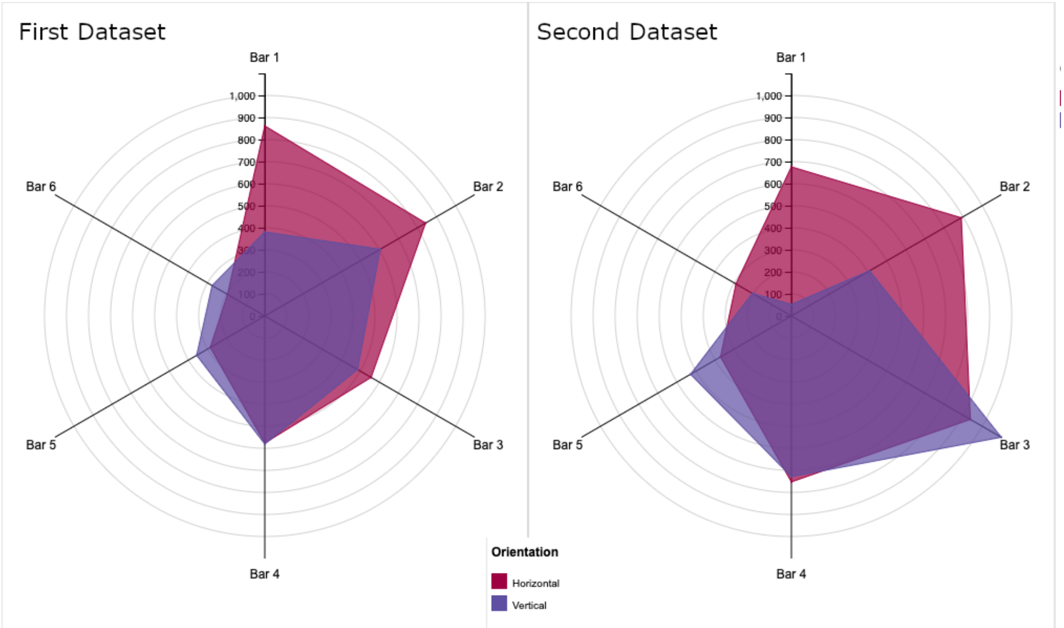


Fig. 5.17 Bars Average Fixations. Attention behavior with the first and second datasets. The colors represent the bar behavior in a different orientations, Purple vertical and Red is horizontal.

points more. On the other hand, it can also be seen that bar 4 is relatively constant in its average fixations regardless of orientation change. This change may be because it is the most central bar in the plot. Thus, the data collected show a noticeable change in the fixation points depending on the orientation of the graph. Such a change can drastically influence the saliency to either increase or decrease depending on the element size.

Attention varies on the bars according to their color.

Our theory is that, apart from the bar size and graph orientation, the graph's color also influences the number of fixations on each element. This hypothesis arises from what we observed from the behavior of the current saliency models evaluated in the first experiment. The evaluated models did not indicate significant variations if we changed the color palette. For this reason, we investigated the influence of color on the saliency of data-containing elements.

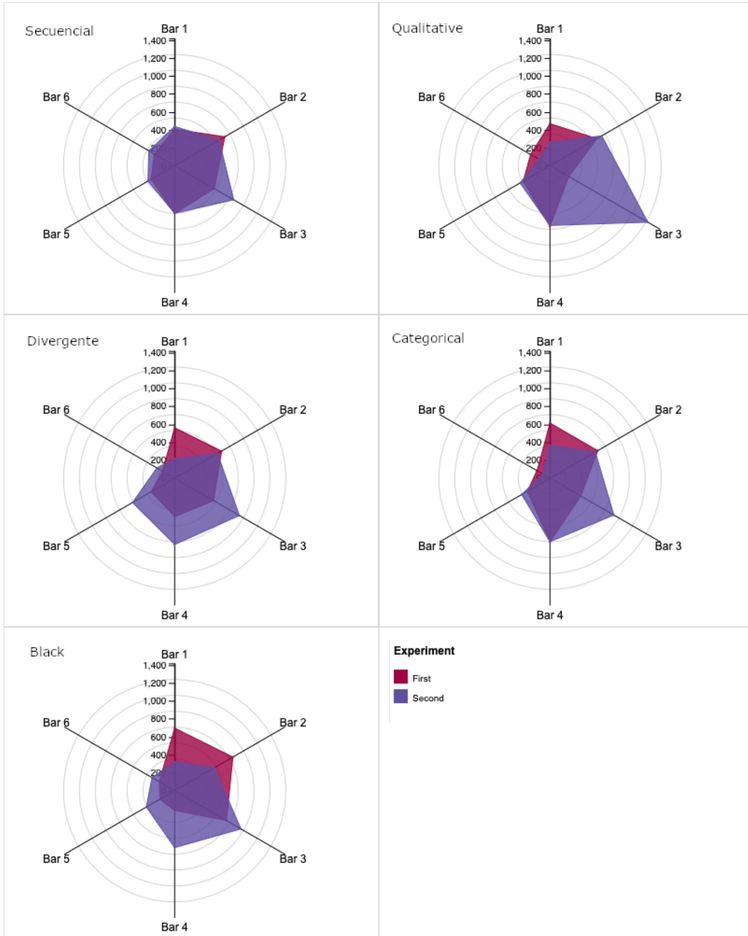


Fig. 5.18 Bars average fixation in each color palette. Attention behavior with the first and second dataset

In Fig. 5.18, we present each bar's average number of fixing points and their variation in each color palette used. Also, we show the attention behavior for each data set.

With the first dataset, the changes in the fixations are most prominent in the central bars (3 and 4). When the color palette is qualitative, bar 3 significantly loses attention, dropping from 600 points on average to minus 200. This decrease may be because the bar is a light color between two dark colors (given the palette handles' degradation). On the other hand, on the same dataset, bar 4 is another bar that changes significantly when the palette is diverging (degrade between two colors) from a maximum point of 700 fixations on average to only 230. In this palette, that bar has a light color, just like bar 3.

In the second data set, the behavior is roughly similar. However, in this case, *bars 3* and *bars 6* had a significant change. When the palette is qualitative, contrary to the first dataset, the bar reaches an average of 1400 fixation points. However, we could not say that is only due to the color palettes, given the opposite change between the first and the second dataset. On the other hand, bar 6 firmly declines, from 200 to less than 100 points, with the categorical and qualitative colors. It showed that the attention is less distributed because it does not reach the lowest value (the smallest bar in this case).

In general, we do not see drastic changes in the saliency of the bars according to their color type; there are no significant variations, and the maximum is 200-400 points. However, it is noteworthy that the Sequential and Divergent color palettes showed a relatively symmetrical dispersion behavior. That is, as shown in Fig. 5.18, both color palettes maintain a more or less equal distribution over most of the data bars. Also, the data show that the changes are notable between datasets and not only between colors.

5.7 Models Comparison

We observed how the models changed their performance based on data collected in the three experiments. Some improved while others declined, depending on the type of visual elements included in the graph.

Table 5.8 Experiments comparison using t-test method.

Model	**Experiments**	**CC**		**NSS**		**AUC**	
		t	**p**	**t**	**p**	**t**	**p**
Bylinskii	*MASSVIS vs. Clean Graph*	7.50	1.06e-09	5.26	2.30e-06	6.13	9.81e-08
	Clean Graph vs. Data-Contained	6.28	9.31e-08	5.94	7.44e-07	5.60	3.08e-06
Matzen	*MASSVIS vs. Clean Graph*	0.66	0.51	0.49	0.63	2.74	0.008
	Clean Graph vs. Data-Contained	0.097	0.92	0.74	0.46	-1.24	0.23
Itti-Koch	*MASSVIS vs. Clean Graph*	-1.90	0.063	-2.40	0.020	-2.70	0.01
	Clean Graph vs. Data-Contained	-3.12	0.003	-2.03	0.052	-1.05	0.30

In order to get a clearer idea of these changes in performance, we use a *t-test* to analyze the changes in each model for each experiment. The *t-test* is a statistical test that indicates how similar two samples are. This test outputs two values: ***t*** and ***p***. The ***t*** value represents the number of times the data samples are different from each other [108]. The p-value (***p***) is the probability that the results from the sample data occurred by chance [108]. If $p < 0.05$ there is a statistically significant difference, and if $p > 0.05$, there is no statistically significant difference.

The hypothesis is that if the samples are similar, the model has a uniform behavior regardless of the type of graph. Otherwise, this *t-test* will show how different the samples are. Table 5.8 shows the results of the t-test (using RStudio). For each model, we compare the first and second experiments and the second and third experiments. We performed the comparisons this way because the experiments ranged from graphs with more contextual elements to graphs with no contextual elements.

Starting with the Bylinskii model, all three metrics and comparing the three experiments' p-values are largely less than 0.05. This means that the values obtained by Bylinskii in each experiment are significantly different from each other. On the other hand, we can see that the proportion of the difference between each experiment (the t values) is above five. These t values show that Bylinskii's performance in the Clean and Data-Contained experiments was at least 5 times lower than in the MASSVIS experiment. With these results, it is evident that this type of model (deep learning) is strongly attached to the training images,

On the other hand, the Matzen model, between the first and second experiments, maintained overall similar results. The t-test showed that in the NSS and CC metrics, its p-value was greater than 0.05, which may denote that the two samples are similar in statistical terms. Only the AUC metric showed different behavior. The above behavior may be because the Matzen model, as a result, shows less saliency distribution, the attention points are more evident, and this causes fewer pixels to show saliency. To a certain extent, this behavior is penalized by AUC since this metric counts positives and negatives without measuring the number of attention points over a region.

Comparing the second and third experiments, the Matzen model showed similar behavior. The CC and NSS metrics maintained their performance, but the AUC metric improved. The t-value showed a negative behavior, which means that the second sample entered in the test, in this case, Data-Contained elements, is larger than the first one (Clean Graph). In this case, Matzen revealed a slight improvement in this metric, which may be due to the lack of text in the graphs, making the saliency more distributed over the data area and not limited to the texts.

Finally, Itti's model was the only one that showed a constant improvement among the three experiments. Unlike the other models, all the t-values obtained were negative. Additionally, in almost every metric, the p-value is less than 0.05, meaning there is a statistical change between the samples. The above evidence suggests that the Itti-Koch model performs better if the graph has fewer contextual elements. The behavior of Itti could be since it is a model developed with the basic concepts of vision, it does not take into account top-down aspects, such as the importance of the texts in the graph. Therefore, the less context information (distractors), the better the saliency prediction.

Altogether, it is evident that Matzen is the model that demonstrated the most stable behavior during the three experiments. Additionally, this model obtained the highest scores among the three metrics. Its minimum values reached a limit of -0.07 when Bylinskii obtained scores as low as -0.35. The only metric Matzen did not achieve the highest value was AUC. As explained above, Matzen may perform saliency prediction in a more focused manner, especially on text, leaving out pixels less important to the algorithm. This could affect the values given by the AUC metric.

It is important to emphasize that we know that *text* is a fundamental part of the graph and that its removal is not proof that the current models do not perform well. Our premises were based on the fact that although the text has the most attention, the other visual elements show us the data, knowing their attention behavior is also essential.

Since **Matzen's model** was the one that kept the best scores constant in all three experiments, we will use this model for the next part of the research. In the following two chapters, we will present two approaches to how this model can be used in the graph design process.

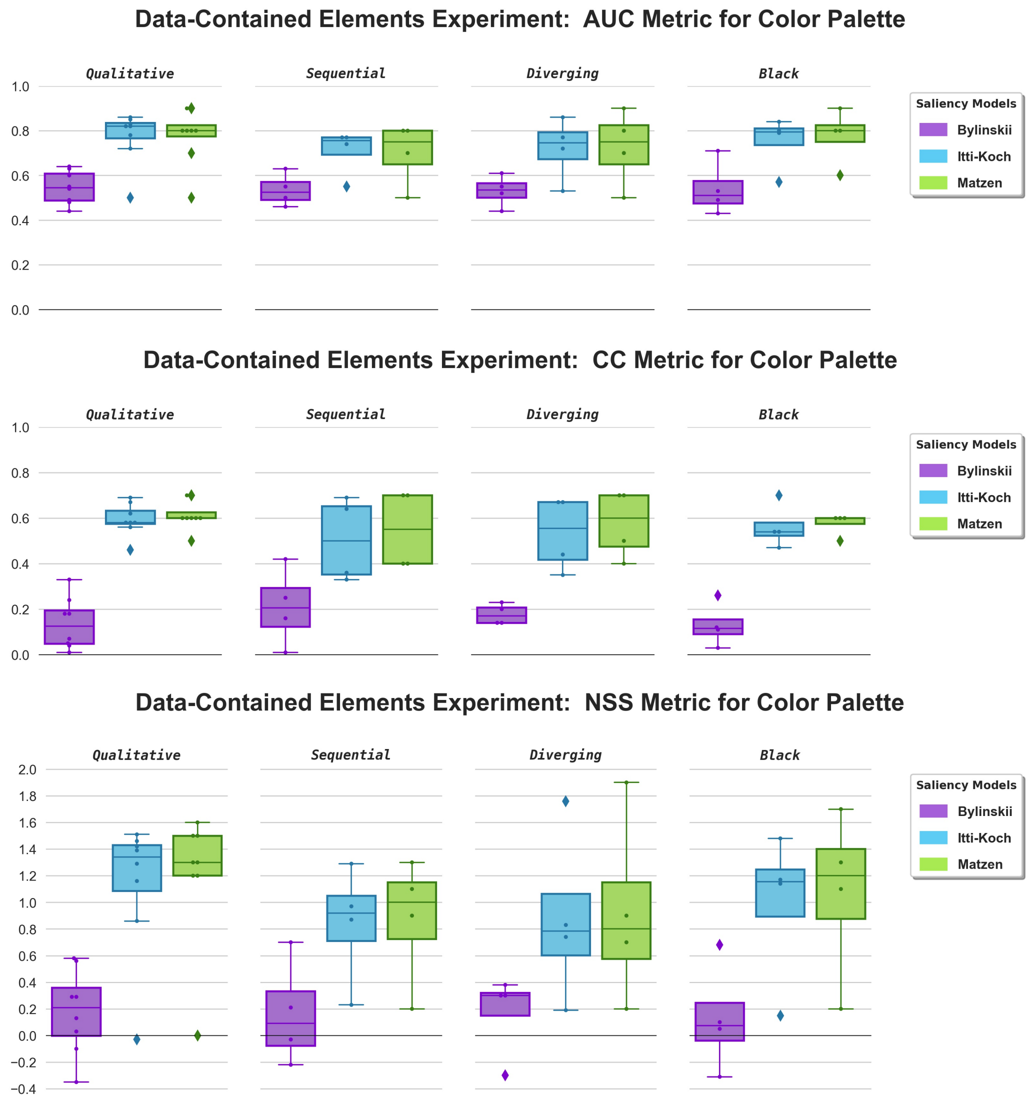


Fig. 5.19 Color Palette in Data-Contained elements - metrics results

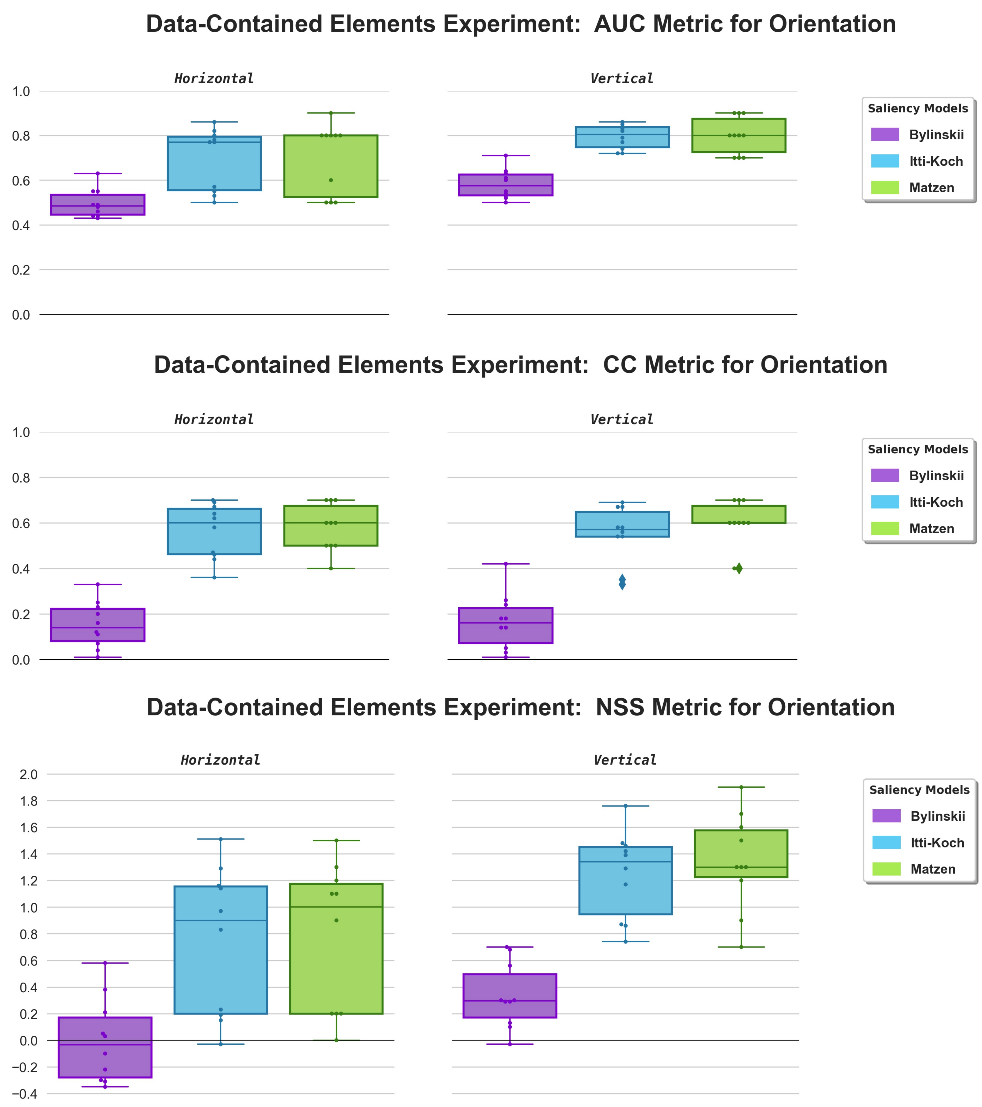


Fig. 5.20 Orientation in Data-Contained elements - metrics results

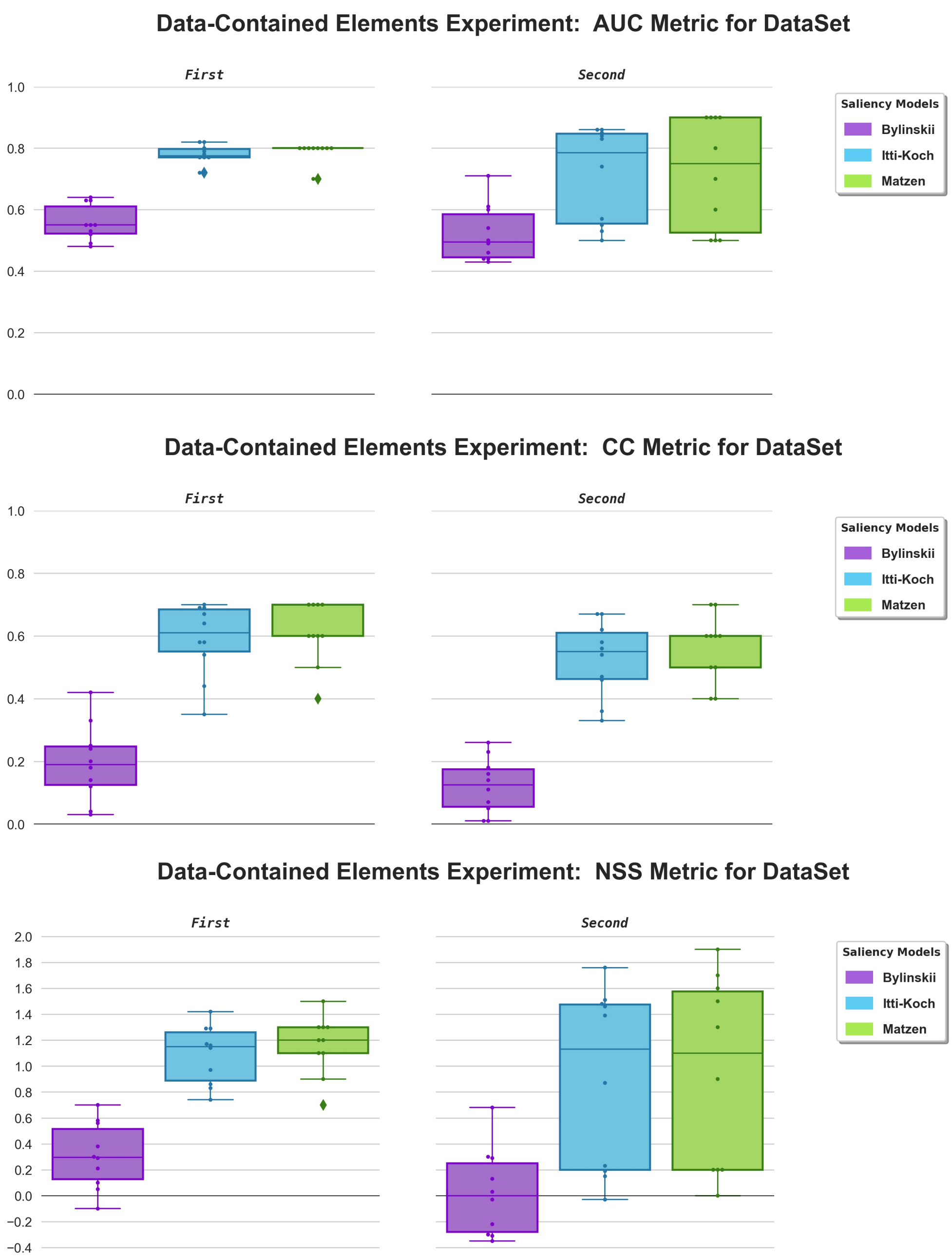


Fig. 5.21 Data set in Data-Contained elements - metrics results

Chapter 6

Using Saliency Prediction as a Design Tool

At this point in our research, and taking into account the results presented in the previous chapters, we have three important insights regarding saliency prediction and graphs design:

1. There is a vast amount of knowledge about how human vision operates and how it affects the cognitive process and, therefore, decision-making in graph design (chapter 3).
2. There is a gap between the graphic designer and the knowledge about how the design features influence saliency and, therefore, the observer's performance(chapter 4).
3. Existing saliency models for DataViz can be used as a tool to validate visual attention on a graph with reasonably good performance (chapter 5).

Bringing these three insights together, we developed two approaches for using saliency prediction within the graph design process. The first approach is used saliency prediction as a **Design Tool** to help the graph designer find the combination of visual elements (color, orientation) that will make the most relevant data more noticeable to the observer. The second approach is used saliency prediction as **Measurement Tool** to validate how each design decision can impact the attention points in the graph. These approaches are proofs-of-concept that demonstrate how a saliency model might assist iterative design and automate particular graph design activities.

This chapter will describe the **Design Tool** approach oriented to draw the observer's attention to specific relevant data selected by the graph designer at design time. For the sake of simplicity, we will call this approach as **Design Tool** meaning the use of saliency prediction as a design tool. The second approach,**Measurement Tool**, will be described in Chapter 7.

6.1 Motivation

Data visualization has become a fundamental part of decision-making in any company or industry. A large amount of daily data has created a need to develop design tools that different user profiles can use. Those in charge of creating graphics, *graph designers*, can have different profiles, such as statisticians, engineers, scientists, and administrators, among others. However, due to the broad profile spectrum that a graph designer can be, the graphic validation process becomes complex from the design point of view. Nowadays, in companies, the process of validating the design and clarity of a graphic is done by presenting the chart to a group of experts who give a final judgment based on their expertise (see section 6.5).

On the other hand, many researchers have studied the impact of visualization on an observer's decision-making [16, 12, 11, 63]. One of the most recent is the study presented by Milutinović *et al.* [11], they demonstrated that *"specific combinations of saliency form and visualization method seem to be favorable in terms of gained decision quality and attribute attachment."* They suggest that saliency in visualization should be considered. Designers might disclose potential risks for biases in DataViz by raising awareness about saliency effects in visualizations using saliency analysis, as with the Matzen model [11]. Finally, they highlighted the importance of development tools for visualization evaluation, which should examine the effects of spatial layout and the usage of visual factors on saliency in visualizations from a broader viewpoint.

Furthermore, according to Jänicke *et al.* [2], *'every visualization creator would be curious about how a visualization is perceived and what the observers learn from it".* Visual saliency models could predict which areas attract the observer's visual attention. With this information about *where* the observer's attention will be focused, the graph designer could know beforehand where his chart will start to be read and also could know how well the chart is aligned with the relevant data. As we explained in section 2.2, *relevant data* is the data that is important to the object of study.

Bringing the Milutinović *et al.* [11] study about specific design combinations to improve decision-making, Jänicke *et al.* statement about the graph interest in aligning the graph with the relevant data, with what has been established from our research results, three insights presented before, we proposed an exploratory method to draw the observers attention to relevant data systematically. We developed two prototype tools to gain further information about this exploratory method. In both, the graph designers use the tool to select the data subset their want the observer's attention to focus on. Then the tools create several graph variations, preserving the same graph semantics but modifying: a) color palettes (using sequential and diverging palettes); b) texture (by creating a black grille with different widths); position (by shifting and reordering data subsets in the X-axis); and c) orientation (by swapping the X and Y axes).

The first tool, made in Matlab, varied in color, texture, and orientation. Regarding determining the graph with the highest salience region in the initially selected data subset, all graph alternatives are analyzed with a saliency map algorithm *et al.* [38]. We developed a second prototype in Python for two reasons: firstly, Matlab's response times were considerably longer than those given by Python; secondly, Python allowed us to design a more natural user interface for a graph designer. Therefore, we combine the color, position, and orientation features in a Python version. Compared to the Matlab version, the python version changes one of the features: Texture to Position. The decision was based on the model's validation results, which showed that position variation could change the saliency points. In addition, due to language limitations, the libraries selected for the Python version did not allow textures.

6.2 Related Works in Saliency Prediction for Graph Design

Some contemporary InfoVis tools (e.g., Tableau) use the preattentive process knowledge in an implicit manner. For instance, they present charts with sequential color palettes to highlight, in a natural way, the most important points (e.g., low and high subset) [109]. However, these tools do not provide a mechanism to know if the selected chart (e.g., subset position, shape, color palette) is the one that better emphasizes the relevant data or how it could be perceived by the observer.

In the literature, we found some studies oriented to assist in the graph design process, in particular in highlighting relevant data, using saliency prediction. For instance, Feng *et al*. [110] developed a methodology to draw the observer's attention toward values of high certainty while not calling attention to uncertain values in scatter plots and parallel coordinate plots. To highlight the relevant points in these two types of plots, they basically superimposed the plot on a gray-scale saliency map. The saliency map used is in fact a probability density function, which *"describes the likelihood of each value, it naturally deemphasizes unreliable data points in the original data"*. This makes the representative outliers visually represented as discrete glyphs. However, this proposed technique was intended only for large volumes of data, the more information (more lines, or more points) the denser the map and therefore the more visible the relevant points.

Similar results were obtained in Jänicke and Chen in [2], who showed that is possible to attract attention to relevant data using a saliency algorithm. In this study, the graph designer can choose the areas that represent the relevant data. Then, they used the Itti-Koch model to generate the saliency maps, balancing the weights of each preattentive feature. Both maps are combined to calculate a quality metric which allows the graph designer to see whether the graph is good o bad based on the graph designer's purpose. Additionally, Jänicke and Chen in [2] include a *contribution map* that shows the contribution that each of the features has on the salience map. Nevertheless, this method only shows how close or far the graph focus attention is from the end user's expectations, but does not give any recommendations on how to achieve the objective.

A study by Shive *et al*. [111] showed the development of a method to support the color selection for data markings in maps using a statistical saliency model. They estimate the importance of an item in relation to the distribution of all features in a particular display. They had a color palette that they assigned to each pushpin on the map and categorized according to their prominence (minimum, median, and maximum). Finally, they developed three versions of each map to see if the statistical saliency model can determine colors on cluttered and uncluttered maps that reduce search time for things. Shive *et al*. [111] showed that search time has a slight decrease. However, because the only color was examined, this method had a restriction on the amount of saliency features that could be used. Furthermore, the saliency model does not include attention across the visual field, which has an impact on search performance.

Finally, an interesting finding was presented in a study presented in Fosco *et al*. [40] that closely correlates with our tool. They created two tools based on their importance saliency prediction model (see section 4.2.1), a deep learning model that used the Bylinskii saliency model. In the first tool, the user is allowed to edit design elements on a canvas and receive immediate feedback about each element's predicted relevance, which means the tool adjusted the relative importance of design elements. Additionally, the users can interact with each design element's importance scores, allowing them to establish parameters to enhance or reduce the importance of design elements. Similar to the previous tool, the second uses the element's importance score to automatically adjust the locations and sizes of design elements to fit new aspect ratios. This second tool scaled and repositioned visual elements based on predetermined graphic layouts to achieve new aspect ratios while maintaining the visual prominence of the input design. Even though the saliency model they employ, as they claim, works for any graphics and DataViz images, both developments were built and tested specifically for posters and infographic designs. Furthermore, as we previously established, the Bylinskii models, that was also utilized for these tools, showed a significant performance decline with images with simple structures.

The mentioned papers employ saliency prediction as a design tool in different approaches to improve the presentation of the graphs in a variety of ways. However, some of them are created for complex visualizations, such as parallel coordinates plots [110], infographics [42] or flow visualizations [2], on which salient area depends more on the quantity of data than the preattentive features. On the other hand, only one study gives recommendations about which should be the best accurate combination of preattentive features to highlight relevant information [111]. However, they only varied one feature, when saliency models generally operate on at least three of them.

In the next section, we will describe our approach which includes the graph designer's desired attention objectives, validated saliency models for DataViz images, and design recommendations varying at least three preattentive attributes. Then, in section 6.4, we will present the two developed prototypes.

6.3 Approach Process

Our main goal is to explore a mechanism for assisting graph designers in drawing the observer's attention to specific relevant data they can choose at design time. To test the approach's viability, we created a tool that allows us to conduct systematic graph variations and then evaluate their impact using a saliency map model [112]. We only consider bar charts for this initial approximation, and we only evaluate three preattentive properties: color, position, texture, and orientation. These attributes were decided based on the results of a data-contained elements experiment described in section 5.6. The saliency map algorithm proposed by Matzen *et al.* [38], which, as previously stated, has the highest performance, is used to analyze graph changes.

The Design Tool proposed approach is composed of five main phases, as illustrated in Fig. 6.1: 1) Load initial data, 2) Charts Systematic Construction, 3) Saliency Maps Generation, 4) Salience Rating, and 5) Plot Final Charts.

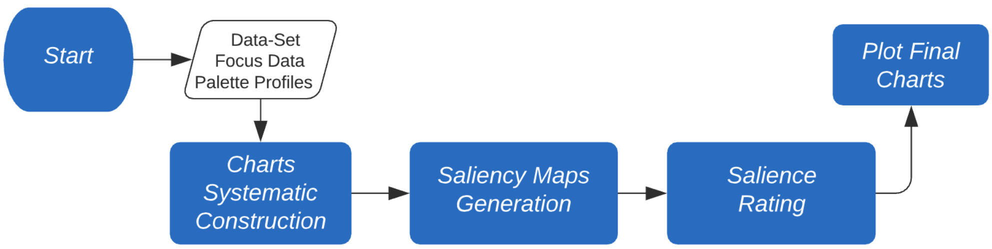


Fig. 6.1 Saliency Maps as Design Tool - Exploratory Approach

6.3.1 Load initial data

Our approach begins with three inputs: the data set to be plotted, the specific data to be highlighted, and a set of palette profiles. In this first approximation, the data to be focused on should be categorical data chosen by the graph designer.

The graph designer could choose a palette profile from a list of pre-set color palettes. These palettes, like other visualization tools, can be of multiple kinds, including sequential, divergent, and qualitative. However, we limited the research qualitative palettes taken from COLOURLovers [113], only, and filtered by those that were related to the keyword “data”. We decided on categorical palettes because the color placements can be varied without changing the meaning of the colors, unlike a sequential palette where the colors must be “arranged.” For each palette, a profile was constructed to determine which color in the palette has the most salience, and whether the salience increases using textures or changing the position on the chart (see Fig. 6.2).

The Matzen *et al*. [38] salience algorithm was used to construct these profiles. The selected palettes were applied and then processed through the salience algorithm using a bar chart as a baseline. Thirty-four color palettes were analyzed. To characterize each palette we identified the dominant and secondary colors. If the color palette showed a constant distribution of saliency, that is, it did not have a dominant color, we changed the position of the colors. If neither of the above two options worked to obtain at least one dominant color, we added texture. We also tested changing the orientation to see if there was any change in color dominance. In some limited cases, we found that there was a saliency change due to orientation. This was a long procedure, given that many combinations were made in order to understand and profile the colors of each palette.

Based on the generated profiles, the colors of each palette were sorted for predominance at the end of this step. The texture and orientation attributes were used in the same way. All these values can be pre-computed and are not dependent on the specific data set.

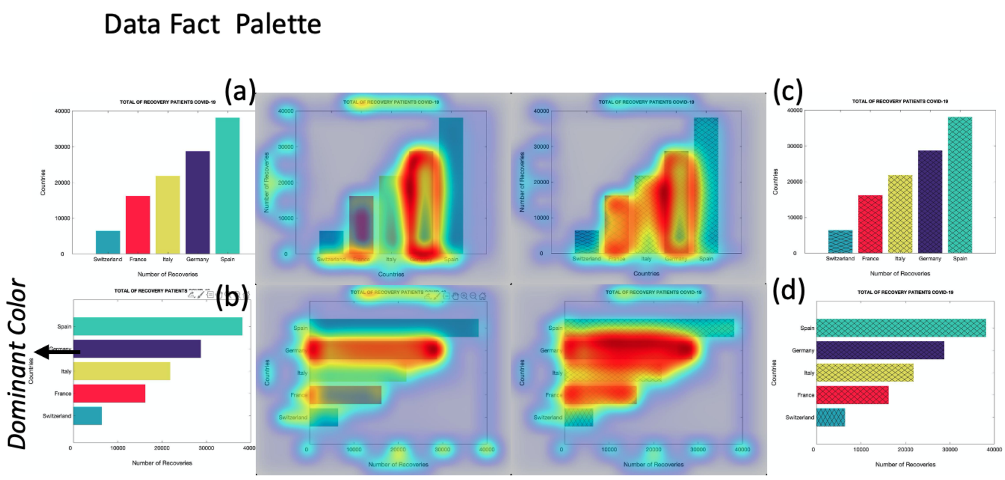


Fig. 6.2 Example of Palette Profiling

In Fig. 6.2, the dominant color is number four (4) in the palette, on the vertical (Fig. 6.2.a) and horizontal (Fig. 6.2.b) position, both without texture. For the prototype, we use the palettes with a clean saliency color in one or more of the combinations (e.g., dominant color and orientation), the others were removed.

Table 6.1 Example Color Palette Profile: "Data Fact" Palette

Property	**Value**
Dominant color	4
Vertical	1
Horizontal	1
Vertical with texture	0
Horizontal with texture	0

Table 6.1 presents an example of a palette profile. "Dominant color" represents the color position in the palette of the most salient color. The other properties are marked with a one (1) if their presence affects salience in the specified palette. For example, the Data Fact palette (see Table 6.1) has a *clean saliency* with vertical and horizontal orientation, but without texture. "Clean saliency" means that the color palette has one dominant color, only, on a specific position (vertical or/and horizontal), and with or without texture. Fig. 6.3 shows the salience map of the previous example dominant color, in different combinations.

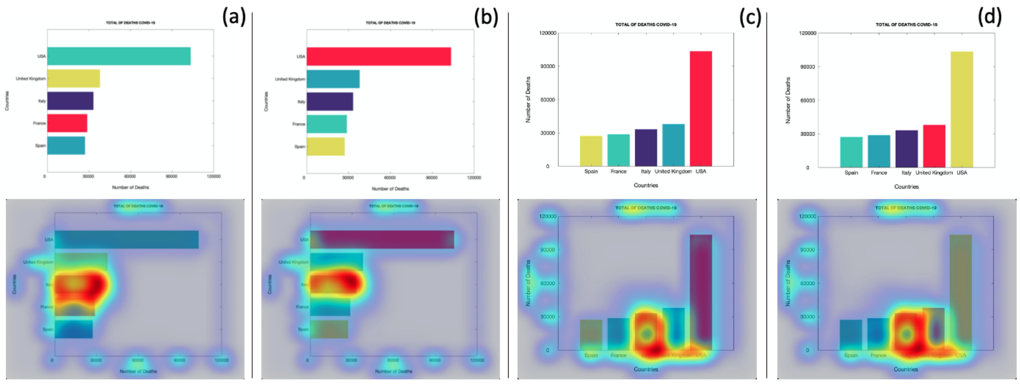


Fig. 6.3 Color Palette saliency changes

6.3.2 Graphs Systematic Construction

The purpose of this stage is to construct graphs methodically based on the graph designer's specifications (color palette and data to be highlighted). The graph designer chooses a variety of color palettes to apply to the graph. A systematic graph construction is carried out using the color palette profiles as input. At each phase, the graph is constructed by depicting the data to be emphasized with the color that has the maximum salience (see Fig. 6.3). The palette's remaining colors are then shifted through the remaining data subsets. The palette profile determines whether salience rises with texture and/or orientation. Therefore several candidate graphs are created.

6.3.3 Saliency Maps Generation

In this step, we generate salience maps for each chart variation created in the previous stage. To accomplish this, as reported in the study on salience map models and algorithms (see Chapter 5), the selected algorithm was DVS, Matzen's model [38], implemented in Matlab language. The algorithm output is the image of the chart, enriched by a salience map layer.

6.3.4 Salience Rating

In this step, each salience map is evaluated to identify which charts have the highest saliency values on the area corresponding to the data subset to highlight. The DVS algorithm has a function that supports this 'proximity' evaluation. This function compares a given set of coordinates with the same coordinates on the salience map. The coordinates of the data subset to be highlighted are sent to this function; the function returns a value between 0.0 and 1.0 that measures the weighted overlap of the data subset coordinates and the chart salience areas. The final result of this step is a list with the percentage of proximity value per each graph variation.

6.3.5 Plot Final Charts

The last step is the selection of the best chart. For each palette the graph designer selects, the chart with the best proximity percentage is chosen (see Section 6.3.4). The winning charts are shown to the graph designers, one for each color palette selected, and they can choose the preferred one.

6.4 Prototype

To validate our approach, we developed a tool that implements the process described in the previous section.

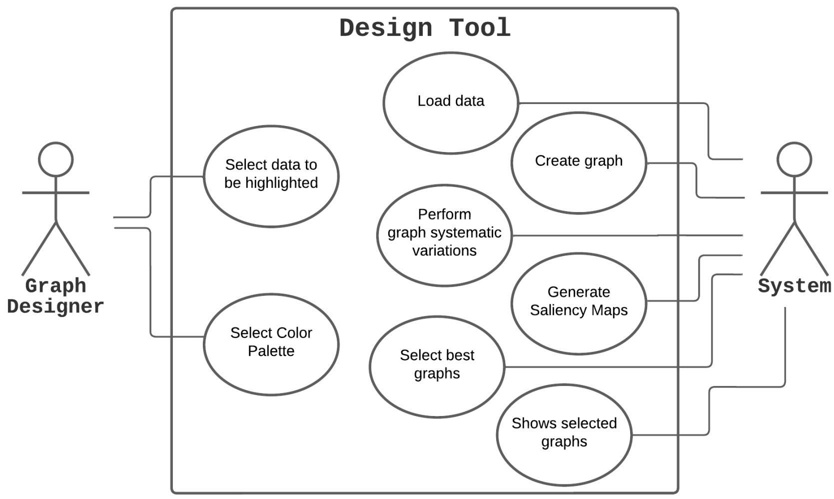


Fig. 6.4 Use Case Diagram for Design Tool, for both versions

Fig. 6.4 shows the tool Uses Cases diagram. The graph designer can select the data to be highlighted and the preferred color palette. The system is responsible for creating the graph, making the variations to the visual elements, generating the saliency maps, and evaluating which graphs are more accurate representations of the graph designer's objective. Each use case represents the phases described in section 6.3.

6.4.1 Development

We first had to make some changes to the original Matzen algorithm to develop the prototype. These changes made the algorithm accessible via web protocol (JSON standard and code packaging). With these changes, it was possible to upload the algorithm to the Matlab Web App Server, making it accessible to other applications external to Matlab.

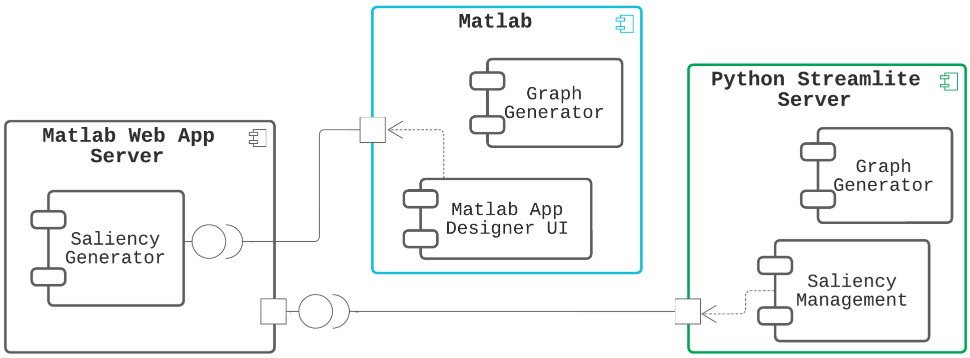


Fig. 6.5 Design Tool - Components Diagram

Therefore, we implemented a first version in Matlab (see Fig. 6.5, blue component). This version uses the Matzen algorithm directly through a Matlab App Designer User Interface responsible for the whole interaction with the graph designer. In addition, in this first version, we developed a component that was in charge of all the system functionalities (use cases) presented in Fig. 6.4.

Due to response times and interface flexibility, we decided to develop a second version, a web Python application (more details in section 6.4.2). This Python version managed two components (see Fig. 6.5, green component): the "Saliency Manager" is in charge of connecting to the Matlab server to generate the saliency maps, and a web component is responsible for graphing and generating the user interface. This second version runs on the streamlite server (python library) [114].

6.4.2 Matlab Version

The first version of the prototype was the Design Tool in Matlab. We used common libraries for data analysis and visualization (e.g., App Designer, Plots), and we generated salience maps with the DVS algorithm [38]. The goal here is to understand the behavior of the preattentive attributes and the feasibility of moving the attention area starting from a baseline chart. In this first round of experiments, we use *bar charts* as a baseline on which to apply the systematic variations.

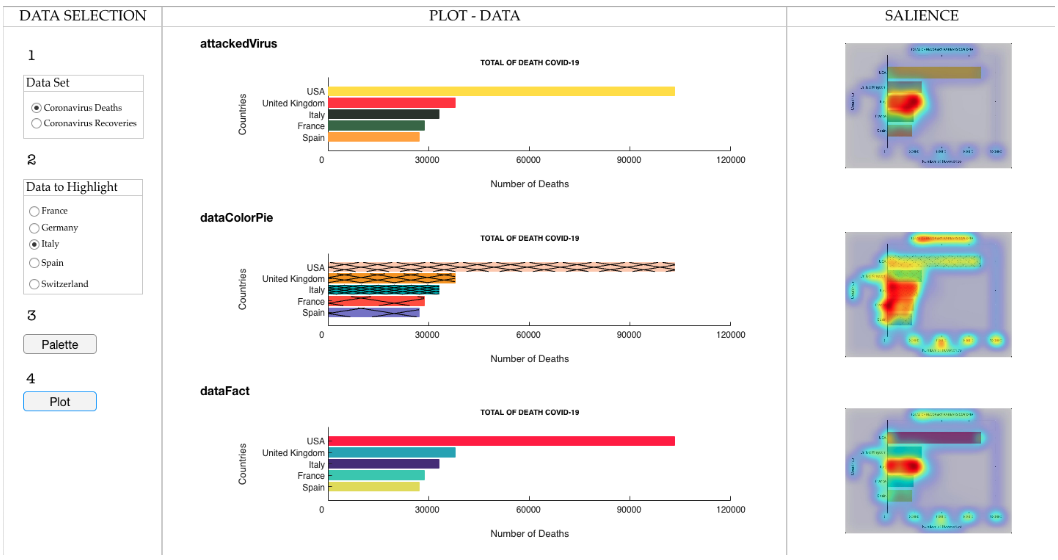


Fig. 6.6 Prototype User Interface

As the case study, we selected the Coronavirus (COVID-19) dataset [115], which consists of data from COVID-19 Deaths Worldwide as of May 29, 2020. We plotted the five countries with the most casualties (USA, UK, Italy, France, and Spain) vs. the number of deaths (103,330, 37,837, 33,142, 28,662, and 27,119, respectively). This combination was selected because it generates a graph with a considerable variation between subsets. The other countries have a difference, in the number of deaths, of at least 17,000 with respect to the selected five, which would make their bars too tiny compared with the others. As the bars are small, they saw as a line and not as a bar, and the changes in color and texture could be difficult to perceive.

Fig. 6.6 shows the prototype interface, where the four steps are controlled in the left column. The graph designers must first select the database for which they want to plot the chart. Based on the selection of the dataset, the graph designer must then select the data subset to be highlighted (step 2). The next step is to choose some possible color palettes that the graph designer wants to use on his chart (see section 6.3.1). The interface shows a pop-up with a color palette list, and the user can choose three of them. After that, the graph designer can execute the algorithm presented in section 6.3 (step 4, button "plot"). For each color palette, the graph designer selects, a minimum of one and a maximum of 8 charts are generated. Four of these charts correspond to the states in which the palette has a dominant color (see section 6.3.2). The other four correspond to variations on the palette's remaining colors (not predominant). Fig. 6.3 shows an example of those variations, with the same palette profile presented in Table 6.1, and *Italy* as a the data selected to be highlighted (third data point in the chart).

Chart *b* in Fig. 6.3 is the chart with the best ranking. The final result comprises the charts selected by the algorithm, one per palette, whose salience map had the highest percentage of proximity to the data selected by the user (Fig. 6.6). Additionally, the interface displays the salience maps of each of those charts.

Insights

We discovered some insights on the impact of feature attributes based on the prototype results. Although we only worked with palettes with a dominant color, we discovered that this predominance might be greater or weaker based on the nearby colors when it came to the preattentive characteristic *Color*. As shown in Fig. 6.3, the salience is more precise in the color combination **b** and less vivid in **d**. Due to a lack of evaluation tools, it is usual for viewers to choose the "strongest" hue to highlight certain data. However, it is feasible to improve accuracy by knowing the occurrence of nearby hues.

Concerning to preattentive attribute *Texture*, in most of the revised color palettes, the texture made the salience region wider. Also, in some palettes, it is only possible to obtain a salience focus on the selected data if that one has a texture. Otherwise, the palette has only a salience in the text on the charts. In the "data Color Pie" color palette chart on Fig. 6.6, the palette has a texture because the saliency would be on the title without a texture, but with the texture is on the data. Nevertheless, the presence of the texture tends to increase the saliency area excessively. Even so, the salience is not as neat as with the other graphs because the saliency region is on three data subsets. A future step could be to systematically handle the textures to achieve a more precise salience.

Regarding the preattentive attribute Orientation, we found that the salience is consistent, with the same predominant color, when the orientation is changed. However, since the Matzen *et al.* algorithm has a text recognition component when some of the data have a long name, the salience region is enlarged. Fig. 6.3.c has the same color order as b, but the saliency area is more expanded at the bottom. The previous behavior happens because the data name (United Kingdom) is large, creating a black text area that will be more prominent for the top-down part of the Matzen algorithm.

This prototype has a performance disadvantage. The program takes at least 20 to 25 minutes to display the results described above. These times were due to the amount of processing that had to be done to evaluate where the highest percentage of saliency was found since it was done pixel by pixel (Saliency Rating step, section 6.3.4).

We conclude that moving the attention area to a certain data subset is possible. The graph designers can choose which data to emphasize and which color palettes to use. Finally, they learn about the potential impact of each color choice on the chart perception during the design stage. That is why multiple color profiling helps the graph designer evaluate more than one alternative in positioning the colors in the graph. The present visualization tools do not provide this functionality.

6.4.3 Python Version

In the first tool version, we had some problems in the development process, specifically with the complexity of making the graph and knowing the positions of each visual element, for instance, the position of each bar within the graph. Furthermore, the processing time took 20 to 25 minutes.

For these reasons, we decided to look for an alternative and more efficient way to test the feasibility of our first approach. So we developed a web application in Python, which also made the interface simpler. This Python version makes the systematic variation of three preattentive features: color, position, and orientation. In contrast to the Matlab version, the texture was left out because the selected Python graphics design libraries are missing this property. Besides, the color palettes used are the ones already established by python libraries, which showed us better behavior when studying their saliency than those selected for the Matlab prototype. Moreover, in this version, we have full control over the position feature, which means that we can change the bars' position and always know their position.

As a result of this release, we realized that it is feasible to help the graph designer establish a focal point in a way that is closer to the common DataViz process.

Python Prototype Process

This prototype integrates three main libraries: Streamlite, an open-source Python library to create and share custom web apps for machine learning and data science [114]; Seaborn, a Python DataViz library; and a Matlab Production Server, for the connection with DVS (Matzen model algorithm).

Fig. 6.7 shows the process implemented in this version. As in the Matlab version, the graph designer chooses the data to highlight and a set of color palettes. Regarding the process carried out by the Python program, there are six main tasks:

- **Generate a set of graphs**. In this task, the program makes the variation of the bar color, position, and orientation to be highlighted. Firstly, the user selects three color palettes as the Matlab version. Then, the tool makes the position variations for each color palette selected. An example of bar position variation is presented in Fig. 6.8. If we wanted to highlight the data PHP, the bar that represents is in the second place (see Fig. 6.8a), then it goes to the fourth place (Fig. 6.8b), and successively until passing through all of the possible positions in the graph. Additionally, a horizontal version of the same graph is also generated for each of these position shifts. In Fig. 6.8c, PHP is in the fourth place but with horizontal orientation. Finally, two other position swings were added, placing all the bars in ascending or descending order, and horizontal and vertical shifts as well (see Fig. 6.8d). This is included because this type of arrangement is one of the most commonly used.
- **Send Encoded Graph and Decode Saliency Map**. Since the saliency algorithm is developed for Matlab, we use a Matlab server we use to render the algorithm into a web service to be accessed from Python. Each image was encoded and sent using *base64* and *json* formats. The Matlab algorithm returns

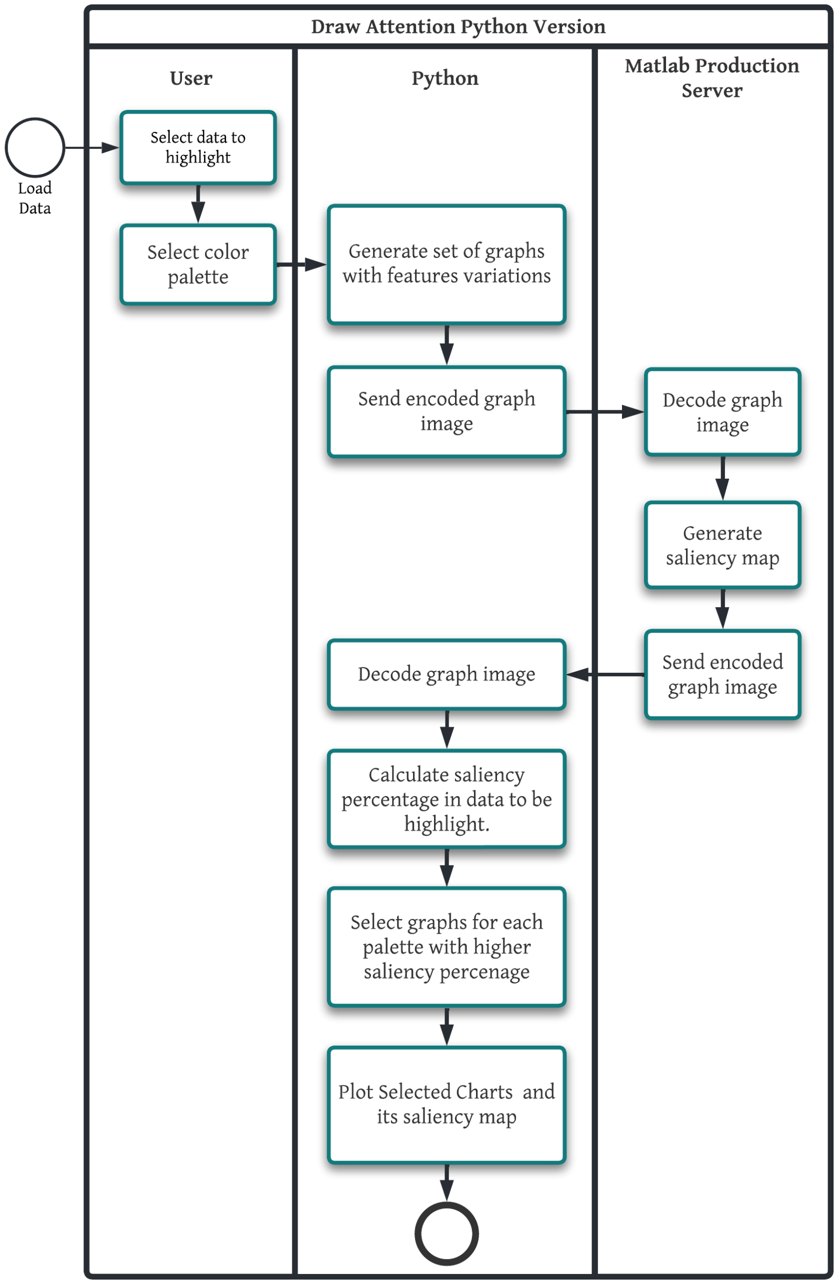


Fig. 6.7 Python Design Tool Process

two images: the saliency map overlaid on the graph image and the saliency map as a binary matrix. Then, both maps are converted to OpenCV images.

- **Calculate Saliency Data Percentage**. First, the received saliency map is transformed to grayscale. This transformation was done because we considered only the intensity of each pixel. Then, we chose a range from 190 to 255, representing the highest (most intense) saliency points and where the saliency is most concentrated.

$$mask[gray > 190] = 1 \tag{6.1}$$

Finally, we calculated the area of saliency over the bar representing the data selected by the graph designer using the equation below 6.2).

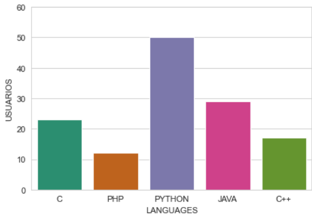
$$area = (((bar_in == 1)\&(mask == 1)).sum())/((mask == 1).sum()) \tag{6.2}$$

- **Selected Best Graph and Plot**. Based on the results of the *Saliency Data Percentage*, the graph with the highest color range is selected and then plotted for each color palette chosen by the graph designer. The selection criteria are the same as the Matlab version.

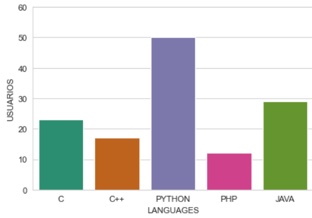
Regarding Matlab Production Server, it is an application part of Matlab. We package the Matzen algorithm using MATLAB Compiler SDK, then deploy it to MATLAB Production Server in a *json* format to receive the image, execute the algorithm, and return the saliency maps.

Results and Tests

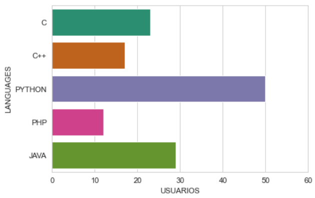
In Fig. 6.9, the interface of the Python version of our Design Tool approach to drawing the observer's attention can be seen. In the first part (see Fig. 6.9a), the program shows the data and the option to select which data will be highlighted. Then, for each group of color palettes (Qualitative, Sequential, and Diverging), the graph designer can choose one from the list. Finally, the program executes the process described in the previous section.



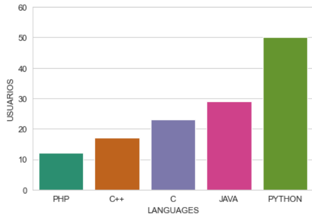
(a) PHP bar in second position



(b) PHP bar in the fourth position

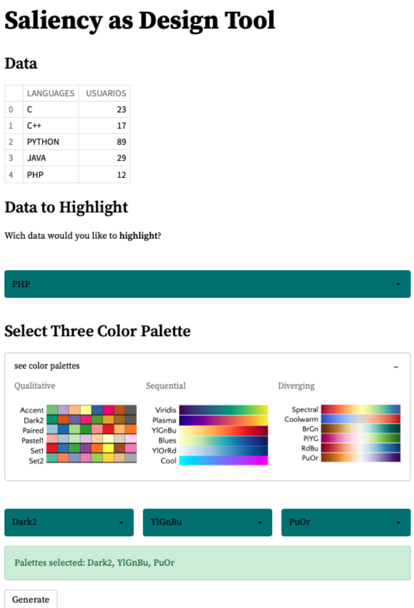


(c) PHP bar third position and horizontal.

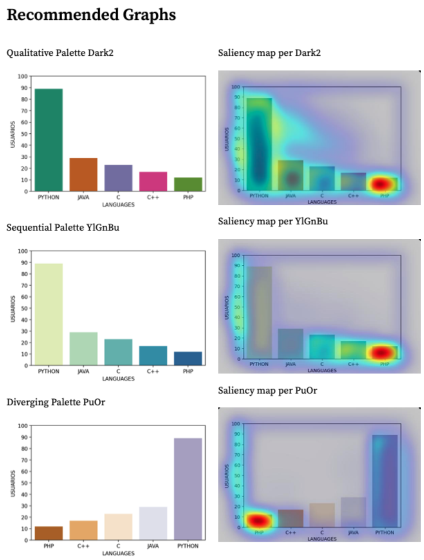


(d) PHP bar in the first position, ordered and vertical.

Fig. 6.8 Design Tool, Python prototype variations example



(a) Python prototype - PHP highlight



(b) Python prototype - PHP highlight results

Fig. 6.9 Python Graph Variations example. PHP was the data selected to be highlighted

We performed two tests with Covid-19 data in Latin American countries and Programming Languages Usage to see the tool's general behavior. We used data from the six (6) Latin American countries with the most Covid-19 deaths for the first test (see Fig. 6.10). The characteristic of this data set is that the countries have a similar number of deaths, which makes the difference between bars relatively small. On the other hand, for the second test, we use data related to the use of programming languages. In this case, the numerical difference between each data is significant (Fig. 6.8 shows an example).

For the first test, we ran 108 combinations for each color palette (18) and looked for combinations that would highlight each of the countries (6). For example, the bar representing Brazil was tested in which position and with which orientation it was most salient. Appendix D has some examples of the winner for two data and four color palettes.

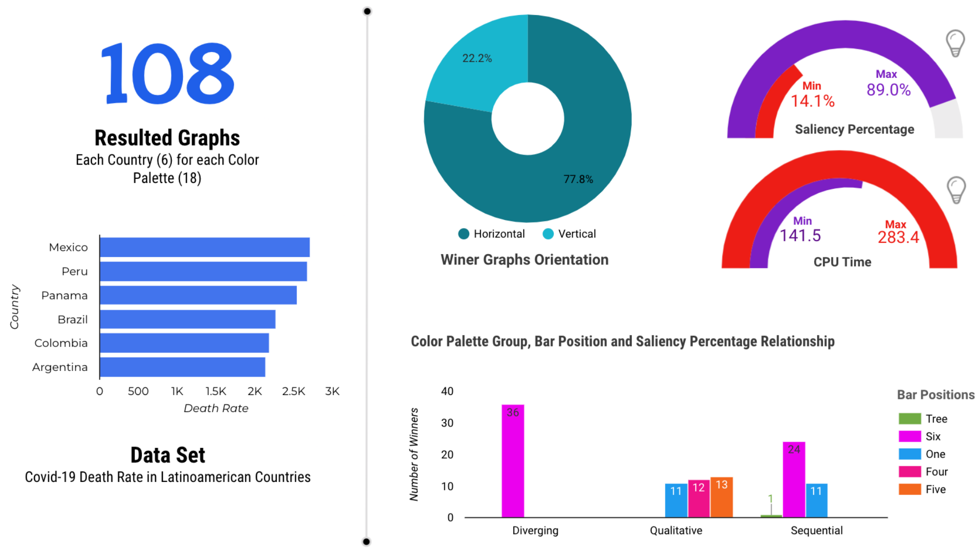


Fig. 6.10 Desing Tool, Python Dashboard with the results from Covid-19 death rate in Latinoamerican Countries. Link to interactive dashboard [1].

Fig. 6.10 shows a dashboard with the resulting data. As can be seen, from this data, most of the winner graphs have a horizontal orientation (77,8%). This is a behavior we discovered in the model validation: the saliency changed and, generally, was more centered when the data orientation was horizontal (see Chapter 5). Regarding the percentage of saliency, it is clear that it is quite variable since it is possible to obtain graphs that have 89% of the total saliency on the selected data. However, on other data, only 14.1% of saliency is achieved. These numbers are not only related to the position, color, or orientation of the highlighted data but also depend on the data size since some data have a smaller area than others.

With respect to the CPU time, as seen in the Fig. 6.10, the tool's maximum time to find a suitable combination for each color palette, position, and orientation, is 283.4 seconds. These times are still highly dependent on the response of the Matzen algorithm when there are more colors or visual elements. Finally, we found that some bar positions have a strong relationship with the type of color palette. For instance, in Diverging color palettes, the winner was always the data in the sixth position (the last one in the graph). On the other hand, positions such as the *second one* did not have any winning graphs with these types of data.

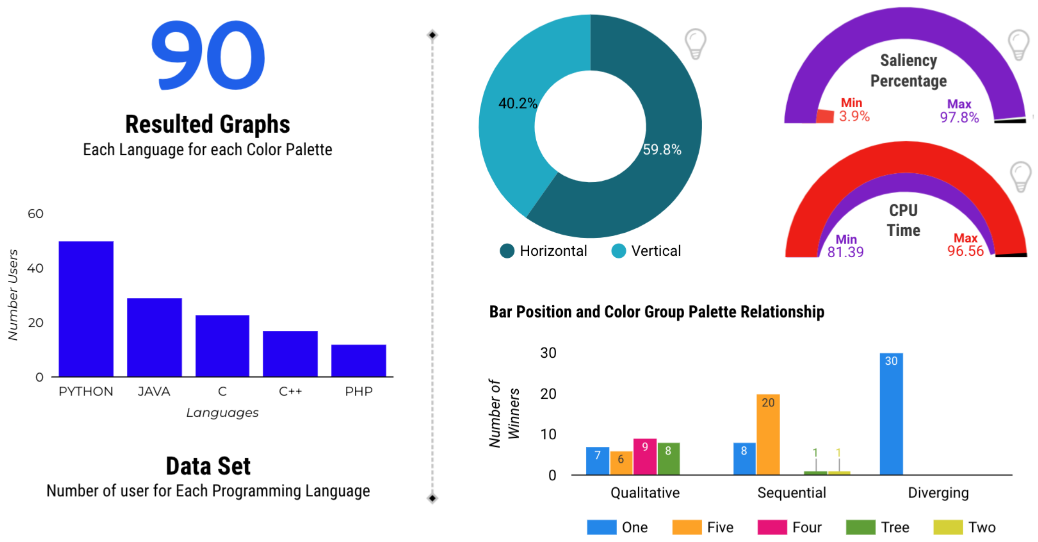


Fig. 6.11 Desing Tool, Python Dashboard with the results from Programming Languages usage. Link to the interactive dashboard[2]

To conduct the second test, we generated 90 permutations for each color palette (18) and looked for combinations that would draw attention to each programming language (5) (see Fig. 6.11). For instance, it was determined which location and orientation made the bar representing the C language the most evident. (see Appendix D. Fig. D.3 and D.4).

The second test showed several differences in behavior from the previous one (see Fig. 6.11). The response times were significantly shorter and more consistent in this test, probably because the graph had greater white space and the bars were smaller. Regarding the orientation, both options are closer in percentage. However, the horizontal orientation is still predominant. One feature that varied quite a bit was the saliency percentages on the data to be highlighted. However, given that there is a large gap between the maximum (97.8%) and minimum (3.9%), it can be noted that it is due to the size of the bars, which in this second test had a smaller area.

Finally, similar to the behavior of the first test, with some color palettes, the data can only be highlighted if it is in an exact position. For example, the Diverging palette type only achieves winning graphs if the data is at position one (at the beginning of the x-axis). Based on these results, we conclude that a systematic change of the design elements can indeed change the possible attention to the data.

Insights

As in the previous development, we can state that it is possible to draw attention to relevant data previously defined. However, in this development, we realized that the number of variations created can be smaller and still achieve the goal. We could also see that using the already established palettes without changing their order makes finding a feature combination where the data can be highlighted more feasible.

One of the most relevant improvements is in time, passing from 20-25 minutes to 8-9 minutes per chart. This decrease in time is due to the saliency search process in each bar since it is now a more straightforward and faster process. Although it is still a considerable waiting time, it is essential to highlight the time of the saliency algorithm and the number of graphs that have to be sent. Possibly, parallel processing would make it more agile.

6.4.4 Limitations

One of our significant limitations was access to the image components individually. Precisely, we needed to know in which pixels of the graph each component was located. This location data is required in order to contrast the resulting saliency map and the graph image. Although this limitation was more complex in the Matlab version, Python also depends on each library's functions. This limitation forced us to work with only one type of graph, bar charts because both programming languages made it simpler to access the positions of each of its components.

The Matzen algorithm was developed in Matlab, which forced us to create a user interface in the same language. This meant that the usability and appearance of the interface were not the best. However, for the Python version, we were able to package the algorithm as a web service so that it can be accessed from any other application.

6.5 Discussion

We developed a tool with two versions in different programming languages, in which we implemented the graph design process assisted by a saliency model. With this tool, it was possible to make systematic modifications to the graph to meet the attention goals given by an end user (graph designer).

Because design decisions are subjective, there is never an objectively best result, i.e., a graph that is the best for every observer or graph designer's purposes. Furthermore, determining the quality of a design decision is challenging because there are many variables related to the final observer, such as level of expertise or visual disabilities and graph designer design level of knowledge. However, the proposed tools described in this section aim to bring graphic designers closer to the visual impact of their design decisions and bring them closer to the final observer.

With the use of tools like these, the designer will be able to know what could be the most appropriate combination of design elements to achieve the highlighting of the data considered most relevant in the graphic. Additionally, the tool also brings the designer closer to the final observer due to the widely studied attention cognitive impact, that is to say, the combination of preattentive attributes, on the performance in a task solution [11] (defined in chapter 3).

Finally, it is essential to note that shifting attention can be extended systematically to different types of graphs. It specifically depends on the tool's flexibility to generate the graphs, the kind of graph, and the possible variations. For example, if a graph shows a timeline, changes in position could not be made because they would alter the main objective of the graph. They are keeping in mind this need to extend the use of saliency prediction to other types of graphs.

Chapter 7

Using Saliency Prediction as a Measurement Tool

The previous chapter presented an exploratory approach (Design Tool) that varies visual graph elements to find the best combination to highlight relevant data selected by the graph designer. The Design Tool approach supports the graph designer in the process of focusing attention on specific data. Regardless of the Design Tool development, saliency was used to quantify the amount of salient in each data-contained element but not to validate design decisions. Furthermore, the graph designer was limited by the data visualization technique (bar chart) and the variations of the visual elements (by selection only color).

On the other hand, in the results of the Literature Review study (see Chapter 3), saliency can also be used as a **design validation instrument**. Based on that study, we established that the use of saliency prediction as a measurement attention method could be helpful for the graph designer who could know the impact that each of her design decisions would have at the visual-cognitive level on the final observer.

The development of this second approach, called **Measurement Tool**, aims to integrate saliency prediction into a classical data visualization system to explore its capabilities as a measuring tool. In addition, this development allows the graph designer to measure how each of her design decisions would impact the observer's attention with some degree of accuracy.

7.1 Motivation

Ware, in his book *"Information Visualization, perception for design"* [6] claims that graph designers have to "*design graphic representations of data by taking into account human sensory capabilities in such a way that important data elements and data patterns can be quickly perceived.*" These *human sensory capabilities*, for Ware, represent the knowledge about the visual impact that each design decision has in the graph observer interpretation. Ware in [6] highlights the importance of attention in DataViz due to its relation with the visual working memory. This type of memory "holds the visual objects of immediate attention, and the contents of working memory can be drawn from long-term memory" [6].

In short, every graph designer must be aware of these relationships between design decisions and their cognitive impact on the observer. However, all mentioned knowledge about attention behavior is information that is not necessarily apprehended in all contexts. Moreover, graph designers currently have different profiles and specialties that are not always related to the design field. For this reason, it is important to develop a tool that can provide insights into the behavior of those human sensory capabilities, limited to attention behavior for this thesis, at the design time.

Besides, saliency prediction has long been used to understand the cognitive process of humans interpreting a graph, as stated in the section on saliency prediction. However, we note that saliency prediction has not been available to graph designers in their daily application toolkit. For example, the classical DataViz systems have not included saliency prediction models as part of the graph design process. Thus, the use of saliency prediction gives the graph designer a general idea of how attention will be distributed in the graph.

Another motivation for this development was expanding the vision of the use of saliency in other DataViz techniques. In the previous approach, *Saliency Prediction as a Design Tool*, our tool was limited to bar charts due to software limitations and the goal of trying to make modifications with features, such as a position, that can only be used in certain graphs. Therefore, extending our development to a system with the entire chart design cycle and a wide range of DataViz techniques expanded the tool capacity to a more natural context for a graph designer.

Altogether, we proposed an integrated approach to bring closer the understanding of the observer's attentional behavior to the graph designer and to use the salience prediction instrument as an explicit function in the graph design process. In this integration, a saliency prediction algorithm is used as a new function in a classical DataViz system, as a **validation tool**.

Finally, with the development of this tool, we noticed that for the most inexperienced designer in the area of graph design, it could become a learning tool as well since novice graph designers can learn visual decision behaviors on each interaction with the measurement tool.

7.2 Related Works in DataViz Evaluation Process

Considering that the goal is to use saliency prediction as a validation tool within the design process, we performed a short search of which methods are generally used for graph design validation. In DataViz, the evaluation process, according to Wall *et al*. [116], could be made by several adapted human-computer interaction usability assessment methodologies. Measuring accuracy and duration in a study where participants do benchmark activities is one method of evaluating a visualization's utility. In addition, these studies can help establish if individuals can correctly understand data by manipulating the user interface and interpreting the visualization [117].

The insight-based visualization evaluation methodology is another approach [118]. Experts in data visualization use this method to determine if the visualization gives valuable information to the end-users. Experts must decide how many new insights about data collection were gained as a result of the visualization. Complex, profound, qualitative, unexpected, and meaningful values are defined as insights. Alternatively, Task performance techniques can be applied [116]. These methods set up a sequence of tasks that observers should be able to complete by using the provided visualization. The visualization performance is then assessed using task completion time and accuracy measures.

However, the most studied and used methods to evaluate a visualization are the *heuristics methodologies*. For constructing visualizations, Zuk and Carpendale [119] propose a set of ten "Cognitive and Perceptual Heuristics." Within these heuristics, four are dedicated to color handling, another four to data positioning and sizing, and the remaining are general recommendations such as using Gestalt Laws or providing multiple levels of detail. Hearts *et al.* [120] instead established only three heuristics: the visualization makes important information visually salient, uses visual components appropriately, and successfully presents multiple relevant facts into a single visual pattern. The above heuristics were also validated with novice observers. Instead, Forsell and Johansson [121] gathered 63 existing heuristics, assessed them on a set of usability issues from prior information visualization evaluations, and then selected 10 top heuristics. Among the heuristics proposed in [121], only 3 are specific to manipulating visual features. The others are targeted to the nature of the graph data, e.g., data reduction or consistency.

Additionally, some studies adapted the heuristics from HCI usability validation. However, these heuristics are reasonably high level, subjective, and provide limited guidance on improving a visualization tool's specific visual or interactive aspects. These heuristics are high-level, suggestive, and offer only a limited amount of direction on how to validate specific visual or interactive components of a visualization tool [116]. In addition, heuristics commonly be misinterpreted by different experts. Another critical issue is that heuristics are created to validate usability performance, but DataViz is deeply rooted in each element's impact on observer attention.

Another technique from the HCI area is assessing observer engagement. A study by Hung *et al.* [122] was about how meaningful user engagement is in InfoVis. They presented a questionnaire with 11 engagement characteristics, where only two of which are related to the graph design (Aesthetics and Attention). The remaining characteristics are related to basic visualization tasks such as exploration and discovery, and user behavior as interest or captivation [122]. A similar methodology is proposed by Ware [123] called a *cognitive walkthrough*. This methodology includes selecting a potential final observer and having them talk aloud through the procedures required to complete a series of activities. Wherever observers become confounded by the information shown or fail to use the DataViz system effectively, it is an improvement point. This method can also be used to identify cognitive bottlenecks, such as excessive memory load or situations where repeated tasks can be delegated to a computer.

We noticed in the exploration of related work that many of the evaluation methods are oriented towards data visualization systems [124–129]. In the papers cited before, the authors, in general, presented standard usability validation methodologies for

InfoVis systems/tools: paper prototype, focus group, and log file analysis [124]; domain experts validation [126, 129]; interview technique with focus groups [127]; and also heuristic evaluation [128]. These studies are oriented to DataViz systems, how the user interacts with them, and the fulfillment of the business objectives with the created graphs.

It is essential to highlight that related works described above are primarily guidelines for evaluating graphs from a high-level point of view. Those guidelines are mainly oriented to answer questions (heuristics) that lead to finding current or future problems in the interaction and design of the graph. In the validation with experts that we carried out at the end of the developments, we noticed that the most common process to validate the graphs in companies is to show them to another group of experts. On the other hand, those studies that focus on expert validation also express that accustomed problems could be an essential bias. The used problem is when experts perform validations in the same context, and this causes a narrow view about patterns that they traditionally apply or that one worked best. The previous statement is not intended to detract from the relevance and usefulness of heuristic validations, only that they can be time-consuming and rigged because of their complexity. In addition, since most of them are subjective, they can lead to problems of understanding.

In short, for a graph designer to perform any validation over the design decisions, he has three options: to complete a heuristic validation process, present his graph to experts for evaluation, or both. The saliency prediction could be an additional tool, which does not replace heuristic or expert assessments, but can complement them and, in some cases, accelerate their process. In section 3.3.2, we discussed how saliency models had been used to validate the observer's graph comprehension, find patterns, or improve the graph design. For these reasons, we have developed this integration between saliency prediction and a DataViz system to bring this validation tool closer to the graph designer. This approach and the resulting tool are tangible validation of how each design decision can affect an image's attention centers or the image layout.

7.3 Measurement Tool Integration Process

7.3.1 Process Definition

The main objective is to integrate saliency at the end of a common graph design process so that every decision the graph designer makes can be seen how it affects the attention distribution in the graph. Fig. 7.2 shows a typical graph design process and the point at which saliency was integrated. The data visualization process mentioned before has four sub-process:

1. **Data Input**. In this process, the graph designer can copy-paste data, upload a file with tabular data or load an example dataset.
2. **Graph Type Selection**. The provided graphs are displayed in a grid along with the visual preview, the visual model name, and the category to which it belongs. The graph designer can select one based on the graph data nature and visual objectives.
3. **Data Mapping**.The data dimensions can be mapped to the graph by users. The column headings identify the dataset dimensions on the screen's left side and the possible chart dimensions on the right. In this step, the system employs symbols to distinguish the type of data (text, number, date).
4. **Customize Visualization**. Users can modify the graph features provided on the left side of the screen and preview the chart in real-time to fine-tune the chart.

The integration of the saliency prediction model is reflected in a new sub-process into the original DataViz process: "Saliency" (see Fig. 7.2 fifth process). In this sub-process, the graph designer can see the saliency map of the graph in its current state. Finally, suppose the saliency map result does not satisfy the designer's objective. In that case, the graph designer can make design modifications to the graph, return to sub-process four (4), and the saliency map will be generated again.

7.3.2 Development Process

To develop the process described in the previous section, we selected an open-source framework for data visualization design as a basis. This system was developed by the DensityDesign Research Lab (Politecnico di Milano) and is called RawGraphs [130]. This framework is a web application primarily built using two different libraries: AngularJS for the visual interface and D3.js for data visualization.

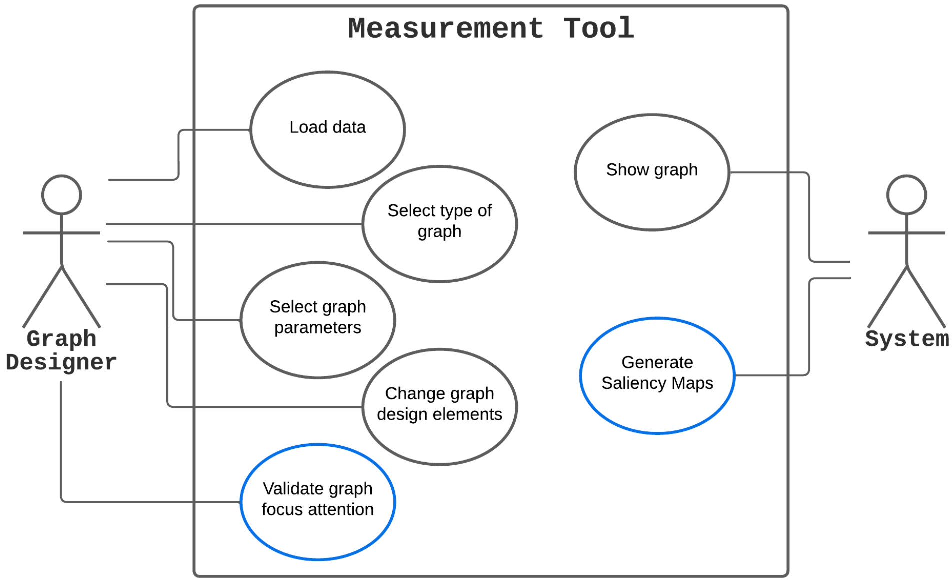


Fig. 7.1 Measurement Tool - Use Case Diagram

In the use case diagram shown in Fig. 7.1, it can be seen the base functions that the graph designer could execute, together with the functions of the system. The end user, the graph designer, is in charge of loading the data, selecting the desired type of graph, mapping the data (graph parameters), and making visual changes to the graph (color, size, text). It should be noted that the use cases in blue are the new saliency functions. After performing all the other actions, the user can also request the saliency map generation to validate their previous design decisions.

The original RawGraphs framework can be accessed at https://app.rawgraphs.io. The RawGraphs design flow is shown in Fig. 7.2 from sub-process one (Data Input) until sub-process four (Customization). Some changes were made to the original design. The aim was to improve the design process and render it clearer than the original version.

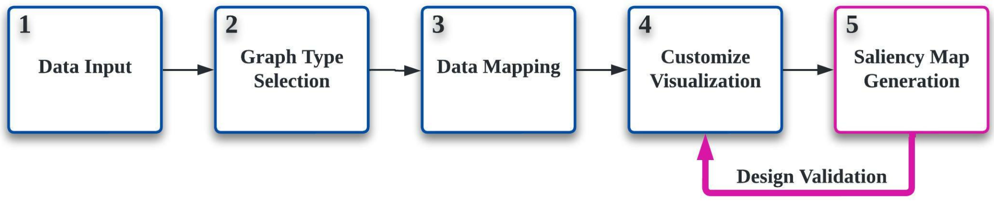


Fig. 7.2 Measurement Tool Process. RawGraphs framework supports the process between one (1) and four (4). We integrated the process five (5).

The main change to RawGraphs was to place each process within a horizontal tab that could be expanded and contracted. Thanks to this modification, the graph designer can better visualize the graph design steps. This change also makes it easier for the graph designer to go back to an earlier point in the process without going through the intermediate stages. Fig. 7.4a shows the first screen of the original framework, next to it in the Fig. 7.4b, the new startup interface can be seen, where all process steps are shown.

Regarding the saliency prediction model integration, we add a new component to the original framework. The step is called “Saliency” (see Fig. 7.4b green tab) and shows the saliency map of the graph created with the framework. The Matzen Model generates the saliency map by web connection (by JSON protocol) between the RawGraphs framework and the algorithm running in a Matlab Production Server (for more details, see section 6.4.3).

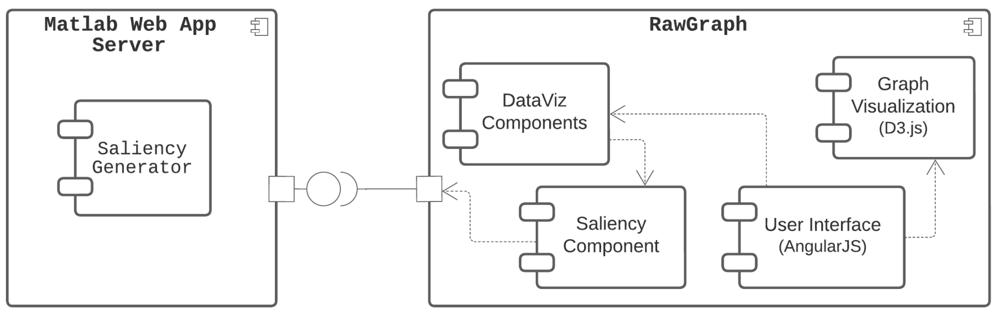
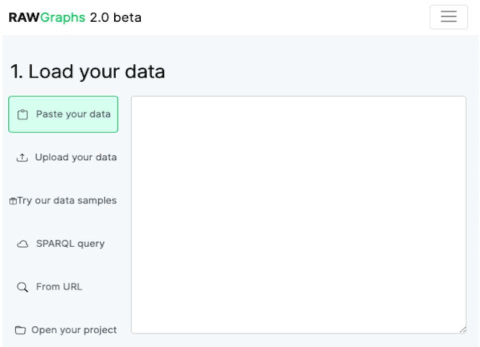
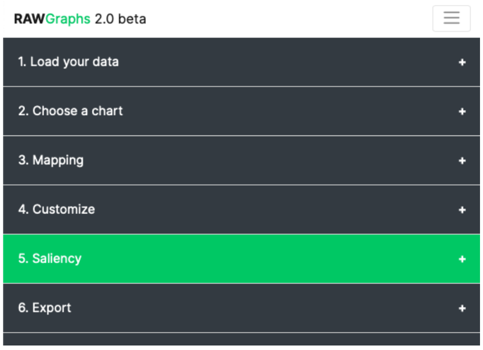


Fig. 7.3 Measurement Tool - Component Diagram

An overview of the software architecture of this tool can be seen in Fig. 7.3. RawGraphs has four subcomponents. The first one is "Graph Visualization," which uses the D3.js library to render and display the graphs on the web. The subcomponent "User Interface" uses AngularJS to construct the user interface and interact with the end user. The DataViz component is responsible for characterizing each graph type and mapping the data. Finally, the Saliency component is the one we added and is responsible for connecting with the saliency algorithm and generating saliency maps.



(a) RawGraphs framework original interface.



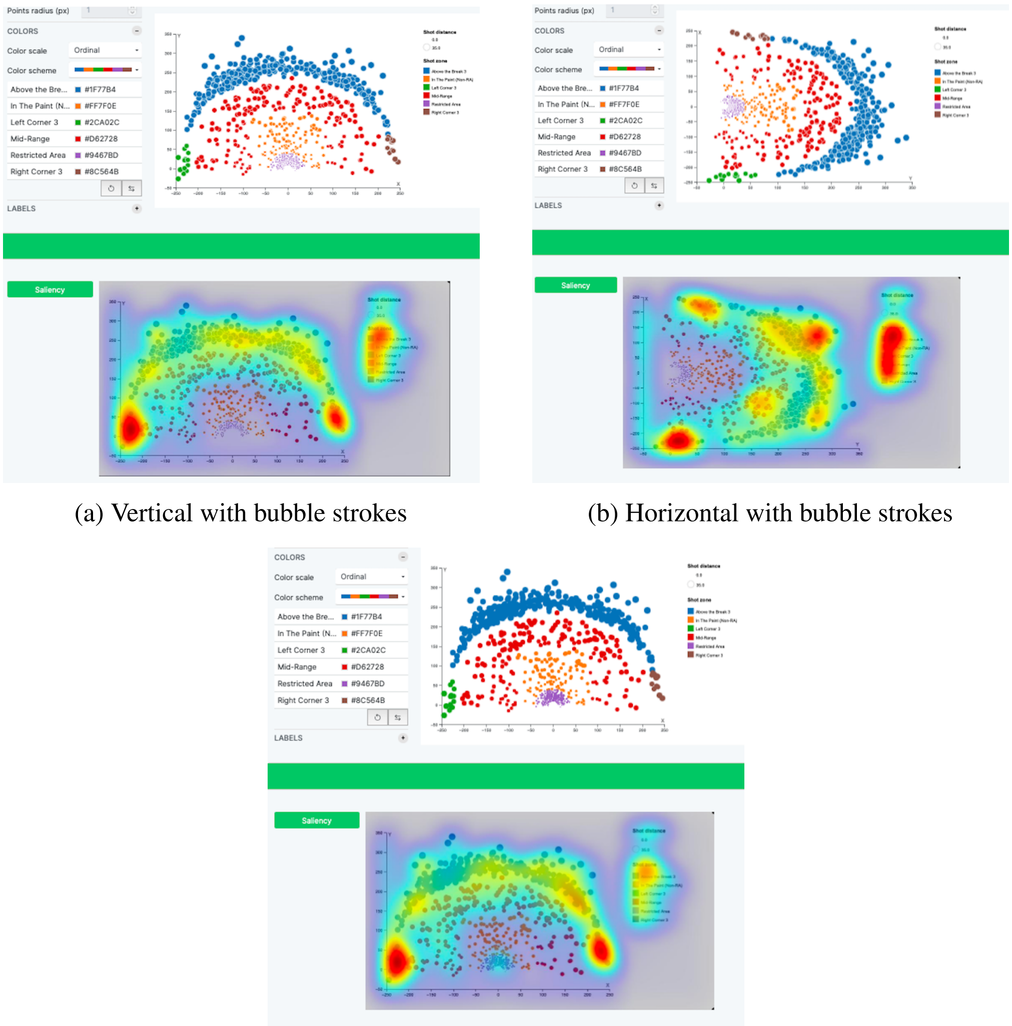
(b) RawGraphs framework interface changes.

Fig. 7.4 Measurement Tool user interface original and changes

As a result, the integration of the saliency maps into a standard graph design framework was successfully achieved. The Fig. 7.5 presents an example of the tool's usage. The graph designer created a *Bubble* chart with the goal of highlighting the initial and final data group (those closest to the x-axis). In the first design, Fig. 7.5a, the desired data, according to the saliency map, has high attention. However, the legend also draws a large part of the attention. If the designers want to change the focus of attention and the graph orientation, as shown in Fig. 7.5b, they change the focus of attention again. This change moves attention slightly longer to the legend and attracts attention to the data farther away from the graph (the farthest from the x-axis).

In the third modification example, in Fig. 7.5b, it is possible to observe how making changes that seem visually not so striking, such as removing the bubble stroke (white border), causes a more distributed saliency by moving the saliency also to the center of the graph. The third modification is more close to the graph designer's attention objective.

On the other hand, unlike the previous development (see chapter 6), the time to see the results of the salience is between 30 and 70 seconds, instead of the 8 minutes that the other ones require. This difference in time is essential; although both developments have different objectives, the fact that the functionality time is shorter helps the user to get engaged more easily in using the tool.



(a) Vertical with bubble strokes

(b) Horizontal with bubble strokes

(c) Vertical without bubble strokes

Fig. 7.5 Measurement Tool usage examples. Bubble chart with different design modifications and the saliency map showing the attention impact.

With this new functionality, the graph designer can see where the final observer's attention may focus. As stated earlier, graph designers need a validation tool to facilitate the validation process of their design decisions to a meaningful degree. Additionally, there is a learning component. Each time the graph designers change the design, there can see how the attention changes and learn how each design decision influences the highlighted visual elements.

In this chapter, we presented how saliency prediction may be integrated as a first-class feature into a DataViz system. This feature allows an exploratory approach through which the designer personally explores and evaluates (with the assistance of the tool). However, to see if this development has any potential among expert graphic designers, we performed a set of experimental evaluations and validation that will be described in the next chapter.

Chapter 8

Experts Evaluation

In order to validate the developments presented in chapters 6 and 7, we carried out validations with experts in data visualization but with different perspectives, one from academia and the other from the industry.

The main objective was to validate with Data Visualization experts the feasibility of use of Saliency Prediction in the DataViz design process. This section will present the expert's validation protocol and results. We made the validation with six (6) experts in the area of DataViz. The results demonstrated that integrating saliency prediction would be useful and can improve design task performance. In addition, although some of the experts disagree that the use of saliency prediction can make their jobs easier, it can help to shorten graph evaluation time.

In this chapter, for the sake of simplicity, we will call the Saliency Prediction as a Design Tool approach as **Design Tool**, and the Saliency Prediction as a Measurement Tool approach as **Measurement Tool**.

8.1 Fundamentals

The expert's validation protocol is based on the *heuristics validation methodology*. This methodology involves examining an interface and attempting to form a judgment about what is excellent and what is terrible about it [131]. There are a variety of heuristics that can be used to make subjective measurements on a tool [132]. However, most of them are oriented to interaction and usability. Since the validation applies to proof-of-concept applications, only, we decided to adopt some parts of each heuristic, only, that were related to perceived usefulness, satisfaction, overall reactions, and ease of use.

The heuristics methodology is commonly performed through the use of questionnaires. Perlman presented in [132] a list of a few standard questionnaires for user interface evaluation. From this list of usability questionnaires, we choose parts of USE (Usefulness, Satisfaction, and Ease of use) and QUIS (Questionnaire for User Interface Satisfaction) for the first approach (Design Tool). To validate the second approach, the Measurement Tool, we chose TAM (Technology Acceptance Model)

The USE instrument [133] measures the subjective usability of a product or service. Furthermore, because it is non-proprietary and technology-agnostic, this metric can be used in a variety of usability evaluation contexts. USE comprises 30 questions divided into four segments: usefulness, ease of use, learning ease, and satisfaction. Each question has a 7-point Likert rating scale [133]. According to Gao *et al.* [134] the reliability of USE has been confirmed with Cronbach's $alpha = 0.98$ (maximum value is 1.0), which means USE could be a valid and reliable instrument to measure usability.

On the other side, the QUIS instrument [135] focuses on how an interface is evaluated, and it was created based on Shneiderman's list of five different types of dependent measures [136]. QUIS has in total 27 questions, that are grouped in five dependent measures: overall satisfaction, screen, terminology and information, learning, and system capabilities. The questions have a 10-point scale. The reliability of QUIS was confirmed with Cronbach's $alpha = 0.98$ according to [137].

The last selected instrument was TAM, also known as *Perceived Usefulness (PU) and Ease of Use (PEU)*. The model intends to forecast the future use of a product (expected usefulness and ease-of-use as viewed before any use) rather than to rate the actual user experience. In the TAM, PU denotes how confident a person is that technology will improve job performance, whereas PEU denotes how confident a person is that using the technology will be simple [138]. The TAM instrument has 12 items in total, six for the PU and six for PEU, and also is a generalizable instrument across different systems, user groups, and research settings [137]. The reliability of TAM instrument was confirmed with Cronbach's $alpha = 0.98$ according to Lah *et al.* [138]

Other instruments, that according to recent studies such as Hodrien and Terrence [137], are more widely used, such as SUS (System Usability Scale), were considered to be beyond our scope. The instruments mentioned above are oriented to validate the interface design and its interaction by executing specific task completion scenarios. These validation characteristics presented a limitation for our expert's evaluation since the developed tools were intended to be assessed only at the conceptual level rather than as a complete system. Additionally, some of these instruments have been built only for specific contexts, unlike QUIS, USE, and TAM, which are not connected to specific technologies.

After selecting the measurement instruments, we took the aspects we were interested in evaluating. Besides, from QUIS, we took only the questions related to "Overall Reaction to the Software." These questions would help us know the general reaction that the evaluation would have to the *Design Tool* approach. In addition, regarding the USE instrument, we chose three groups of questions: Perceived Usefulness, Perceived Easy to Use, and Perceived Satisfaction. This instrument was also used for *Design Tool*. On the other hand, from the TAM instrument, we use all questions, and it was used for the *Measurement Tool* approach.

Finally, in the selection of the appropriate number of experts for this proof-of-concept, we took as a reference what was established by Nielsen [131]. Nielsen study established that *"the number of usability results found by aggregates of evaluators grows rapidly in the interval from one to five evaluators but reaches the point of diminishing returns around the point of ten evaluators."*. For these reasons, Nielsen claims that it is "*reasonable to recommend the use of about five evaluators*" to get an acceptable reliability validation. Based on this statement, we performed the validation with six experts, three with academic backgrounds and three from the industry.

8.2 Protocol

Firstly, we established an **expert's profile** in which the expert has at least these characteristics:

- *Expertise areas*: informatics, statistics, data visualization or data analysis.
- *Years of expertise*: minimum 5 years working in data visualization.

- *Current Job*: academic or industry.

The above characteristics ensured that the evaluators had a broad background in data visualization. We also considered looking for experts in the business sector because the tools developed are intended for inexperienced users in the visual-cognitive area of graph design. In addition to the minimum profile established, we looked for experts from different countries so that we would have a globalization component in the evaluation. Finally, in the search for experts, we sought a gender balance.

Concerning to the validation protocol, we established eight steps:

1. Introduce the saliency prediction process, DataViz design process, and Relevant Data. Also, explain the project aim and the integration of the previous concepts in the graph design process.

2. Collection of basic data from the expert.

3. Open questions about highlighting relevant data and the conscious selection of design elements.

4. Demonstration of the first approach, *Design Tool*, with the explanation of how it performs the variations of the design elements. Both developed versions (Matlab and Python) are shown.

5. Expert compiles USE and QUIS selected questions regarding the Design Tool prototype.

6. Inspection question about expert's ability to detect points of attention on a graph.

7. Demonstration of the second prototype, *Measurement Tool*, with the explanation of how it is connected with the saliency algorithm. The result of the graph shown in the previous step and the realization of a completely different graph is shown.

8. Experts must validate the feasibility of the *Measurement Tool* integration by TAM instrument.

The protocol was carried out with six experts, having separate meetings with each expert and in a virtual format. Thus, the evaluation protocol begins with a brief explanation of the preattention process, the concepts of salience prediction, and its relationship with the cognitive process. Afterward, basic data was collected from the expert, such as years of experience, age, gender, and country where he/she works.

In the protocol steps number three (3) and six (6), to inquire about the expert's knowledge of the visual-cognitive processes that may influence design decision making, we generated some **open-ended questions**:

- *Generally, when you design a graph, do you have in mind that one or more data points should be highlighted?*
- *When you choose visual elements such as color or orientation, do you think you are consciously or unconsciously selecting them?*
- *Generally, when you design a graph, do you have any kind of tool to validate your design decisions?*
- *Do you know the visual-cognitive impact that design decisions have on the observer? do you think it would be useful*?

The first two questions correspond to the protocol step four (4). The purpose of these questions was to understand if it was natural for the expert that some data had to be highlighted and its visual attention implications. In addition, we asked about the level of conscience with which the expert chooses the design elements. The intention of the other two questions for the protocol step six (6) was to inquire about the expert's process to validate the graph design. Also, we wanted to know how aware the expert was of the design decisions she made.

During the demonstration steps of the prototypes, steps four (4) and seven (7), a full working sample was shown, along with an explanation of how it worked. Initially, a video of the Matlab version was shown for the first prototype because it takes more than 20 minutes to generate a result. On the other hand, the Python version was presented with an established data set in real-time. Finally, for the second prototype, a real-time sample was also demonstrated, first with a data set and then with the test data provided by the framework. For this second prototype, it was essential to highlight that the functionality was also shown with other graphs, not only the bar graph, which is the base graph in the first prototype.

Finally, in order to obtain information about the experts' perception of the presented prototypes, we used the validation methods explained in section 8.1. For the first prototype, we used both QUIS and USE questionnaires. Both questionnaires were chosen because the “Design Tool” prototype is independent. This means that the full prototype is developed entirely, contrary to the Measurement Tool where we integrated the salience into an existing application. Due to the above, it was also necessary to know what the general perception of this development is, and QUIS has a dedicated section of questions, “Overall Reaction,” that would give us information about the general vision of having an application like this (see Table B.1). About the USE instrument, the used questions can be seen in Table B.2.

Regarding the perception validation of the second prototype, we only used the TAM instrument to know if the inclusion of saliency could mean an improvement in its design processes (see Table B.3). We omitted an interface validation or general responses since we did not develop the base application. We only made some minor improvements. The idea was to evaluate if integrating a saliency map viewing functionality in a typical data visualization application could be helpful in their work. Appendix B presents the complete list of questions chosen for each validation instrument.

The first part of the protocol was to know the experts' profiles who participated in the validation. In Table 8.1, the country of origin, age, gender, areas of expertise (expressed by each of them), and years of experience working in data visualization are presented.

The Table shows a balanced group of experts, half were women, and half were men. On the other hand, in the column "Expertise Area," those working in academia have an (A) and those from industry an (I). The previous parameter is also balanced, three in academia and 3 in industry. Finally, with this group of experts, we also seek to have a more geographical vision on the subject, which is reflected in the four countries where the experts work.

In the next sections, we will present the results of each development tool according to two views of a point, academic and industrial.

Table 8.1 Experts Profiles. Those with the **A** are working in academia. Those with the **I** are working in the industry.

Country	Age	Gender	Expertise Area	Years
Colombia	29	Woman	(I) Data Visualization Expert	6
Canada	45	Man	(A) Higher Education/Research Information Visualization, HCI	10
France	37	Man	(I) Software Development/Data Visualization	5
Italy	41	Man	(I) Project Manager in Business Intelligence	11
Italy	32	Woman	(A) Computational Biology	7
Colombia	44	Woman	(A) Informatics	8

8.3 Design Tool Results

In this section, we will show the results obtained divided according to the two main points of view: academic and industry.

8.3.1 Academic Experts

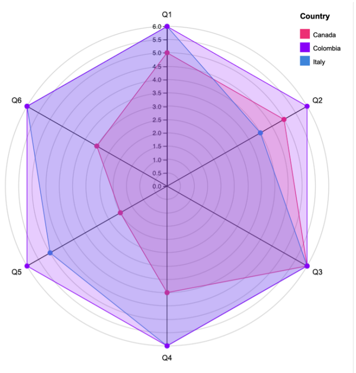
From the results of the protocol step tree, open questions, we noticed that, in general, everyone is aware that each graph has one or more pieces of information that should be highlighted. Moreover, that data is already known by them, i.e., in all cases, after analyzing the data, they already know which data they want to highlight. Regarding the selection of design elements such as colors or orientation, they all agreed that they consider themselves to be aware of their use. Most of them already have a basic knowledge of color management. They also know what colors work for their graphics and that these elements should also be selected according to the data context (observers and data meaning). Another essential point to mention is that most experts spoke only about color attributes. However, only one expressed the importance of making a good selection of the graph orientation and position or the graph within others (as in a dashboard).

In Fig. 8.1 shows the results of the three groups of questions we chose from the USE for each expert measurement instrument. The first group of questions, *Usefulness* (see Fig. 8.1a), are oriented to validate if for the users the presented tool could help them to be more productive, effective, and in general, save time. As we can see in the Fig. 8.1a graph, two of the experts consider that this tool could help them to be more effective in performing their visualization tasks. The expert who provided the lowest scores is one of those who has the most experience designing and interacting with graphs. Because he already has extensive experience in this area, it might not help with task execution time. However, the evaluator emphasized the importance of these tools for more inexperienced people and felt that they would be helpful in less academic contexts.

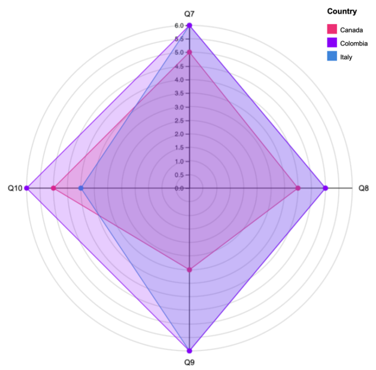
In the second group of USE questions, Ease of Use (see Fig. 8.1b), we can see that everyone agrees that the tool is not entirely “easy to use”(Q8). However, in “I can use it without written instructions.”(Q9), two evaluators expressed that they could use it without instructions. The third expert explained that it was not easy to use without instructions because the saliency concept had to be explained, along with the process the tool goes through to select the winning graphs. Without this prior explanation, the reliability of the process is not clear, and for someone in the academia, many doubts about the process would arise. About the third group, Satisfaction (see Fig. 8.1c), we can notice that there is general Satisfaction with the tool because all experts would recommend it to a friend. However, some do not feel that they need to have it.

Regarding the QUIS validation results, Table 8.2 presents the average rating and the median scores for each question scale. The overall reactions to the tool are positive in general. It can be emphasized that the tool was considered “Wonderful” and “Satisfying” by all of them, which means they see it as a tool that they would feel motivated to use. However, one of the lowest scores was seen in the “Flexible” criterion, this could be because the tool currently only works for one type of graph, and it is not interactive graph (e.g., zooming to see more details). The output is a static image.

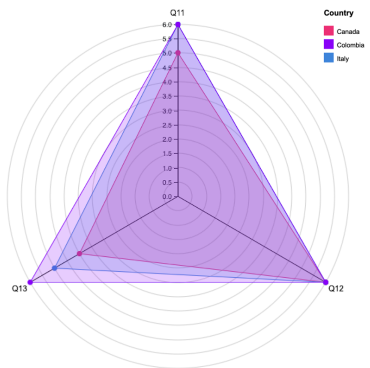
In general, all three surveyors perceived the Design tool as an exciting development, which can be extended and which they find very useful, especially for support in design decision making.



(a) USEFULNESS results



(b) EASE OF USE



(c) SATISFACTION

Fig. 8.1 Results of USE validation instrument from Academic Experts

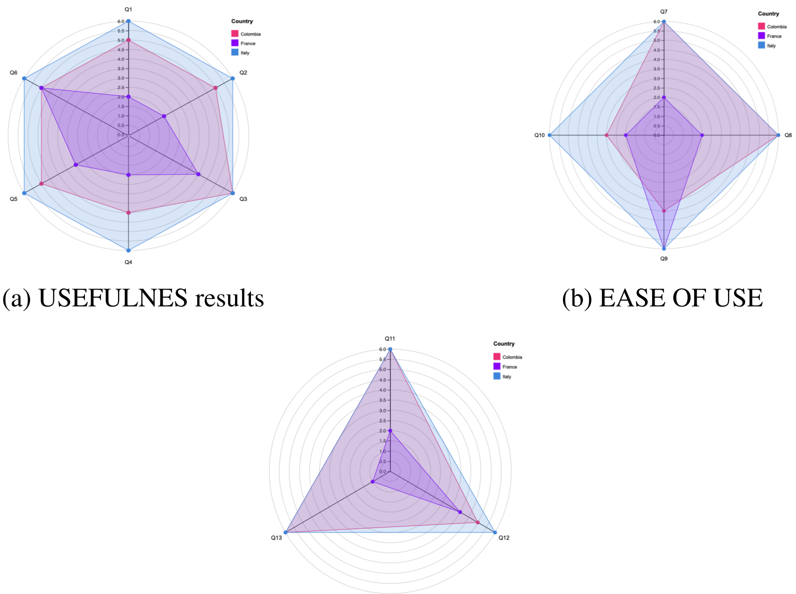
8.3.2 Industrial Experts

On the other hand, the results obtained from industrial experts slightly differ from the perception of academics. The main points of difference are the Usefulness and the Ease of Use of the tool. Regarding the USE instrument results presented in Fig. 8.2, we can see that, in general, two of the experts perceived the tool as useful, easy to use, and were satisfied with it. One of these evaluators has the amplest experience in the visualization area in a software development company. However, another of the evaluators expressed that even though he found the tool useful and could help reduce some validation process time, he did not feel that such a tool was really indispensable to him in his job tasks (travel data visualizations).

One of the questions with the lowest scores was "Both occasional and regular users would like it."(Q10). This low score on the Q10 question was because the experts thought that the tool is very specialized, and someone who does not do

Table 8.2 QUIS Academia Experts

Scale	Average Rating	Median Score
Terrible *0* . . . *9* Wonderful	8,3	8,0
Difficult 0 . . . 9 Easy	7,3	7,0
Frustrating 0 . . . 9 Satisfying	8,3	8,0
Inadequate 0 . . . 9 Adequate	7	8,0
Dull 0 . . . 9 Stimulating	7,7	9,0
Rigid 0 . . . 9 Flexible	5,3	8,0



(a) USEFULNES results

(b) EASE OF USE

(c) SATISFACTION

Fig. 8.2 Results of USE validation instrument from Industrial Experts for Design Tool

visualizations all the time may not need the tool. Another question whose answers caught our attention was Q4. As shown in Fig. 8.2a, each expert gave a very different score. The question Q4 was, "It gives me more control over the activities in my daily DataViz tasks." Regarding this, some of the experts expressed that the tool would not give them more control due to the fact that they would still have to go through a validation process within the company focus group. However, they said that the tool could reduce these validation times.

One of the comments made about the tool's design was that it should give the possibility to choose each of the graph colors. The above is because many of the

Table 8.3 QUIS Industrial Experts results for Design Tool validation

Scale	Average Rating	Median Score
Terrible 0 ... 9 Wonderful	7,3	8
Difficult 0 ... 9 Easy	6	7
Frustrating 0 ... 9 Satisfying	6,3	8
Inadequate 0 ... 9 Adequate	6,3	8
Dull 0 ... 9 Stimulating	7	9
Rigid 0 ... 9 Flexible	7,3	8

companies have their own color palettes, and they have to play with these to maintain the client's parameters.

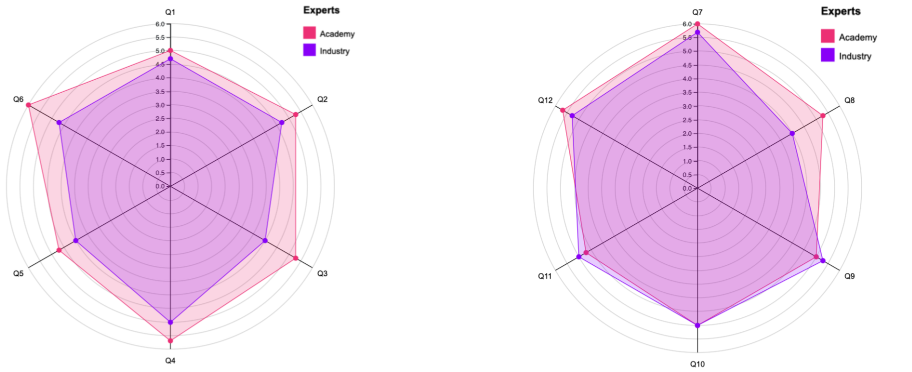
The results obtained from the QUIS questions were as divided as those of USE. Two of the experts scored the majority of the items above seven points. However, the other expert gave values below five. The above difference can be seen in Table 8.3. For instance, the "Inadequate - Adequate" scale has an average of 6.3, but the median is 8,0, demonstrating that at least two experts selected a score over seven.

This validation showed us a very relevant aspect of our development concerning its usefulness in different contexts. We see it reflected in the fact that experts working on more general developments, such as dashboards, perceived the tool as very useful, a bit complicated to use but with great potential. On the other hand, in contexts where the graph design is a bit more specialized, working on the same data (maps graph), they found the tool useful, and well developed but did not consider it a fundamental tool to improve or support their daily tasks.

8.4 Measurement Tool

In the evaluation of this Measurement Tool, the evaluators had to answer the questions of the TAM survey.

Initially, we asked open-ended questions about what design validation tool each expert uses and if they have some ideas about the visual impact of each design decision. The industry experts expressed that the design validation is done with members of the same development team or with a panel of experts within the same company. Additionally, they said that cross-validation is a time-consuming process. One of them expressed that he does not use any validation tool, her decisions are



(a) TAM Perceived Usefulness (PU) (b) TAM Perceived Ease of Use (PEU)

Fig. 8.3 Results of TAM validation instrument from Academic and Industry Experts

based on experience. The academic experts indicated that their only validation tool is their knowledge and experience. Concerning whether the experts are aware of the design decisions' visual impact, more than half of the experts said that they were not fully aware of the topic. Only one of the experts, who has the most experience in visualization design, expressed that this knowledge is acquired with time, although it is only from academia.

Finally, the experts answered the TAM questions in the last part of the validation. Fig. 8.3 shows the results, in pink the average of the academic's answers and purple those of the industry experts. In Fig. 8.3b, the scores were between five and six, with six being the highest possible value, especially in the "perceived Ease of Use" part. In this TAM-PEU part, we can see a consensus among the expert groups. The only question that had a low value (4.0 on average), especially for the industry experts, was, "I would find it easy to get Saliency Prediction Tool to do what I want it to do." This can be interpreted, along with some comments they made to us, that they work more on dashboards in their daily tasks, and the tool presented is for single charts.

On the other side, in the TAM-PU questions part (see Fig. 8.3a), a little more difference can be seen between the two groups of experts, but not more than one score point difference. Academics have the highest usability perception of having saliency in the application, while industry experts think it is slightly less useful. In the lowest-scoring question, "Using Saliency Prediction in my job would increase my productivity" (Q3), industry experts expressed that although the tool would certainly help them to reduce design time, they do not believe it would increase productivity.

Overall, the final expert's reaction was enthusiasm for integrating saliency in a data visualization system. They said that it is a relevant tool and that they would like to be able to access it. On the other hand, especially those who work in the industry said that they wished the tool was also intended to be used for dashboards and show more information about the saliency percentages seen on the map.

We will discuss the results obtained on the two tools developed in the following section.

8.5 Discussion

Design Tool Prototype

In general, we could notice that all experts know that one piece of data is always more important than the others. Besides, it highlights that data is also part of their visualization tasks. These statements confirmed that this task is fundamental in the graph design process, so making careful design decisions to highlight data is also relevant.

The academia experts expressed that it was a tool with much potential and would be very useful. In addition, they expressed that although the saliency model did not have 100% reliability, it was clear to them that this design tool is a visual support for decision-making only. From the industry experts' perspective, they had many questions about the development, specifically how it had been implemented for future use.

Finally, both expert groups agreed that the tool should provide more information about saliency. For example, the tool may show the percentage of attention on each graph element. Furthermore, it should be a bit more ample with other visualization types, such as dashboards, taking advantage of the position change function that the tool already handles. Regarding the two tool versions (Matlab and Python), both expert groups indicated that the tool made in Python was much better in terms of appearance and usability than the one developed in Matlab.

Measurement Tool Prototype

During the interviews with each expert, we noticed that the visualization validation process is done in two ways: showing the visualization to a focus group or by using the experience. Showing the visualization to a focus group is used in companies, and the graphs are shown to colleagues in the same area or a specialized User Experience group. However, that process can take a considerable amount of time. On the other hand, the rest of the experts validate their designs based on their experience and theoretical design knowledge. Based on these statements, we can state that a tool is needed to support the validation process of a graph designer's intent.

The experts highlighted two complementary advantages that saliency prediction integration could offer. The first is that it is a learning tool. This statement is because specific attention behavior is learned if the tool is used for an extended period. Therefore, the graph designer is acquiring knowledge while validating the graph's attention. The second benefit is expert-oriented since there may be an acquired data visualization bias over time. For example, an expert who has been working on the same type of data for a long time, who has seen it for a long time, may have a bias since it is already clear to him what the relevant data is. However, it may not be the same for a final observer who sees the data for the first time.

Like the other tool, the experts said it would be interesting if it were more extensive and could be used for dashboards. Additionally, the academia experts mentioned again that they would like more information about saliency reflected in the resulting saliency map. For example, divide the graph into four quadrants and indicate the salience percentages in each quadrant.

In conclusion, most results demonstrated that integrating saliency prediction into the design process, both as a Measurement Tool and Design Tool, is a relevant and valuable approach. It is noteworthy that the experts mostly expressed that they had not seen this type of support in other tools, and they have significant potential.

Chapter 9

Conclusions

> *"The greatest value of a picture is when it forces us to notice what we never expected to see."* John Tukey (U.S. statistician)

The thesis objective was to integrate visual-cognitive concepts (particularly saliency prediction) into the DataViz design process to bring them closer to the graph designer and support the design decisions. To achieve this objective, firstly, we conducted a **Literature Review** to understand how the preattentive process, the first step in human attention, has been used in InfoVis in the last ten years. We found that understanding the impact of preattentive attributes (e.g., colors, orientation) is mainly used implicitly, which means they were used to generate design patterns and incorporated into design systems. However, graph designers are oblivious to their visual impact. On the other side, we found that the effects of the preattentive process are commonly used for graph validation. However, graph validation takes place in laboratories with controlled environments and takes an extended period. Finally, we were unable to locate any tool that could show the graph designer the influence of their design decisions on graph attention during the design and validation phases. To sum, we identified a gap between the existing knowledge about the visual-attention impact and the graph designer.

We realized that we could bring graph designers closer to the concepts of human visual impact through **saliency prediction**. Saliency prediction models have been developed over 30 years, mainly oriented to object detection. However, in the last five years, we found that some models have been designed specifically for attention analysis in InfoVis images. One of the essential factors in the limited use of saliency prediction models in the field of InfoVis is the difference between commonly natural images used (e.g., landscapes) and graph images: InfoVis images are born digital; they have meaningful small objects; different color scales; the white space between visual elements is higher than in natural images; and an essential present of textual elements. Furthermore, these new models specialized for InfoVis have been developed with classical theories (Bottom-up and Top-down) and with more modern methods such as deep learning.

Given the relative novelty of the InfoVis saliency models, we decided to perform a **experimental validation** of the accuracy of such saliency models. As a result, we found that models based on classical (bottom-up) saliency prediction obtained the best results under three different scenarios. First, using MASSVIS, a commonly used dataset with more than 300 InfoVis images, together with the saliency data. In the second scenario, we used another dataset but with *clean graphs*, created by us, with only essential visual elements (titles, axes, and values), and with ground data collected from 80 observers using an eye-tracking device. Finally, the third scenario with graphs without textual elements only focuses on the attention behavior in data-contained elements. In addition, we found some insights about the saliency behavior on specific features, e.g., the position, and orientation of the data-contained elements (bars in a bar chart) have an apparent influence on saliency behavior. Finally, based on these results, the Matzen model, which used a classical saliency model (Bottom-up) combined with the text saliency model (Top-Down), obtained the best result on average in each scenario.

> *“To find signals in data, we must learn to reduce the noise - not just the noise that resides in the data, but also the noise that resides in us. It is nearly impossible for noisy minds to perceive anything but noise in data.”*
> – Stephen Few (founder and principle of Perceptual Edge)

In order to use the saliency model selected above and make it part of the graph design process in an explicit way, we decided to address the two steps of the graph design process: the initial design and the validation step. For the initial design step, we use saliency prediction as a **Design Tool**. We developed a tool in which the designers can choose the data they want to highlight, and it presents three options of graphs that would highlight that specific piece of data. First, the graphs are generated by systematically varying color, position, orientation, and texture. Then, the tool applies the saliency prediction to choose the combinations showing the highest percentage of saliency in the selected data. This tool has Matlab and Python versions.

For the second selected design step, the graph validation, we developed a second tool that uses saliency prediction as a **Measurement Tool**. This development aimed to allow the graph designer to measure how each of her design decisions would impact the observer's attention. The tool is a data visualization web application integrated with a saliency map generation function. As a result, the graph designer can choose from a variety of graphs, make design changes, and interactively and iteratively see how each of these choices will impact the observer's attention on the data (visual graph elements).

Finally, we conducted validations with experts in data visualization from distinct viewpoints, three from academia and three from industry, to validate the two advanced tools. In general, the academic experts perceived the *Design Tool* as an exciting development that can be extended and which they find very useful, especially for support in design decision-making. On the other hand, the industrial experts slightly differ in the usefulness and the ease of use criteria because they perceive the tool as a bit complicated to use but with great potential.

The expert's reaction was enthusiasm for integrating saliency in a data visualization system for the *Measurement Tool*. They said it is a relevant tool that could help reduce validation times as it is generally a lengthy cross-validation process. In addition, the experts highlighted two complementary advantages that saliency prediction integration could offer: learning tools and bias-breaker. First, the tool can be used as a learning instrument because the designer will have acquired a lot of information about which design decisions have the most impact after a period of use. Secondly, the tool can be a bias-breaker because if an expert has been working on a visualization for a long time, the relevant data is already clear to her. Still, it may be perceived differently by the final observer.

The experts said it would be interesting if both tools were more extensive and could be used for dashboards. Additionally, they mentioned that they would like more information about saliency reflected in the resulting saliency map (saliency percentage per area).

> *"How we visualize data will evolve into more complex forms that better communicate uncertainty and complexity."* – Amanda Makulec (Data viz designer, teacher & speaker)

Although our approaches are useful, they are limited by the saliency prediction model's weaknesses. Firstly, saliency models for DataViz are for "explorative tasks," but in DataViz, there are more tasks (e.g., correlate, find anomalies, retrieve value). Due to the relationship between human vision and the human cognitive process, the nature of the task should also be included in how it influences attention. Another limitation of our approach is the response time. Although we improved the response time by making the tools web-based, it is still a shortcoming, especially in the design tool. Although, according to our understanding, this could be improved with a better machine capacity or parallel computing.

> *"Visualization gives you answers to questions you didn't know you had."* by Ben Schneiderman (Computer scientist and notable referent in HCI)

According to Schneiderman in the above quote, visualization can provide a lot of information. Evidently, the construction of visualizations focused on properly showing that information to the observer. However, the research studies ignore the graph designers, who are responsible for combining many visual elements so that their visualizations are clear to the observer. In this thesis, we presented an investigation of how the same visualizations can also answer questions the graph designers do not even know they have. Furthermore, we prove that information can be gleaned from the same graph to support the graph design process with saliency prediction.

> *"The purpose of visualization is insight, not pictures."* by Ben Schneiderman

Future Work. Researchers are now working on customized saliency prediction algorithms for InfoVis and DataViz, so much work is still to be done in this field. With the insights found during the validation of these models, it would be possible to start working on improving saliency models. For example, InfoVis Saliency models could have two levels of detail: a general one on the whole graph (titles, captions, context images); and a second level that would predict the attention behavior specifically on the data-containing elements. The presented developments could be extended and used in InfoVis by adding a study on how to work with saliency in dashboards.

Regarding the developed tools, we consider it possible to extend the graph types to be used. First, additional experiments should be carried out on graphs other than bar graphs to validate the saliency model's performance with them. Then, within the developed tools, we should have a profile of each graph to know which visual elements can be varied without affecting the consistency of the graph.

Finally, it is also open to future work on the amount of detail that can be given on the saliency in the graph, for example, showing saliency percentages by zones or making static calculations on saliency variations from one design to another.

> *"The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them."* by William Lawrence Bragg, a British physicist.

References

[1] Rahul Hooda, Ajay Mittal, and Sanjeev Sofat. An efficient variant of fully-convolutional network for segmenting lung fields from chest radiographs. *Wireless Personal Communications*, 101, 08 2018. DOI: https://doi.org/10.1007/s11277-018-5777-3.

[2] H. Jänicke and M. Chen. A salience-based quality metric for visualization. *Computer Graphics Forum*, 29(3):1183–1192, 2010. DOI: https://doi.org/10.1111/j.1467-8659.2009.01667.x.

[3] Yanyu Xu, Shenghua Gao, Junru Wu, Nianyi Li, and Jingyi Yu. Personalized saliency and its prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(12):2975–2989, 2019. DOI: https://doi.org/10.1109/TPAMI.2018.2866563.

[4] Simone Frintrop, Erich Rome, and Henrik I. Christensen. Computational visual attention systems and their cognitive foundations: A survey. *ACM Trans. Appl. Percept.*, 7(1), jan 2010. DOI: https://doi.org/10.1145/1658349.1658355.

[5] Ali Borji and Laurent Itti. State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):185–207, 2013. DOI: https://doi.org/10.1109/TPAMI.2012.89.

[6] Colin Ware. *Information visualization: Perception for design*. Morgan Kaufmann, 4 edition, 2020.

[7] Riccardo Mazza. *Introduction to Information Visualization*, volume I. Springer-Verlag London, 2009. DOI: https://doi.org/10.1007/978-1-84800-219-7.

[8] Jia Li and Wen Gao. *Visual Saliency Computation: A Machine Learning Perspective*. Springer Publishing Company, Incorporated, 2014. ISBN: 978-3-319-05641-8.

[9] Amy Alberts and Andy Cotgreave. Eye Tracking: What it teaches us about dashboard design, June 2017. https://tabsoft.co/2lBMGpT.

[10] Michael J. Haass, Andrew T. Wilson, Laura E. Matzen, and Kristin M. Divis. Modeling human comprehension of data visualizations. In Stephanie Lackey and Randall Shumaker, editors, *Virtual, Augmented and Mixed Reality*, pages 125–134, Cham, 2016. Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-39907-2_12.

[11] Goran Milutinović, Ulla Ahonen-Jonnarth, and Stefan Seipel. Does visual saliency affect decision-making? *Journal of Visualization*, 24(6):1267–1285, 2021. DOI: https://doi.org/10.1007/s12650-021-00760-4.

[12] Milica Milosavljevic, Vidhya Navalpakkam, Christof Koch, and Antonio Rangel. Relative visual saliency differences induce sizable bias in consumer choice. *Journal of Consumer Psychology*, 22(1):67–74, 2012. DOI: https://doi.org/10.1016/j.jcps.2011.10.002.

[13] Anne M. Treisman and Garry Gelade. Feature-integration theory of attention. *Cognitive Psychology*, 12(1):97–136, January 1980. DOI: https://doi.org/10.1016/0010-0285(80)90005-5.

[14] C. Healey and J. Enns. Attention and Visual Memory in Visualization and Computer Graphics. *IEEE Transactions on Visualization and Computer Graphics*, 18(7):1170–1188, July 2012. DOI: https://doi.org/10.1109/TVCG.2011.127.

[15] Jose Fernando Rodrigues, Andre G. R. Balan, Agma J. M. Traina, and Caetano Traina. The Visual Expression Process: Bridging Vision and Data Visualization. In Andreas Butz, Brian Fisher, Antonio Krüger, Patrick Olivier, and Marc Christie, editors, *Smart Graphics*, Lecture Notes in Computer Science, pages 207–215. Springer Berlin Heidelberg, 2008. ISBN: 978-3-540-85412-8.

[16] Evanthia Dimara, Anastasia Bezerianos, and Pierre Dragicevic. The attraction effect in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):471–480, 2017. DOI: https://doi.org/10.1109/TVCG.2016.2598594.

[17] Anne M. Treisman and S. Gormican. Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95(1):15–48, January 1988. DOI: https://doi.org/10.1037/0033-295X.95.1.15.

[18] Christine H. Chih and Douglass S. Parker. The persuasive phase of visualization. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 884–892. ACM, August 2008. DOI: https://doi.org/10.1145/1401890.1401996.

[19] Colin Ware. Visual Salience and Finding Information. In *Information Visualization, Perception for Design*. Morgan Kaufmann, 225 Wyman Street, Waltham, MA 02451, USA, 3rd edition edition, 2013. ISBN: 978-0-12-381464-7.

[20] Stephen Few. Tapping the Power of Visual Perception, September 2004. DOI: https://www.perceptualedge.com/articles/ie/visual_perception.pdf.

[21] J. M. Findlay and I. D. Gilchrist. Eye guidance and visual search. In G. Underwood, editor, *Cognitive processes in eye guidance.*, pages 259–281. Oxford University Press, Oxford, July 2005. DOI: https://doi.org/10.1093/acprof:oso/9780198566816.001.0001.

[22] Dirk Kerzel. Perception in Vision. In Marc D. Binder, Nobutaka Hirokawa, and Uwe Windhorst, editors, *Encyclopedia of Neuroscience*, pages 3098–3102. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. DOI:https://doi.org/10.1007/978-3-540-29678-2_4455.

[23] Marc D. Binder, Nobutaka Hirokawa, and Uwe Windhorst, editors. *Visual Salience*, pages 4336–4336. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. DOI: https://doi.org/10.1007/978-3-540-29678-2_6371.

[24] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998. DOI: https://doi.org/10.1109/34.730558.

[25] Anne Treisman. How the deployment of attention determines what we see. *Visual Cognition*, 14(4-8):411–443, 2006. DOI: https://doi.org/10.1080/13506280500195250.

[26] Tobii. What is eye tracking? https://www.tobii.com/group/about/this-is-eye-tracking/. Accessed: 2022-02-07.

[27] Interaction Design Foundation. Information visualization. https://www.interaction-design.org/literature/topics/information-visualization. Accessed: 2022-02-07.

[28] Gartner Glossary. Data visualization. https://www.gartner.com/en/marketing/glossary/data-visualization. Accessed: 2022-02-07.

[29] D. Alves, K. Obraczka, and R. Lindberg. Identifying relevant data center telemetry using change point detection. In *2020 IEEE 9th International Conference on Cloud Networking (CloudNet)*, pages 1–4, 2020. DOI: https://doi.org/10.1109/CloudNet51028.2020.9335800.

[30] Z. Dzolkhifli, H. Ibrahim, L. S. Affendey, and P. Madiraju. A framework for caching relevant data items for checking integrity constraints of mobile database. In *2008 Third International Conference on Digital Information Management*, pages 476–482, 2008. DOI: https://doi.org/10.1109/ICDIM.2008.4746801.

[31] Manuela Waldner, Alexey Karimov, and Eduard Gröller. Exploring visual prominence of multi-channel highlighting in visualizations. In *Proceedings of the 33rd Spring Conference on Computer Graphics*, page 8. ACM, May 2017. DOI: https://doi.org/10.1145/3154353.3154369.

[32] J. F. Rodrigues, A. J. M. Traina, M. C. F. de Oliveira, and C. Traina. Reviewing Data Visualization: An Analytical Taxonomical Study. In *Tenth International Conference on Information Visualisation (IV'06)*, pages 713–720, July 2006. DOI: https://doi.org/10.1109/IV.2006.94.

[33] Richard Veale, Ziad M. Hafed, and Masatoshi Yoshida. How is visual salience computed in the brain? Insights from behaviour, neurobiology and modelling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1714), February 2017. DOI: https://doi.org/10.1098/rstb.2016.0113.

[34] Dereck Toker and Cristina Conati. Eye Tracking to Understand User Differences in Visualization Processing with Highlighting Interventions. In Vania Dimitrova, Tsvi Kuflik, David Chin, Francesco Ricci, Peter Dolog, and Geert-Jan Houben, editors, *User Modeling, Adaptation, and Personalization*, Lecture Notes in Computer Science, pages 219–230. Springer International Publishing, 2014. DOI: https://doi.org/10.1007/978-3-319-08786-3_19.

[35] Luisa Barrera-Leon, Fulvio Corno, and Luigi De Russis. How the preattentive process is exploited in practical information visualization design: a review. *International Journal of Human–Computer Interaction*, pages 1–14, 2022. DOI: http://dx.doi.org/10.1080/10447318.2022.2049137.

[36] J. Liang and M. L. Huang. Highlighting in Information Visualization: A Survey. In *2010 14th International Conference Information Visualisation*, pages 79–85, July 2010. DOI: https://doi.org/10.1109/IV.2010.21.

[37] SCImago. Scimagojr, August 2021.

[38] Laura E. Matzen, Michael J. Haass, Kristin M. Divis, Zhiyuan Wang, and Andrew T. Wilson. Data Visualization Saliency Model: A Tool for Evaluating Abstract Data Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):563–573, January 2018. DOI: https://doi.org/10.1109/TVCG.2017.2743939.

[39] Zoya Bylinskii, Nam Wook Kim, Peter O'Donovan, Sami Alsheikh, Spandan Madan, Hanspeter Pfister, Fredo Durand, Bryan Russell, and Aaron Hertzmann. Learning visual importance for graphic designs and data visualizations. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, page 57–69, New York, NY, USA, 2017. Association for Computing Machinery. DOI: https://doi.org/10.1145/3126594.3126653.

[40] Camilo Fosco, Vincent Casser, Amish Kumar Bedi, Peter O'Donovan, Aaron Hertzmann, and Zoya Bylinskii. Predicting visual importance across graphic design types. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20, page 249–260, New York, NY, USA, 2020. Association for Computing Machinery. DOI: https://doi.org/10.1145/3379337.3415825.

[41] Patrik Polatsek, Manuela Waldner, Ivan Viola, Peter Kapec, and Wanda Benesova. Exploring visual attention and saliency modeling for task-based visual analysis. *Computers & Graphics*, 72:26–38, 2018. DOI: https://doi.org/10.1016/j.cag.2018.01.010.

[42] Mark A. Livingston, Laura E. Matzen, Andre Harrison, Alex Lulushi, Mikaila Daniel, Megan Dass, Derek Brock, and Jonathan W. Decker. A study of perceptual and cognitive models applied to prediction of eye gaze within statistical graphs. In *ACM Symposium on Applied Perception 2020*, SAP '20, New York, NY, USA, 2020. Association for Computing Machinery. DOI: https://doi.org/10.1145/3385955.3407931.

[43] Matthias Kümmerer, Thomas S. A. Wallis, and Matthias Bethge. Saliency benchmarking made easy: Separating models, maps and metrics. In Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, and Yair Weiss, editors, *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part XVI*, volume 11220 of *Lecture Notes in Computer Science*, pages 798–814. Springer, 2018. DOI: https://doi.org/10.1007/978-3-030-01270-0_47.

[44] Saturnino Luz and Masood Masoodian. Improving focus and context awareness in interactive visualization of time lines. In *Proceedings of the 24th BCS Interaction Specialist Group Conference*, BCS '10, page 72–80. BCS Learning & Development Ltd., 2010. DOI: https://doi.org/10.14236/ewic/HCI2010.11.

[45] M. Steinberger, M. Waldner, M. Streit, A. Lex, and D. Schmalstieg. Context-Preserving Visual Links. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2249–2258, December 2011. DOI: https://doi.org/10.1109/TVCG.2011.183.

[46] P. Craig, A. Cannon, R. Kukla, and J. Kennedy. MaTSE: The microarray time-series explorer. In *2012 IEEE Symposium on Biological Data Visualization (BioVis)*, pages 41–48, October 2012. DOI: https://doi.org/10.1109/BioVis.2012.6378591.

[47] C. Tominski, H. Schumann, G. Andrienko, and N. Andrienko. Stacking-Based Visualization of Trajectory Attribute Data. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2565–2574, December 2012.

[48] Matthew Dunlop, William Urbanski, Randy Marchany, and Joseph Tront. Leveraging cognitive principles to improve security visualization. In Rachid Benlamri, editor, *Networked Digital Technologies*, pages 262–276, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. DOI: http://dx.doi.org/10.1007/978-3-642-30567-2_22.

[49] R. Sadana, T. Major, A. Dove, and J. Stasko. OnSet: A Visualization Technique for Large-scale Binary Set Data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1993–2002, December 2014. DOI: https://doi.org/10.1109/TVCG.2014.2346249.

[50] Zhuang Cai, Yi-Na Li, Xianjun Sam Zheng, and Kang Zhang. Applying feature integration theory to glyph-based information visualization. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, pages 99–103, April 2015. DOI: https://doi.org/10.1109/PACIFICVIS.2015.7156363.

[51] Amy L. Griffin and Anthony C. Robinson. Comparing Color and Leader Line Highlighting Strategies in Coordinated View Geovisualizations. *IEEE Transactions on Visualization and Computer Graphics*, 21(3):339–349, March 2015. DOI: https://doi.org/10.1109/TVCG.2014.2371858.

[52] Marie-Christin Ostendorp, Thomas Friedrichs, and Andreas Lüdtke. Supporting supervisory control of safety-critical systems with psychologically well-founded information visualizations. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, NordiCHI '16, New York, NY, USA, 2016. Association for Computing Machinery. DOI: https://doi.org/10.1145/2971485.2971512.

[53] Richard Brath and Ebad Banissi. Using typography to expand the design space of data visualization. *She Ji: The Journal of Design, Economics, and Innovation*, 2(1):59–87, 2016. DOI: https://doi.org/10.1016/j.sheji.2016.05.003.

[54] Wen Bo Wang, Mao Lin Huang, and Quang Vinh Nguyen. A space optimized scatter plot matrix visualization. In Yuhua Luo, editor, *Cooperative Design, Visualization, and Engineering*, pages 382–385, Cham, 2016. Springer International Publishing. DOI: https://link.springer.com/chapter/10.1007%2F978-3-319-46771-9_50.

[55] John King, Kathy Sonderer, and Kevin Lynch. Cognitive benefits of a simple visual metrics architecture. In Fiona Fui-Hoon Nah and Chuan-Hoo Tan, editors, *HCI in Business, Government, and Organizations: eCommerce and Innovation*, pages 319–329, Cham, 2016. Springer International Publishing. DOI: http://dx.doi.org/10.1007/978-3-319-39396-4_29.

[56] Rich ard Brath and Ebad Banissi. Font attributes enrich knowledge maps and information retrieval. *International Journal on Digital Libraries*, 18(1):5–24, March 2017. DOI: https://doi.org/10.1007/s00799-016-0168-4.

[57] M. A. Yalçın, N. Elmqvist, and B. B. Bederson. Keshif: Rapid and Expressive Tabular Data Exploration for Novices. *IEEE Transactions on Visualization and Computer Graphics*, 24(8):2339–2352, August 2018. DOI: https://doi.org/10.1109/TVCG.2017.2723393.

[58] G. Richer, R. Bourqui, and D. Auber. Corfish: Coordinating emphasis across multiple views using spatial distortion. In *2019 IEEE Pacific Visualization Symposium (PacificVis)*, pages 1–10, 2019. DOI: https://doi.org/10.1109/PacificVis.2019.00009.

[59] Ji Hwan Park, Saad Nadeem, and Arie Kaufman. GeoBrick: Exploration of spatiotemporal data. *The Visual Computer*, 35(2):191–204, February 2019. DOI: https://doi.org/10.1007/s00371-017-1461-y.

[60] Younho Seong, Joseph Nuamah, and Sun Yi. Guidelines for cybersecurity visualization design. In *Proceedings of the 24th Symposium on International Database Engineering & Applications*, IDEAS '20, New York, NY, USA, 2020. Association for Computing Machinery. DOI: https://doi.org/10.1145/3410566.3410606.

[61] L. Zhou, M. Rivinius, C. R. Johnson, and D. Weiskopf. Photographic high-dynamic-range scalar visualization. *IEEE Transactions on Visualization and Computer Graphics*, 26(6):2156–2167, 2020. DOI: https://doi.org/10.1109/TVCG.2020.2970522.

[62] Michael Correll, Danielle Albers, Steven Franconeri, and Michael Gleicher. Comparing averages in time series data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, page 1095–1104, New York, NY, USA, 2012. Association for Computing Machinery. DOI: https://doi.org/10.1145/2207676.2208556.

[63] S. Kim, Z. Dong, H. Xian, B. Upatising, and J. S. Yi. Does an Eye Tracker Tell the Truth about Visualizations?: Findings while Investigating Visualizations for Decision Making. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2421–2430, December 2012. DOI: https://doi.org/10.1109/TVCG.2012.215.

[64] Dereck Toker, Cristina Conati, Ben Steichen, and Giuseppe Carenini. Individual User Characteristics and Information Visualization: Connecting the Dots Through Eye Tracking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 295–304, Paris, France, 2013. ACM. DOI: https://doi.org/10.1145/2470654.2470696.

[65] Giuseppe Carenini, Cristina Conati, Enamul Hoque, Ben Steichen, Dereck Toker, and James Enns. Highlighting interventions and user differences: Informing adaptive information visualization support. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, page 1835–1844, New York, NY, USA, 2014. Association for Computing Machinery. DOI: https://doi.org/10.1145/2556288.2557141.

[66] Dereck Toker, Ben Steichen, Matthew Gingerich, Cristina Conati, and Giuseppe Carenini. Towards Facilitating User Skill Acquisition: Identifying Untrained Visualization Users Through Eye Tracking. In *Proceedings of the 19th International Conference on Intelligent User Interfaces*, IUI '14, pages 105–114, Haifa, Israel, 2014. ACM. DOI: https://doi.org/10.1145/2557500.2557524.

[67] Kouhei Yasuda, Shigeo Takahashi, and Hsiang-Yun Wu. Enhancing Infographics Based on Symmetry Saliency. In *Proceedings of the 9th International Symposium on Visual Information Communication and Interaction*, VINCI '16, pages 35–42, Dallas, TX, USA, 2016. ACM. DOI: https://doi.org/10.1145/2968220.2968224.

[68] Dereck Toker, Sébastien Lallé, and Cristina Conati. Pupillometry and Head Distance to the Screen to Predict Skill Acquisition During Information Visualization Tasks. In *Proceedings of the 22Nd International Conference on Intelligent User Interfaces*, IUI '17, pages 221–231, Limassol, Cyprus, 2017. ACM. DOI: https://doi.org/10.1145/3025171.3025187.

[69] E. Dimara, G. Bailly, A. Bezerianos, and S. Franconeri. Mitigating the Attraction Effect with Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):850–860, January 2019. DOI: https://doi.org/10.1109/TVCG.2018.2865233.

[70] Lillian Cornish, Andrew Hill, Mark S. Horswill, Stefanie I. Becker, and Marcus O. Watson. Eye-tracking reveals how observation chart design features affect the detection of patient deterioration: An experimental study. *Applied Ergonomics*, 75:230–242, 2019. DOI: https://doi.org/10.1016/j.apergo.2018.10.005.

[71] João Moreira, Daniel Mendes, and Daniel Gonçalves. Incidental visualizations: Pre-attentive primitive visual tasks. In *Proceedings of the International Conference on Advanced Visual Interfaces*, AVI '20, New York, NY, USA, 2020. Association for Computing Machinery. DOI: https://doi.org/10.1145/3399715.3399841.

[72] Michael Raschke, Tanja Blascheck, Marianne Richter, Tanja Agapkin, and Thomas Ertl. Visual analysis of perceptual and cognitive processes. In Robert S. Laramee, Andreas Kerren, and José Braz, editors, *Proceedings of the 5th International Conference on Information Visualization Theory and Applications, IVAPP 2014, Lisbon, Portugal, 5-8 January, 2014*, pages 284–291. SciTePress, 2014. DOI: https://doi.org/10.5220/0004687802840291.

[73] Sebastian Elsa and Daniel Neenu. A Survey on Various Saliency Detection Methods. *International Journal of Computer Applications*, 161(5):5–8, March 2017. DOI:https://doi.org/10.5120/ijca2017913187.

[74] Laurent Itti and Christof Koch. Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3):194–203, March 2001. DOI: https://doi.org/10.1038/35058500.

[75] K. Dinkla, M. A. Westenberg, and J. J. van Wijk. Compressed Adjacency Matrices: Untangling Gene Regulatory Networks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2457–2466, December 2012. DOI: https://doi.org/10.1109/TVCG.2012.208.

[76] H. E. Egeth and S. Yantis. Visual attention: Control, representation, and time course. *Annual Review of Psychology*, 48:269–297, 1997. DOI: https://doi.org/10.1146/annurev.psych.48.1.269.

[77] Cynthia A. Brewer. Prediction of simultaneous contrast between map colors with Hunt's model of color appearance. *Color Research & Application*, 21(3):221–235, 1996. DOI: https://doi.org/10.1002/(SICI)1520-6378(199606)21:3<221::AID-COL4>3.0.CO;2-U.

[78] Fei Yan, Cheng Chen, Peng Xiao, Siyu Qi, Zhiliang Wang, and Ruoxiu Xiao. Review of visual saliency prediction: Development process from neurobiological basis to deep models. *Applied Sciences*, 12(1), 2022. DOI: https://doi.org/10.3390/app12010309.

[79] Laurent Itti. Modeling primate visual attention. In J. Feng, editor, *Computational Neuroscience: A Comprehensive Approach*, pages 635–655. CCRC Press, 1st edition edition, 2003. DOI: https://doi.org/10.1201/9780203494462.

[80] Charles E. Connor, Howard E. Egeth, and Steven Yantis. Visual attention: Bottom-up versus top-down. *Current Biology*, 14(19):R850–R852, 2004. DOI: https://doi.org/10.1016/j.cub.2004.09.041.

[81] Laurent Itti and Ali Borji. Computational Models. *The Oxford Handbook of Attention*, January 2014. DOI: https://doi.org/10.1093/oxfordhb/9780199675111.013.026.

[82] Laurent Itti and Ali Borji. Computational models: Bottom-up and top-down aspects. *CoRR*, abs/1510.07748, 2015. http://arxiv.org/abs/1510.07748.

[83] Laurent Itti and Christof Koch. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10):1489–1506, 2000. DOI: https://doi.org/10.1016/S0042-6989(99)00163-7.

[84] Christof Koch and Shimon Ullman. Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry. *Human neurobiology*, 4(4):219–227, 1985. url: https://pubmed.ncbi.nlm.nih.gov/3836989/.

[85] A. Borji. Saliency prediction in the deep learning era: Successes and limitations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 43(02):679–700, feb 2021. DOI: https://doi.org/10.1109/TPAMI.2019.2935715.

[86] Qi Zhao and Christof Koch. Learning saliency-based visual attention: A review. *Signal Processing, Special issue on Machine Learning in Intelligent Image Processing*, 93(6):1401–1407, 2013. DOI: https://doi.org/10.1016/j.sigpro.2012.06.014.

[87] Shoaib Azam, Syed Omer Gilani, Moongu Jeon, Rehan Yousaf, and Jeong-Bae Kim. A benchmark of computational models of saliency to predict human fixations in videos. In Nadia Magnenat-Thalmann, Paul Richard, Lars Linsen, Alexandru C. Telea, Sebastiano Battiato, Francisco H. Imai, and José Braz, editors, *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*

(VISIGRAPP 2016), volume 4, pages 134–142. SciTePress, 2016. DOI: https://doi.org/10.5220/0005678701340142.

[88] Ali Borji and Laurent Itti. Cat2000: A large scale fixation dataset for boosting saliency research. *CVPR 2015 workshop on "Future of Datasets"*, 2015. arXiv preprint arXiv:1505.03581.

[89] Ming Jiang, Shengsheng Huang, Juanyong Duan, and Qi Zhao. Salicon: Saliency in context. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1072–1080, 2015. DOI: https://doi.org/10.1109/CVPR.2015.7298710.

[90] Zoya Bylinskii, Michelle A. Borkin, Nam Wook Kim, Hanspeter Pfister, and Aude Oliva. Eye fixation metrics for large scale evaluation and comparison of information visualizations. In Michael Burch, Lewis Chuang, Brian Fisher, Albrecht Schmidt, and Daniel Weiskopf, editors, *Eye Tracking and Visualization*, pages 235–255, Cham, 2017. Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-47024-5_1.

[91] Laura E., Michael J. Haass, Kristin M. Divis, and Mallory C. Stites. Patterns of attention: How data visualizations are read. In Dylan D. Schmorrow and Cali M. Fidopiastis, editors, *Augmented Cognition. Neurocognition and Machine Learning*, pages 176–191, Cham, 2017. Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-58628-1_15.

[92] Mark A. Livingston, Laura Matzen, Derek Brock, Andre Harrison, and Jonathan W. Decker. Testing the value of salience in statistical graphs. In *IS&T Int'l. Symp. on Electronic Imaging: Visualization and Data Analysis*, pages 329–1 – 329–11, 2021. DOI: https://doi.org/10.2352/ISSN.2470-1173.2021.1.VDA-329.

[93] J. Zhang and S. Sclaroff. Saliency detection: A boolean map approach. In *2013 IEEE International Conference on Computer Vision*, pages 153–160, Dec 2013. DOI: https://doi.org/10.1109/ICCV.2013.26.

[94] E. Vig, M. Dorr, and D. Cox. Large-scale optimization of hierarchical features for saliency prediction in natural images. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2798–2805, 2014. DOI: https://doi.org/10.1109/CVPR.2014.358.

[95] Stephanie Elzer, Nancy L. Green, Sandra Carberry, and James Hoffman. A model of perceptual task effort for bar charts and its role in recognizing intention. *User Model User-Adap Inter.*, 16(1):1–30, 2006. DOI: https://doi.org/10.1007/s11257-006-9002-9.

[96] Jonathan Harel, Christof Koch, and Pietro Perona. Graph-based visual saliency. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems*, volume 19. MIT Press, 2006. https://proceedings.neurips.cc/paper/2006/file/4db0f8b0fc895da263fd77fc8aecabe4-Paper.pdf.

[97] Jeremy Wolfe and Todd Horowitz. What attributes guide the deployment of visual attention and how do they do it? *Nature reviews. Neuroscience*, 5:495–501, 07 2004. DOI: https://doi.org/10.1038/nrn1411.

[98] Laurent Itti and Christof Koch. Comparison of feature combination strategies for saliency-based visual attention systems. In Bernice E. Rogowitz and Thrasyvoulos N. Pappas, editors, *Human Vision and Electronic Imaging IV*, volume 3644 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 473–482, May 1999. DOI: https://doi.org/10.1117/12.348467.

[99] Zbigniew Zdziarski. Visual attention using 2d & 3d displays. In *University of Dublin, Trinity College*, 2015. https://bit.ly/Visualattentionusing2D3Ddisplays.

[100] Nam Wook Kim, Zoya Bylinskii, Michelle A. Borkin, Krzysztof Z. Gajos, Aude Oliva, Fredo Durand, and Hanspeter Pfister. Bubbleview: an interface for crowdsourcing image importance maps and tracking visual attention. *ACM Transactions on Computer-Human Interaction*, 24(5):1–40, Oct 2017. DOI: http://dx.doi.org/10.1145/3131275.

[101] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3431–3440, 2015. DOI: https://doi.org/10.1109/CVPR.2015.7298965.

[102] Hannah I. Weller and Mark W. Westneat. Quantitative color profiling of digital images with earth mover's distance using the r package colordistance. *PeerJ*, 7:e6398, February 2019. DOI: https://doi.org/10.7717/peerj.6398.

[103] Anna Vital, Mark Vital, and Alexander Vushkan. Visualization universe, 2021. http://visualizationuniverse.com.

[104] John Elflein. Coronavirus (covid-19) disease pandemic, statistics & facts, 2020. https://bit.ly/statistaCoronavirus.

[105] Matthias Kümmerer, Zoya Bylinskii, Tilke Judd, Ali Borji, Laurent Itti, Frédo Durand, Aude Oliva, and Antonio Torralba. Mit/tübingen saliency benchmark, 2012. https://saliency.tuebingen.ai/.

[106] Zoya Bylinskii, Tilke Judd, Aude Oliva, Antonio Torralba, and Frédo Durand. What do different evaluation metrics tell us about saliency models? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(3):740–757, 2019. DOI: https://doi.org/10.1109/TPAMI.2018.2815601.

[107] Sebastiaan Mathôt, Daniel Schreij, and Jan Theeuwes. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2):314–324, June 2012. DOI: https://doi.org/10.3758/s13428-011-0168-7.

[108] Stephanie Glen. T test (student's t-test): Definition and examples, Apr 2022.

[109] Michael Waskom. Choosing color palettes, 2019. Seaborn: statistical data visualization.

[110] David Feng, Lester Kwock, Yueh Lee, and Russell Taylor. Matching visual saliency to confidence in plots of uncertain data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):980–989, 2010. DOI: https://doi.org/10.1109/TVCG.2010.176.

[111] Joshua Shive, Sharra Rosichan, Sherika Davis, Christena Wade, James Ellison, and Santos Santoni-Sanchez. The statistical saliency model can choose colors for items on maps. *IEEE Transactions on Human-Machine Systems*, 49(6):569–578, 2019.

[112] Luisa Barrera-Leon, Fulvio Corno, and Luigi De Russis. Systematic variation of preattentive attributes to highlight relevant data in information visualization. In *2020 24th International Conference Information Visualisation (IV)*, pages 74–79, 2020. DOI: http://dx.doi.org/10.1109/IV51561.2020.00022.

[113] Darius A. Monsef IV. Colourlovers, 2020. https://www.colourlovers.com/palettes/.

[114] Adrien Treuille, Thiago Teixeira, and Amanda Kelly. Streamlite, data web application library, 2022. https://streamlit.io.

[115] John Elflein. Coronavirus (covid-19) disease pandemic, statistics & facts, 2020. https://bit.ly/statistaCoronavirus.

[116] Emily Wall, Meeshu Agnihotri, Laura Matzen, Kristin Divis, Michael Haass, Alex Endert, and John Stasko. A heuristic approach to value-driven evaluation of visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):491–500, 2019. DOI: https://doi.org/10.1109/TVCG.2018.2865146.

[117] CHAOMEI CHEN and YUE YU. Empirical studies of information visualization: a meta-analysis. *International Journal of Human-Computer Studies*, 53(5):851–866, 2000. DOI: https://doi.org/10.1006/ijhc.2000.0422.

[118] P. Saraiya, C. North, and K. Duca. An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):443–456, 2005. DOI: https://doi.org/10.1109/TVCG.2005.53.

[119] Torre Zuk, Lothar Schlesier, Petra Neumann, Mark S. Hancock, and Sheelagh Carpendale. Heuristics for information visualization evaluation. In *Proceedings of the 2006 AVI Workshop on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization*, BELIV '06, page 1–6, New York, NY, USA, 2006. Association for Computing Machinery. DOI: https://doi.org/10.1145/1168149.1168162.

[120] Marti A. Hearst, Paul Laskowski, and Luis Silva. Evaluating information visualization via the interplay of heuristic evaluation and question-based scoring. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 5028–5033, New York, NY, USA, 2016. Association for Computing Machinery. DOI: https://doi.org/10.1145/2858036.2858280.

[121] Camilla Forsell and Jimmy Johansson. An heuristic set for evaluation in information visualization. In *Proceedings of the International Conference on Advanced Visual Interfaces*, AVI '10, page 199–206, New York, NY, USA, 2010. Association for Computing Machinery. DOI: https://doi.org/10.1145/1842993.1843029.

[122] Ya-Hsin Hung and Paul Parsons. Assessing user engagement in information visualization. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, page 1708–1717, New York, NY, USA, 2017. Association for Computing Machinery. DOI: https://doi.org/10.1145/3027063.3053113.

[123] Colin Ware and John C. Beatty. Using color dimensions to display data dimensions. *Human Factors*, 30(2):127–142, 1988. DOI: https://doi.org/10.1177/001872088803000201.

[124] Simone Kriglstein, Margit Pohl, Nikolaus Suchy, Johannes Gärtner, Theresia Gschwandtner, and Silvia Miksch. Experiences and challenges with evaluation methods in practice: A case study. In *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*, BELIV '14, page 118–125, New York, NY, USA, 2014. Association for Computing Machinery. DOI: https://doi.org/10.1145/2669557.2669571.

[125] Tanja Merčun. Evaluation of information visualization techniques: Analysing user experience with reaction cards. In *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*, BELIV '14, page 103–109, New York, NY, USA, 2014. Association for Computing Machinery. DOI: https://doi.org/10.1145/2669557.2669565.

[126] Michael Sedlmair, Petra Isenberg, Dominikus Baur, and Andreas Butz. Evaluating information visualization in large companies: Challenges, experiences and recommendations. In *Proceedings of the 3rd BELIV'10 Workshop: BEyond Time and Errors: Novel EvaLuation Methods for Information Visualization*, BELIV '10, page 79–86, New York, NY, USA, 2010. Association for Computing Machinery. DOI: https://doi.org/10.1145/2110192.2110204.

[127] Riccardo Mazza and Alessandra Berre. Focus group methodology for evaluating information visualization techniques and tools. In *2007 11th International Conference Information Visualization (IV '07)*, pages 74–80, 2007. DOI: https://doi.org/10.1109/IV.2007.51.

[128] Beatriz Sousa Santos, Beatriz Quintino Ferreira, and Paulo Dias. Using heuristic evaluation to foster visualization analysis and design skills.

IEEE Computer Graphics and Applications, 36(1):86–90, 2016. DOI: https://10.1109/MCG.2016.7.

[129] Victor M. González, Héctor Hugo Huipet, Cassandra Pratt, and Guillermo Garduño. A case study of usability evaluation of an infovis tool for analyzing twitter. In *Proceedings of the XVIII International Conference on Human Computer Interaction*, Interacción '17, New York, NY, USA, 2017. Association for Computing Machinery.

[130] Michele Mauri, Tommaso Elli, Giorgio Caviglia, Giorgio Uboldi, and Matteo Azzi. Rawgraphs: A visualisation platform to create open outputs. In *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter*, CHItaly '17, New York, NY, USA, 2017. Association for Computing Machinery.

[131] Jakob Nielsen and Rolf Molich. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '90, page 249–256, New York, NY, USA, 1990. Association for Computing Machinery.

[132] Perlman Gary. User interface usability evaluation with web-based questionnaires, 2021. https://garyperlman.com/quest/.

[133] Arnold M Lund. Measuring usability with the use questionnaire. *Usability interface*, 8(2):3–6, 2001. https://bit.ly/Measuring_Usability_USE.

[134] Meiyuzi Gao, Philip Kortum, and Frederick Oswald. Psychometric evaluation of the use (usefulness, satisfaction, and ease of use) questionnaire for reliability and validity. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62(1):1414–1418, 2018. DOI: https://doi.org/10.1177/1541931218621322.

[135] John P. Chin, Virginia A. Diehl, and Kent L. Norman. Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '88, page 213–218, New York, NY, USA, 1988. Association for Computing Machinery. DOI: https://doi.org/10.1145/57167.57203.

[136] Ben Shneiderman. *Designing the User Interface (Videotape)*. University of Maryland at College Park, USA, 1987. DOI: https://doi.org/10.5555/61961.

[137] Andrew Hodrien and Terrence Fernando. A review of post-study and post-task subjective questionnaires to guide assessment of system usability. *J. Usability Studies*, 16(3):203–232, may 2021. DOI: https://doi.org/10.5555/3532758.3532762.

[138] Urška Lah, James R. Lewis, and Boštjan Šumak. Perceived usability and the modified technology acceptance model. *International Journal of Human–Computer Interaction*, 36(13):1216–1230, 2020. DOI: https://doi.org/10.1080/10447318.2020.1727262.

Appendix A

Saliency Models Validation Results

All the results of the three experiments performed in chapter 5 can be downloaded from the following link: https://bit.ly/eyeTracking_ImagesResults

The folder is organized as indicated below:

- **First Experiment**. This folder contains a sub-folder with selected MASSVIS images. A second sub-folder has the saliency maps of each image for each validated saliency model. This folder also has a .mat file (Matlab) that contains the metadata and the fixation points for each image.
- **Second Experiment**. This folder includes a sub-folder with the created Clean Graph set of images. A second sub-folder contains the saliency maps of each image for each validated saliency model. The last sub-folder has the ground truth: observer's data, graph description, fixations, and the heatmap.
- **Third Experiment**. This folder includes a sub-folder with the second group of Clean Graphs. A second sub-folder contains the image fixations and the heatmaps. The other sub-folder has the resulted saliency maps for each model. Also, this folder has the eye-tracking collected data. The last sub-folder has the images with the number of fixations.
- **Metrics Results**. For each experiment, there is an excel file where the values obtained from the metrics (three excel files in total)

- **Results Power Point**. This file shows all the images resulting from the second and third experiments. This file provides an easy way to view the results obtained.

Appendix B

Experts Evaluation Questionnaires

Questionnaire for User Interface Satisfaction (QUIS)

The scale used for this validation instrument was from zero (0) to nine (9).

Table B.1 QUIS selected questions

USEFULNESS	
Question Number	**Scale**
1	Terrible to Wonderful
2	Difficult to Easy
3	Frustrating to Satisfying
4	Inadequate Power to Adequate Power
5	Dull to Stimulating
6	Rigid to Flexible

Usefulness, Satisfaction, and Ease of use (USE)

The used scale for this validation instrument is: DISAGREE (0) to AGREE (6)

Table B.2 USE selected questions

Question Number	Question
USEFULNESS	
1	It could help me be more effective.
2	It could help me be more productive.
3	It is useful.
4	It gives me more control over the activities in my daily DataViz tasks.
5	It makes the things I want to accomplish easier to get done.
6	It could save me time when I use it.
EASY TO USE	
7	It is simple to use
8	I can use it without written instructions.
9	I don't notice any inconsistencies as I use it.
10	Both occasional and regular users would like it.
SATISFACTION	
11	I am satisfied with it.
12	I would recommend it to a friend.
13	I feel I need to have it.

Technology Acceptance Model (TAM)

The used scale for this validation instrument was one (1) to six (6).

Table B.3 TAM list of questions

Question Number	**Question**
	PERCEIVED USEFULNESS
1	Using Saliency Prediction in my job would enable me to accomplish tasks more quickly.
2	Using Saliency Prediction would improve my job performance.
3	Using Saliency Prediction in my job would increase my productivity.
4	Using Saliency Prediction would enhance my effectiveness on the job.
5	Using Saliency Prediction would make it easier to do my job.
6	I would find Saliency Prediction useful in my job.
	PERCEIVED EASE OF USE
7	Learning to operate Saliency Prediction would be easy for me.
8	I would find it easy to get Saliency Prediction to do what I want it to do.
9	My interaction with Saliency Prediction would be clear and understandable.
10	I would find Saliency Prediction to be flexible to interact with.
11	It would be easy for me to become skillful at using Saliency Prediction.
12	I would find Saliency Prediction easy to use.

Appendix C

Eye-Tracking Process Screenshots

The following images show the process developed in the OpenSesame application for data collection with the eye-tracker.

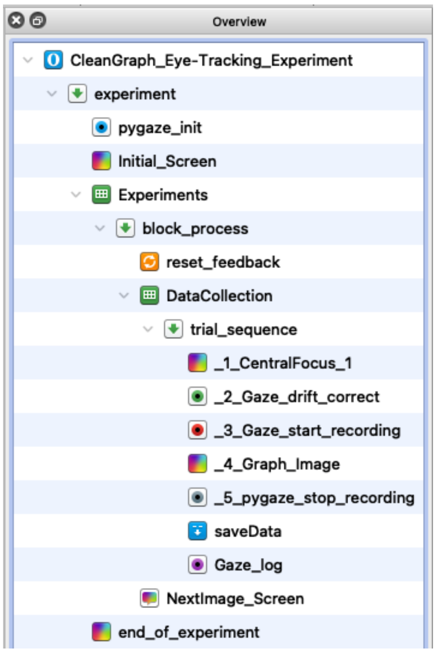


Fig. C.1 General Eye-Tracking developed flow.

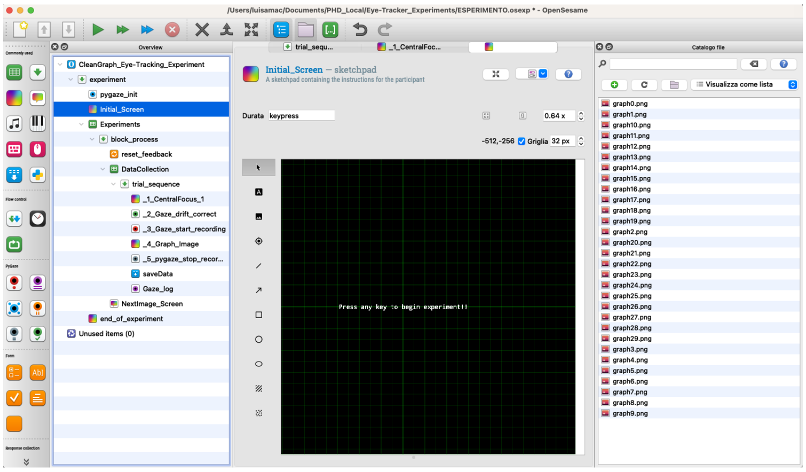


Fig. C.2 Initial black screen

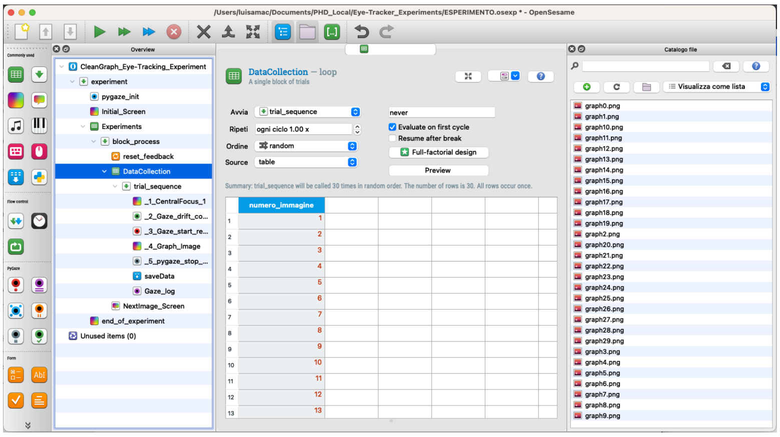


Fig. C.3 Image List. In this part of the process, the program chooses the images randomly and takes the numbers from this list.

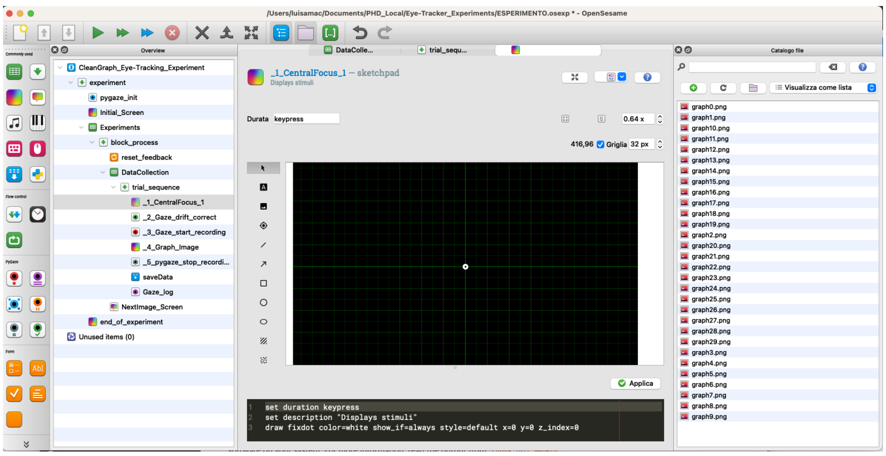


Fig. C.4 Black point screen to focus the observer's attention previous to the graph image.

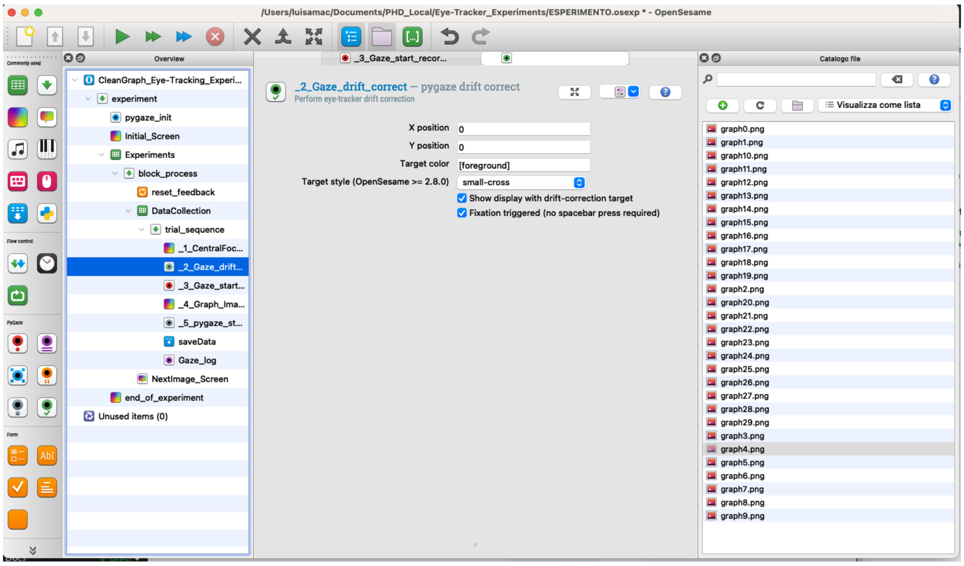


Fig. C.5 The 0,0 point for the Eye-tracker is also located, thus synchronizing both the Eye-tracker and the first fixation point of the observer.

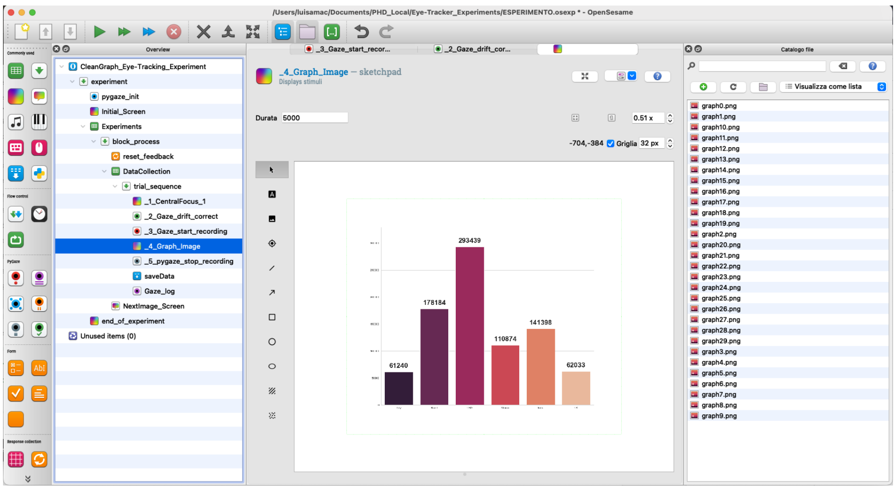


Fig. C.6 Show clean graph (randomly)

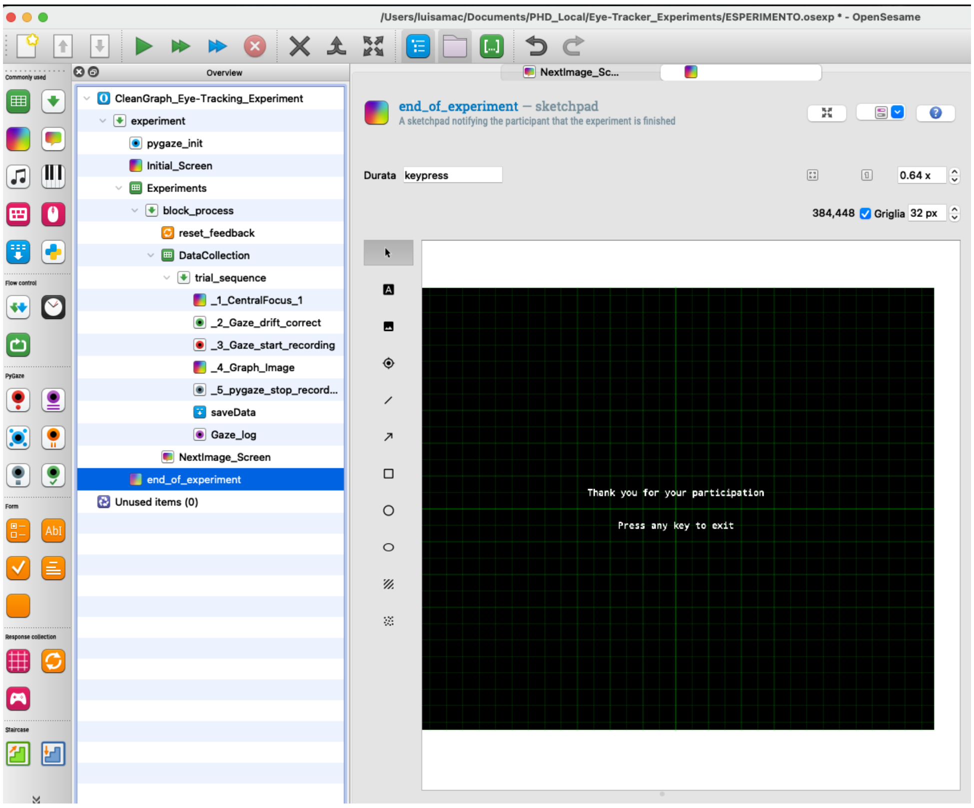
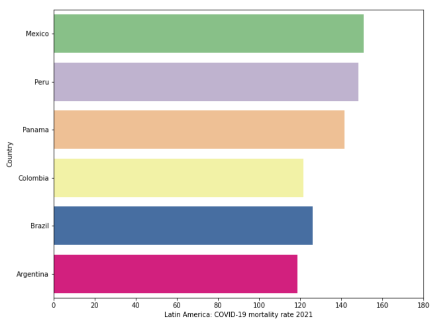


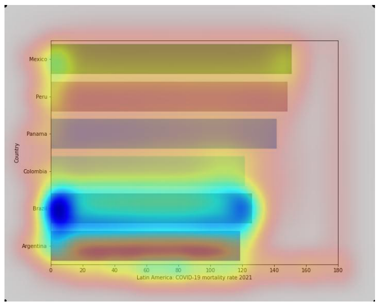
Fig. C.7 Final Screen

Appendix D

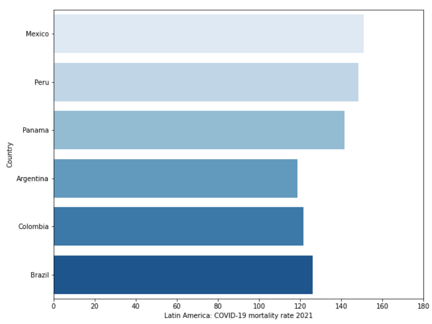
Design Tool - Python Examples



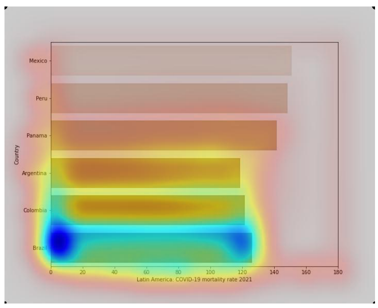
(a) Accent color palette, horizontal and Brazil



(b) Accent Saliency Map

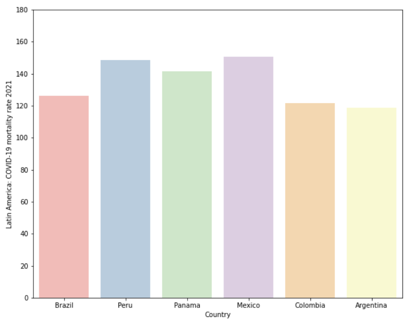


(c) Blue color palette, horizontal and Brazil

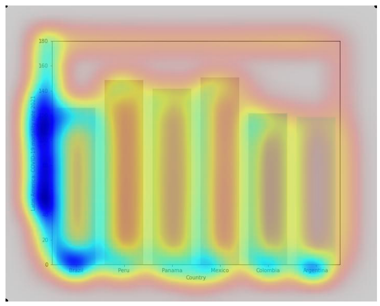


(d) Blue Saliency Map

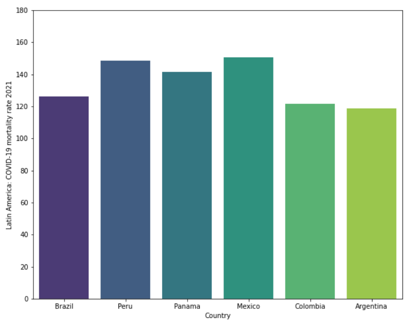
Fig. D.1 Example Design Tool (python). Data to be highlighted is Brazil, winners per palette



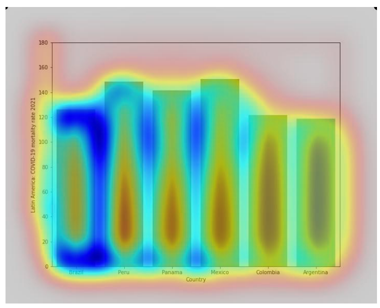
(a) Pastel1 color palette, vertical and Brazil



(b) Pastel1 saliency map

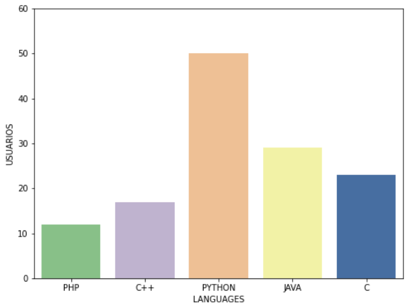


(c) Viridis color palette, vertical and Brazil

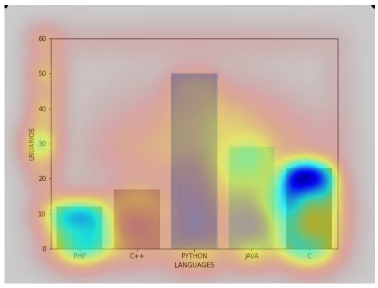


(d) Viridis saliency map

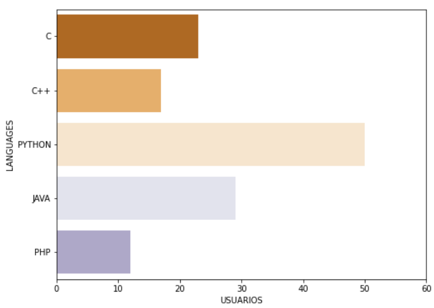
Fig. D.2 Example Design Tool (python). Data to be highlighted is Brazil, winners per palette



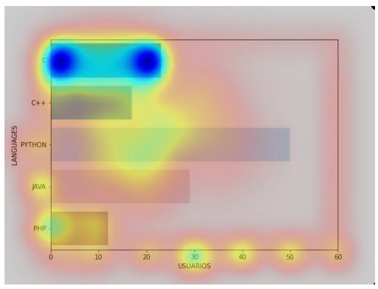
(a) Accent color palette, horizontal and C



(b) Accent Saliency Map

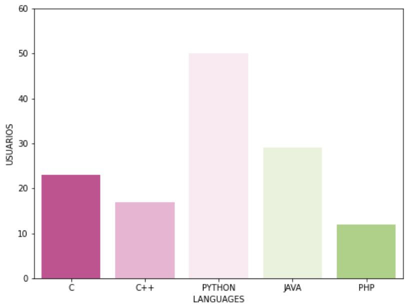


(c) PuOr color palette, horizontal and C

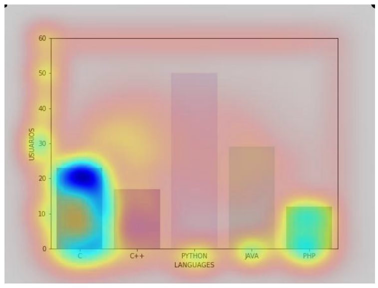


(d) PuOr Saliency Map

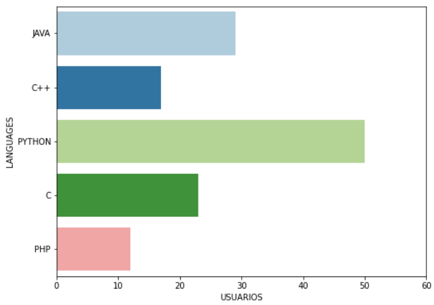
Fig. D.3 Example Design Tool (python). Data to be highlighted is C, winners per palette



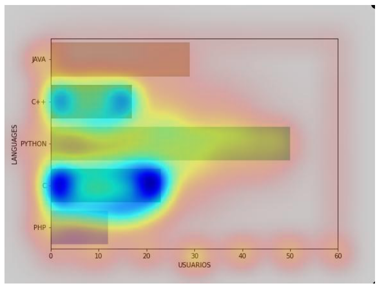
(a) PiYG color palette, vertical and C



(b) PiYG saliency map



(c) Paired is the color palette, vertical and C



(d) Paired saliency map

Fig. D.4 Example Design Tool (python). Data to be highlighted is C, winners per palette

Appendix E

Publications

1. Luisa Barrera-Leon, Fulvio Corno, and Luigi De Russis. Systematic variation of preattentive attributes to highlight relevant data in information visualization. In 2020 24th International Conference Information Visualisation (IV), pages 74–79, 2020. DOI: http://dx.doi.org/10.1109/IV51561.2020.00022.

2. Luisa Barrera-Leon, Fulvio Corno, and Luigi De Russis. How the preattentive process is exploited in practical information visualization design: a review. International Journal of Human–Computer Interaction, 0(0):0, 2022. DOI: http://dx.doi.org/10.1080/10447318.2022.2049137

3. Luisa Barrera-Leon, Fulvio Corno, and Luigi De Russis. Saliency Models in Statistical Graphs: Validation, Analysis, and Insights. IEEE Transactions on Visualization and Computer Graphics (TVCG)(submitted)