UNIVERSIDADE FEDERAL DO CEARÁ

CENTRO DE CIÊNCIAS

DEPARTAMENTO DE FÍSICA

PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

DOUTORADO EM FÍSICA

DÉBORA TORRES

EYE-TRACKING AS A PROXY FOR COHERENCE AND COMPLEXITY OF TEXTS, AND FOR THE INTERPRETATION OF CYTOPATHOLOGICAL IMAGES

FORTALEZA

2024

DÉBORA TORRES

EYE-TRACKING AS A PROXY FOR COHERENCE AND COMPLEXITY OF TEXTS, AND FOR THE INTERPRETATION OF CYTOPATHOLOGICAL IMAGES

PhD thesis presented to the Post-Graduation Course in Physics of the Federal University of Ceará as part of the requisites for obtaining the Degree of Doctor in Physics.

Advisor: Prof. Dr. José Soares de Andrade Jr..

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

T644e Torres, Débora.
Eye-tracking as a proxy for coherence and complexity of texts, and for the interpretation of cytopathological images / Débora Torres. – 2024.
150 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Física , Fortaleza, 2024.
Orientação: Prof. Dr. José Soares de Andrade Junior.

1. eye-tracking. 2. text complexity and coherence. 3. Maximum Entropy Model. 4. visual attention. 5. saliency models. I. Título.

CDD 530

DÉBORA TORRES

EYE-TRACKING AS A PROXY FOR COHERENCE AND COMPLEXITY OF TEXTS, AND FOR THE INTERPRETATION OF CYTOPATHOLOGICAL IMAGES

PhD thesis presented to the Post-Graduation Course in Physics of the Federal University of Ceará as part of the requisites for obtaining the Degree of Doctor in Physics.

Aprovada em: 07/03/2024

BANCA EXAMINADORA

Prof. Dr. José Soares de Andrade Jr. (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. André Auto Moreira
Universidade Federal do Ceará (UFC)

Prof. Dr. Humberto de Andrade Carmona
Universidade Federal do Ceará (UFC)

Prof. Dr. Luciano Rodrigues da Silva
Univeridade Federal do Rio Grande do Norte (UFRN)

Prof. Dr. Rilder de Sousa Pires
Universidade de Fortaleza (UNIFOR)

A mis padres, Pedro y Alicia.

ACKNOWLEDGEMENTS

I extend my heartfelt gratitude to Professor Soares, my advisor, whose guidance and support have been invaluable since my arrival. Throughout my academic journey, his expertise and encouragement have been instrumental in shaping this work. I'm truly grateful for your patience and belief in me, especially during challenging times.

A sincere thank you to my colleagues and professors at the Complex Systems research group for their camaraderie and support each day. Each one of you has enriched my work and experience here, and I'm deeply grateful for the contributions of everyone I've had the privilege to work with. Special gratitude goes to Laísa for your trust and companionship over the past year.

I also want to express my appreciation to all the faculty members and staff at the Physics Department and Post-Graduation Program for their care and support. To everyone at Universidade Federal do Ceará, thank you for making this institution feel like home.

I am indebted to the CNPq and Petrobras agencies for their financial support, which has allowed me to pursue my academic goals and make this research possible.

Lastly, to my friends and family, thank you for your unconditional love and support. Your presence has been my anchor through every high and low.

We meet here in a new light the old truth that in our description of nature the purpose is not to disclose the real essence of the phenomena but only to track down, so far as it is possible, relations between the manifold aspects of our experience. (Niels Bohr, from the Introductory Survey of "Atomic Theory and the Description of Nature" (1934), written in 1929.)

ABSTRACT

Eye movements offer a unique insight into cognitive processes, bridging the mechanics of vision with underlying neural activities. Through swift movements called "saccades", our eyes direct attention to areas of interest, while "fixations" provide moments for detailed visual information gathering. In tasks like reading, fixation patterns reveal cognitive demands and processing depth. Eye tracking, a non-invasive method, captures these movements with precision, enabling the study of visual behavior across tasks and stimuli complexities. By analyzing eye movement data using statistical physics and machine learning tools, we gain insights into collective behaviors and cognitive processing patterns. Equipped with a cutting-edge eye tracker, our research focused on eye-tracking experiments across various visual tasks, resulting in two main projects published in reputable journals. Our primary project, "Eye-tracking as a Proxy for Text Complexity and Coherence", involved experiments where participants read diverse texts. Using a binary model for fixation sequences akin to an Ising system integrated into the Maximum Entropy Model, we examined fixation patterns. Leveraging results from an extensive survey, we compared these findings with properties derived from eye-tracking experiments. Our analysis revealed a positive correlation between text complexity and average fixation density, measured by magnetizations. Additionally, statistical modeling and Pairwise Maximum-Entropy analysis demonstrated the effectiveness of critical point distance in assessing text coherence, indicating a potential for robust cognitive measures compared to conventional survey responses. In parallel, a collaborative project with the "Engenheria em Teleinformatica" post-graduate program focused on visual attention among cytopathologists analyzing Pap smear slides. This research offered insights into visual mechanisms during Pap smear analysis, potentially enhancing current practices and informing the development of innovative analysis techniques. Our investigation encompassed a review of visual attention dynamics, exploration of saliency learning models for medical diagnosis, and evaluation of eye-tracking applications in medical imaging. Through the comparison of eye fixation maps with saliency prediction algorithms, the reliability of these algorithms in detecting clinically relevant cells was assessed. The findings from our primary project demonstrate the robustness of the maximum-entropy approach for probing diverse complex systems effectively. Additionally, our observations highlight how humans respond cohesively to coherent textual information, reflecting advanced language formation and reading prediction mechanisms. Moreover, our secondary project revealed a significant agreement in gaze patterns among participants, supporting the identification of clinical

Regions of Interest in Pap smear images. Further investigation into saliency models provided valuable insights into their predictive accuracy and potential implications for medical image interpretation. Lastly, our interdisciplinary research, spanning neuroscience, linguistics, and computer science, provided valuable insights into applying physical models to study cognitive activity.

Keywords: eye-tracking; text complexity and coherence; Maximum Entropy Model (MEM); visual attention, saliency models.

RESUMO

Os movimentos dos olhos proporcionam uma visão única dos processos cognitivos, conectando a mecânica da visão com atividades neurais subjacentes. Através de movimentos rápidos chamados de “sacadas”, nossos olhos direcionam a atenção para áreas de interesse, enquanto as “fixações” oferecem momentos para a coleta de informações visuais detalhadas. Em tarefas como a leitura, os padrões de fixação revelam demandas cognitivas e a profundidade do processamento. O rastreamento ocular, um método não invasivo, captura esses movimentos com precisão, permitindo o estudo do comportamento visual em diferentes tarefas e estímulos visuais. Ao analisar dados de movimento ocular usando física estatística e ferramentas de aprendizado de máquina, obtemos insights sobre comportamentos coletivos e padrões de processamento cognitivo. Equipada com um rastreador ocular de ponta, nossa pesquisa focou em experimentos de rastreamento ocular em várias tarefas visuais, resultando em dois projetos principais publicados em revistas renomadas. Nosso projeto principal envolveu experimentos nos quais os participantes liam textos diversos. Usando um modelo binário para sequências de fixação semelhante a um sistema Ising integrado ao Modelo de Entropia Máxima, examinamos padrões de fixação. Realizando uma extensa pesquisa, comparamos esses resultados com propriedades derivadas dos experimentos de rastreamento ocular. Nossa análise revelou uma correlação positiva entre a complexidade do texto e a densidade média de fixação, medida pelas magnetizações. Além disso, a modelagem estatística e a análise de Máxima Entropia Par a Par demonstraram a eficácia da distância do ponto crítico na avaliação da coerência do texto, indicando um potencial para medidas cognitivas robustas em comparação com respostas de pesquisa convencionais. Paralelamente, um projeto colaborativo com o programa de pós-graduação em Engenharia de Teleinformática focou na atenção visual de citopatologistas analisando lâminas de esfregaço de Pap. Esta pesquisa ofereceu insights sobre mecanismos visuais durante a análise de esfregaços de Pap, potencialmente aprimorando práticas atuais e informando o desenvolvimento de técnicas de análise inovadoras. Nossa investigação abrangeu uma revisão da dinâmica da atenção visual, exploração de modelos de aprendizagem de saliência para diagnóstico médico e avaliação de aplicações de rastreamento ocular em interpretação de imagens médicas. Através da comparação de mapas de fixação ocular com algoritmos de previsão de saliência, a confiabilidade desses algoritmos na detecção de células clinicamente relevantes foi avaliada. Os resultados de nosso projeto principal demonstram a robustez da abordagem de máxima-entropia para sondar sistemas complexos de forma eficaz. Além disso, nossas observações destacam como os humanos respondem de forma coesa a informações textuais

coerentes, refletindo mecanismos avançados de formação de linguagem e previsão de leitura. Além disso, nosso projeto secundário revelou um acordo significativo nos padrões de olhar entre os participantes, apoiando a identificação de Regiões de Interesse Clínico em imagens de esfregaços de Pap. Investigações adicionais sobre modelos de saliência forneceram insights valiosos sobre sua precisão preditiva e implicações potenciais para a interpretação de imagens médicas. Por fim, nossa pesquisa interdisciplinar, abrangendo neurociência, linguística e ciência da computação, proporcionou insights valiosos sobre a aplicação de modelos físicos para estudar a atividade cognitiva.

Palavras-chave: rastreamento ocular; complexidade e coerência de textos; Modelo de Entropia Máxima (MEM); atenção visual, modelos de saliência.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

Eye movements provide a unique window into cognitive activity, since they represent an observable link between the mechanics of vision and the intricate neural processes that underlie it. As we interact with visual stimuli, whether analyzing an image, reading text, or engaging in any other visual task, our gaze gravitates towards particular areas of interest enabling our brain to process the captured information. In this process, our eyes execute swift movements, referred to as "saccades", which direct attention and influence perceptual mechanisms. In between these saccadic movements are brief pauses, referred to as "fixations", during which our eyes momentarily gather detailed visual information. The duration and location of these fixations provide insights into the cognitive demands of the ongoing task, as well as the need for deeper information processing. This is particularly noticeable when reading, where our gaze scans the text through saccades, successively fixating on words. The dynamics of these fixation patterns are closely linked to word recognition and the processing of textual information (Clifton *et al.*, 2016). Likewise, eye movements play a significant role in visual search tasks, unveiling the various strategies individuals employ while exploring a visual scene to identify a target object. Furthermore, aspects of eye movement behavior can be suggestive of the decision-making process. Notably, revisitations to specific visual cues can be indicative of uncertainty or the necessity for additional cognitive processing.

Central to the study of these processes is the tool of eye tracking. Unlike methods for assessing neural activity, which often require invasive procedures and extensive resources, eye tracking provides a non-invasive and comparatively more accessible means of observing cognitive activity. Modern eye trackers can measure eye position within milliseconds with high accuracy, enabling precise identification of fixations and saccades, and thus providing a clear and detailed measurement of eye movement. This data allows us to examine how visual stimuli of varying complexity levels influence eye movement behavior. It also enables us to explore whether cohesive patterns exist in the eye movements of different individuals when exposed to similar stimuli.

For instance, we ask ourselves whether individuals reading a particular text display comparable fixation patterns, signaling similar cognitive processing. This leads us to further inquire about how we can quantify and characterize such measures of "similarity" in fixation configurations. Similarly, when shifting our focus to image analysis, we question whether consistent scanning patterns exist among different individuals. To answer these questions, we

turn to tools from complex systems research, which provide robust methods for modeling this type of data. Complex systems, characterized by numerous interacting components, often exhibit emergent behavior not solely determined by individual component properties. For example, in a fluid system, while individual particle behavior follow classical mechanics, the collective fluid behavior can give rise to phenomena like turbulence or phase transitions, which are not apparent at the level of individual particles. We model these systems using approaches from statistical physics, treating them as ensembles of interacting components. By describing these interactions in terms of probabilities, we can derive macroscopic properties of the system. These interactions can also reveal potential collective behaviors. Meanwhile, machine learning tools for visual pattern analysis enable us to study complex systems from a physics perspective by recognizing, modeling, and interpreting intricate visual dynamics. These algorithms, trained on large datasets of labeled visual data, such as images or videos, learn to identify specific features of interest.

In our research group, equipped with a state-of-the-art eye tracker, one of our goals is studying visual behavior across diverse cognitive tasks. This involves analyzing descriptive data such as gaze location, timings, reactions, and task-related responses. By applying physical modeling techniques to eye-tracking data, we gain valuable analytical tools for interpretation. During the course of this doctoral research, our eye-tracking experiments on different visual tasks have led to two main projects, both of which have been published in reputable peer-reviewed journals. With the increasing interest in eye-tracking research in reading and the complexities of reading-data modeling becoming apparent, our attention focused on developing methodologies for analyzing this data, forming the basis of our primary project. Here, we study how individuals process different types of texts, hypothesizing that varying levels of cognitive engagement while reading different texts would be reflected in distinct eye-movement patterns. We found the Maximum Entropy Model to be a fitting framework for modeling eye-tracking data. This choice was informed by its demonstrated effectiveness in studying neural activity and other interconnected systems. In parallel, a collaborative project with a research group from the “Engenheria em Teleinformatica” post-graduate program at UFC University presented an opportunity to contribute to a study on visual attention among cytopathologists evaluating Pap smear slides —a crucial task considering the significance of Pap smear testing in women’s healthcare. Investigating visual mechanisms during Pap smear analysis not only holds implications for enhancing current practices but also for developing innovative analysis techniques.

This thesis is dedicated to the exploration and synthesis of findings from these two projects, organized as follows. The introductory sections provide a comprehensive foundation, delving into fundamental concepts of human vision and a historical overview of eye-tracking research. The eye-tracker employed in this study is meticulously introduced, presenting its functioning and setup. In Chapter 2 we look into our primary project, "Eye-tracking as a Proxy for Text Complexity and Coherence", wherein we conducted experiments with participants reading diverse texts, employing a binary model for fixation sequences akin to an Ising system integrated into the Maximum Entropy Model. This approach unveiled potential indexes as proxies for text complexity and coherence, later validated through an extensive internet reading survey. In Chapter 3, we present our secondary project, "Visual Attention Dynamics in Medical Image Interpretation: A Study of Cytopathologists Using Eye Tracking". This chapter includes a comprehensive review of visual attention, exploration of saliency learning models for medical diagnosis, and an investigation into the applications of eye-tracking in medical imaging with learning models. By recording eye fixation maps of cytopathologists and comparing them with various saliency prediction algorithms, we assessed the reliability of algorithms in identifying clinically relevant cells. The thesis concludes in the final chapter, offering insights, summarizing findings, and outlining future perspectives.

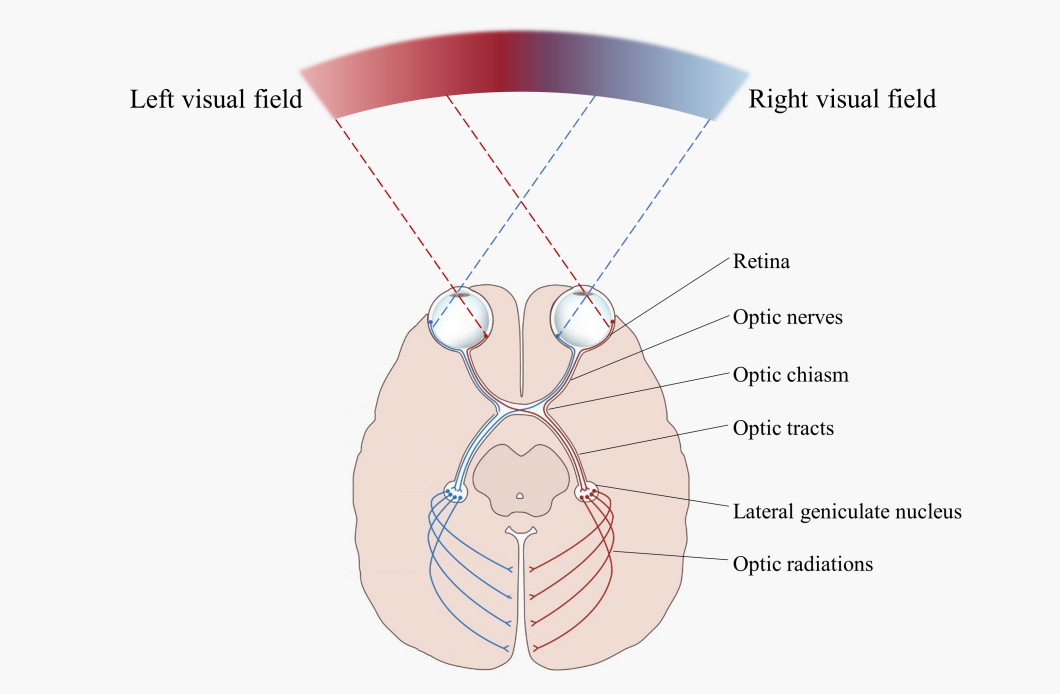
1.1 Foundations of human vision and visual processing

Our understanding of the world and our memories relies largely on sight, yet discerning the mechanisms that underlie the visual process is highly intricate. Essential problems, such as perceiving form, movement, identifying objects, and decoding signs, pose computational challenges that even today, artificial vision systems struggle to overcome. Vision extends beyond object recognition, encompassing the interpretation of diverse information in our surroundings and guiding our actions in complex visual environments. This constructive and dynamic process of visual perception unfolds across three hierarchical levels—low, intermediate, and high (Kandel *et al.*, 2013). It spans from the initial detection of light to the formation of complex neural associations for object recognition. Despite these diverse features converging into a unified percept, their integration is facilitated by multiple areas in the brain, nourished by at least two major interacting neural pathways.

1.1.1 Neural pathways of visual perception: from retina to cortical processing

The process of visual perception involves the geniculostriate pathway, which initiates with visual information captured by the retinas (see Fig. 1). Axons from retinal ganglion cells form the optic nerve, leading to the optic chiasm where fibers cross over. This ensures that information from each hemifield is processed in the contralateral hemisphere of the brain's visual processing center. The optic tract, composed of axons from nasal and temporal hemiretinas, proceeds to the lateral geniculate nucleus in the thalamus, organized into six layers corresponding to eye input. The geniculostriate pathway continues to the primary visual processing center, characterized by the preservation of the spatial arrangement of retinal inputs.

Figure 1 – Overview of the geniculostriate visual pathway

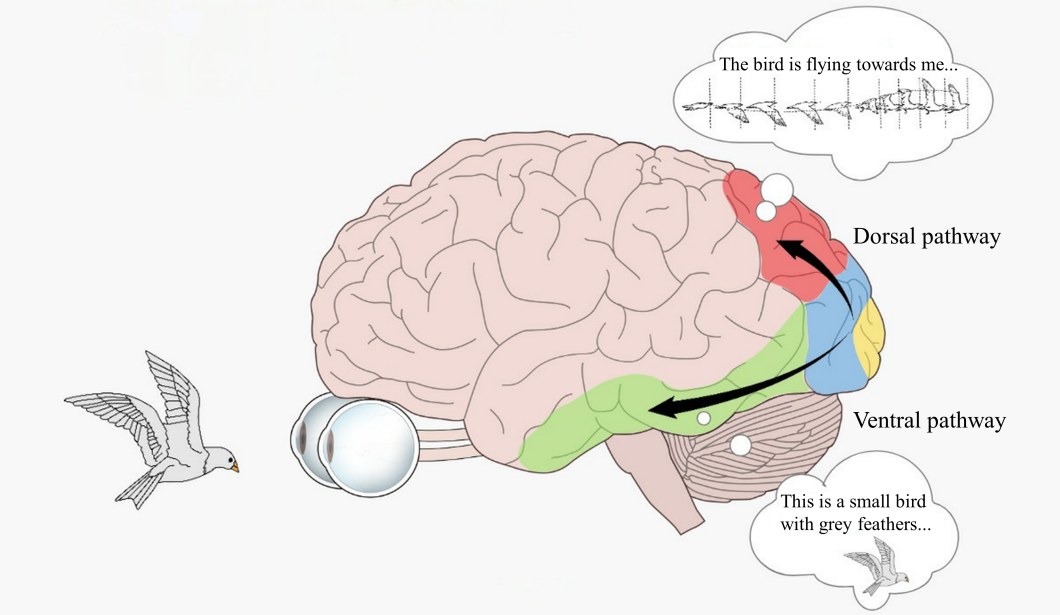


Source: Adapted by the author from original material from (The University of British Columbia, 2023).
Temporal retinal fibers in each eye convey information from the contralateral medial visual field, while medial retinal fibers carry information from the contralateral temporal visual field. Optic nerves aggregate and transmit all information from the ipsilateral eye. At the optic chiasm, nasal retinal fibers cross, while temporal retinal fibers remain on the same side. Optic tracts then contain fibers originating from the contralateral visual field. The subsequent optic radiations extend from the lateral geniculate nucleus in the thalamus, projecting to the primary visual cortex.

Beyond the primary visual processing center, areas form neural maps of the visual field. Information then diverges into two major pathways: a ventral route to the temporal lobe for stimulus identification and a dorsal route to the parietal lobe for spatial localization critical in guiding movement (see Fig. 2). The corpus callosum, a significant fiber bundle

connecting hemispheres, plays a role in unifying the perception of objects across the vertical meridian, where hemifield representations overlap. This pathway, known as the geniculostriate pathway, constitutes the initial processing level for visual information, demonstrating the intricate organization of visual perception in the human brain.

Figure 2 – High order visual processing

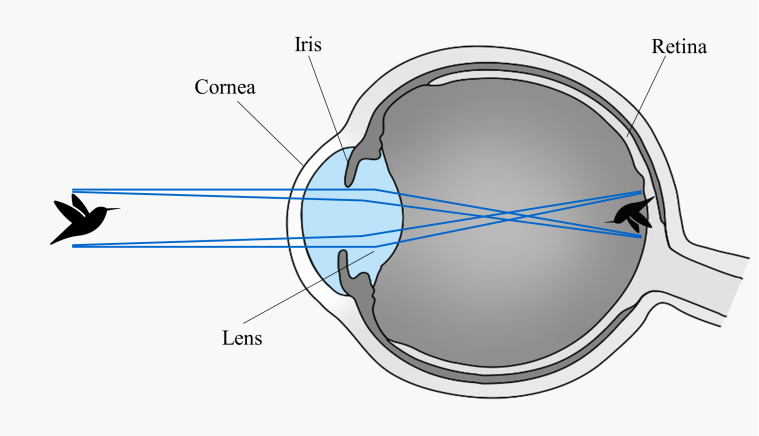


Source: Adapted by the author from original material from (The University of British Columbia, 2023).
The Primary Visual Cortex integrates information from parallel pathways, projecting to distinct layers. The Dorsal Pathway focuses on "where" and "how" in visual input, interpreting spatial information such as location and motion. It projects to the parietal lobe. Meanwhile, the Ventral Pathway emphasizes "what", interpreting object characteristics such as color, shape, and patterns, and directs projections to the temporal lobe. This depiction outlines the specialized roles of the dorsal and ventral pathways in high-order visual processing.

1.1.1.1 Low-level visual processing

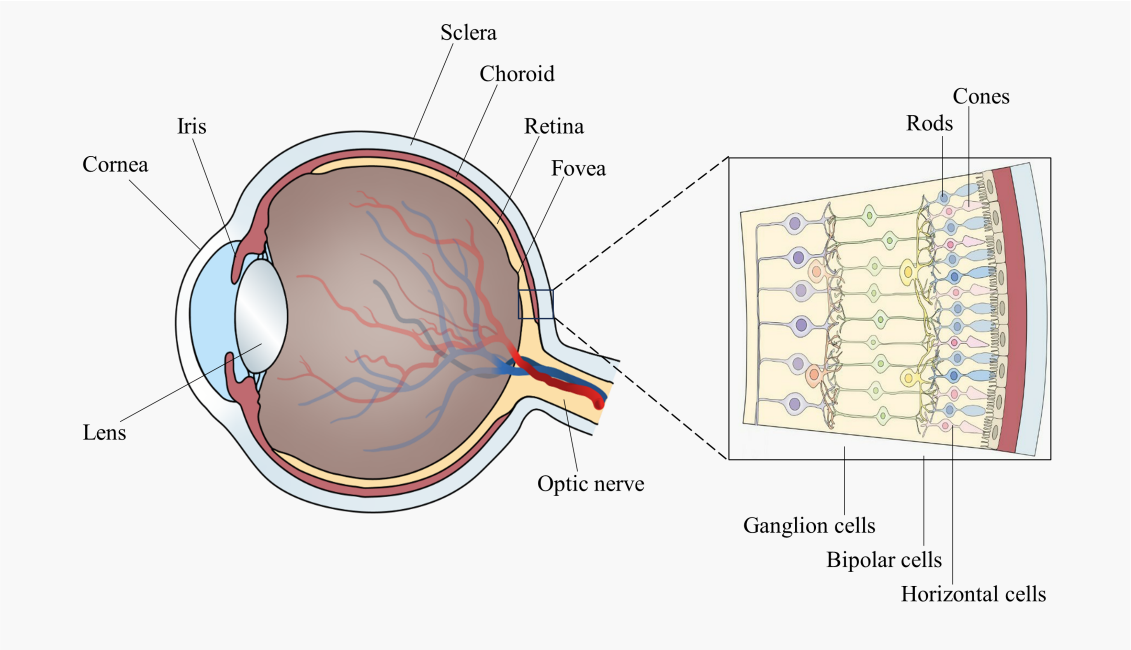
When a visual stimulus is perceived, light passes through the pupil, refracts through the cornea and lens, and reaches the retina, a thin neural sheet at the back of the eyeball. Fig. 3 illustrates the convex lens properties of the human cornea and lens, resulting in the convergence of light rays. This refraction leads to an inverted and reversed image formation on the retina. The retina comprises five major cell types organized into three cellular layers, as illustrated in Fig. 4. In its outermost layer, photoreceptor cells, including two types—cones and rods, undertake phototransduction, converting absorbed light into neural signals. These signals are then transmitted synaptically to bipolar cells, which connect to retinal ganglion cells, the retina's output neurons. The axons of these ganglion cells collectively form the optic nerve. During this

Figure 3 – Projection of an object onto the retina



Source: Adapted by the author from original material from (The University of British Columbia, 2023).
Light from an object in the visual field is refracted by the cornea and lens and focused onto the retina's photoreceptors.

Figure 4 – Human eye anatomy and retinal structure



Source: Adapted by the author from original material from (The University of British Columbia, 2023).
Left: Overview of the eye's posterior layers – retina (site of phototransduction), choroid (vascular supply), and sclera (structural support). The fovea, situated within the retina, marks the region of maximal visual acuity. Right: A detailed section of the retina shows the cellular components involved in visual processing. Photoreceptor cells (rods and cones) transform incident light into neural signals. These signals undergo synaptic transmission to bipolar cells, which in turn connect to retinal ganglion cells. The axons of these ganglion cells collectively form the optic nerve, serving as the conduit for visual information from the retina to the brain.

phase, the retina filters and extracts specific low-level features from the visual field, discarding much of the stimulus information available at the receptor level. It emphasizes intensity gradients, such as object edges, and enhances temporal changes over unchanging scene portions. This

processing, governed by adaptable rules, allows the retina to adjust sensitivity to changing illumination conditions by the contraction of the iris. This mechanism ensures relatively stable vision despite varying light intensities. Remarkably, despite transmitting data to the brain through a mere one million optic nerve fibers, almost half of the cerebral cortex is dedicated to processing these signals.

1.1.1.2 Intermediate-level visual processing

Moving on to intermediate-level visual processing, this stage involves deciphering the boundaries and surfaces within a visual scene. This process includes determining object boundaries, discriminating surface lightness and color based on the intensity and wavelength of reflected light, and assembling local elements into a coherent arrangement of objects and their background. The brain utilizes inherent logic to make assumptions about likely spatial relationships between elements, overcoming ambiguity in signals from the retina. Context plays a fundamental role, influencing how a visual feature is perceived; the response of neurons in the visual cortex is context-dependent, influenced by contours and surfaces beyond its receptive field.

Moreover, the dynamic adaptability of visual cortex neurons, influenced by visual experience and perceptual learning, contributes to the brain's ability to refine its interpretation of visual information over time. Cognitive functions, such as attention, expectation, and perceptual task engagement, further influence visual processing in the cortex. The interplay between visual context, experience-driven changes in cortical circuitry, and cognitive factors is vital for the visual system's analysis of complex scenes.

1.1.1.3 High-level visual processing

In the last stage of visual processing, the brain engages in high-level functions, recognizing objects and assigning meaning to them. Meaning, in this context, is associated with the brain's capacity to identify objects consistently, even when presented with different views. This remarkable ability is attributed to specific neurons in the inferior temporal cortex, responding to various retinal images of the same object. Neurons in the prefrontal cortex respond not only to objects that look different but are semantically related. Assigning meaning also involves notions of function, utility, or intention, shaped by the observer's past sensory experiences, leading to improved perceptual and neuronal selectivity for commonplace objects. This implies that the

visual system becomes particularly attuned to recognizing familiar objects.

Additionally, the brain establishes connections or links between the neural representations of objects. In other words, in the visual recognition process, contextual information about a visual stimulus is integrated to produce a meaningful interpretation. This complex interplay between sensory experiences, semantic associations, and contextual integration emphasizes the nuances of meaning formation in visual cognition.

1.1.2 Visual field dynamics: foveal precision and peripheral awareness

The human visual field represents the entirety observable with fixed eyes. Although human visual fields span approximately 220°, this expansive range is not uniformly registered. When our gaze shifts towards a specific region within an image or scene, the focal point aligns with the fovea, a minute central pit in the retina densely populated with cone cells (see Fig. 4). This small but critical region, covering a mere $1-2^{\circ}$ of the visual field, is instrumental in capturing detailed data and facilitating color perception, thus providing sharp central vision. Surrounding the fovea, the parafoveal and peripheral regions contribute to the broader visual field, albeit with reduced detail. The parafovea, covering an area of approximately 5° of the visual field, extends the range of reasonable detailed vision. Beyond this, the peripheral vision encompasses the outer regions of the visual field, extending from approximately $30-60^{\circ}$ or more from the center. It excels in detecting motion and functions effectively in low light conditions, primarily owing to the prevalence of rod cells, which are more light-sensitive.

Despite the varying detail across regions, the brain efficiently utilizes strategic eye movements to gather visual information. Each fixation contributes to the collection of data, guiding the brain in determining the next focal point. This dynamic process enables the construction of a comprehensive and coherent image from the diverse pieces of visual information gathered during the exploration of the visual field.

1.2 Evolution of eye movement research: bridging mechanics and cognitive processes

The history of eye movement research has transformed from its 18th-century origins, where early scholars, lacking precise tools, could only observe the persistent motion of the eye. Throughout the 19th century, the focus was primarily on understanding the mechanics and control of eye movements, often likened to studying the “wires and pulleys” of the oculomotor

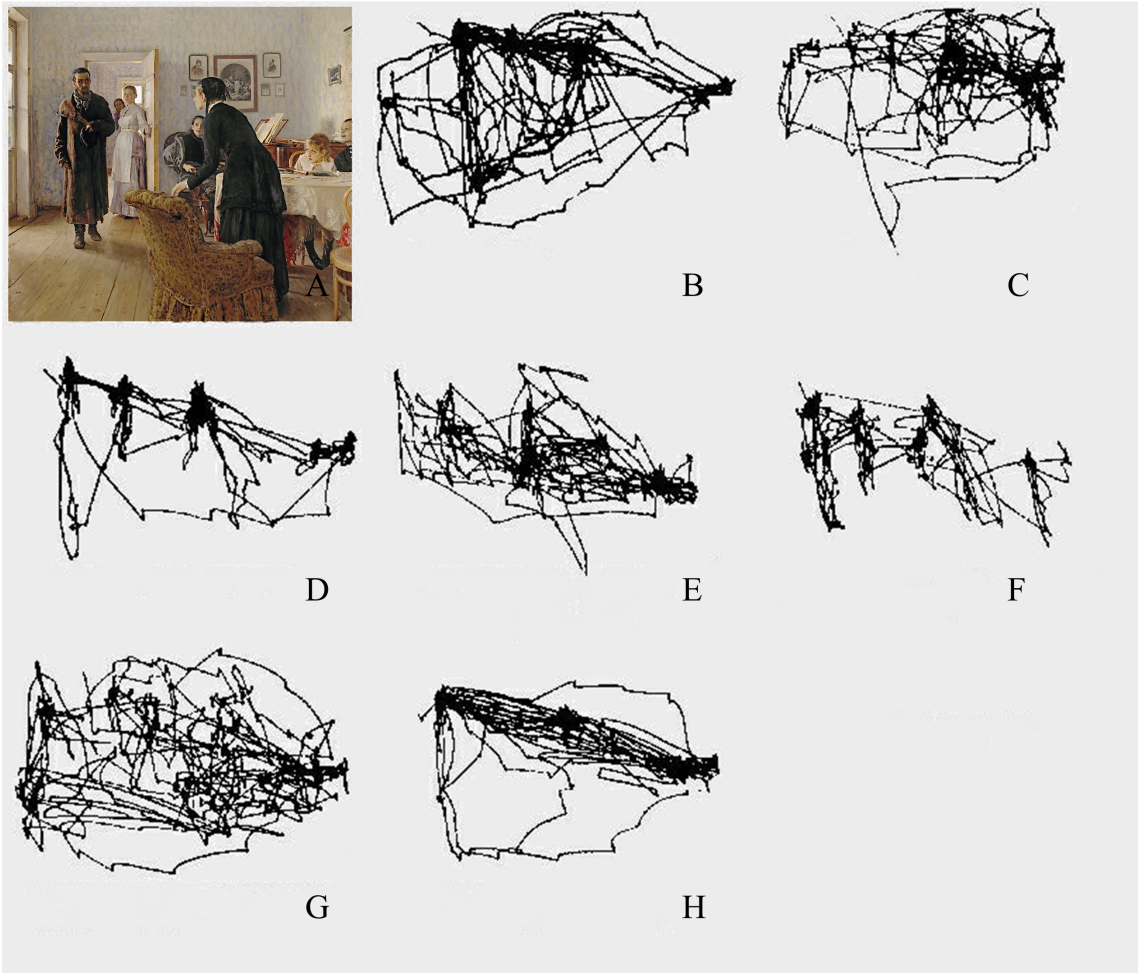
system (Wade, 2010; Findlay, 2003). Advancements in technology, notably photographic techniques in the late 19th century, allowed deeper investigations into the nature of eye movements. Javal's 1878 work on eye movements during reading, *Essai sur la Physiologie de la Lecture*, marked a turning point (Javal *et al.*, 1990). He quantitatively studied eye movements during reading, introducing the concept of fixations and saccades. This paradigm forms the basis for understanding visual text processing and marked the shift towards a more integrated approach that considers both the oculomotor system and cognitive factors, like attention and visual processing.

In the early 20th century, Guy Buswell contributed significantly to the field by focusing on eye movements during picture viewing. His monograph *How people look at pictures* emphasized the role of fixations in complex picture viewing and introduced the concept of "centers of interest" where fixations concentrate (Buswell, 1935). In the 1950s and 1960s, Alfred Yarbus conducted an influential experiment demonstrating the influence of observer instructions on eye movements (Yarbus, 1967). Participants directed their gaze to different areas of the same visual stimulus based on the specific task, highlighting the dynamic nature of eye movements influenced by cognitive processes and task demands (see Fig. 5). As Yarbus observed:

> *"Records of eye movements show that the observer's attention is usually held only by certain elements of the picture. (...) Eye movements reflect the human thought processes; so the observer's thought may be followed to some extent from records of eye movements (the thought accompanying the examination of the particular object). It is easy to determine from these records which elements attract the observer's eye (and, consequently, his thought), in what order, and how often."*
>
> *"Eye movements and vision" (Alfred L. Yarbus, 1967, p.190)*

The late 20th century witnessed a surge in eye movement research during reading, aided by modern eye trackers. Leading researchers like Keith Rayner and Reinhold Kliegl explored word recognition mechanisms and visual attention allocation during reading. They also investigated factors influencing fixation durations, shedding light on how word properties, sentence context, and reading speed interact (Rayner, 1998; Kliegl *et al.*, 2004). At the same time, researchers developed mathematical models to describe both fixation durations and saccade lengths, aiming to better model eye movements during reading (Reichle *et al.*, 1998; Engbert *et al.*, 2002).

Figure 5 – Visual Exploration of Repin's "They Did Not Expect Him"



Source: Adapted by the author from original material from (Repin, Between 1884 and 1888; Lucs-kho, 2007).
A) "They did not expect him", oil on canvas by Ilya Repin, 1884-1888. The painting depicts the unexpected return of an exile to his family after many long years of penal servitude (Yudenkova, 2019). In Yarbus's study "Eye movements and vision" (Yarbus, 1967), a series of figures and eye movement records are presented, illustrating observer gaze patterns during B) free examination, and specific tasks: C) estimate material circumstances of the family, D) give the ages of the people, E) surmise what the family had been doing before the arrival of the unexpected visitor, F) remember the clothes worn by the people, G) remember positions of people and objects in the room, and H) estimate how long the visitor had been away from the family. Gaze paths in each scenario reveal the major impact of altering observer instructions, and consequently their viewing task, on inspection behavior.

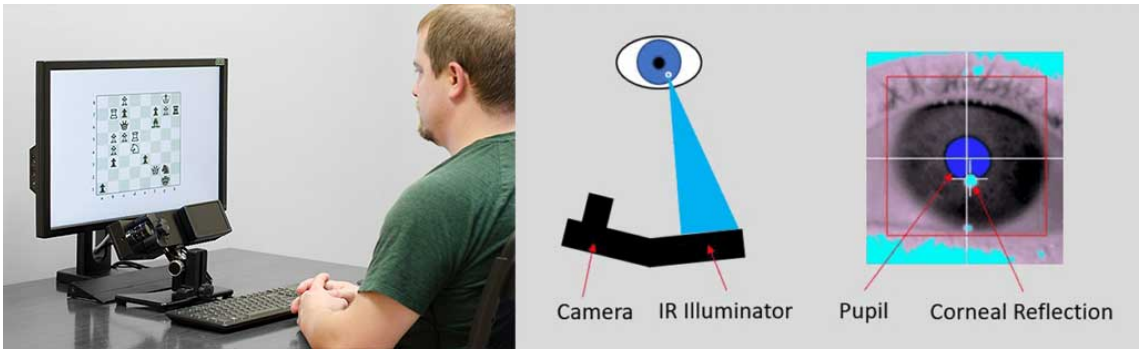
In recent decades, our understanding of the interaction between cognitive processes and eye movement has grown, and this progress has had a notable impact on other fields. Jeremy Wolfe and Anne Treisman's work in the 1980s and 1990s uncovered systematic scanning of visual scenes based on task relevance, significantly impacting fields such as visual search and human-computer interaction (Treisman e Gelade, 1980; Wolfe, 1994; Wolfe e Horowitz, 2004). Further research explored how top-down cognitive factors influence eye movements and examined the interaction between cognitive load, eye movements, and working memory (Torralba *et al.*, 2006; van Gog e Scheiter, 2010; Sweller *et al.*, 2011; Lin *et al.*, 2015). This revealed that

cognitive load significantly influences fixation choices and memory. Moreover, eye-tracking has found clinical applications, including diagnosing ophthalmological conditions and assessing neurological disorders such as Parkinson's disease, autism, and schizophrenia (Sweeney *et al.*, 1994; Chen *et al.*, 2018; Ahmed *et al.*, 2022). Advancements in machine learning techniques have enabled sophisticated analysis of eye-tracking data, predictive modeling for processes like visual attention prediction, gaze behavior analysis, and gaze-based interaction in applications like assistive technology and virtual reality. This research has deepened our understanding of human visual cognition and behavior with practical applications across multiple domains.

1.3 Eye-Tracking methodology using SR Research EyeLink 1000

In our study, we employed the SR Research EyeLink 1000 eye tracker, a state-of-the-art infrared video-based system designed for precise capture of eye movements (SR Research Ltd., 2023). Operating on advanced image processing algorithms, the EyeLink 1000 swiftly determines the participant's gaze on the screen, achieving an impressive timeframe of up to 2ms. This eye tracker relies on the Pupil Center Corneal Reflection (P-CR) system, known for its precision and non-intrusiveness in eye-tracking (Poole e Ball, 2006; Raney *et al.*, 2014). When presenting stimuli, near-infrared light from a fixed illuminator beside the camera is directed onto participants' eyes. A specialized camera captures two key reflections: the center of the pupil and the center of the corneal reflection, manifested as a small, sharp glint (the "first Purkinje image") and a distinct, well-defined pupil (the "bright pupil" effect), respectively (see Fig. 6). As the eye moves, the pupil's position on the camera sensor changes, while the corneal reflection remains stable, assuming the head is stabilized. To compensate for inevitable head movements, the EyeLink 1000 utilizes P-CR tracking, "subtracting" changes in the corneal reflection from pupil position changes, distinguishing genuine eye rotations from head shifts. The vector connecting the pupil and corneal reflections is used to precisely calculate the gaze location for each sample.

Figure 6 – Gaze detection in video-based eye-tracking system



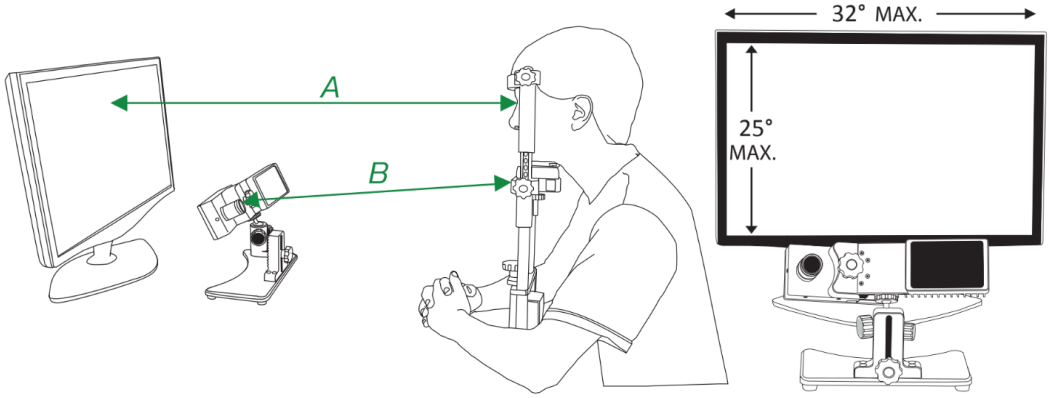
Source: Image retrieved from SR Research EyeLink website, with permission (SR Research Ltd, 2023a). Left: A participant engaged in an eye-tracking experiment using the SR Research EyeLink 1000. The video-based eye-tracking system employs a camera located below the display monitor, capturing a series of images of the eye at a frequency of 1 kHz in our setup. Image processing algorithms identify key locations in these images—the center of the pupil and the center of the corneal reflection. Gaze position is calculated based on these identified locations. Right: A close-up of the eye-tracking monitor displaying the detected eye with the pupil and corneal reflections. The corneal reflection, represented by a small, sharp glint (the "first Purkinje image"), originates from light directly reflected in the cornea. Simultaneously, the pupil reflection, a distinct, well-defined disc (the "bright pupil" effect), results from light reaching the retina and reflecting back.

The EyeLink configuration ensures seamless coordination between data collection and stimulus presentation. A typical setup consists of two computers: the Host PC, dedicated to data collection, and the Display PC, responsible for presenting stimuli to the participant. These computers are connected through an Ethernet link, enabling the exchange of critical information between them. The Host PC shares data such as eye events, gaze position, and camera images during participant setup with the Display PC. Conversely, the Display PC can communicate with the Host PC, allowing applications on the Display PC to control data collection. This bidirectional communication also facilitates the accurate recording of events from I/O devices on the Display PC, like keyboard presses and mouse clicks, in the data file.

In our experimental setup, we employed the Desktop Mount Participant Setup, featuring a chin and head rest to ensure participant comfort and minimize head movement during the trials. To maintain consistent eye positions across participants, the chin support height was individually adjusted according to each participant's height. The geometrical arrangement of the camera, participant, and screen adhered to recommended guidelines provided by SR Research Ltd., as illustrated in Fig. 7 (SR Research Ltd, 2023a). Specifically, participants positioned their eyes to align with the top 25% of the monitor, and the monitor itself was placed at a distance (A) of at least 1.75 times its width. The camera, equipped with a 35mm lens, was optimally positioned at a distance (B) ranging between 50 to 55 cm, following ideal recommendations. This setup was configured with a sampling frequency of 1 kHz for monocular recording, a

specification sufficient for our reading and image interpretation experiments, obviating the need for higher frequencies or tracking both eyes. Adjustments for each participant included setting the camera focus appropriately and determining illumination thresholds for pupil and corneal reflections. These thresholds varied among participants, with particular attention to those wearing glasses, which typically necessitate higher illumination intensity.

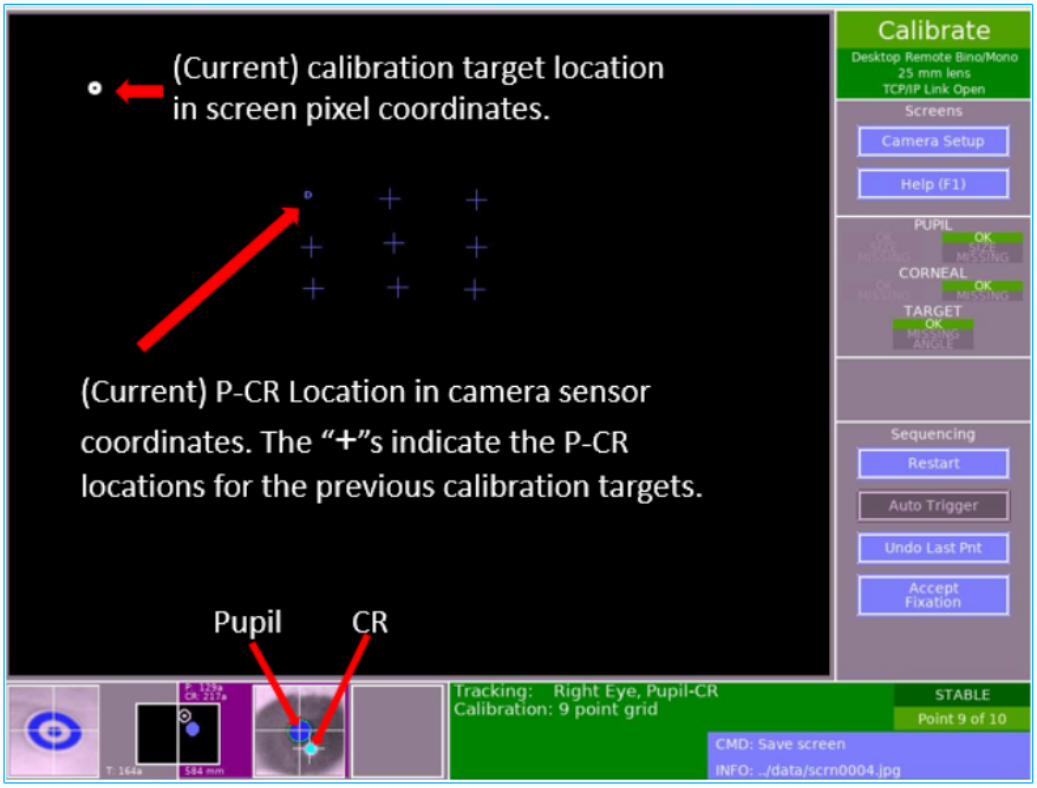
Figure 7 – Eye-tracking experimental setup



Source: Image retrieved from SR Research online Support Forum, with permission (SR Research Ltd, 2023a). Left: Experimental setup showcasing a participant using the SR Research Eye-Tracking Desktop Mount Participant Setup with chin and head support. Distances A (monitor to eyes) and B (camera to chin support) are indicated. Right: Specifications of the display monitor.

The calibration process in eye-tracking is an essential step to ensure the accuracy of gaze position data (SR Research Ltd, 2023b). In calibration, participants are instructed to focus on a series of predetermined target points displayed on the monitor. These targets are presented one at a time, and as the participant looks at each point, the eye tracker collects samples of pupil and corneal reflections data. Subsequently, the eye tracker software utilizes a regression model to map the locations of the pupil and corneal reflection centers, represented in camera sensor pixel coordinates, into gaze positions on the display screen, measured in pixel coordinates (see Fig. 8). To validate the calibration, the software automatically compares the predicted gaze positions with the known target locations, offering diagnostics. Our configuration adopts a 9-point calibration, consistent with recommended practices for reading experiments. Following calibration, the Validation step assesses the accuracy of the calibration. The validation test confirms precision within the predetermined error range of 0.25° to 1°. This thorough calibration process is performed before each participant's trial, ensuring precise eye-tracking measurements throughout the experiment.

Figure 8 – Eye-tracking calibration process



Source: Image retrieved from SR Research online Support Forum, with permission (SR Research Ltd, 2023b).
The calibration process integrates two inputs: Pupil-Corneal Reflection data and the coordinates of calibration targets on the screen in pixel units. Both of these inputs are visually represented in the Host PC software during the calibration procedure. The 'D' symbol, denoting the current eye location, and the + symbols, indicating the past fixation locations for each target, represent the P-CR data within the camera sensor coordinate space. This integration ensures the precise mapping of gaze positions onto the display screen.

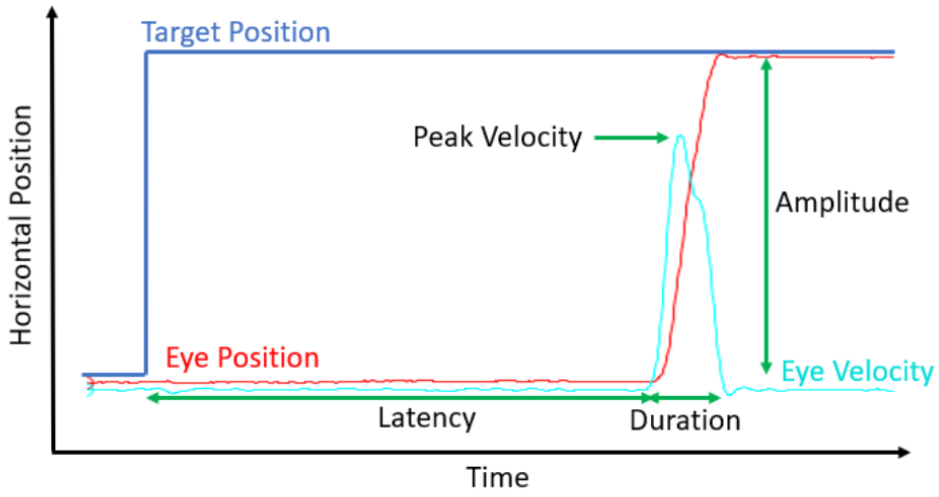
In implementing our eye-tracking experiments, we utilized the SR Research Experiment Builder software (EB), a graphical programming environment specifically designed for psychology and neuroscience studies, with a focus on eye-tracking research. Our experimental designs were structured as flow diagrams, incorporating stages for calibration, validation, and image display. Configuration settings included calibration parameters, trial randomization, and triggers for keyboard and mouse inputs. Additionally, we programmed the software to record data related to keyboard presses, mouse clicks, and pointer localization.

Upon concluding an eye-tracking session, the software generates a data file including eye-tracking data such as gaze coordinates, timing, and blink occurrences, alongside event messages configured within the EB experiment. Furthermore, the EB experiment is configured to promptly compute fixation and saccade coordinates and timing. A supplementary file containing this derived data is also generated. The criteria for identifying these variables are outlined below.

1.3.1 Saccadic detection and fixation identification

To identify fixations in our eye-tracking experiments, we used the EyeLink On-Line Parser, employing velocity and acceleration-based methods to interpret eye position data and distinguish between saccades and fixations. The parser operates with a "saccade picking" approach during data collection on the Host PC, distinct from other methods as it avoids specific duration or dispersion thresholds for fixations. Instead, it recognizes intervals between saccades to identify fixations. For each sampled data point, the parser calculates instantaneous velocity and acceleration, comparing them against predetermined thresholds. A signal for saccade onset is generated when either velocity or acceleration surpasses a specified threshold, while a signal for saccade offset occurs when either drops below the threshold. To define fixations, the parser identifies the conclusion of a saccade, marking the start of the ensuing fixation, and vice versa (see Fig 9).

Figure 9 – Saccadic detection and fixation identification



Source: Image retrieved from SR Research EyeLink website, with permission (SR Research Ltd, 2023b).
The graph illustrates the temporal dynamics of an individual's eye movements while observing a specific point and subsequently shifting gaze to a target on the right. The red curve traces the eye position over time, distinguishing fixations, where the eye remains steady, from saccades, characterized by rapid changes in eye position. The light blue curve represents eye velocity, registering zero during fixations and reaching peak values during saccades. Saccade onset and offset are identified based on velocity thresholds. The end of a saccade marks the beginning of a new fixation, as evidenced by the rapid transition to a new eye position (target position).

In our experimental configuration, we fine-tuned the saccadic detection process with specific parameters tailored to cognitive research:

- Saccade Velocity Threshold: 30°/*sec*
- Saccade Acceleration Threshold: 8000°/sec^2
- Saccade Motion Threshold: 0.1°

These parameters, optimized for reading and cognitive experiments where fixation durations are critical, operate with a conservative bias, effectively reducing sensitivity to noise and disregarding most saccades smaller than 0.6°. The higher velocity threshold of 30°/*sec*, results in fewer saccades and extended fixation durations. The motion threshold of 0.1° strategically delays saccade onset until a significant eye movement occurs. This configuration enhances sensitivity to larger saccades relevant to cognitive processes, while minimizing the influence of microsaccades and short fixations in the data. Typically, fixation durations range from 200 to 300 milliseconds but can vary significantly depending on the task context. For instance, fixations tend to be shorter during reading compared to viewing scenes.

Blinks are identified by the eye-tracking software when the pupil size is notably small or obscured due to eyelid occlusion. The software accurately marks the initiation and conclusion of each blink, signifying potential periods of data loss. It is important to note that the recorded position and velocity data during blink events are not valid. Therefore, any data captured between the onset and conclusion of blinks is deemed unreliable and should be excluded. To ensure data integrity, we examined fixations immediately preceding and following blinks, considering their potential influence from the blink process. In line with recommended practices for this setting, fixations with durations shorter than 100 milliseconds before or after blinks were discarded. By configuring precise parameters for eye event detection and removing possible artifacts in the data, we enhance the reliability of our fixation identification, in line with the specific goals of our research.

2 EYE-TRACKING AS A PROXY FOR COHERENCE AND COMPLEXITY OF TEXTS

2.1 Maximum-Entropy Modeling of eye fixations: a framework for analyzing reading processes

In the late 20th century, a new area of linguistic research emerged, focusing on the examination of eye movements during reading. These studies employed diverse strategies to analyze how words are fixated based on specific linguistic factors. It was found that both the number and duration of the fixations on each word constitute plausible measures for quantifying word processing and language comprehension (Clifton *et al.*, 2016; Rayner e Duffy, 1986; Kliegl *et al.*, 2006; Staub e Rayner, 2007). Generally accepted factors significantly correlated with word processing include a word's frequency in the language (Rayner e Duffy, 1986; Kliegl *et al.*, 2006; Rayner e Raney, 1996; Kliegl *et al.*, 2004; Ashby *et al.*, 2005; Rayner *et al.*, 2006), length (Kliegl *et al.*, 2006; Kliegl *et al.*, 2004; Juhasz *et al.*, 2008) and predictability in context (Kliegl *et al.*, 2006; Kliegl *et al.*, 2004; Rayner *et al.*, 2006; Ehrlich e Rayner, 1981; Schad *et al.*, 2010). Within this framework, eye-tracking experiments, when combined with mathematical models, have successfully provided formal descriptions of how individuals control eye movements during reading, with attention shifting from word to word (Reichle *et al.*, 1998; Engbert e Kliegl, 2001; Engbert *et al.*, 2002; Engbert *et al.*, 2004).

The complexity and coherence of a text are recognized as key linguistic attributes for evaluating reading comprehension and learning difficulties (Rothman, 2012; McNamara *et al.*, 1996). Text complexity measurement has been a focus of linguistic research in recent decades (Fisher *et al.*, 2012), particularly due to the necessity of selecting appropriate texts for different educational levels to facilitate progressive development of reading and text comprehension skills (Rothman, 2012). Mathematical expressions and metrics have been devised to assess the readability of texts, facilitating the quantification of their complexity and the categorization of reading materials (Nelson *et al.*, 2012). Commonly employed variables include the average length of words and their frequency in the language, both accounting for semantic difficulty, and sentence length, which is linked to syntactic complexity. The underlying premise of these empirical expressions is evident: texts characterized by uncommon, lengthy words and extensive sentences pose a greater cognitive challenge than those featuring familiar vocabulary and concise sentence structures. In contrast, the coherence of a text is related to its meaningfulness, a

concept rooted in semantics rather than grammatical structure (Reinhart, 1980). A coherent text possesses a logical flow wherein ideas are seamlessly interconnected, contributing to the overall consistency of the narrative. Each sentence not only holds significance independently but, more importantly, contributes successively to the evolving meaning of the text.

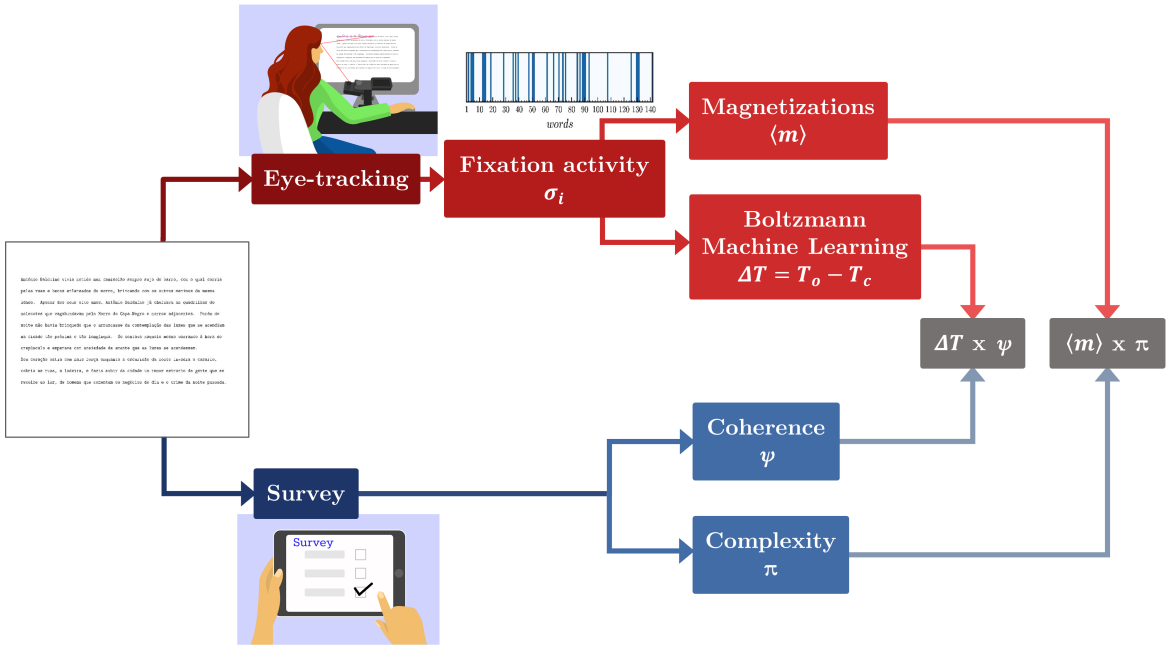
It is expected that features of a text such as genre and style may be reflected in the eye movement patterns of individuals when reading text passages. Different types of texts may elicit different reading responses in terms of fixation configurations and, consequently, different cognitive reactions. To investigate this interplay, a phenomenological modeling approach based on eye-movement data, specifically fixation patterns, should capture the underlying cognitive processes of reading. In this context, models from Statistical Physics, such us the Maximum-Entropy Model (MEM) developed in information theory, provide a statistical framework to understand natural processes in terms of "interactions" among elementary units using experimentally obtained statistical data (Shannon, 1948; Jaynes, 1957a; Nguyen *et al.*, 2017). The principle of maximum entropy posits that the probability distribution best representing the state of a system is the one that maximizes its entropy, while conforming to specific constraints. This principle encapsulates the solution to the *Inverse Ising Problem*, in which the Hamiltonian (*i.e.*, the interactions) in a complex system can be inferred from observed statistical correlations among its components. This statistical analysis is often referred to as the *Boltzmann-machine*, since it uses the Boltzmann distribution in its core.

The MEM approach has demonstrated versatility through its application to a range of systems that can be mapped to Ising-like models. In this conceptual framework, elements exhibiting interactions can exist in either an active or inactive state, mirroring an Ising-type system. To elaborate, consider an Ising system as a lattice of dipole moments where spins adopt either an "up" state ($+1$) or a "down" state (-1), influenced by external fields. In the context of neuronal networks, the MEM methodology has been instrumental in inferring interactions among neurons under specific stimuli, leveraging data derived from the observed firing patterns (Schneidman *et al.*, 2006; Cocco *et al.*, 2009; Shlens *et al.*, 2006; Tang *et al.*, 2008). On a larger scale, interactivity among regions of the human brain has been investigated using data from nuclear magnetic resonance (Watanabe *et al.*, 2013). An important application of the MEM is the characterization of protein- protein interaction benefiting from large protein databases (Morcos *et al.*, 2011; Weigt *et al.*, 2009). Specifically, this approach has been employed to infer genetic interaction networks based on known gene expression patterns (Stein *et al.*, 2015;

Lezon *et al.*, 2006; Locasale e Wolf-Yadlin, 2009). The MEM has also been instrumental in examining collective behaviors, such as the coordinated movement of bird flocks (Bialek *et al.*, 2012; Bialek *et al.*, 2014). Furthermore, it has been applied to understand the emergence of collective behavior from the eye movement patterns of individuals watching commercial videos, where pairwise correlations among series of instantaneous eye velocities were employed to capture the collective response and correlate it with video popularity (Burleson-Lesser *et al.*, 2017). Beyond the realm of biology, the MEM approach has demonstrated effectiveness in diverse fields. For instance, in the study by Bury et al. (Bury, 2013), the intricate micro-structure of network interactions in the stock market was unveiled through pairwise correlations calculated from extensive databases of stock variability. This versatility evidences the broad applicability of the MEM methodology outside the biological domain.

As illustrated in Fig. 10, in this chapter we address the quantitative characterization of text complexity and coherence through a two-fold approach. First, we conduct eye-tracking experiments with a limited participant group to directly analyze their fixation data while reading diverse texts -children stories, literary excerpts and random word generated texts (see Table 1). This is achieved by expressing experimental results in terms of a binary model for fixation sequences, analogous to an Ising system, duly embedded in the MEM. This enables the disclosure of two indexes that can serve as potential proxies for complexity and coherence: magnetization (a measure of density of the fixation sequences) and the distance between the "operating temperature" of the system and its critical temperature (a measure of cohesion among fixation sequences). Second, our experimental approach is validated through an extensive Internet reading survey, with access to a vast respondent sample, to categorize the same texts according to different complexity and coherence levels, allowing a direct comparison with the obtained eye-tracking indexes.

Figure 10 – Diagram of the research route



Source: Prepared by the author.
We take a twofold approach to characterize text complexity and coherence through cognitive reading activity. On one side, we conduct an eye-tracking experiment to collect fixation data from a participant group. Fixation activity for each subject, while reading a specific text, is computed by binarizing word states: positive $(+1)$ if fixated at least twice, and negative (-1) if not fixated or fixated only once. The text reading "magnetization" is then computed per subject, averaged to obtain $\langle m \rangle$. Utilizing pairwise cross-correlations among fixation sequences, we infer a "Hamiltonian" for each text via the maximum entropy principle, employing a Boltzmann learning algorithm. Thermodynamic analysis of energy fluctuations aids in determining the text's proximity to a "critical point". In parallel, we collect reading-comprehension data from an an extensive Internet survey involving 400 participants to quantify text complexity ($\langle \pi \rangle$) and coherence ($\langle \psi \rangle$).

Table 1 – Information on selected texts

Symbol	Title	Author	Year	Country
GAU	O Gaúcho	José de Alencar	1870	Brazil
GSV	Grande Sertão: Veredas	João Guimarães Rosa	1956	Brazil
HCL	História do Cerco de Lisboa	José Saramago	1989	Portugal
JUB	Jubiabá	Jorge Amado	1935	Brazil
MEL	A Mão e a Luva	Machado de Assis	1874	Brazil
QUI	O Quinze	Rachel de Queiroz	1930	Brazil
RT1	Random text 1	-	-	-
RT2	Random text 2	-	-	-
ST1	Story 1: A patinha Esmeralda	-	-	Brazil
ST2	Story 2: A menina do leite	-	-	Brazil

Source: Prepared by the author.
Basic information about the 10 texts used in the eye-tracking experiments. All texts are written in Portuguese. RT1 and RT2 texts are generated with an online random word generator (Random Text Generator, 2018). ST1 and ST2 texts are popular children stories of unknown author and year.

2.2 Maximum-Entropy Model

In exploring complex systems, the challenge often involves managing extensive experimental data on system variables without explicit knowledge of the underlying system function that represents its state. In such cases, computing quantities such as expectation values or correlation functions among system components serves as a strategic approach to address the inverse problem and infer the system's Hamiltonian.

Here, we derive the maximum-entropy inference method grounded on the principle of maximum entropy. This principle posits that, given prior known data, the probability distribution most accurately representing the system's state is the one maximizing entropy. We begin by reviewing the historical definitions and concepts of entropy. Then, we lay the foundation for the Maximum Entropy Model employed in this work.

2.2.1 Entropy revisited: integrating Statistical Mechanics and Information Theory

In the 1870's Ludwig Boltzmann set the foundations for statistical mechanics, introducing a novel statistical interpretation of entropy. Traditionally conceived on a macroscopic scale, entropy was viewed as an extensive property of thermodynamic systems, linked to the inherent dissipation of usable energy in energy-to-work transformations. The second law of thermodynamics established that the total entropy of an isolated system cannot decrease over time, reflecting the irreversibility of spontaneous processes.

The evolution of statistical mechanics shifted the focus to the microscopic nature of systems, adopting a probabilistic treatment for system properties. At the microscopic scale, variables of all particles constituting a system (such as velocity, spin, etc.) define numerous possible configurations (microstates) corresponding to a single macroscopic state of the system (macrostate) with specific parameters (such as temperature and magnetization). Consequently, it is possible to define a probability distribution p over all possible microstates of the system, and characterize the system based on this distribution. In particular, Boltzmann's entropy formulation,

$$S = -k \sum_i p_i \log p_i \tag{2.1}$$

serves as a quantification of the system's possible microstates, for the corresponding macrostate. Here, p_i represents the probability of the system being in the i-th microstate, and k is the Boltzmann constant. Statistical mechanics, offering profound insights into the underlying nature

of macroscopic observables, expanded the interpretation of quantities such as entropy. The second law was reinterpreted; of all possible states the system can evolve to, the most probable is the one corresponding to a larger number of microstates available, that is, a state of larger entropy, and is this probability that determines the direction of the system evolution.

Decades later, the concept of entropy resurfaced with the development of information theory. In 1948, Claude Shannon's seminal work "A Mathematical Theory of Communication", looked into the statistical structure of messages and factors influencing information transmission (Shannon, 1948). To describe how a source generates a message, Shannon conceptualized the source as selecting successive symbols based on certain probabilities. The central question was: within a set of possible events $\{x_i\}$ with probabilities of occurrence $\{p_i\}$, is it possible to quantify the uncertainty over the outcome of each event x_i? Shannon introduced a function of uncertainty, denoted as H, over the probabilities $\{p_i\}$ and established fundamental conditions for uniqueness and consistency:

1. H must be a continuous function of the probabilities p_i
2. If all p_i are equal ($p_i = 1/n$), then H should be a monotonically increasing function of n. This means that, if all events are equally likely, greater uncertainty corresponds to a greater number of possible events.
3. Composition law. If a choice is decomposed into two choices, the original H should be the sum of weighted values of H for each component.

For instance, considering a set of events $\{x_1, x_2, x_3\}$ with probabilities of occurrence $\{1/2, 1/3, 1/6\}$, decomposing the choice into $\{x_1, x_2'\} = \{x_1, \{x_2, x_3\}\}$ with probabilities $\{1/2, 1/2\}$ and $\{x_2, x_3\}$ with probabilities $\{1/3, 1/6\}$, H should satisfy $H(1/2, 1/3, 1/6) = H(1/2, 1/2) + 1/2, H(1/3, 1/6)$, where the factor $1/2$ accounts for the frequency of the second choice.

Shannon then demonstrated that the sole function satisfying the given conditions is of the form,

$$H = -\sum_i p_i \log p_i \,, \tag{2.2}$$

which mirrors Boltzmann's entropy (see Eq. (2.1)). This function, known as Shannon's entropy, will be referred to simply as entropy in this work and denoted by S, following historical arguments as reported in (Tribus e McIrvine, 1971):

> *'My greatest concern was what to call it. I thought of calling it "information", but the word was overly used, so I decided to call it "uncertainty". When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, 'You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage.'*
>
> *Conversation between Claude Shannon and John von Neumann regarding what name to give to the "measure of uncertainty" or attenuation in phone-line signals (Tribus and McIrvine, 1971, p.180).*

Shannon's formulation enabled the application of maximum-entropy inference to various statistical problems. The rationale is that when selecting a probability distribution to describe a system's state, the most representative choice maximizes uncertainty and aligns with any prior information. This unbiased selection ensures no arbitrary assumptions are considered. This postulate, introduced in (Jaynes, 1957a), is commonly known as the *maximum entropy principle*.

The alignment between Eqs. (2.1) and (2.2) does not inherently establish a direct relationship between the two concepts. It was in follow-up works (Jaynes, 1957a; Jaynes, 1957b) that the link between information theory and statistical mechanics was thoroughly examined. These studies revealed that established physical relations from statistical mechanics could be independently derived through the maximum entropy principle. Importantly, this derivation did not necessitate strict physical assumptions such as ergodicity, metric transitivity, or equal *a priori* probabilities. Boltzmann's entropy, along with other thermodynamic quantities, can then be straightforwardly identified. Consequently, statistical mechanics transcends its traditional scope as a theory of thermodynamics, emerging as a broader framework for statistical inference applicable to a diverse range of systems.

2.2.2 Formulation of the Maximum-Entropy Model

We now formally derive the probability model employed in this study (Jaynes, 1957a), grounded in the maximum entropy principle elucidated in the preceding section. Our analysis involves general variables; in the subsequent subsection we will establish their correspondence with experimental quantities.

Consider a random variable x with potential values $\{x_1, ..., x_n\}$ and associated probabilities $\{p_1, ..., p_n\}$. Our knowledge is confined to the expectation value of a function $f(x)$:

$$\langle f(x)\rangle = \sum_{i=1}^{n} p_i f(x_i)\,. \tag{2.3}$$

Our goal is to determine the discrete probability distribution $\{p_1,...,p_n\}$ consistent with this information, without imposing further assumptions.

The initial requirement is that the unknown probability distribution is normalized:

$$\sum_{i=1}^{n} p_i(x) = 1\,. \tag{2.4}$$

On the basis of this partial information, and drawing from insights in the previous subsection, we posit that the probability distribution should maximize entropy while adhering to the constraints defined by Eqs. (2.3) and (2.4). We introduce auxiliary functions g_0 and g_1, each corresponding to a condition:

$$\begin{array}{l} g_0 = \sum_{i=1}^{n} p_i - 1 \\ g_1 = \sum_{i=1}^{n} p_i f(x_i) - \langle f(x)\rangle \end{array}\,, \tag{2.5}$$

leading to the need to solve:

$$\text{maximize } S = -\sum_{i=1}^{n} p_i\, ln\, p_i \text{ subject to } \begin{cases} g_0 = 0 \\ g_1 = 0 \end{cases} \tag{2.6}$$

We recall that solving for extrema subject to equality constraints can be accomplished using the Lagrange multipliers method. To begin, we define the Lagrangian $\mathscr{L}$ as a linear combination of the function to be maximized, S, and functions g_0 and g_1:

$$\mathscr{L}(\{p_i\},\lambda_0,\lambda_1) = S(\{p_i\}) + \lambda_0 g_0 + \lambda_1 g_1 \tag{2.7}$$

where coefficients λ_1 and λ_2 are Lagrange multipliers. The probability distribution maximizing entropy among all possible discrete distributions on variable x must satisfy:

$$\frac{\partial \mathscr{L}}{\partial p_k}(\{p_i\},\lambda_0,\lambda_1) = 0\,. \tag{2.8}$$

This yields n equations, solved for $k = 1,...,n$. Substituting Eqs. (2.5), (2.6), (2.7) into Eq. (2.8), we obtain:

$$\frac{\partial}{\partial p_k}\left(-\sum_{i=1}^{n} p_i\, ln\, p_i\right) + \lambda_0 \frac{\partial}{\partial p_k}\left(\sum_{i=1}^{n} p_i - 1\right) + \lambda_1 \frac{\partial}{\partial p_k}\left(\sum_{i=1}^{n} p_i f(x_i) - \langle f(x)\rangle\right) = 0 \;\Rightarrow \tag{2.9}$$

$$-ln(p_k) - 1 + \lambda_0 + \lambda_1 f(x_k) = 0 \;\Rightarrow \tag{2.10}$$

$$p_k = e^{-1+\lambda_0+\lambda_1 f(x_k)} = e^{-1+\lambda_0} e^{\lambda_1 f(x_k)} . \tag{2.11}$$

Substituting into Eq. (2.4), we determine λ_0:

$$-1+\lambda_0 = -ln\left(\sum_{i=1}^{n} e^{\lambda_1 f(x_i)}\right). \tag{2.12}$$

Substituting this back into Eq. (2.11) yields:

$$p_k = \frac{e^{\lambda_1 f(x_k)}}{Z(\lambda_1)} , \tag{2.13}$$

for $k = 1, ..., n$. Eq. (2.13) represents the probability distribution with maximum entropy, where Z is the partition function:

$$Z(\lambda_1) = \sum_{i=1}^{n} e^{\lambda_1 f(x_i)} . \tag{2.14}$$

Finally, incorporating Eq. (2.3), we arrive at:

$$\langle f(x) \rangle = \frac{\partial}{\partial \lambda_1} lnZ(\lambda_1) . \tag{2.15}$$

We can extend this procedure to any number m of functions, corresponding to multiple constraints for which we possess information. For each constraint $r = 1, ..., m$, we have:

$$\langle f_r(x) \rangle = \sum_{i=1}^{n} p_i f_r(x_i) . \tag{2.16}$$

Similar to Eq. (2.14), the partition function is given by:

$$Z(\lambda_1, ..., \lambda_m) = \sum_{i=1}^{n} e^{\lambda_1 f_1(x_i) + ... + \lambda_m f_m(x_i)} , \tag{2.17}$$

and the probability distribution can be expressed as:

$$p_k = \frac{e^{\lambda_1 f_1(x_i) + ... + \lambda_m f_m(x_i)}}{Z(\lambda_1, ..., \lambda_m)} , \tag{2.18}$$

for $k = 1, ..., n$. Additionally, Lagrange multipliers λ_0, λ_r, $r = 1, ..., m$, satisfy:

$$-1+\lambda_0 = -ln(Z(\lambda_1, ..., \lambda_m)) , \tag{2.19}$$

and,

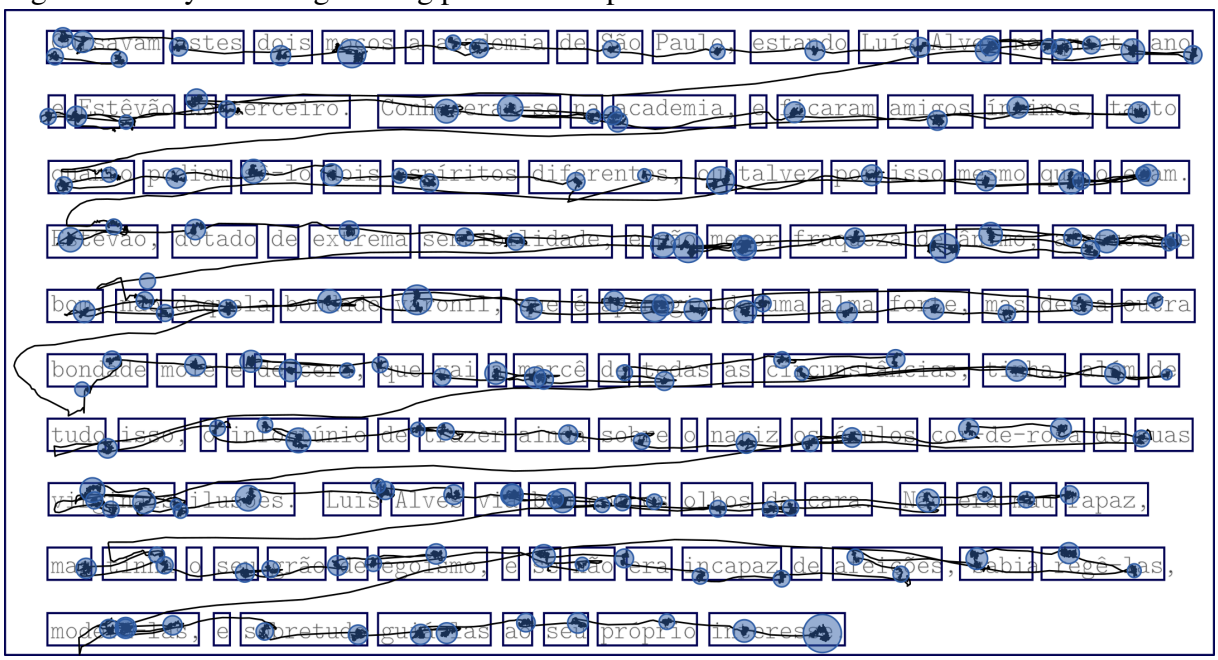
$$\langle f_r(x) \rangle = \frac{\partial}{\partial \lambda_r} lnZ(\lambda_r) . \tag{2.20}$$

2.3 Eye-Tracking experiment design: participants, materials, and procedures

We recruited 20 participants among physics and engineering graduate and postgraduate students, aged 17 to 34, all native Brazilian Portuguese speakers. The reading material comprised 10 diverse texts in Portuguese, including two children's stories, two randomly generated word texts (maintaining standard grammatical structure but featuring random content) (Random Text Generator, 2018), and six excerpts from literary works (refer to Table 1). All texts were presented in a 12-point monospaced font. In our equipment configuration, these specifications align with a visual angle of 1° spanning a length of 3 characters, ensuring word position accuracy (Raney *et al.*, 2014). The letters were displayed in light cyan against a dark gray background, ensuring high color contrast and moderate brightness to enhance readability while improving eye-tracking accuracy. Ethical approval for the eye-tracking experiment procedures was obtained from the Research Ethics Committee of the Federal University of Ceará (COMEPE, Universidade Federal do Ceará, Brazil), and all participants provided written informed consent. To encourage conscious reading, participants were informed before the experiment that they would be asked a simple question after reading each text. During the experiment, participants were seated in front of a display screen with their heads stabilized using an adjustable head and chin rest. The texts were sequentially presented on the screen, intercalated with corresponding questions. No time limit was imposed on the reading session. Throughout the reading session, the eye tracker recorded gaze location data at a sample rate of 1000 Hz, resulting in an average temporal error of 0.5 ms (approximately half the duration between samples) (Raney *et al.*, 2014). The collected data were de-identified to protect participants' privacy. The experiment was designed using the SR Research Experiment Builder program (version 1.10.1630), and the collected data were displayed and filtered using the Eye-Link Data Viewer software (version 1.10.1).

In the data processing phase, rectangles were used to delimit each word in the texts. The eye tracker provided spatial coordinates of fixations for each subject reading a given text, and the number of fixations within each rectangle was quantified, as exemplified in Fig. 11. At the conclusion of the experiment, an array was generated for each text, with elements n_i^r signifying the number of fixations a specific word r received from subject i. Additionally, the duration of each reading session was concurrently recorded during every eye-tracking experiment.

Figure 11 – Eye-tracking reading pattern example



Source: Prepared by the author.
Eye-tracking experiment plot illustrating gaze sequences and fixations during typical reading, using the example of the MEL text. Blue circles denote fixations, with sizes representing respective durations. Solid lines connecting circles depict the gaze trajectory. For a given text, the frequency n_i^r quantifies how often subject i fixates within the rectangle delimiting word r throughout the entire reading session.

2.4 Theoretical-Experimental synthesis: Maximum-Entropy Model for eye fixations in text reading

To establish an analogy with the Ising system, we define the fixation activity $\sigma_i = \{\sigma_i^1, ..., \sigma_i^M\}$ for each subject i reading a given text with M words based on the state of each word r, denoted by $\sigma_i^r = \pm 1$ according to the following rule,

$$\sigma_i^r = \begin{cases} +1 & \text{if } n_i^r \geq 2 \\ -1 & \text{if } n_i^r < 2 \ , \end{cases} \tag{2.21}$$

Here, n_i^r represents the number of times subject i fixates on word r during reading. With each subject having two possible states ($+1$ or -1), the system exhibits 2^N potential states for each word in the text. The threshold parameter of 2 fixations per word is adopted, considering that nearly every word in our eye-tracking experiments was fixated at least once by all subjects. This aligns with previously reported reading patterns (Clifton *et al.*, 2016), allowing meaningful distinctions in fixation activities by considering words with one fixation versus those with two or more.

The fixation activities obtained from eye-tracking experiments are illustrated in Fig. 12, presenting raster plots for all texts and subjects. Clear distinctions emerge in reading

patterns, notably in the significantly lower density of active states for ST1 and ST2 compared to RT1 and RT2.

Fixation activity density is quantified in terms of the "magnetization" m_i for each subject $i = 1, ..., N$. This metric represents the average fixation states σ_i^r over the M words of the text, expressed as:

$$m_i = \langle \sigma_i \rangle = \frac{1}{M} \sum_{r=1}^{M} \sigma_i^r \ . \tag{2.22}$$

Consequently, an overall magnetization for each text is defined as:

$$\langle m \rangle = \frac{1}{N} \sum_{i=1}^{N} m_i \ . \tag{2.23}$$

A relevant measure readily obtained from the conducted experiments, being certainly indicative of text processing demand, is the average reading time per word of the text, calculated as:

$$\langle t \rangle = \frac{1}{N} \sum_{i=1}^{N} t_i / M \ , \tag{2.24}$$

where t_i represents the reading time of subject i. Table 2 presents the values of both the average magnetization $\langle m \rangle$ and average reading time per word $\langle t \rangle$ obtained from the experiments for each text. Notably, the reading of ST1 and ST2 texts yielded the lowest values for both measures,

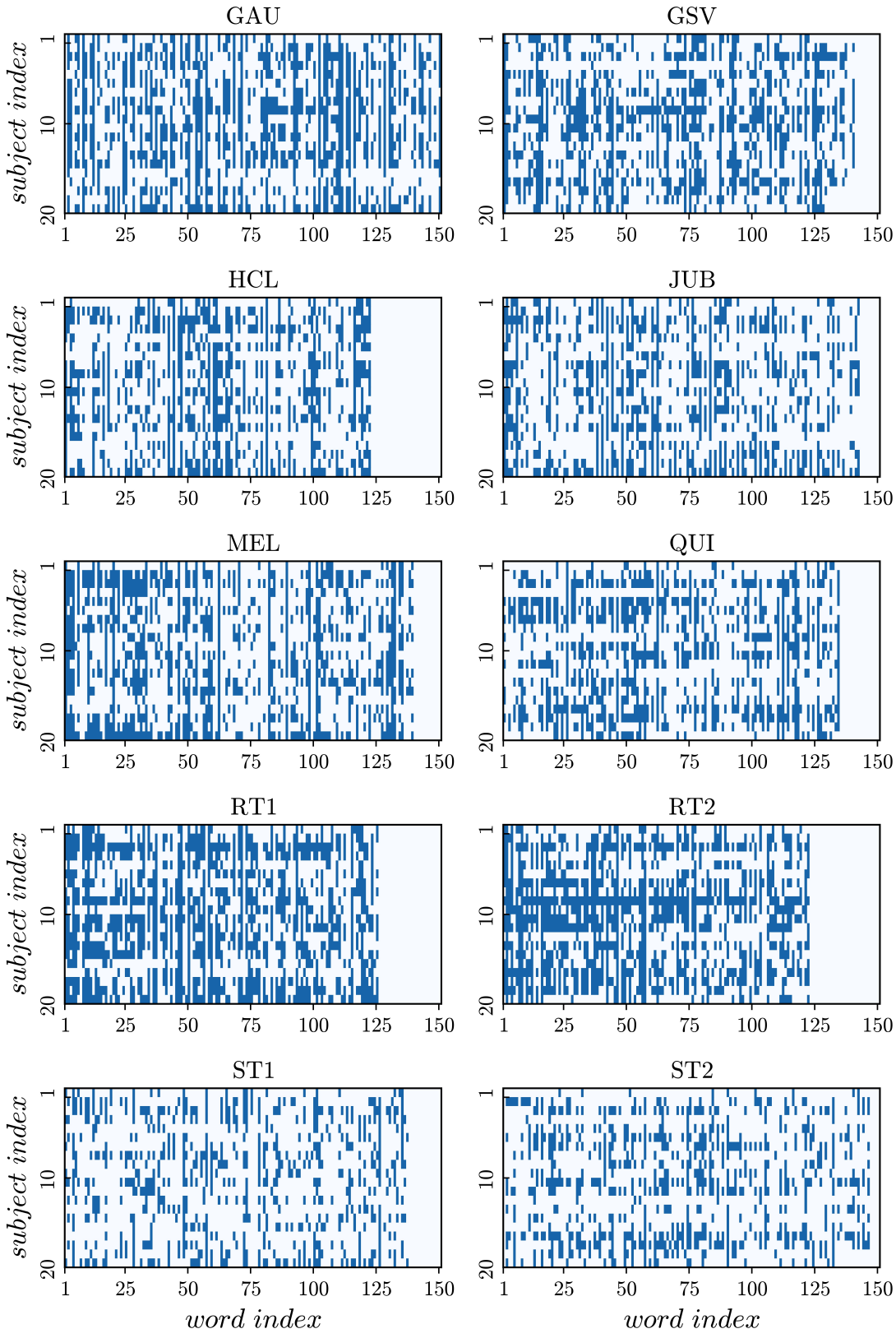
Table 2 – Average magnetizations and reading times per word across texts

Text	$\langle m \rangle$	$\langle t \rangle$ (ms)
GAU	-0.2147	744.82
GSV	-0.2657	611.12
HCL	-0.2918	595.08
JUB	-0.3697	552.34
MEL	-0.2460	641.70
QUI	-0.3022	513.37
RT1	-0.0464	923.98
RT2	-0.0664	793.13
ST1	-0.5190	398.38
ST2	-0.4966	394.21

Source: Prepared by the author.
The magnetization ($\langle m \rangle$) and reading time per word ($\langle t \rangle$) for each text represent the average values of the reading fixation activity and time per word, respectively. These averages are computed across all subjects.

suggesting efficiency in processing. Additionally, the resemblance in values between RT1 and

Figure 12 – Fixation activities



Source: Prepared by the author.
Raster plots depict fixation activities for all subjects during text reading. Each word state σ_i^r is considered active ($+1$ in blue) if $n_i^r \geq 2$, or inactive (-1 in white) if $n_i^r < 2$ for each subject i.

RT2, significantly higher than those of other texts, suggest a degree of correlation between the two measures. Despite these similarities, $\langle m \rangle$ and $\langle t \rangle$ are not perfectly aligned. Subsequent analysis will reveal that $\langle m \rangle$ captures nuanced information on cognitive activity during reading more effectively than $\langle t \rangle$.

We proceed by calculating the covariance C_{ij} between fixation activities for pairs of subjects i and j across the M words of the text:

$$C_{ij} = \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle \ , \tag{2.25}$$

where

$$\langle \sigma_i \sigma_j \rangle = \frac{1}{M} \sum_{r=1}^{M} \sigma_i^r \sigma_j^r \ , \tag{2.26}$$

and $\langle \sigma_i \rangle$ is defined by Eq (2.22). We derive P under the condition of reproducing experimental magnetization and covariance values for all subjects. Firstly, P needs to be normalized::

$$\sum_{\{\sigma\}} P(\{\sigma\}) = 1 \tag{2.27}$$

Next, we require the first moment of P to match the sample average for $i = 1, ..., N$:

$$\langle \sigma_i \rangle_P = \sum_{\{\sigma\}} P(\{\sigma\}) \sigma_i = \langle \sigma_i \rangle = \frac{1}{M} \sum_{r=1}^{M} \sigma_i^r \tag{2.28}$$

Similarly, the second moment of P should equal the measured covariances values for $i, j = 1, ..., N$, $i \neq j$:

$$\langle \sigma_i \sigma_j \rangle_P = \sum_{\{\sigma\}} P(\{\sigma\}) \sigma_i \sigma_j = \langle \sigma_i \sigma_j \rangle = \frac{1}{M} \sum_{r=1}^{M} \sigma_i^r \sigma_j^r \tag{2.29}$$

Seeking to maximize the entropy, $S = -\sum_{\{\sigma\}} P(\{\sigma\}) \ln P(\{\sigma\})$, we employ the method of Lagrange multipliers with equality constraints from Eqs. (2.27)-(2.29). The Lagrangian is expressed as:

$$\mathscr{L} = S + \lambda_0 (\langle 1 \rangle_P - 1) + \sum_{i=1}^{N} \lambda_{1_i} (\langle \sigma_i \rangle_P - \langle \sigma_i \rangle) + \sum_{i,j=1}^{N} \lambda_{2_{ij}} (\langle \sigma_i \sigma_j \rangle_P - \langle \sigma_i \sigma_j \rangle) \tag{2.30}$$

This Lagrangian must satisfy:

$$\frac{\partial \mathscr{L}}{\partial P(\{\sigma\})} = 0 \ \Rightarrow \ -\ln P(\{\sigma\}) - 1 + \lambda_0 + \sum_{i=1}^{N} \lambda_{1_i} \sigma_i + \sum_{i,j=1}^{N} \lambda_{2_{ij}} \sigma_i \sigma_j = 0 \tag{2.31}$$

where λ_0, $\{\lambda_{1_i}\}$, and $\{\lambda_{2_{i,j}}\}$ are the Lagrange multipliers. The solution to Eq.(2.31) is given by:

$$P(\{\sigma\}) = \frac{1}{Z} \exp\left(\sum_{i=1}^{N} \lambda_{1_i} \sigma_i + \sum_{i,j=1}^{N} \lambda_{2_{ij}} \sigma_i \sigma_j \right) \tag{2.32}$$

where Z is the partition function and serves as the normalization constant:

$$Z = \sum_{\{\sigma\}} \exp\left(\sum_{i=1}^{N} \lambda_{1_i} \sigma_i + \sum_{i,j=1}^{N} \lambda_{2_{ij}} \sigma_i \sigma_j \right) = \exp\left(1 - \lambda_0\right) \tag{2.33}$$

Observing Eq.(2.32), we note that P follows Boltzmann's probability distribution (Jaynes, 1957a):

$$P(\{\sigma\}) \sim \mathrm{e}^{-E/T}, \tag{2.34}$$

where T represents a temperature, and E corresponds to a Hamiltonian. For simplicity, here we arbitrarily set the "operating temperature", namely, the reading temperature, $T_o = 1$. This distribution emerges as the least biased representation for an Ising-type system, like ours, with known first and second moments. Moreover, the energy term aligns with the random field Ising Hamiltonian (Belanger e Young, 1991; Nguyen *et al.*, 2017):

$$E = -\sum_{i=1}^{N} h_i \sigma_i - \sum_{i,j=1}^{N} J_{ij} \sigma_i \sigma_j. \tag{2.35}$$

with the correspondence:

$$\begin{cases} \lambda_{1_i} = h_i \\ \lambda_{2_{ij}} = J_{ij} \ . \end{cases} \tag{2.36}$$

The inverse Ising problem involves deducing unknown couplings J_{ij} and fields h_i from known observables $\langle \sigma_i \sigma_j \rangle$ and $\langle \sigma_i \rangle$. From Eq. 2.33, it follows that

$$\langle \sigma_i \rangle = \frac{\partial}{\partial h_i} \ln Z(\mathbf{J}, \mathbf{h}) \tag{2.37}$$

$$\langle \sigma_i \sigma_j \rangle = \frac{\partial}{\partial J_{ij}} \ln Z(\mathbf{J}, \mathbf{h}) \tag{2.38}$$

where $\mathbf{J} = \{J_{ij}\}$ and $\mathbf{h} = \{h_i\}$. Solving Eqs. 2.37 and 2.38 enables the inference of couplings and fields, as depicted in Fig. 22 and Fig. 23 in Appendix A, demonstrating precise agreement between theoretical values $\langle \sigma_i \rangle_{th}$ and $\langle \sigma_i \sigma_j \rangle_{th}$ and experimental values.

The mathematical correspondence presented in Eq. 2.36 naturally suggests interpreting h_i as the action of a local external stimulus (text) on subject i, akin to a "random field", and J_{ij} as the "coupling coefficients" between subjects i and j. While direct communication doesn't occur among participants in our eye-tracking experiments, conceptualizing the text as a

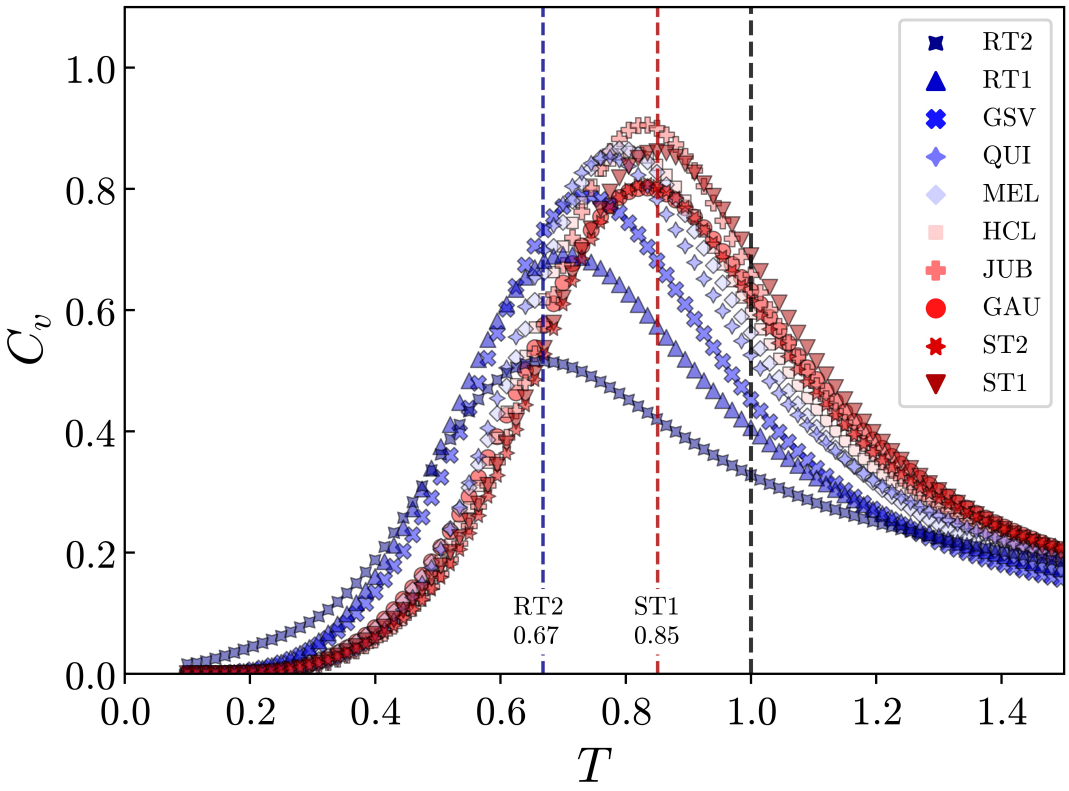
medium allows us to conceive an "interaction" between subjects i and j. Though individuals read texts independently, similarities in their cognitive responses to text characteristics may result in correlated fixation activities. These pairwise couplings or interactions contribute to observed correlations, potentially giving rise to emergent collective effects, offering valuable insights into the system under study.

With knowledge of the system's probability distribution, additional quantities can be estimated. The heat capacity C_v, defined as the rate of change of the system's average energy with temperature, can be calculated in terms of the energy mean square deviation (Newman e Barkema, 1999):

$$C_v = \frac{\beta}{N}\langle E^2\rangle - \langle E\rangle^2 \tag{2.39}$$

via the partition function. At a "critical temperature" T_c, C_v is maximal, indicating a phase transition: if $T_c < T_o$, the system is in a "liquid" or random state; conversely, $T_c > T_o$ signifies a more "ordered" condition. The critical temperature is identified as the point where C_v reaches its maximum. In Fig. 13, we illustrate the variation of C_v with respect to T for all texts. Regardless of the text, the operating temperature $T_o = 1$ consistently surpasses the critical point T_c. Notably, the distance to criticality $T_o - T_c$ varies depending on the text. Larger values, observed particularly for RT1 and RT2 in Table 3, indicate a greater distance compared to other texts. This distinction suggests that $T_o - T_c$ can serve as an index for discerning meaningful texts from random ones. Indeed, we will demonstrate that $T_o - T_c$ is intricately linked to language processing, particularly in terms of the perceived coherence during text reading.

Figure 13 – Heat capacity as a function of temperature for the system of fixation activities



Source: Prepared by the author.
Heat capacity curves for all texts, where C_v peaks at the critical temperature T_c. The operating temperature, denoted as $T = T_o = 1$, represents the reading temperature. Notably, the system operates above and in proximity to the critical point for all texts, with RT1 and RT2 exhibiting the greatest distance from the critical point.

Table 3 – Distance to criticality across texts

Text	$T_o - T_c$
GAU	0.169
GSV	0.262
HCL	0.192
JUB	0.170
MEL	0.207
QUI	0.229
RT1	0.296
RT2	0.332
ST1	0.149
ST2	0.167

Source: Prepared by the author.
The table presents distances to criticality, $T_o - T_c$, calculated for all texts. Here, $T_o = 1$ denotes the reading operating temperature, and the critical temperature T_c corresponds to the point where the heat capacity C_v for a given text reaches its maximum.

2.5 Quantifying text complexity and coherence

2.5.1 Methodological framework

To corroborate our eye-tracking findings, we engaged in an Internet survey facilitated by the MindMiners services company (MindMiners, 2019) in São Paulo, Brazil. This agency offers a digital platform for comprehensive research projects, encompassing questionnaire creation, data collection, and access to a respondent panel boasting over 400 thousand engaged users nationwide (MeSeems (MeSeems, 2019)). The methodology involves two key stages: the targeted selection of respondents based on study-specific criteria and the meticulous development and review of the survey questionnaire.

For our study, we carefully selected two demographically diverse groups, each comprising 200 individuals, ensuring diversity in age, gender, and geographic location from the extensive pool of 400 thousand respondents within the survey agency's panel. Stratification details are outlined in Table 4. A prerequisite for participation was a minimum of a high school degree to ensure mature reading-comprehension skills. Socio-economic distinctions were not considered. Each group received a questionnaire related to five out of the total ten texts. The distribution of texts was as follows:

– Group A: O Quinze, Grande Sertão: Veredas, História do Cerco de Lisboa, Story 2, Random text 2
– Group B: Jubiabá, O Gaúcho, A mão e a Luva, Story 1, Random text 1

The questionnaire comprises three parts, featuring exclusively multiple-choice questions. In the initial section, participants read texts sequentially, responding to simple content-related questions, aiming to encourage conscious reading. In the second part, each text is presented again, and participants rate its complexity on a scale of 1 to 5, ranging from "very simple" to "very complex". Finally, the texts are presented once more for respondents to evaluate coherence on a scale from 1 to 5, spanning "not coherent at all" to "very coherent". To mitigate response bias, texts are randomized for each participant.

The MindMiners company employs a dual approach for user validation, utilizing both manual and automated processes. User information is verified through external sources, including social networks and the Brazilian Department of Revenue. The company also employs algorithms to detect atypical behaviors, such as non-compliance with questionnaire instructions and abnormal response speeds. To ensure panel integrity, identifiers are embedded across all

Table 4 – Demographic stratification of respondents panel data

	GroupA		GroupB	
	Respondents	Percentage	Respondents	Percentage
Gender				
Male	86	43.0%	90	45.0%
Female	114	57.0%	110	55.0%
Age				
≤ 17	2	1.0%	3	1.5%
18-24	48	24.0%	46	23.0%
25-30	26	13.0%	35	17.5%
31-40	70	35.0%	61	30.5%
≥ 41	54	27.0%	55	27.5%
Education				
High school	99	49.5%	83	41.5%
University	101	50.5%	117	58.5%
Place of residence				
Central-West	20	10.0%	20	10.0%
Northeast	46	23.0%	46	23.0%
North	10	5.0%	9	4.5%
Southeast	94	47.0%	98	49.0%
South	30	15.0%	27	13.5%

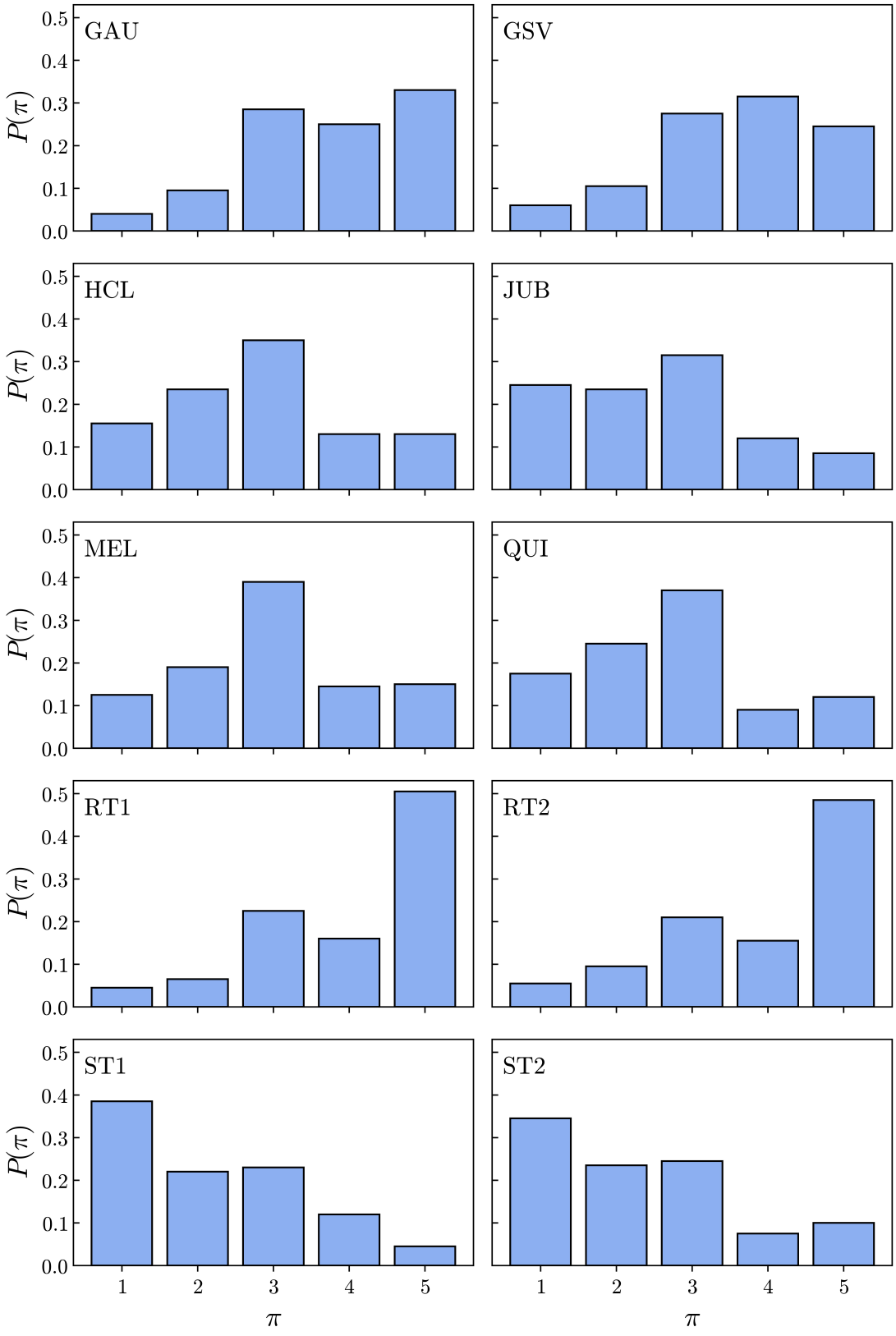
Source: Prepared by the author.
Stratification of respondents based on gender, age, education level (requiring a high school degree), and geographical location in Brazil (categorized by regions).

respondent devices, resulting in an exclusive panel comprised of unique users.

2.5.2 Insights from survey responses

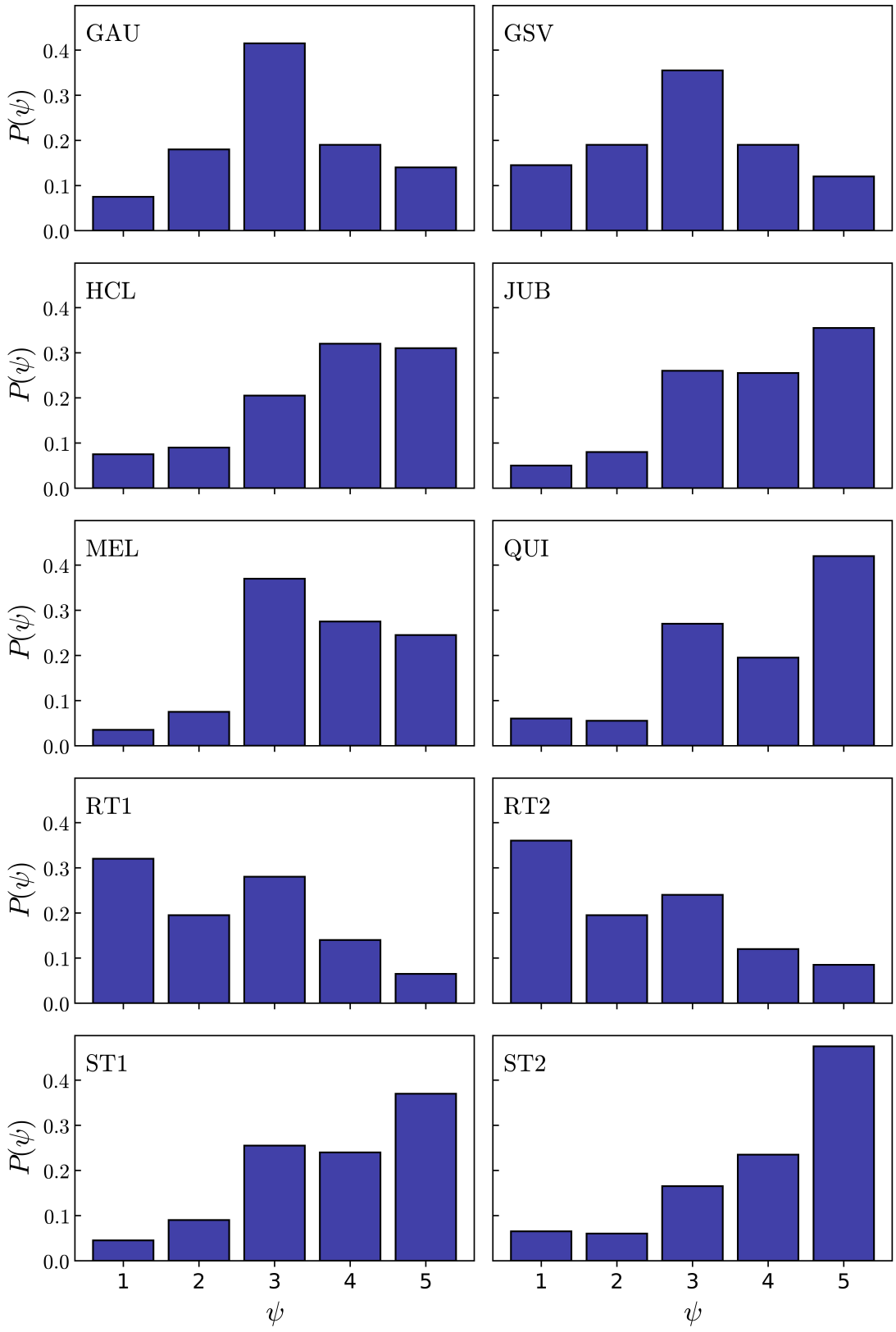
In Fig. 14 we present the distribution $P(\pi)$, representing the fractions of individuals assigning complexity ratings π to each text. We observe that approximately 40% of respondents deemed ST1 and ST2 as "very simple", contrasting with nearly 50% who considered RT1 and RT2 as "very complex". Other texts received diverse ratings spanning intermediate complexity values. Mean complexity values, $\langle\pi\rangle$, are detailed in Table 5. Fig. 15 displays the distributions $P(\psi)$, illustrating the fractions of individuals rating texts with coherence values ψ. The responses for ST1 and JUB exhibit a remarkably similar pattern. Approximately 60% of participants rated these texts as coherent or very coherent (grades 4 and 5), around 25% assigned an intermediate coherence level (grade 3), and a minority (less than 15%) ranked the texts with a low level of

Figure 14 – Distributions of complexity ratings



Source: Prepared by the author.
Distributions of complexity ratings among survey participants for all read texts. The values $\pi = 1, 2, 3, 4, 5$ represent a scale from "very simple" ($\pi = 1$) to “very complex” ($\pi = 5$).

Figure 15 – Distributions of coherence ratings



Source: Prepared by the author.
Distributions of coherence ratings among survey participants for all examined texts. The values $\psi = 1, 2, 3, 4, 5$ represent a scale encompassing perceptions of text coherence, ranging from "not coherent" ($\psi = 1$) to “very coherent” ($\psi = 5$).

coherence (grades 1 and 2). Conversely, for the random texts RT1 and RT2, the majority of ratings lean toward "not coherent" (approximately 50%), with about 20% indicating an intermediate coherence level, and just over 15% ranking them as coherent or very coherent. Notably, the GSV text yielded a distinctive response, with the largest group of respondents (approximately 40%) assigning an intermediate grade of 3. This suggests a notable indecision among participants regarding the coherence of the GSV text.

Table 5 – Average Complexity and coherence ratings across surveyed texts

Text	$\langle\pi\rangle$	$\langle\psi\rangle$
GAU	3.74 ± 0.08	3.14 ± 0.08
GSV	3.58 ± 0.08	2.95 ± 0.09
HCL	2.85 ± 0.09	3.70 ± 0.09
JUB	2.57 ± 0.09	3.79 ± 0.08
MEL	3.01 ± 0.09	3.62 ± 0.07
QUI	2.74 ± 0.09	3.86 ± 0.09
RT1	4.02 ± 0.08	2.44 ± 0.09
RT2	3.92 ± 0.09	2.38 ± 0.09
ST1	2.22 ± 0.09	3.80 ± 0.08
ST2	2.35 ± 0.09	4.00 ± 0.09

Source: Prepared by the author.
Mean values of complexity $\langle\pi\rangle$ and coherence $\langle\psi\rangle$ calculated for all texts in the survey.

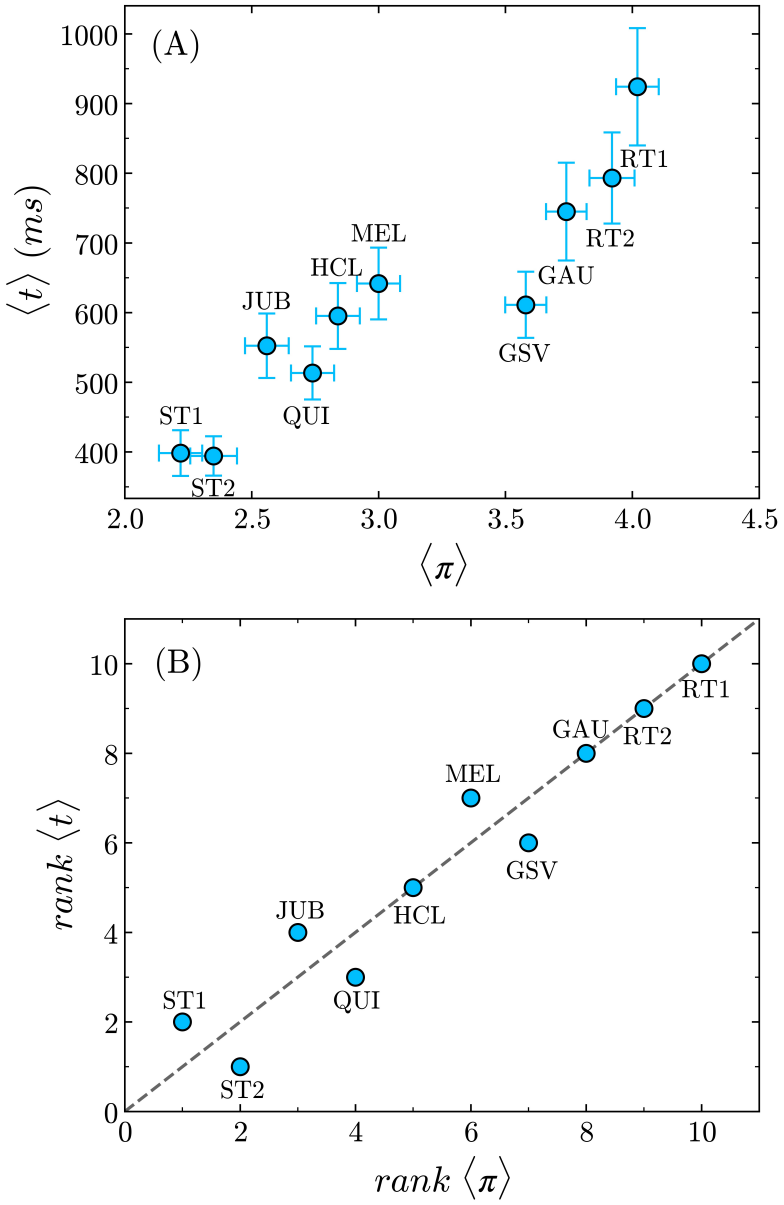
As indicated in Table 5, coherence mean values ($\langle\psi\rangle$) from the survey categorize texts into three groups. RT1 and RT2 exhibit low coherence ($\langle\psi\rangle < 2.75$), while ST1, ST2, JUB, HCL, MEL, and QUI yield high coherence ratings ($\langle\psi\rangle > 3.25$). GAU and GSV texts fall into an intermediate coherence range ($2.75 \leq \langle\psi\rangle \leq 3.25$).

2.6 Results

2.6.1 The average magnetization of the fixation activity reflects the level of text complexity

As previously emphasized, the reading time span holds significance as a measure of text processing, establishing a logical correlation with text complexity. In Fig. 16 A, we illustrate the average reading times per word against average complexity values. While $\langle t\rangle$ generally increases with $\langle\pi\rangle$, the relationship is not strictly monotonic. This is evidenced in Fig. 16 B, where the relative ranks of these variables are plotted against each other. Despite occasional

Figure 16 – Reading times against text complexity

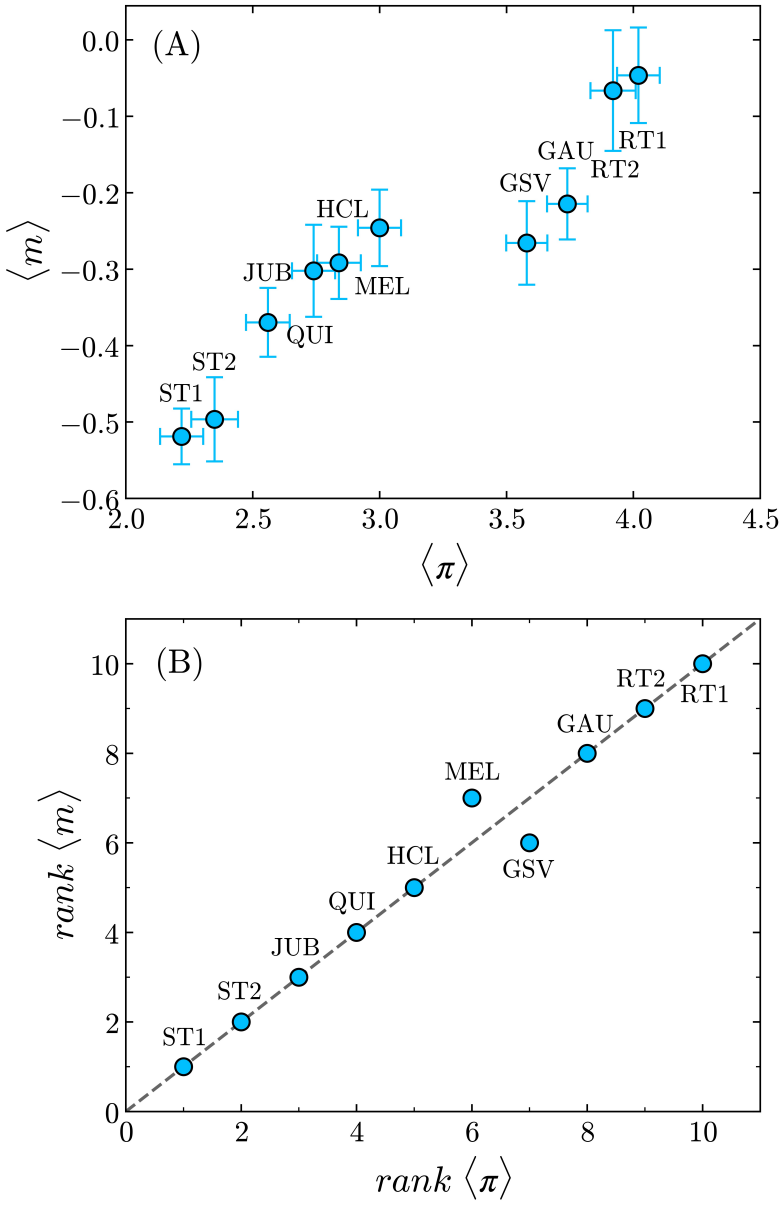


Source: Prepared by the author.
(A) The average reading time per word ($\langle t \rangle$) exhibits a general increase with $\langle \pi \rangle$, with a non-monotonic trend. (B) The plot of the rank of $\langle t \rangle$ against the rank of $\langle \pi \rangle$ reveals several discordant pairs. The dashed line represents the function $y = x$. The Kendall rank correlation coefficient is $\tau = 0.87$ ($p = 0.0001$).

discordant pairs in rank order among the ten comparable pairs, a non-parametric statistic, the Kendall rank correlation coefficient τ, reveals a strong correlation of $\tau = 0.87$ ($p = 0.0001$), signifying a reasonably high degree of positive monotonicity between $\langle t \rangle$ and $\langle \pi \rangle$ for the texts (see Appendix B). The relationship between the average magnetization ($\langle m \rangle$) and average text complexity ($\langle \pi \rangle$) is presented in Fig. 17 A. The two measures exhibit a highly correlated, albeit nonlinear, association, with $\langle m \rangle$ generally increasing with $\langle \pi \rangle$. An exception is a capricious local

minimum observed at the average complexity of GSV. In Fig. 17 B, the relative ranks of $\langle m \rangle$ and $\langle \pi \rangle$ for each text are plotted against each other, revealing that eight out of ten texts share identical positions in both lists. Comparing with reading time per word, the higher Kendall rank correlation coefficient in this instance ($\tau = 0.96$, $p = 5 \times 10^{-6}$) verifies that $\langle m \rangle$ serves as a more robust proxy for ranking text complexities ($\langle \pi \rangle$).

Figure 17 – Average magnetization against text complexity



Source: Prepared by the author.
(A) The average magnetization, $\langle m \rangle$, of fixation activities shows an almost monotonic increase with $\langle \pi \rangle$, except for a local minimum observed at the complexity of GSV. (B) Ranking both measures in ascending order and plotting the ranks of $\langle m \rangle$ against the ranks of $\langle \pi \rangle$ for all texts reveals that eight out of ten texts share identical positions in the two lists. The dashed line represents the $y = x$ function. The Kendall rank correlation coefficient $\tau = 0.96$ ($p = 5 \times 10^{-6}$) signifies a high degree of monotonicity between the two variables.

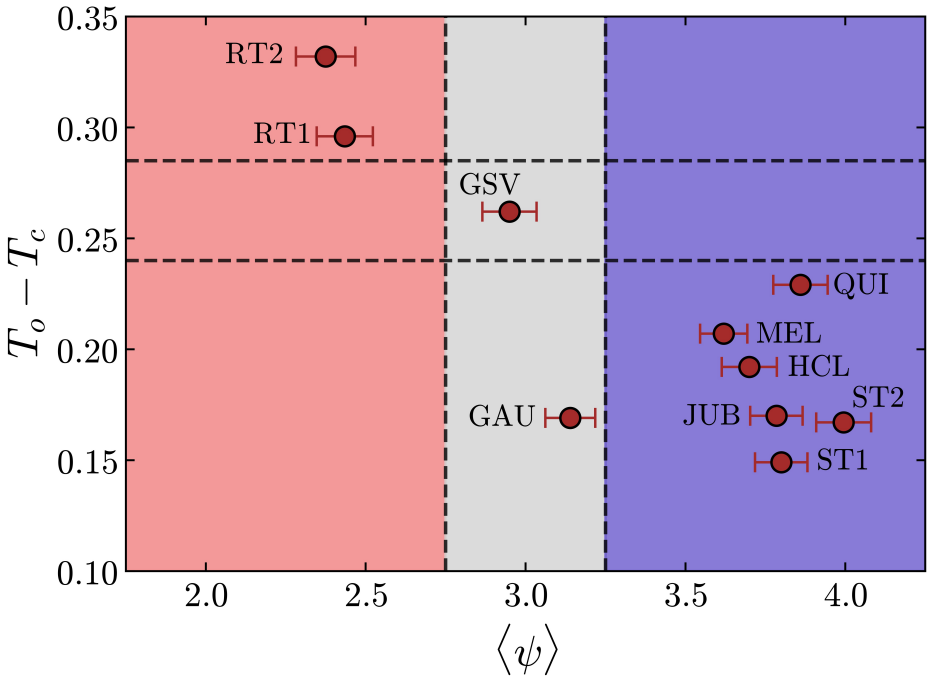
While one might anticipate that more complex texts would necessitate longer analysis times and, consequently, more fixations, an examination of average reading times per word reveals nuances. For instance, the HCL text, despite being similarly complex to the QUI text according to the survey, takes approximately 80 milliseconds longer per word on average, reflecting a 15% difference. A similar trend is observed for the GSV and GAU texts, with subjects spending an additional 133 milliseconds per word on average for the latter (a 20% difference), even though their relative complexity difference is under 5% (refer to Table 5). A comparative analysis between Figs. 16 B and 17 B, along with their corresponding Kendall coefficients, unequivocally suggests that average magnetization serves as a more reliable indicator of perceived complexity than reading time per word.

2.6.2 Text coherence perception evidenced by distance to criticality

The results depicted in Fig. 18 demonstrate a consistent correlation between large distances to criticality $(T_o - T_c)$ and the low-coherence levels ($\langle\psi\rangle < 2.75$) of both random texts (RT1 and RT2). Additionally, texts characterized by high coherence $\langle\psi\rangle > 3.25$ (ST1, ST2, JUB, HCL, MEL, and QUI) cluster in the bottom-left corner of the plot, exhibiting correspondingly small values of $(T_o - T_c)$. This alignment suggests that coherent texts elicit more correlated responses in readers' fixation activity, implying implicit cohesive interactions among them. Notably, the two texts rated with intermediate coherence values $2.75 < \langle\psi\rangle < 3.25$ in the survey (GAU and GSV) elicit divergent responses in terms of the distance to criticality obtained from the eye-tracking readings. To comprehend this disparity, a more nuanced analysis is necessitated, focusing on the literary styles and linguistic aspects of the source books from which these texts have been extracted, as elucidated in section 2.7.

In interpreting the observed numerical distinctions in $(T_o - T_c)$, it is important to emphasize that reliance is solely on the specific characteristics of cross-correlations derived from fixation activity series between pairs of readers for a given text (refer to Fig. 12). To test this hypothesis, additional calculations were conducted on fixation activities for a given text while preserving the mean magnetization $\langle\sigma_i\rangle$. The shuffling of σ_i values among randomly selected word pairs in the text was implemented to eliminate strong correlations, if present. Following data shuffling, the same sequence of calculations ensued: pairwise correlation determination C_{ij}, computation of fields h_i and couplings J_{ij}, and assessment of heat capacity C_v at varying temperatures T. The impact of suppressing strong correlations is a notable reduction in interactions J_{ij},

Figure 18 – Distance to criticality and text coherence



Source: Prepared by the author.
Correlation between the distance to criticality $T_o - T_c$ and the average coherence $\langle\psi\rangle$ of the texts. Texts with low coherence $\langle\psi\rangle < 2.75$ exhibit elevated values of $T_o - T_c$ (RT1 and RT2). Conversely, coherent texts $\langle\psi\rangle > 3.25$ (ST1, ST2, JUB, HCL, MEL, QUI) are situated near criticality, implying a cohesive reading response among individuals. Interestingly, texts with intermediate $\langle\psi\rangle$ values (GAU and GSV) elicit diverse responses in terms of $T_o - T_c$.

consequently lowering the T_c values for all texts and augmenting their $(T_o - T_c)$ distances (see Appendix C)

2.7 Discussion

Linguists have extensively explored text complexity and coherence. Recent attempts to quantify these aspects involve readability formulas, employing metrics such as word and sentence lengths and word frequency. However, the adequacy of these formulas in determining text complexity is debatable for two main reasons. Firstly, straightforward texts may employ numerous infrequent words, as exemplified in (Rothman, 2012): “Any text on raccoons would use ‘raccoon’ a lot, as well as ‘nocturnal’ and ‘foraging’ ”. This scenario illustrates a common pitfall wherein most readability expressions may erroneously overestimate the complexity score of the text. Second, texts conveying intricate concepts may employ simple vocabulary and feature concise sentence structures. This occurs, for instance, in texts featuring abstract narratives, metaphors, and obscure allusions, resulting in potential undervaluation by complexity scores.

A well known example in the literature is Ernest Hemingway's "The Old Man and the Sea", often underrated in complexity by readability formulas despite its depth and narrative richness. Consequently, alongside quantitative metrics, linguists frequently rely on qualitative analyses for a more nuanced categorization of reading material.

In our study, instead of applying empirical mathematical expressions, we opted for a data-driven approach to quantify text complexity and coherence. Leveraging results from a comprehensive survey involving 400 participants, we compared these findings with properties derived from eye-tracking experiments. The calculated magnetizations, a fundamental metric representing fixation density, were employed. Our hypothesis, confirmed through the survey, posits that text complexity correlates positively with average magnetization. These findings suggest that the fixation activity, as defined here, encapsulates sufficient cognitive information to describe reading patterns in terms of active and inactive states. The chosen threshold for the number of fixations on a given word r ($n_i^r = 2$) succeeded in delimiting the minimum criteria for characterizing words in the text requiring deeper cognitive processing. This threshold not only accommodates the impact of repetitive fixation influenced by word size and low frequency but also normalizes variations in fixation patterns accounting for different reading skills among the participants and their ability to predict word occurrences while reading. This assertion was validated by testing larger threshold values. Even at a threshold of 3, fixation activities become markedly influenced by word size and frequency, introducing substantial dissimilarities among subjects. In this context, our selected threshold emerges as a straightforward yet remarkably effective approach for mitigating inter-subject differences.

The observation of low coherence ratings for the randomly generated texts (RT1 and RT2), contrasting with the generally perceived coherence of literary texts in the survey, was expected. The intermediate coherence ratings for the literary texts GSV and GAU, however, are worth of nuanced analysis. Guimarães Rosa, the acclaimed Brazilian author of "Grande Sertão: Veredas" (GSV), from which the fragment GSV was extracted, is renowned for a distinctive writing style often likened to James Joyce's, particularly in terms of linguistic innovation and experimentation (Almino, 2018). In Rosa's literary works, we often find unconventional punctuation and grammar in story-rich writing, as well as created neologisms from erudite and popular expressions, regionalisms, archaic words and inventive use of prefixes and suffixes (Almino, 2018; Silviano, 2017). In fact, several linguistic features like these are found in the GSV excerpt used in this study, with which the book opens. The very first word,

"nonada", though existing in Portuguese, is archaic and hardly used in literature. Its usage, even in context, remains enigmatic, leading to different interpretations over the years (Zilly, 2017; de Castro, 1970; de Castro, 1976). The second sentence exhibits a non-traditional syntactic structure, reminiscent of regional orality, as we learn from Garcia's study (García, 2015). We observe this linguistic feature reoccurring in the fifth and fifteenth sentences. Furthermore, the author employs incomplete suggestive expressions in three sentences (second, fifth and seventh sentences), a linguistic construction characteristic of Rosa's writing. In addition to this, the text introduces two neologisms coined by the author, namely, "erroso" and "prascovio", appearing in the eighth and thirteenth sentences, respectively. Without delving into a more exhaustive analysis, it is reasonable to assert that the GSV fragment deviates from conventional literary norms, posing challenges for comprehension, particularly when isolated from the broader context of the book's narrative. Consequently, readers may experience confusion and uncertainty, encountering difficulty in categorizing the narrative as coherent.

The text GAU, in contrast, is a transcription from the 1870 novel "O Gaúcho" by José de Alencar. This excerpt is drawn from the end of chapter one, specifically describing the story's setting. The writing style is marked by a profound, philosophical portrayal of the scenario (Araripe Júnior, circa 1880). The narrative's abstract tone might impart a sense of vagueness, potentially leading to uncertainty in evaluating its coherence. Unlike GSV and GAU, other literary excerpts employed, such as HCL and MEL, adopt descriptive, straightforward writing, while QUI, JUB, ST1, and ST2 employ linear, plain storytelling-attributes. This divergence in style likely contributes to their ease of interpretation and, consequently, their perceived coherence.

The outcomes derived from a straightforward statistical model and the Pairwise Maximum-Entropy analysis reveal the effectiveness of the critical point distance in segregating the texts into three main groups. The random generated texts (RT1 and RT2) exhibit the greatest distance from the critical point, characterized by operating temperatures T_o significantly higher than T_c. Following closely is the GSV text, while the remaining texts cluster much nearer to T_c. In line with our prior discussions on critical phenomena in physics, when $T_o > T_c$, weak interactions prevail, leading to a disordered system. Our findings indicate that fixation activities for texts with low coherence (RT1 and RT2) exhibit a certain degree of randomness, indicating that the reading response to these stimuli does not promote robust virtual connections among individuals. When T_o approaches T_c, heightened "interactions" among elements facilitate the propagation

of local effects throughout the entire system. This phenomenon renders the system susceptible to global changes, potentially leading to the emergence of collective behavior. We observe this effect in texts rated with high coherence levels (ST1, ST2, JUB, HCL, MEL, and QUI), as well as in the GAU text, which, despite receiving an intermediate coherence rating, exhibits similar characteristics. We reason that a high degree of coherence in a text likely induces a cohesive reading response, as evidenced by the proximity to its critical point. While readers do not directly interact, their cognitive responses exhibit similarity when confronted with consistent text content. An intriguing question arises regarding the ambiguous relationship between the average coherence rating $\langle\psi\rangle$ and $T_o - T_c$ for the GAU and GSV texts. Despite GAU aligning with the cluster of texts near T_c and GSV being distant from it, both texts received average coherence ratings indicating intermediate coherence levels. In our earlier discussion, we elaborated on these texts characteristics, highlighting linguistic features that set them apart from other literary fragments examined in this study. Specifically, in the case of the GAU excerpt, readers engage with an intelligible yet abstractly styled text, potentially leading to ambiguous interpretations. While assigning a precise coherence value to such text might be challenging, it is reasonable to assert that the content remains internally coherent, meaning its narrative maintains consistency. In contrast, the GSV text deviates from typical readability due to its highly technical language and unconventional linguistic elements. It can be regarded as an intermediate text, bridging the gap between a concrete narrative and a randomly incongruous one. Considering this and the low $T_o - T_c$ value induced by GAU, the results depicted in Fig. 18 suggest that the distance to the critical point effectively segregates texts based on a coherence measure. This measure, rooted in an inherent cognitive mechanism, may be less susceptible to extrinsic influences compared to questionnaire responses in a digital survey protocol.

3 VISUAL ATTENTION DYNAMICS IN MEDICAL IMAGE INTERPRETATION: A STUDY OF CYTOPATHOLOGISTS USING EYE TRACKING

3.1 Navigating visual attention in medical diagnosis

Medical imaging specialists play a crucial role in diagnosing anomalies or potential health concerns within diagnostic images such as X-rays, MRIs, and tissue samples. Despite the advancements in medical screening, diagnoses can be compromised by factors such as image limitations, poor resolution, and human bias. Concerns about misdiagnoses, particularly false positives, add another layer of complexity. Several countries implement cancer screening programs, addressing the imperative need for early detection, such as in lung and breast cancer. [1] While guidelines aim to standardize image interpretation, variations in diagnostic accuracy persist (McCreadie e Oliver, 2009; Lee *et al.*, 2013). False positives lead to patient anxiety and unnecessary procedures, while false negatives increase the risk of delayed treatment. Over-diagnosis and misdiagnosis also impose substantial financial burdens on states.

The integration of eye-tracking technology in medical imaging offers a unique perspective into the cognitive processes of these specialists, enabling precise recording and analysis of their eye movements and fixations during image evaluation (van der Gijp *et al.*, 2017; Harezlak e Kasprowski, 2018; Wu e Wolfe, 2019). Analyzing eye-tracking data provides insights into the cognitive processes underlying the interpretation of medical images, potentially uncovering areas of common attention or critical details that specialists may overlook. This understanding not only optimizes the design of medical imaging technology but also improves specialists' training, ultimately enhancing the accuracy and efficiency of medical diagnoses. The healthcare field recognizes the potential of eye-tracking technology to enhance medical training, diagnostic precision, and patient care. Once confined to specialized labs, recent trends have democratized eye tracking with affordable, user-friendly devices. Moreover, the integration of machine learning models with eye-tracking data analysis holds promise in aiding medical image interpretation by identifying abnormalities, parsing images into diagnostic Regions of Interest (ROIs), and providing real-time feedback to diagnosticians.

[1] Lung cancer is the most common type of cancer among men, and second most common among women, following breast cancer (Forman e Ferlay, 2014). These are also leading causes of cancer-related death worldwide, emphasizing the urgency for health organizations to develop prevention and early-detection strategies. Major medical organizations recommend screening programs in high-risk populations with the aim to identify the disease at earlier stages, when treatment can be more successful (von Karsa *et al.*, 2014a; von Karsa *et al.*, 2014b).

In collaboration with the postgraduate program "Engenheria em Teleinformatica" at UFC University, this study employs eye-tracking to examine the visual attention of proficient cytopathologists specializing in cervical care while assessing Pap smear slides. The chapter centers on the analysis of eye-tracking data collected from these experiments. The subsequent translation of eye fixations into binary matrices and the generation of attention heat maps lay the foundation for studying visual attention during the analysis of cervical cell images. The exploration extends to the evaluation of the consistency in visual focus among participants, a detailed examination of center bias within the CRICVA database, and a comprehensive review of saliency models. These efforts provide insights into predictive accuracy and the potential implications of saliency models for medical image interpretation. Addressing the work in (Ferreira *et al.*, 2019b), this study is guided by two fundamental hypotheses. The first hypothesis (H1) suggests that bottom-up methods, designed to simulate low-level human visual attention, can yield results akin to predictions derived from expert visual attention. The second hypothesis (H2) posits that contemporary saliency models exhibit effectiveness in detecting clinically relevant ROIs within Pap smear images.

Ultimately, this study conducts a thorough review of image interpretation and analysis of eye-tracking data in medical imaging. By assessing the alignment between computational models and expert visual attention, the chapter contributes to the discourse on integrating eye tracking and machine learning in complex medical image interpretation.

3.1.1 The cognitive foundations of image interpretation: examining visual attention

Visual perception is a fundamental process that underlies our ability to understand and interpret the visual information we encounter in our environment. This intricate process begins with the reception of visual stimuli through our eyes, where light is converted into neural signals. These signals are then meticulously processed within the brain to create a coherent and meaningful representation of the world around us. Central to this process is the concept of visual attention, which pertains to the selective allocation of cognitive resources to specific aspects of the visual input. Although the term “attention” is widely used, it is not always explicitly defined. Attention serves as our mental filter, enabling us to navigate the vast stream of visual information by prioritizing certain elements while disregarding others (Carrasco, 2011). This selective mechanism transforms the act of looking into the meaningful act of seeing by choosing relevant information and dismissing irrelevant noise. Our understanding of visual attention has evolved

significantly, thanks to a range of research methods. Psychophysical studies involving humans, neurophysiological investigations in primates, neuroimaging techniques like fMRI and ERPs, eye-tracking technology, and computational models have all played important roles in characterizing the relationship between attention and perception. For example, neurophysiological research has identified distinct neural networks and specific brain areas associated with various aspects of attention, such as alerting, orienting, and executive control.

Visual perception is influenced by both top-down and bottom-up processes (Zhang e Lin, 2013). Top-down attention is a cognitive process guided by our goals, expectations, and prior knowledge, where we direct our attention to specific areas or features based on our intentions. In the context of eye-tracking experiments, this may be related to the tasks participants are assigned and the instructions they receive. Conversely, bottom-up attention is driven by the inherent properties of visual stimuli, capturing our attention through salient or novel features, often without conscious control. In eye-tracking research, this can manifest as sudden shifts in fixation points due to unexpected, attention-grabbing stimuli. Frequently, visual perception and attention operate sequentially. Attention plays a pivotal role in directing the allocation of cognitive resources during perception, as observed in eye-tracking experiments, where participants initially perceive an image and then focus their attention on specific regions or objects within it.

Importantly, it is observed that the impact of attention on visual processing intensifies as we ascend the hierarchy of cortical visual areas in the brain, thereby influencing conscious perception (Carrasco, 2011). Attention can amplify the processing of attended stimuli, rendering them more vivid and memorable, while diminishing the significance of unattended stimuli. This selective perception is a fundamental aspect of visual cognition, exemplifying how attention dynamically shapes our understanding of specific images, a significant focus in research.

3.1.2 Integration of eye-tracking data in medical image analysis: enhancing accuracy and interpretability

Initial research in eye-tracking for medical image interpretation primarily focused on examining gaze patterns among both experts and novices as they analyzed complex images, such as X-rays, histology slides, and computed tomography (CT) scans. For instance, (Manning *et al.*, 2006) conducted a study that assessed radiologists with varying expertise in detecting pulmonary nodules on chest X-rays and observed distinct differences in search strategies between experienced radiologists and novices. Performance metrics, such as fixation count and scrutiny

time per decision per image, indicated higher values for experienced radiologists, suggesting potential diagnostic challenges. A more detailed study by (Lévêque *et al.*, 2019) focused on radiologists analyzing mammogram images for breast lesions. They evaluated and quantified the spatial gaze distribution among different groups and observed deviations in the gaze patterns of trainees compared to expert radiologists. Furthermore, trainees exhibited significantly longer mean fixation durations, which may indicate increased uncertainty or difficulty in interpreting specific image regions (Brunyé *et al.*, 2019). These studies laid the foundation for understanding how medical professionals visually analyze images and uncovering potential sources of diagnostic errors. Researchers also explored whether specialists follow consistent gaze patterns or employ unique approaches when analyzing images. They analyzed transition behaviors, including measures like saccade amplitude and direction, to assess how attention shifts from one image area to another, potentially revealing predetermined scanning strategies or areas of heightened concern.

In general, eye-tracking studies have identified three types of diagnostic interpretation errors: search errors, recognition errors, and decision errors (Brunyé *et al.*, 2019). Search errors occur when critical regions or diagnostic features are not fixated, often due to factors like low target prevalence, satisfaction of search, or low visual salience (Wolfe e van Wert, 2010; Berbaum *et al.*, 2001; Itti e Koch, 2000). Recognition errors result from an inability to correctly identify or recognize critical information, often reflected in fixation durations (Mello-Thoms *et al.*, 2005). Decision errors occur when misjudgments or incorrect conclusions are made based on the observed information. Decision errors, which can be challenging to detect through existing eye-tracking metrics, may be associated with disorganized eye movements over the image (Fabio *et al.*, 2015). Experienced diagnosticians are generally less prone to making interpretation errors compared to novices. They develop refined search strategies and richer knowledge, guiding their attention towards diagnostically relevant regions and potentially reducing search errors (Kundel e La Follette Jr, 1972). For example, studies involving neurologists examining brain CT scans for cerebrovascular accidents and pathologists interpreting digitized breast biopsy images revealed that novices tend to fixate on visually salient but diagnostically irrelevant regions, potentially leading to search errors (Matsumoto *et al.*, 2011; Brunyé *et al.*, 2014).

The insights gained from studies like the ones mentioned have prompted the integration of specialists' eye-tracking data into medical image analysis workflows employing machine learning models. Machine learning algorithms aid in extracting critical information from medical

images through a multi-step process (Litjens *et al.*, 2017). It begins with the collection and preprocessing of medical image data, involving tasks like noise reduction, contrast enhancement, and geometric correction. After preprocessing, the next critical step is feature extraction. Features are distinctive patterns or characteristics within the medical images, such as texture, shape descriptors and pixel intensity values. Traditionally, feature extraction required expert knowledge and was a time-consuming process. However, deep learning models can automatically extract relevant features through a process known as feature learning. Convolutional Neural Networks (CNNs), a class of deep learning models, have proven highly effective in this regard. CNNs are designed to mimic the human visual system by applying a series of convolutional and pooling layers to identify spatial hierarchies of features. In the context of medical image analysis, CNNs can recognize complex patterns, such as tumors, lesions, or anomalies, within images. Training a CNN involves providing it with labeled images and adjusting the model's internal parameters (weights and biases) to minimize the difference between the predicted output and the actual labels. Additionally, transfer learning is a valuable strategy where pre-trained CNN models are fine-tuned for medical image analysis tasks. This approach makes use of the learned features from a vast dataset to improve the performance on the specific medical imaging task of interest. To assess the performance of deep learning models, various metrics are used, including sensitivity, specificity, precision, and recall. Cross-validation and split datasets are employed to validate the models and ensure they generalize well to new, unseen data.

Deep learning models have been successfully applied to various medical imaging tasks, including the early detection of diseases, segmentation of organs or lesions, and even predicting patient outcomes. For instance, in the domain of cancer marker detection, deep learning architectures have emerged as powerful tools for detecting cancer markers in lung CT scans and mammograms, often demonstrating performance on par with or exceeding that of experienced radiologists (Ardila *et al.*, 2019; McKinney *et al.*, 2020). In ophthalmology, Optical Coherence Tomography (OCT) scans, consisting of three-dimensional views of retinal layers, have been subjected to innovative deep learning techniques for the detection of sight-threatening retinal diseases (de Fauw *et al.*, 2018). These models have excelled in providing referral recommendations that are comparable to expert assessments. Deep learning has also found application in dermatology, specifically for automated melanoma recognition in dermoscopy images, which offer high-resolution views of both the skin's surface and subsurface features. This approach has outperformed shallower models, enhancing both segmentation and classification tasks (Yu *et*

al., 2016). Furthermore, in cytology, deep learning was introduced for segmenting overlapping cervical cells in Pap smear images, yielding successful results (Song *et al.*, 2016). These examples highlight the diverse applications of machine learning in medical image analysis, addressing various tasks across different medical specialties and consistently achieving remarkable outcomes (for a comprehensive review of the literature in this topic, please refer to (Litjens *et al.*, 2017)).

Integrating experts' gaze data, which signifies regions of interest within images, holds the potential to enhance the learning efficacy of machine learning models. This gaze data is subjected to a process of feature engineering, where it is transformed into informative attributes, including fixation heatmaps and statistical measures that reflect fixation duration and frequency within specific regions of the image. This augmented eye-tracking dataset complements the original image data, enriching the information accessible to the machine learning model. The model, when trained on this combined dataset, learns to associate areas of expert fixation with pertinent features in medical images. This approach amplifies the model's accuracy in localizing features and enhances its interpretability. It becomes finely tuned to the regions of interest identified by experts, thereby improving its capacity to identify and analyze key features and anomalies within the images. Additionally, it offers interpretability by delineating the regions that influence its predictions, thus facilitating more transparent decision-making processes.

In a study by (Khosravan *et al.*, 2019), a common challenge in computer-aided diagnosis (CAD) systems was addressed, particularly the frequent occurrence of false positive findings and missed tumors in early-stage lung cancer screening. To enhance the efficiency and accuracy of the screening process, a novel algorithm was introduced that incorporates gaze data from radiologists during the screening process. This integration effectively reduces false positive findings while simultaneously enhancing true positive results, thus reducing the occurrence of previously overlooked cases. In a related study by (Stember *et al.*, 2019), the utility of eye tracking data was explored for generating segmentation masks, particularly in the fields of radiology and pathology. The research findings demonstrated that CNNs trained on eye-tracking-generated masks perform nearly as well as those trained on manually annotated masks, exhibiting a high degree of similarity between the two methods. These findings imply the potential for leveraging eye-tracking technology to enhance the efficiency of deep learning semantic segmentation in radiology workflows. Overall, the integration of advanced machine learning models with eye-tracking data offers substantial promise for streamlining medical image interpretation, potentially raising the level of clinical competence, and providing valuable support

for medical education and training.

3.1.3 Saliency prediction methods with eye-tracking data: the case of Pap smear analysis

To computationally model visual attention, researchers have developed saliency maps. These maps provide a topographical representation of the saliency, or conspicuousness, of individual pixels in an image. Various algorithms have been proposed and applied across diverse domains, including computer vision, robotics, and medical image analysis (Murabito *et al.*, 2018; Loukas *et al.*, 2018; Nguyen *et al.*, 2018). These models often fall into two distinct categories. Bottom-up approaches often focus on modeling low-level image features like color, contrast and orientation, and may employ cognitive concepts, probabilistic frameworks, spectral analysis, or other techniques to generate saliency maps (Borji e Itti, 2012). These models typically aim to identify image regions that stand out from their surroundings without necessarily addressing the cognitive phenomena that render these regions relevant. In contrast, top-down approaches aim to account for image semantics, taking into consideration the demands of specific tasks and expectations (Judd *et al.*, 2009).

The continuous evolution of deep learning techniques (Zhang *et al.*, 2018) and the increasing availability of large-scale annotated datasets (Winkler e Subramanian, 2013; Jiang *et al.*, 2015) have led to significant advancements in the field of saliency modeling, enabling models to perform end-to-end learning and achieve substantial improvements [14]. The introduction of models such as eDN (ensembles of Deep Networks) (Vig *et al.*, 2014), multiresolution CNN (Mr-CNN) (Liu *et al.*, 2016), SALICON (Saliency in Context) (Huang *et al.*, 2015), DeepFix (Kruthiventi *et al.*, 2017), and others has broadened the scope of saliency prediction, each offering novel approaches to the task.

While existing saliency prediction models have predominantly been focused on natural images, often representing everyday scenes (Borji e Itti, 2012; Borji *et al.*, 2015), there is a growing interest in the application of these models to medical image analysis, particularly within the domain of cell analysis (Matsumoto *et al.*, 2011; Li *et al.*, 2016; Lévêque *et al.*, 2018). However, the intersection of visual attention models and cell analysis remains relatively unexplored, particularly in the context of identifying relevant clinical areas within images, such as Pap smear slides (Zhang *et al.*, 2013). Such an approach could effectively guide the focus of classification, image compression, and other routine tasks. It could also address the current challenges associated with real-time Pap smear image analysis under adverse conditions,

including noise, artifacts, and occlusions (William *et al.*, 2018).

Prior research, such as the work of (Coombes e Culverhouse, 2003), has employed visual attention theory to analyze cells, utilizing eye-tracking devices to collect cytopathologists' visual data. Their findings suggest that saliency maps have the potential to reduce diagnostic disparities in routine cell screening. Furthermore, they underline the strong correlation between expert eye fixations and cell staining within the images. Additionally, (Zhang *et al.*, 2013) also explored bottom-up attention mechanisms for abnormal cervical cell detection but limited their study to liquid-based cytological images (Zhu *et al.*, 2007).

In contrast to prior approaches, such as those discussed in (Coombes e Culverhouse, 2003) and (Zhang *et al.*, 2013), this work investigates the feasibility of employing saliency prediction methods to support the screening of cervical cells within conventional Pap smears. Conventional Pap smears pose challenges primarily due to their subjective interpretation, which can lead to inter-observer variability. The reliance on visual examination introduces subjectivity and the potential for false positives and false negatives. Additionally, the quality of collected samples can vary, affecting the accuracy of the test. Nevertheless, they are widely used in numerous countries because they offer a cost-effective and well-established method for cervical cancer screening, contributing to reduced mortality rates. Existing healthcare infrastructure and expertise further support their widespread adoption, ensuring accessibility even in resource-limited settings.

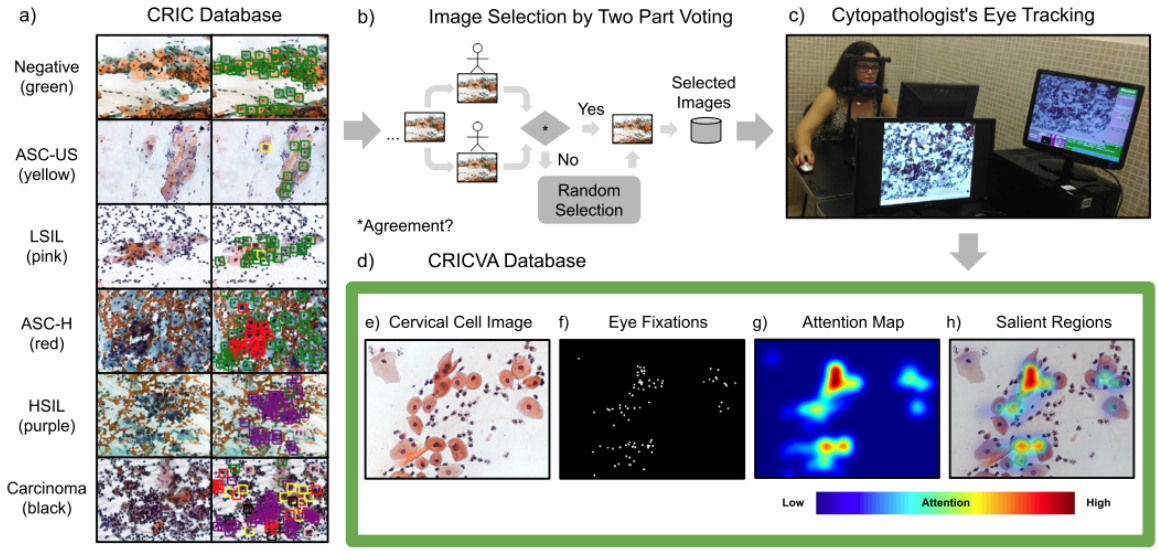
3.2 Experimental methodology: database structure and Eye-Tracking data collection

3.2.1 Databases: CRIC and the emerging CRICVA

The primary data source for this study is the Center for Recognition and Inspection of Cells (CRIC) database. This database contains digitized Pap smear images that were captured using a Carl Zeiss microscope equipped with a Zeiss AxioCam MRc camera, magnifying at 40×. These images have a pixel size of 0.255 μm and a resolution of 1392×1040 pixels at an 8-bit color depth. The dataset encompasses images of specimens prepared through conventional Pap smears, featuring cervical cells alongside common artifacts typically found in such tests. Within the CRIC dataset, cervical cells are classified as either normal or abnormal, with the latter further subdivided into specific pathological patterns, including Atypical Squamous Cells of Undermined Significance (ASC-US), Atypical Squamous Cells of High Significance (ASC-H),

Low-grade Squamous Intraepithelial Lesion (LSIL), High-grade Squamous Intraepithelial Lesion (HSIL), and Carcinoma (CA) (Nayar e Wilbur, 2015). For clarity, Fig. 19 a provides an overview of the organization of the CRIC dataset.

Figure 19 – CRICVA dataset creation overview



Source: Retrieved from (Ferreira *et al.*, 2019b).
Stages in CRICVA dataset construction: (a) CRIC dataset images featuring labeled cells. (b) Image selection methodology designed to mitigate selection bias. (c) Overview of the eye-tracking experiment capturing the gaze patterns of cytopathologists. (d) Visual attention data embedded in the CRICVA dataset. (e) RGB cervical cell image. (f) Binary map depicting cytopathologist's eye fixations. (g) Attention map generated by placing a Gaussian at each fixation position. (h) Overlay between the input image and attention map, accentuating salient regions. Dark red denotes the most conspicuous areas, while dark blue indicates less attention.

A new dataset named CRICVA (CRIC Visual Attention) was introduced, consisting of 232 images selected from the CRIC dataset (Ferreira *et al.*, 2019a), as depicted in Fig. 19 d. The curation of this dataset adhered to specific criteria, ensuring that images displayed a uniform distribution of cells across the entire visual field and featured diverse artifacts, such as blood cells, inflammatory cells, distorted cells, overlapping objects, and mucus. In addition to this, the selection process involved a two-party voting system and the random inclusion of 80 images from the remaining images in the CRIC database, as illustrated in Fig. 19 b. The inclusion of random images was aimed at mitigating potential selection bias in the voting process, specifically any bias that specialists might exhibit towards certain types of images during the selection process.

3.2.2 Collecting visual attention data

We conducted a comprehensive eye-tracking study involving three experienced cytopathologists, specialists in cervical care and interpretation of conventional Pap smear slides, to investigate their visual attention during the analysis of cervical cell images. The participants, with varying degrees of expertise, held either Ph.D. or Master's degrees in cytopathology, with ages ranging from 27 to 46 years and corresponding career lengths of 3 to 20 years in cytopathology laboratories. All participants exhibited normal or corrected-to-normal vision using corrective lens glasses. Prior to their involvement, participants provided informed written consent, and the study received approval from the Ethics Committee of Universidade Federal do Ceará.

The eye-tracking experiments were conducted using the EyeLink 1000 eye tracker, introduced in Section 1.3, operating at a sampling rate of 1000 Hz and recording data from the participants' right eye (Fig. 19 c). We employed the SR Research Experiment Builder V4 software to set up the experiments, consisting of eight trials, each with a different number of images for the experts to analyze. Trials intentionally varied the number of images (26 for the first two, 25 for trials 3 to 6, and 40 for trials 7 to 8) to capture fluctuations in eye-tracking data reflective of the cytopathologists' evolving screening performance over time. This deliberate variation facilitated training an algorithm on the CRICVA database to understand the nuances of performance fluctuations during sequential image analysis in cytological screening.

After calibration, participants were tasked with examining Pap smear images and identifying abnormal cells through mouse clicks. Participants had unrestricted time for image analysis, and mouse clicks were discreetly recorded. The cervical cell images, resized to 1280 × 1024 pixels while preserving the original aspect ratio, were presented in a randomized order to minimize potential bias. Participants advanced to the next image by pressing the space bar, with no option to revisit the previous one. Following each trial, participants had a break to relax, report any discomfort, and undergo re-calibration before starting the next trial.

The positions of eye fixations were translated into a binary matrix, as illustrated in Fig. 19 f. Subsequently, attention heat maps were generated (Fig. 19 g) by convolving the binary matrix with a two-dimensional isotropic Gaussian distribution characterized by a 1° visual angle spread. This choice approximates the foveal size (Torralba *et al.*, 2006; Le Meur e Baccino, 2013), providing a reference to how a person's visual attention is distributed around each fixation point. Salient regions, indicating areas of attention, were determined by superimposing the

experts' attention maps onto the input image, color-coded to represent attention levels (Fig. 19 h), where dark red denotes the most attention-attracting areas, and dark blue signifies the least conspicuous regions.

3.3 Visual attention analysis

3.3.1 Assessing consistency in participants' visual focus

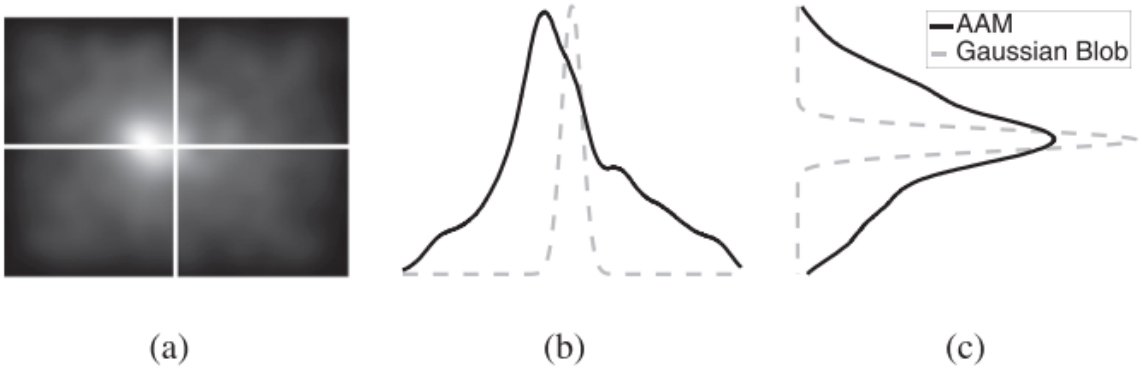
To evaluate gaze agreement among cytopathologists (Volokitin *et al.*, 2016; Bylinskii *et al.*, 2018), the Area Under the Receiver Operating Characteristic Curve (AUC-Judd) was utilized as a metric assessing the efficacy of binary classification models (Judd *et al.*, 2009). For each cytopathologist examining an image, a fixation map representing the spatial distribution of fixations was generated. These maps were then used to predict the fixations of one cytopathologist based on those of others. In this context, the AUC quantifies the alignment between predicted and actual fixations made by the targeted cytopathologist. The average AUC-Judd across all cytopathologists was calculated at 0.823 (SD=±0.007), with a p-value of 0.117 from a Kruskal-Wallis ANOVA test with $\alpha = 0.05$ (Kruskal e Wallis, 1952). This finding indicates a high level of gaze pattern agreement across all participants when considering the entire image dataset. Additionally, individual average AUC-Judd values were computed for each cytopathologist across different trials, yielding values of 0.818 (SD=±0.016), 0.813 (SD=±0.018), and 0.829 (SD=±0.016). These results suggest consistent fixation patterns for each cytopathologist across trials, supporting the hypothesis (referred to as H2) that well-defined clinical Regions of Interest (ROIs) exist in Pap smear images, as evidenced by the concordance in attention focus among cytopathologists.

3.3.2 Center bias analysis in CRICVA database

The presence of center bias, characterized by a propensity to direct attention towards the central region of an image, is a common phenomenon observed in both human visual perception and computer vision datasets (Tatler, 2007; Borji *et al.*, 2019). This bias is often exacerbated by the photographer's inclination to frame relevant objects within the central portion of an image (Borji *et al.*, 2015). To measure the extent of center bias within the CRICVA database, an analysis employing the Average Annotation Map (AAM) was conducted. The AAM offers the average of the ground-truth annotations representing visual attention across the entire image

dataset (Borji *et al.*, 2015) (see Fig. 20 a). To assess the dispersion of the AAM, the horizontal and vertical midlines were contrasted with those generated from a Gaussian blob centered at the image center, as illustrated in Figs. 20 b and 20 c, respectively. The Gaussian kernel applied for this analysis had a spread of 1° of visual angle. Although the AAM exhibits heightened activation near the image center, a significant variation in the spatial distribution of activations is observed. This variation implies that cytopathologists dedicated considerable attention to objects located away from the central region of the image. This nuanced understanding of center bias in visual attention patterns enhances the interpretability of gaze distribution within the provided dataset.

Figure 20 – Center bias analysis in CRICVA dataset



Source: Retrieved from (Ferreira *et al.*, 2019b).
Illustration of the center bias analysis conducted on a CRICVA image: (a) Average Annotation Map (AAM) representing the spatial distribution of ground-truth visual attention. Comparison of (b) horizontal and (c) vertical mid-lines of the AAM with those derived from a Gaussian blob positioned at the image center

3.4 Review of saliency models and key research findings

While the detailed technical implementation of saliency models falls beyond the scope of this thesis, this section provides an overview of the diverse approaches investigated in (Ferreira *et al.*, 2019b) concerning the study of saliency in cervical cell images. Specifically, we introduce the surveyed saliency models, offering a succinct explanation of their methodologies, and highlight the primary research outcomes. These findings serve to elucidate how eye-tracking data can enhance the development of machine learning methods in the domain of medical image interpretation and contribute to a broader understanding of the functionality of learning models.

3.4.1 Surveyed saliency models

A comprehensive evaluation of both bottom-up and top-down saliency prediction models (Borji e Itti, 2012) was undertaken. The top-down methods included two variations of CNN-based models and an additional model proposed by Zhang et al. (Zhang *et al.*, 2013), specifically tailored for exploring top-down factors in the context of cervical cell images. A concise summary of these methods is presented in Table 6, accompanied by designated model abbreviations for clarity, method description, year of publication, model category (Top-down or Bottom-up), and benchmark. All algorithms, except for the one proposed by (Zhang *et al.*, 2013), underwent validation using publicly accessible datasets. The inclusion criteria for this study were meticulously defined as follows: (1) the input comprises a single image; (2) the source code is publicly available; (3) demonstrated high performance in recognized saliency ranking lists; (4) execution time is within three seconds per image; and (5) the algorithms represent the current state-of-the-art or serve as benchmarks in the existing literature. Adhering to these criteria, the MIT Saliency Benchmark (Bylinskii *et al.*, 2015) was consulted, and after ranking results by Normalized Scanpath Saliency (NSS), the top-performing CNN-based algorithms, SAMv and SAMr (Cornia *et al.*, 2018), were selected. Additionally, five non-deep models (BMS (Zhang e Sclaroff, 2013), LDS (Fang *et al.*, 2016), FES (Tavakoli *et al.*, 2011), SWD (Duan *et al.*, 2011), and UHM (Tavakoli e Laaksonen, 2017)) were chosen. The study was complemented with IT (Itti *et al.*, 1998), GB (Harel *et al.*, 2006), SR (Hou e Zhang, 2007), SS (Hou *et al.*, 2011), and SIM (Murray *et al.*, 2011) bottom-up models from the extensive benchmark introduced by Borji et al. (Borji *et al.*, 2015). For a thorough examination of the implemented methods' structure, consult Annex C.

Table 6 – Surveyed saliency prediction models

#	Literature	Abbrev.	Description	Year	Cat.	Ben.
1	SAMv (Saliency Attentive Model - VGG-16)	SVS, SVC	Employs the VGG-16 CNN as a backbone and integrates an Attentive Convolutional Long Short-Term Memory network (Attentive ConvLSTM) for eye fixation prediction in images. It includes a center prior component to learn the database's inherent center bias. (Cornia *et al.*, 2018)	2018	T	
2	SAMr (Saliency Attentive Model - ResNet-50)	SRS, SRC	Similar to the SAMv model, employing the ResNet-50 network as the backbone (Cornia *et al.*, 2018).	2018	T	
3	BMS (Boolean Map based Saliency)	BMS	Characterizes image using binary maps via random thresholding in a whitened feature space. Employs topological analysis of Boolean maps to infer surrounding regions and estimate saliency. (Zhang e Sclaroff, 2013)	2013	B	
4	LDS (Learning Discriminative Subspaces)	LDS	The saliency map is computed through the learning of discriminative subspaces optimized for target highlighting and artifact suppression. LDS generates candidate subspaces using principal component analysis. (Fang *et al.*, 2016)	2017	B	MIT
5	FES (Fast and Efficient Saliency)	FES	Utilizes a center-surround approach within a Bayesian framework to estimate local feature contrast saliency. Probability distributions are derived through sparse sampling and kernel density estimation. (Tavakoli *et al.*, 2011)	2011	B	
6	SWD (Spatially Weighted Dissimilarity Saliency)	SWD	Integrates dissimilarities and spatial distance among image patches, incorporating a weighting mechanism to address center bias. Spatial distance, considering corresponding dissimilarities, utilizes principal component analysis for dimension reduction. (Duan *et al.*, 2011)	2011	B	

Continued on the next page.

Table 6 – Surveyed saliency prediction models

#	Literature	Abbrev.	Description	Year	Cat.	Ben.
7	UHM (Unsupervised Hierarchical Models)	UHM	The unsupervised multi-scale hierarchical saliency model utilizes independent subspace analysis (ISA), akin to a two-layer neural architecture. It explores both local and global saliency concepts, obtaining a hierarchical representation of the input by stacking ISA networks together, reminiscent of deep models. (Tavakoli e Laaksonen, 2017)	2016	B	
8	IT (Itti's Saliency Model). Implementation by (Harel *et al.*, 2006)	IT	Pioneering purely bottom-up saliency prediction model. Extracts low-level features from local center-surround differences in intensity, color, and orientation across multiple scales. Fusion and normalization of these maps result in three conspicuity maps, combined to form the saliency map. (Itti *et al.*, 1998)	1998	B	
9	GB (Graph-Based Visual Saliency)	GB	The model, akin to IT, extracts low-level features and constructs a fully connected graph using a Markov chain, connecting all grid locations for each feature map. Node weights are based on feature dissimilarity and spatial distance, informing the saliency map through equilibrium distribution. (Harel *et al.*, 2006)	2016	B	
10	SR (Spectral Residual Approach)	SR	Independent of features or prior object knowledge, this approach estimates saliency by exploring background properties. It evaluates the log-spectrum, extracts spectral residual, and transforms it into the spatial domain to generate the saliency map. (Hou e Zhang, 2007)	2007	B	(Borji *et al.*, 2015)
11	SS (Sparse Salient Regions)	SS	Utilizes the sign function of the Discrete Cosine Transform (DCT) on image data to create a signature, primarily capturing details related to the image foreground. This signature is then used by the algorithm to identify specific regions and generate saliency maps. (Hou *et al.*, 2011)	2012	B	

Continued on the next page.

Table 6 – Surveyed saliency prediction models

#	Literature	Abbrev.	Description	Year	Cat.	Ben.
12	SIM (Saliency by Induction Mechanisms)	SIC	The methodology replicates early human visual processing, incorporating color-opponent, luminance channels, and multiscale decomposition (Murray *et al.*, 2011). Employing the unified color induction model by Otazu et al. (Otazu *et al.*, 2010), it simulates visual cortex cell inhibition mechanisms and integrates information across multiple scales using an inverse wavelet transform.	2012	B	
13	ZH (Detection of Abnormal Nuclei in Cervical Smear Images (Zhang *et al.*, 2013))	ZH	The study explores bottom-up attention and a target-driven strategy for abnormal cell detection in liquid-based cervical smear images (Zhang *et al.*, 2013). The model extracts conspicuous regions based on direction and brightness features, modulating this information using an annular template matching model designed from abnormal nuclei statistics, with parameters set following (Bylinskii *et al.*, 2016).	2013	T	*

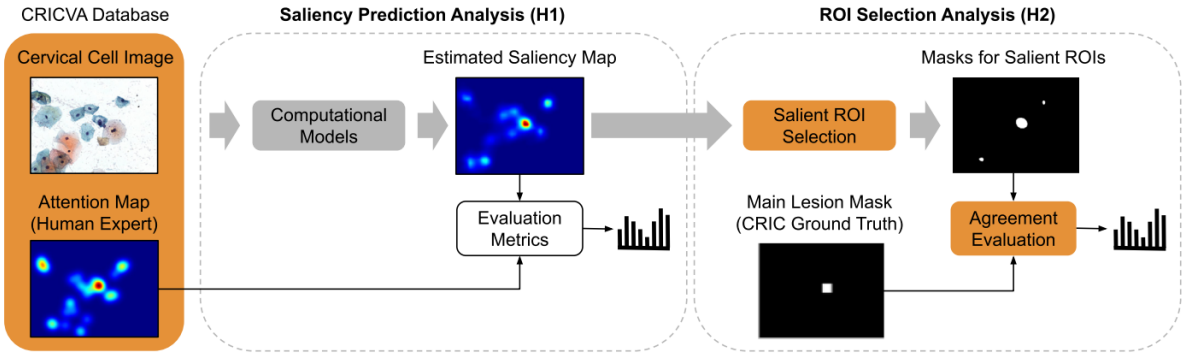
Source: Compiled by the author, informed by the content of (Ferreira *et al.*, 2019b).
A concise summary of saliency prediction models explored in this study, featuring the name in literature (Name), abbreviation used herein (Abbrev.), model description (Description), year of publication (Year), model category (Cat.), denoting T: Top-down model, B: Bottom-up model, and benchmark (Ben.).
* Tailored model for cellular image analysis.

Concludes.

3.4.2 Overview of saliency prediction and ROIs identification: preliminary insights and results

The methodology employed to investigate Hypothesis H1 is illustrated in Fig. 21. The focus of this examination was to assess the predictive accuracy of the surveyed saliency models in delineating areas of interest as observed by cytopathologists during the analysis of cervical cell images. This involved the processing of RGB color cervical cell images through the surveyed saliency models, with the resulting estimated maps being compared to the visual attention data gathered in the eye-tracking experiment. State-of-the-art metrics were then applied to evaluate these results and analyze the impact of top-down factors on the cytopathologist's image analysis.

Figure 21 – Research methodology overview



Source: Retrieved from (Ferreira *et al.*, 2019b).
The study unfolds in two stages: (1) the comprehensive evaluation of surveyed saliency models utilizing state-of-the-art metrics (H1 test), and (2) a meticulous analysis of Region of Interest (ROI) selection, probing the efficacy of highlighting the most pertinent lesions on the estimated saliency map (H2 test).

For the exploration of Hypothesis H2, the probability of the most relevant cell lesions being identified as salient objects by the surveyed models was measured. Following the approach outlined by (Bylinskii *et al.*, 2016), salient objects were identified based on their positions aligning with highlighted regions, i.e., Regions of Interest, in the predicted saliency map. A comprehensive description of the algorithm employed to identify ROIs in the estimated saliency maps can be found in Annex C.

The analysis of classic (non-deep) bottom-up methods revealed their tendency to highlight false positive attention regions on cervical cell images. These methods, designed to simulate low-level human visual attention, demonstrated sensitivity to features like contrast between cells and background, impacting saliency prediction. Concerning top-down models, The ZH model, implementing a target-driven approach, presented a promising strategy for localizing conspicuous areas on cervical cell images. However, saliency maps may be sparse for predicting cytopathologists' eye fixations, emphasizing the need for effective correspondence between abnormal cell templates and cell shapes.

The SRC model demonstrated a robust ability to identify the location of the most relevant cell lesion for various pathologies, suggesting its potential as a ROI extractor for automated cell pre-screening. SVC also showed valuable results, albeit with less robustness for specific lesions. Bottom-up strategies, particularly FES and GB, displayed suitability for ROI selection, providing valuable insights for applications with sparse cell clumps and low dark artifact presence.

3.5 Discussion

The third chapter of this thesis explores the integration of eye-tracking technology and machine learning models for the analysis of medical image interpretation, particularly focusing on cervical cell images. The convergence of these methodologies holds significant promise for advancing diagnostic precision, medical training, and overall patient care within the healthcare domain. Through automated competency assessments and real-time diagnostic feedback, these integrated systems offer substantial support to medical professionals and trainees. The potential to predict diagnostic errors and identify deviations from expert strategies reinforces the substantial role of eye tracking in clinical practice and medical education.

The study's eye-tracking data from experienced cytopathologists specializing in cervical care revealed a high level of gaze pattern agreement. The consistent fixation patterns observed across participants support the hypothesis that well-defined clinical Regions of Interest (ROIs) exist in Pap smear images. This finding has important implications for understanding how medical professionals visually engage with medical images, emphasizing the presence of consistent attention focus across trials.

An analysis of center bias within the CRICVA database highlighted a significant variation in the spatial distribution of cytopathologists' attention. Despite heightened activation near the image center in the Average Annotation Map (AAM), this variation suggests that cytopathologists dedicated considerable attention to objects located away from the central region. This nuanced exploration adds depth to our understanding of visual attention dynamics in medical image interpretation.

The survey of bottom-up and top-down saliency prediction models provided valuable insights into their predictive accuracy in delineating areas of interest observed by cytopathologists. The analysis of classic bottom-up methods revealed their sensitivity to features like contrast between cells and background, impacting saliency prediction. Notably, the ZH model, a top-down model implementing a target-driven approach, presented a promising strategy for localizing conspicuous areas on cervical cell images. However, sparse saliency maps stresses the need for effective correspondence between abnormal cell templates and shapes. The SRC model demonstrated robust ability for identifying the location of the most relevant cell lesions, suggesting its potential as a ROI extractor for automated cell pre-screening.

4 CONCLUSIONS AND FUTURE PERSPECTIVES

In conclusion, the findings presented in Chapter 2 evidence the efficacy of processing and analyzing eye-tracking data as a means to derive reliable proxies for the complexity and coherence of varied texts. To validate this proposition comprehensively, a diverse range of texts, spanning children's stories, randomly generated word texts, and excerpts from literary works, were subjected to validation through an extensive Internet survey involving a substantial cohort of readers. The validity of our findings is affirmed through (*i*) the observed nearly monotonic relationship between the average magnetization $\langle m \rangle$ of fixation activities and the average complexity $\langle \pi \rangle$ of the texts, and (*ii*) the efficacy of the distance $(T_o - T_c)$ in distinguishing random texts (devoid of meaningful content but with preserved grammatical structures) from coherent ones. We recall that the curve $T - T_c$ for each text is computed from the "energy" of the system, obtained through the application of the maximum-entropy learning algorithm to the fixation activities of all eye-tracked readers. This finding holds notable significance for several reasons. Firstly, it serves as an exemplification of the robustness of the maximum-entropy approach as a fundamental yet effective method for probing real complex systems. Simultaneously, our observations shed light on the cohesive response of humans when engaging with coherent and consistent textual information. This coherence is indicative of the advanced language formation and reading prediction mechanisms that individuals have developed over time. Conversely, in the case of nonsensical written information, we discern a dispersed collective cognitive response, underscoring the sensitivity of human cognitive processes to the semantic coherence of textual content.

In reflecting on the scope of our research, it is important to note that while our experimental reading materials were relatively short, there remains potential for extending our method to longer texts. However, this expansion would require participants to scroll for continued reading, which may introduce artifacts in fixation series due to visual accommodation. Addressing this issue would involve adjusting fixation data during scrolling, possibly through the use of an averaging function. Moreover, when dealing with a larger dataset, solving the inverse Ising problem may not be as straightforward as in our study. The Metropolis Algorithm emerged as a valuable tool in our research, proving effective in this context. This method involves iteratively exploring the parameter space of the Ising model, generating candidate parameter sets that are then accepted or rejected based on the energy difference calculated from the model. Ultimately, this iterative process leads to the convergence towards parameter values that most

accurately represent the observed data. Additionally, natural language processing programs available nowadays offer powerful tools for experimenting with texts of diverse characteristics. This approach facilitates the creation and fine-tuning of various aspects of texts, suitable for testing reading material to evaluate perceived complexity, coherence, and other linguistic metrics of interest.

In Chapter 3, we conducted a thorough examination of visual attention dynamics in the analysis of cervical cell images using eye-tracking. Our study involved three highly experienced cytopathologists, specialists in cervical care, aiming to understand the nuances of their cognitive processes. The employed eye-tracking methodology allowed for precise recording and analysis of participants' eye movements and fixations. The translation of these fixations into binary matrices and subsequent generation of attention heat maps provided a visual depiction of the focus areas of the cytopathologists. The results showcased a notable agreement in gaze patterns among participants, supporting the existence of well-defined clinical Regions of Interest (ROIs) in Pap smear images. An additional exploration of center bias within the CRICVA database unveiled substantial variations in the spatial distribution of activations, indicating a notable focus on objects situated away from the central region. Further investigation into saliency models, guided by hypotheses H1 and H2, provided valuable insights into their predictive accuracy and potential implications for medical image interpretation. The evaluation of both bottom-up and top-down saliency prediction models revealed nuanced strengths and weaknesses, contributing to future research and applications in automated cell pre-screening.

Our contribution to this project primarily focused on the initial phases of research, which included devising strategies for analyzing image interpretation using eye-tracking, particularly simulating the workflow of pap smear analysis. We were also responsible for designing and implementing the experiments, as well as compiling raw data. Throughout this process, we gained valuable insights, specially in understanding how eye-tracking data can inform and refine saliency learning algorithms. This provided us with a valuable opportunity to explore the application of eye-tracking technology in this context and sparked considerations for future projects in medical diagnosis and other relevant fields.

Finally, it is important to note that our research has been deeply interdisciplinary, drawing from fields such as neuroscience, linguistics, and computer science. Understanding the cognitive pathway of visual processing was key for implementing our experiments and interpreting the data effectively. Operating the eye tracker and designing experiments required

insights into visual function, while interpreting the data demanded knowledge of eye movement patterns, linguistic features, reading comprehension processes and image interpretation, among other technical aspects. Furthermore, validating our experimental results involved searching innovative strategies, leading to the development of an extensive survey. This practical approach emerged during the analysis phase and proved instrumental in validating our main project's findings. Navigating these challenges proved to be an enriching experience, offering valuable insights into the application of physical models for studying cognitive activity. We hope that our work encourages future students to conduct further research in this area, as well as engaging in other interdisciplinary collaborations.

BIBLIOGRAPHY

AHMED, I. A.; SENAN, E. M.; RASSEM, T. H.; ALI, M. A.; SHATNAWI, H. S. A.; ALWAZER, S. M.; ALSHAHRANI, M. Eye tracking-based diagnosis and early detection of autism spectrum disorder using machine learning and deep learning techniques. **Electronics**, MDPI, v. 11, n. 4, p. 530, 2022.

ALMINO, J. Guimarães Rosa, do Sertão às fronteiras. **Revista Brasileira (Academia Brasileira de Letras)**, v. 96, p. 19–36, 2018. ISSN 0103707-2. Author is occupant of Chair 22 at the Academia Brasileira de Letras. Available at: http://www.academia.org.br/sites/default/files/publicacoes/arquivos/revista_brasileira_096_internet.pdf. Access in: 2 jan. 2024.

ARARIPE JÚNIOR, T. A. **José de Alencar**: perfil literario. Rio de Janeiro: Typ. da Escola de Serafim José Alves, circa 1880. 140-146 p. Available at: https://digital.bbm.usp.br/handle/bbm/5206. Access in: 2 jan. 2024.

ARDILA, D.; KIRALY, A. P.; BHARADWAJ, S.; CHOI, B.; REICHER, J.; PENG, L.; TSE, D.; ETEMADI, M.; YE, W.; CORRADO, G.; NAIDICH, D.; SHETTY, S. R. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. **Nature Medicine**, 2019. Available at: https://rdcu.be/bDpp1. Access in: 2 jan. 2024.

ASHBY, J.; RAYNER, K.; CLIFTON, C. Eye movements of highly skilled and average readers: differential effects of frequency and predictability. **The Quarterly Journal of Experimental Psychology Section A**, v. 58, n. 6, p. 1065–86, 08 2005. ISSN 0272-4987.

BELANGER, D.; YOUNG, A. The random field ising model. **Journal of Magnetism and Magnetic Materials**, v. 100, n. 1, p. 272–291, 1991. ISSN 0304-8853. Available at: https://www.sciencedirect.com/science/article/pii/030488539190825U. Access in: 2 jan. 2024.

BERBAUM, K. S.; BRANDSER, E. A.; FRANKEN, E.; DORFMAN, D. D.; CALDWELL, R. T.; KRUPINSKI, E. A. Gaze dwell times on acute trauma injuries missed because of satisfaction of search. **Academic Radiology**, Elsevier, v. 8, n. 4, p. 304–314, 2001. ISSN 1076-6332. Available at: https://www.sciencedirect.com/science/article/pii/S1076633203804993. Access in: 2 jan. 2024.

BIALEK, W.; CAVAGNA, A.; GIARDINA, I.; MORA, T.; SILVESTRI, E.; VIALE, M.; WALCZAK, A. M. Statistical mechanics for natural flocks of birds. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 109, n. 13, p. 4786–4791, 2012. ISSN 0027-8424. Available at: https://www.pnas.org/content/109/13/4786. Access in: 2 jan. 2024.

BIALEK, W.; CAVAGNA, A.; GIARDINA, I.; MORA, T.; POHL, O.; SILVESTRI, E.; VIALE, M.; WALCZAK, A. M. Social interactions dominate speed control in poising natural flocks near criticality. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 111, n. 20, p. 7212–7217, 2014. ISSN 0027-8424. Available at: https://www.pnas.org/content/111/20/7212. Access in: 2 jan. 2024.

BORJI, A.; CHENG, M.-M.; HOU, Q.; JIANG, H.; LI, J. Salient object detection: A survey. **Computational visual media**, Springer, v. 5, p. 117–150, 2019.

BORJI, A.; CHENG, M.-M.; JIANG, H.; LI, J. Salient object detection: A benchmark. **IEEE transactions on image processing**, IEEE, v. 24, n. 12, p. 5706–5722, 2015.

BORJI, A.; ITTI, L. State-of-the-art in visual attention modeling. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 35, n. 1, p. 185–207, 2012.

BRUNYÉ, T. T.; CARNEY, P. A.; ALLISON, K. H.; SHAPIRO, L. G.; WEAVER, D. L.; ELMORE, J. G. Eye movements as an index of pathologist visual expertise: A pilot study. **PLOS ONE**, Public Library of Science, v. 9, n. 8, p. 1–7, 2014. Available at: https://doi.org/10.1371/journal.pone.0103447. Access in: 2 jan. 2024.

BRUNYÉ, T. T.; DREW, T.; WEAVER, D. L.; ELMORE, J. G. A review of eye tracking for understanding and improving diagnostic interpretation. **Cognitive research**: Principles and Implications, SpringerOpen, v. 4, n. 1, p. 1–16, 2019.

BURLESON-LESSER, K.; MORONE, F.; DEGUZMAN, P.; PARRA, L. C.; MAKSE, H. A. Collective behaviour in video viewing: A thermodynamic analysis of gaze position. **PLOS ONE**, v. 12, n. 1, p. 1–19, jan. 2017. Available at: https://doi.org/10.1371/journal.pone.0168995. Access in: 2 jan. 2024.

BURY, T. Market structure explained by pairwise interactions. **Physica A**: Statistical Mechanics and its Applications, v. 392, n. 6, p. 1375 – 1385, 2013. ISSN 0378-4371. Available at: http://www.sciencedirect.com/science/article/pii/S0378437112009685. Access in: 2 jan. 2024.

BUSWELL, G. T. **How people look at pictures**: a study of the psychology and perception in art. Chicago: University of Chicago Press, 1935. Available at: https://www.journals.uchicago.edu/doi/10.1086/457424. ISBN 9780598831330. Access in: 2 jan. 2024.

BYLINSKII, Z.; JUDD, T.; BORJI, A.; ITTI, L.; DURAND, F.; OLIVA, A.; TORRALBA, A. **MIT Saliency Benchmark**. 2015. Available at: http://saliency.mit.edu/. Access in: 2 jan. 2024.

BYLINSKII, Z.; JUDD, T.; OLIVA, A.; TORRALBA, A.; DURAND, F. What do different evaluation metrics tell us about saliency models? **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 41, n. 3, p. 740–757, 2018.

BYLINSKII, Z.; RECASENS, A.; BORJI, A.; OLIVA, A.; TORRALBA, A.; DURAND, F. Where should saliency models look next? In: **Computer Vision–ECCV 2016**. Cham: Springer International Publishing, 2016. p. 809–824. ISBN 978-3-319-46454-1. 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V. Available at: https://people.csail.mit.edu/recasens/docs/bylinskii_eccv2016.pdf. Access in: 2 jan. 2024.

CARRASCO, M. Visual attention: The past 25 years. **Vision Research**, v. 51, n. 13, p. 1484–1525, 2011. ISSN 0042-6989. Vision Research 50th Anniversary Issue: Part 2. Available at: https://www.sciencedirect.com/science/article/pii/S0042698911001544. Access in: 2 jan. 2024.

CHEN, Z. H.; FU, H.; LO, W. L.; CHI, Z.; XU, B. Eye-tracking-aided digital system for strabismus diagnosis. **Healthcare technology letters**, Wiley Online Library, v. 5, n. 1, p. 1–6, 2018.

CLIFTON, C.; FERREIRA, F.; HENDERSON, J. M.; INHOFF, A. W.; LIVERSEDGE, S. P.; REICHLE, E. D.; SCHOTTER, E. R. Eye movements in reading and information processing: Keith rayner's 40year legacy. **Journal of Memory and Language**, v. 86, p. 1–19, jan. 2016. ISSN 0749-596X. Available at: http://www.sciencedirect.com/science/article/pii/S0749596X15000960. Access in: 2 jan. 2024.

COCCO, S.; LEIBLER, S.; MONASSON, R. Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 106, n. 33, p. 14058–14062, 2009. ISSN 0027-8424. Available at: https://www.pnas.org/content/106/33/14058. Access in: 2 jan. 2024.

COOMBES, L.; CULVERHOUSE, P. Pattern recognition in cervical cytological slide images. In: ICAPR. **Fifth International Conference on Advances in Pattern Recognition (ICAPR 2003)**. Calcutta, India, 2003. Available at: https://www.researchgate.net/publication/268364829_Pattern_Recognition_in_Cervical_Cytological_Slide_Images. Access in: 2 jan. 2024.

CORNIA, M.; BARALDI, L.; SERRA, G.; CUCCHIARA, R. Predicting human eye fixations via an lstm-based saliency attentive model. **IEEE Transactions on Image Processing**, IEEE, v. 27, n. 10, p. 5142–5154, 2018.

DE CASTRO, M. A. **O homem provisório no Grande Sertão**: um estudo de Grande Sertão: Veredas. Rio de Janeiro: Edições Tempo Brasileiro, 1976. 44 p. (Biblioteca Tempo universitário).

DE CASTRO, N. L. **Universo e vocabulário do Grande sertão**. Rio de Janeiro: José Olympio, 1970. ISBN 6586722489.

DE FAUW, J.; LEDSAM, J. R.; ROMERA-PAREDES, B.; NIKOLOV, S.; TOMASEV, N.; BLACKWELL, S.; ASKHAM, H.; GLOROT, X.; O'DONOGHUE, B.; VISENTIN, D.; DRIESSCHE, G. van den; LAKSHMINARAYANAN, B.; MEYER, C.; MACKINDER, F.; BOUTON, S.; AYOUB, K.; CHOPRA, R.; KING, D.; KARTHIKESALINGAM, A.; HUGHES, C. O.; RAINE, R.; HUGHES, J.; SIM, D. A.; EGAN, C.; TUFAIL, A.; MONTGOMERY, H.; HASSABIS, D.; REES, G.; BACK, T.; KHAW, P. T.; SULEYMAN, M.; CORNEBISE, J.; KEANE, P. A.; RONNEBERGER, O. Clinically applicable deep learning for diagnosis and referral in retinal disease. **Nature medicine**, v. 24, n. 9, p. 1342—1350, sep. 2018. ISSN 1078-8956. Available at: https://doi.org/10.1038/s41591-018-0107-6. Access in: 2 jan. 2024.

DUAN, L.; WU, C.; MIAO, J.; QING, L.; FU, Y. Visual saliency detection by spatially weighted dissimilarity. In: IEEE. **CVPR 2011**. Colorado Springs, CO, USA, 2011. p. 473–480.

EHRLICH, S. F.; RAYNER, K. Contextual effects on word perception and eye movements during reading. **Journal of Verbal Learning & Verbal Behavior**, v. 20, n. 6, p. 641–655, dec. 1981. ISSN 0022-5371. Available at: http://www.sciencedirect.com/science/article/pii/S0022537181902206. Access in: 2 jan. 2024.

ENGBERT, R.; KLIEGL, R. Mathematical models of eye movements in reading: a possible role for autonomous saccades. **Biological cybernetics**, v. 85, n. 2, p. 77 – 87, 2001. ISSN 0340-1200. Available at: http://www.sciencedirect.com/science/article/pii/S0042698901003017. Access in: 2 jan. 2024.

ENGBERT, R.; KLIEGL, R.; LONGTIN, A. Complexity of eye movements in reading. **International Journal of Bifurcation and Chaos**, v. 14, n. 2, p. 493 – 503, 2004. ISSN 1793-6551.

ENGBERT, R.; LONGTIN, A.; KLIEGL, R. A dynamical model of saccade generation in reading based on spatially distributed lexical processing. **Vision Research**, v. 42, n. 5, p. 621 – 636, 2002. ISSN 0042-6989. Available at: http://www.sciencedirect.com/science/article/pii/S0042698901003017. Access in: 2 jan. 2024.

FABIO, R. A.; INCORPORA, C.; ERRANTE, A.; MOHAMMADHASNI, N.; CAPRÌ, T.; CARROZZA, C.; DE SANTIS, S.; FALZONE, A. The influence of cognitive load and amount of stimuli on entropy through eye tracking measures. In: **EuroAsianPacific Joint Conference on Cognitive Science**. Torino, Italy: CEUR Workshop Proceedings, 2015. Available at: https://api.semanticscholar.org/CorpusID:3011808. Access in: 2 jan. 2024.

FANG, S.; LI, J.; TIAN, Y.; HUANG, T.; CHEN, X. Learning discriminative subspaces on random contrasts for image saliency analysis. **IEEE transactions on neural networks and learning systems**, IEEE, v. 28, n. 5, p. 1095–1108, 2016.

FERREIRA, D. S.; RAMALHO, G. L. B.; TORRES, D.; TOBIAS, A. H. G.; REZENDE, M. T.; MEDEIROS, F. N. S.; BIANCHI, A. G. C.; CARNEIRO, C. M.; USHIZIMA, D. M. **CRICVA Database**. 2019. Mendeley Data, V1. Available at: https://data.mendeley.com/datasets/bk45c9yxb9/1. Access in: 2 jan. 2024.

FERREIRA, D. S.; RAMALHO, G. L. B.; TORRES, D.; TOBIAS, A. H. G.; REZENDE, M. T.; MEDEIROS, F. N. S.; BIANCHI, A. G. C.; CARNEIRO, C. M.; USHIZIMA, D. M. Saliency-driven system models for cell analysis with deep learning. **Computer methods and programs in biomedicine**, Elsevier, v. 182, p. 105053, 2019. ISSN 0169-2607. Available at: https://www.sciencedirect.com/science/article/pii/S0169260718316870. Access in: 2 jan. 2024.

FINDLAY, J. M. Eye movements and visual information processing. In: **The Mind's Eye**: Cognitive and Applied Aspects of Eye Movement Research. [*S. l.*]: Elsevier, 2003. ISBN 978-0-444-51020-4.

FISHER, D.; FREY, N.; LAPP, D. **Text Complexity**: Raising Rigor in Reading. [*S. l.*]: International Reading Association, 2012. ISBN 9780872074781.

FORMAN, D.; FERLAY, J. The global and regional burden of cancer. In: STEWART, B. W.; WILD, C. P. (Ed.). **World Cancer Report 2014**. Lyon, France: International Agency for Research on Cancer, World Health Organization, 2014. cap. 1.1, p. 16–53.

GARCÍA, M. S. **Grande Sertão: Veredas, de João Guimarães Rosa**. Análise textual da obra e das duas traduções ao espanhol. 2015. Dissertação (Mestrado em Estudos da Tradução) – Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2015.

HAREL, J.; KOCH, C.; PERONA, P. Graph-based visual saliency. In: SCHÖLKOPF, B.; PLATT, J.; HOFFMAN, T. (Ed.). **Advances in Neural Information Processing Systems**. [*S. l.*]: MIT Press, 2006. v. 19. Available at: https://proceedings.neurips.cc/paper_files/paper/2006/file/4db0f8b0fc895da263fd77fc8aecabe4-Paper.pdf. Access in: 2 jan. 2024.

HAREZLAK, K.; KASPROWSKI, P. Application of eye tracking in medicine: A survey, research issues and challenges. **Computerized Medical Imaging and Graphics**, v. 65, p. 176 – 190, 2018. ISSN 0895-6111. Part of a special issue: Advances in Biomedical Image Processing. Available at: http://www.sciencedirect.com/science/article/pii/S0895611117300435. Access in: 2 jan. 2024.

HOU, X.; HAREL, J.; KOCH, C. Image signature: Highlighting sparse salient regions. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 34, n. 1, p. 194–201, 2011.

HOU, X.; ZHANG, L. Saliency detection: A spectral residual approach. In: IEEE. **2007 IEEE Conference on computer vision and pattern recognition**. Minneapolis, MN, USA, 2007. p. 1–8. Available at: http://vigir.missouri.edu/~gdesouza/Research/Conference_CDs/IEEE_CVPR_2007/data/papers/0299.pdf. Access in: 2 jan. 2024.

HUANG, X.; SHEN, C.; BOIX, X.; ZHAO, Q. Salicon: Reducing the semantic gap in saliency prediction by adapting deep neural networks. In: **Proceedings of the IEEE international conference on computer vision**. Santiago, Chile: IEEE, 2015. p. 262–270. Available at: https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Huang_SALICON_Reducing_the_ICCV_2015_paper.pdf. Access in: 2 jan. 2024.

ITTI, L.; KOCH, C. A saliency-based search mechanism for overt and covert shifts of visual attention. **Vision research**, Elsevier, v. 40, n. 10-12, p. 1489–1506, 2000.

ITTI, L.; KOCH, C.; NIEBUR, E. A model of saliency-based visual attention for rapid scene analysis. **IEEE Transactions on pattern analysis and machine intelligence**, Ieee, v. 20, n. 11, p. 1254–1259, 1998.

JAVAL, E.; CIUFFREDA, K. J.; BASSIL, N. Essay on the physiology of reading. **Ophthalmic and Physiological Optics**, v. 10, n. 4, p. 381–384, 1990. Originally published in 1879 as "Essai sur la physiologie de la lecture" in Annales d'Ocilistique, Volume 82, 242–253. Translated by K. J. Ciuffreda and N Bassil. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-1313.1990.tb00885.x. Access in: 2 jan. 2024.

JAYNES, E. T. Information theory and statistical mechanics. **Physical Review**, American Physical Society, v. 106, p. 620–630, may 1957. Available at: https://link.aps.org/doi/10.1103/PhysRev.106.620. Access in: 2 jan. 2024.

JAYNES, E. T. Information theory and statistical mechanics. II. **Physical Review**, American Physical Society, v. 108, p. 171–190, oct. 1957. Available at: https://link.aps.org/doi/10.1103/PhysRev.108.171. Access in: 2 jan. 2024.

JIANG, M.; HUANG, S.; DUAN, J.; ZHAO, Q. Salicon: Saliency in context. In: **Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)**. Boston, MA, USA: IEEE, 2015. p. 1072–1080. Available at: https://www-users.cse.umn.edu/~qzhao/publications/pdf/salicon_cvpr15.pdf. Access in: 2 jan. 2024.

JUDD, T.; EHINGER, K.; DURAND, F.; TORRALBA, A. Learning to predict where humans look. In: IEEE. **2009 IEEE 12th International Conference on Computer Vision**. Kyoto, Japan, 2009. p. 2106–2113. Available at: https://people.csail.mit.edu/tjudd/WherePeopleLook/index.html. Access in: 2 jan. 2024.

JUHASZ, B. J.; LIVERSEDGE, S. P.; WHITE, S. J.; RAYNER, K. Eye movements and the use of parafoveal word length information in reading. **Journal of Experimental Psychology: Human Perception and Performance**, v. 34, n. 6, p. 1560–1579, dec. 2008. Available at: https://doi.org/10.1037/a0012319. Access in: 2 jan. 2024.

KANDEL, E. R.; SCHWARTZ, J. H.; JESSELL, T. M.; SIEGELBAUM, S. A.; HUDSPETH, A. J. Part V: Perception, chapters 25-29. In: **Principles of Neural Science**, Fifth Edition. [*S. l.*]: McGraw-Hill Education, 2013, (McGraw-Hill's AccessMedicine). ISBN 9780071390118.

KENDALL, M. G. A new measure of rank correlation. **Biometrika**, Oxford University Press, Biometrika Trust, v. 30, n. 1/2, p. 81–93, 1938. ISSN 00063444. Available at: http://www.jstor.org/stable/2332226. Access in: 2 jan. 2024.

KHOSRAVAN, N.; CELIK, H.; TURKBEY, B.; JONES, E. C.; WOOD, B.; BAGCI, U. A collaborative computer aided diagnosis (C-CAD) system with eye-tracking, sparse attentional model, and deep learning. **Medical image analysis**, Elsevier, v. 51, p. 101–115, 2019.

KLIEGL, R.; GRABNER, E.; ROLFS, M.; ENGBERT, R. Length, frequency, and predictability effects of words on eye movements in reading. **European Journal of Cognitive Psychology**, Routledge, v. 16, n. 1-2, p. 262–284, 2004. Available at: https://doi.org/10.1080/09541440340000213. Access in: 2 jan. 2024.

KLIEGL, R.; NUTHMANN, A.; ENGBERT, R. Tracking the mind during reading: The influence of past, present, and future words on fixation durations. **Journal of Experimental Psychology**, v. 135, p. 12–35, 02 2006. ISSN 1939-2222. Available at: http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-57225. Access in: 2 jan. 2024.

KRUSKAL, W. H.; WALLIS, W. A. Use of ranks in one-criterion variance analysis. **Journal of the American Statistical Association**, Taylor & Francis, v. 47, n. 260, p. 583–621, 1952.

KRUTHIVENTI, S. S. S.; AYUSH, K.; BABU, R. V. Deepfix: A fully convolutional neural network for predicting human eye fixations. **IEEE Transactions on Image Processing**, IEEE, v. 26, n. 9, p. 4446–4456, 2017.

KUNDEL, H. L.; LA FOLLETTE JR, P. S. Visual search patterns and experience with radiological images. **Radiology**, The Radiological Society of North America, v. 103, n. 3, p. 523–528, 1972.

LE MEUR, O.; BACCINO, T. Methods for comparing scanpaths and saliency maps: strengths and weaknesses. **Behavior research methods**, Springer, v. 45, n. 1, p. 251–266, 2013.

LEE, C. S.; NAGY, P. G.; WEAVER, S. J.; NEWMAN-TOKER, D. E. Cognitive and system factors contributing to diagnostic errors in radiology. **American Journal of Roentgenology**, American Roentgen Ray Society, v. 201, n. 3, p. 611–617, sep. 2013. ISSN 0361-803X.

LÉVÊQUE, L.; BOSMANS, H.; COCKMARTIN, L.; LIU, H. State of the art: Eye-tracking studies in medical imaging. **IEEE Access**, IEEE, v. 6, p. 37023–37034, 2018. Available at: https://orca.cardiff.ac.uk/id/eprint/112911/1/08399735.pdf. Access in: 2 jan. 2024.

LÉVÊQUE, L.; VANDE BERG, B.; BOSMANS, H.; COCKMARTIN, L.; KEUPERS, M.; VAN ONGEVAL, C.; LIU, H. A statistical evaluation of eye-tracking data of screening mammography: Effects of expertise and experience on image reading. **Signal Processing**: Image Communication, v. 78, p. 86 – 93, 2019. ISSN 0923-5965. Available at: http://www.sciencedirect.com/science/article/pii/S0923596518309652. Access in: 2 jan. 2024.

LEZON, T. R.; BANAVAR, J. R.; CIEPLAK, M.; MARITAN, A.; FEDOROFF, N. V. Using the principle of entropy maximization to infer genetic interaction networks from gene expression patterns. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 103, n. 50, p. 19033–19038, 2006. ISSN 0027-8424. Available at: https://www.pnas.org/content/103/50/19033. Access in: 2 jan. 2024.

LI, R.; SHI, P.; PELZ, J.; ALM, C. O.; HAAKE, A. R. Modeling eye movement patterns to characterize perceptual skill in image-based diagnostic reasoning processes. **Computer Vision and Image Understanding**, Elsevier, v. 151, p. 138–152, 2016.

LIN, Y.-C.; LIU, T.-C.; SWELLER, J. Improving the frame design of computer simulations for learning: Determining the primacy of the isolated elements or the transient information effects. **Computers & Education**, Elsevier, v. 88, p. 280–291, 2015.

LITJENS, G.; KOOI, T.; BEJNORDI, B. E.; SETIO, A. A. A.; CIOMPI, F.; GHAFOORIAN, M.; VAN DEL LAAK, J. A.; VAN GINNEKEN, B.; SÁNCHEZ, C. I. A survey on deep learning in medical image analysis. **Medical image analysis**, Elsevier, v. 42, p. 60–88, 2017.

LIU, N.; HAN, J.; LIU, T.; LI, X. Learning to predict eye fixations via multiresolution convolutional neural networks. **IEEE transactions on neural networks and learning systems**, IEEE, v. 29, n. 2, p. 392–404, 2016.

LOCASALE, J. W.; WOLF-YADLIN, A. Maximum entropy reconstructions of dynamic signaling networks from quantitative proteomics data. **PLOS ONE**, Public Library of Science, v. 4, n. 8, p. 1–10, 08 2009. Available at: https://doi.org/10.1371/journal.pone.0006522. Access in: 2 jan. 2024.

LOUKAS, C.; VARYTIMIDIS, C.; RAPANTZIKOS, K.; KANAKIS, M. A. Keyframe extraction from laparoscopic videos based on visual saliency detection. **Computer methods and programs in biomedicine**, Elsevier, v. 165, p. 13–23, 2018.

LUCS-KHO. **The Visitor**. 2007. Licensing: Released into the public domain by the author, Lucs-kho at English Wikipedia. Available at: https://commons.wikimedia.org/w/index.php?title=File: Yarbus_The_Visitor.jpg. Access in: 2 jan. 2024.

MANNING, D.; ETHELL, S.; DONOVAN, T.; CRAWFORD, T. How do radiologists do it? The influence of experience and training on searching for chest nodules. **Radiography**, v. 12, n. 2, p. 134 – 142, 2006. ISSN 1078-8174. Available at: http://www.sciencedirect.com/science/article/pii/S1078817405000131. Access in: 2 jan. 2024.

MATSUMOTO, H.; TERAO, Y.; YUGETA, A.; FUKUDA, H.; EMOTO, M.; FURUBAYASHI, T.; OKANO, T.; HANAJIMA, R.; UGAWA, Y. Where do neurologists look when viewing brain CT images? An eye-tracking study involving stroke cases. **PLOS ONE**, Public Library of Science, v. 6, n. 12, p. 1–7, dec. 2011. Available at: https://doi.org/10.1371/journal.pone.0028928. Access in: 2 jan. 2024.

MCCREADIE, G.; OLIVER, T. Eight CT lessons that we learned the hard way: an analysis of current patterns of radiological error and discrepancy with particular emphasis on CT. **Clinical Radiology**, v. 64, n. 5, p. 491 – 499, 2009. ISSN 0009-9260. Available at: http://www.sciencedirect.com/science/article/pii/S000992600900035X. Access in: 2 jan. 2024.

MCKINNEY, S. M.; SIENIEK, M.; GODBOLE, V.; GODWIN, J.; ANTROPOVA, N.; ASHRAFIAN, H.; BACK, T.; CHESUS, M.; CORRADO, G. C.; DARZI, A.; ETEMADI, M.; GARCIA-VICENTE, F.; GILBERT, F. J.; HALLING-BROWN, M.; HASSABIS, D.; JANSEN, S.; KARTHIKESALINGAM, A.; KELLY, C. J.; KING, D.; LEDSAM, J. R.; MELNICK, D.; MOSTOFI, H.; PENG, L.; REICHER, J. J.; ROMERA-PAREDES, B.; SIDEBOTTOM, R.; SULEYMAN, M.; TSE, D.; YOUNG, K. C.; DE FAUW, J.; SHETTY, S. International evaluation of an AI system for breast cancer screening. **Nature**, Nature Publishing Group, v. 577, n. 7788, p. 89–94, jan. 2020. ISSN 0028-0836.

MCNAMARA, D. S.; KINTSCH, E.; SONGER, N. B.; KINTSCH, W. Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. **Cognition and Instruction**, Routledge, v. 14, n. 1, p. 1–43, 1996. Available at: https://doi.org/10.1207/s1532690xci1401_1. Access in: 2 jan. 2024.

MELLO-THOMS, C.; HARDESTY, L.; SUMKIN, J.; GANOTT, M.; HAKIM, C.; BRITTON, C.; STALDER, J.; MAITZ, G. Effects of lesion conspicuity on visual search in mammogram reading. **Academic radiology**, Elsevier, v. 12, n. 7, p. 830–840, 2005.

MESEEMS. **MindMiners Respondents Panel**. 2019. São Paulo, Brazil. Available at: https://meseems.com.br/. Access in: 2 jan. 2024.

MINDMINERS. **Research Services**. 2019. São Paulo, Brazil. Available at: https: //mindminers.com/. Access in: 2 jan. 2024.

MORCOS, F.; PAGNANI, A.; LUNT, B.; BERTOLINO, A.; MARKS, D. S.; SANDER, C.; ZECCHINA, R.; ONUCHIC, J. N.; HWA, T.; WEIGT, M. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 108, n. 49, p. E1293–E1301, 2011. ISSN 0027-8424. Available at: https://www.pnas.org/content/108/49/E1293. Access in: 2 jan. 2024.

MURABITO, F.; SPAMPINATO, C.; PALAZZO, S.; GIORDANO, D.; POGORELOV, K.; RIEGLER, M. Top-down saliency detection driven by visual classification. **Computer Vision and Image Understanding**, Elsevier, v. 172, p. 67–76, 2018.

MURRAY, N.; VANRELL, M.; OTAZU, X.; PARRAGA, C. A. Saliency estimation using a non-parametric low-level vision model. In: IEEE. **CVPR 2011**. Colorado Springs, CO, USA, 2011. p. 433–440.

NAYAR, R.; WILBUR, D. C. The Pap Test and Bethesda 2014 “The reports of my demise have been greatly exaggerated.” (after a quotation from Mark Twain). **Acta cytologica**, S. Karger AG, v. 59, n. 2, p. 121–132, 2015.

NELSON, J.; PERFETTI, C.; LIBEN, D.; LIBEN, M. Measures of text difficulty: Testing their predictive value for grade levels and student performance. **Council of Chief State School Officers**, Washington, DC, 2012.

NEWMAN, M. E. J.; BARKEMA, G. T. **Monte Carlo methods in statistical physics**. [*S. l.*]: Oxford: Clarendon Press, 1999. Chapters 1-4. ISBN 0-19-851797-1.

NGUYEN, H. C.; ZECCHINA, R.; BERG, J. Inverse statistical problems: from the inverse ising problem to data science. **Advances in Physics**, Informa UK Limited, v. 66, n. 3, p. 197–261, jun. 2017. ISSN 1460-6976. Available at: http://dx.doi.org/10.1080/00018732.2017.1341604. Access in: 2 jan. 2024.

NGUYEN, T. V.; ZHAO, Q.; YAN, S. Attentive systems: A survey. **International Journal of Computer Vision**, Springer, v. 126, p. 86–110, 2018.

OTAZU, X.; PARRAGA, C. A.; VANRELL, M. Toward a unified chromatic induction model. **Journal of Vision**, The Association for Research in Vision and Ophthalmology, v. 10, n. 12, p. 5–5, 2010.

POOLE, A.; BALL, L. J. Eye tracking in HCI and usability research. In: GHAOUI, C. (Ed.). **Encyclopedia of Human Computer Interaction**. Hershey, PA, USA: IGI Global, 2006. p. 211–219. ISBN 9781591405627.

R. TAVAKOLI, H.; LAAKSONEN, J. Bottom-up fixation prediction using unsupervised hierarchical models. In: CHEN, C.-S.; LU, J.; MA, K.-K. (Ed.). **Computer Vision–ACCV 2016 Workshops**. Cham: Springer International Publishing, 2017. p. 287–302. ISBN 978-3-319-54407-6.

R. TAVAKOLI, H.; RAHTU, E.; HEIKKILÄ, J. Fast and efficient saliency detection using sparse sampling and kernel density estimation. In: HEYDEN, A.; KAHL, F. (Ed.). **Image Analysis**. Berlin, Heidelberg: Springer, 2011. p. 666–675. ISBN 978-3-642-21227-7.

RANDOM TEXT GENERATOR. **Texts generator using random words**. 2018. https: //www.randomtextgenerator.com/ (Note: Website no longer available as of 2023. Alternative tools for generating similar random text can be found at https://rosettacode.org/wiki/Rosetta_Code, a platform offering solutions to programming tasks in multiple languages. Access in: 2 jan. 2022.

RANEY, G. E.; CAMPBELL, S. J.; BOVEE, J. C. Using eye movements to evaluate the cognitive processes involved in text comprehension. **Journal of Visualized Experiments**: JoVE, n. 83, p. 641–655, 2014.

RAYNER, K. Eye movements in reading and information processing: 20 years of research. **Psychological bulletin**, American Psychological Association, v. 124, n. 3, p. 372, 1998.

RAYNER, K.; DUFFY, S. A. Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. **Memory & Cognition**, v. 14, n. 3, p. 191–201, may 1986. ISSN 1532-5946. Available at: https://doi.org/10.3758/BF03197692. Access in: 2 jan. 2024.

RAYNER, K.; RANEY, G. E. Eye movement control in reading and visual search: Effects of word frequency. **Psychonomic Bulletin & Review**, v. 3, p. 245–248, 1996. ISSN 1531-5320.

RAYNER, K.; REICHLE, E. D.; STROUD, M. J.; WILLIAMS, C. C.; POLLATSEK, A. The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. **Psychology and Aging**, v. 21, p. 448–465, 2006. ISSN 0882-7974.

REICHLE, E. D.; POLLATSEK, A.; FISHER, D. L.; RAYNER, K. Toward a model of eye movement control in reading. **Psychological Review**, v. 105, p. 125 – 157, 1998. ISSN 1939-1471.

REINHART, T. Conditions for text coherence. **Poetics Today**, Duke University Press, Porter Institute for Poetics and Semiotics, v. 1, n. 4, p. 161–180, 1980. ISSN 03335372, 15275507. Available at: http://www.jstor.org/stable/1771893. Access in: 2 jan. 2024.

REPIN, I. **They Did Not Expect Him**. Between 1884 and 1888. Medium: Oil on canvas. Dimensions: 160.5 × 167.5 cm. Collection: Tretyakov Gallery. Source: www.allpaintings.org. Licensing: Public domain; author died in 1930. Available at: https://commons.wikimedia.org/w/index.php?title=File:Ilya_Repin_Unexpected_visitors.jpg. Access in: 2 jan. 2024.

ROTHMAN, R. The complex matter of text complexity. **Harvard Education Letter**, v. 28, n. 5, 2012. ISSN 8755-3716. Available at: https://www.hepg.org/hel-home/issues/28_5/helarticle/ the-complex-matter-of-text-complexity_544. Access in: 2 jan. 2024.

SCHAD, D. J.; NUTHMANN, A.; ENGBERT, R. Eye movements during reading of randomly shuffled text. **Vision Research**, v. 50, n. 23, p. 2600 – 2616, 2010. ISSN 0042-6989. Vision Research Reviews.

SCHNEIDMAN, E.; BERRY, M. J.; SEGEV, R.; BIALEK, W. Weak pairwise correlations imply strongly correlated network states in a neural population. **Nature**, v. 440, p. 1007–1012, 04 2006. ISSN 1476-4687. Available at: https://doi.org/10.1038/nature04701. Access in: 2 jan. 2024.

SHANNON, C. E. A mathematical theory of communication. **The Bell System Technical Journal**, Nokia Bell Labs, v. 27, n. 3, p. 379–423, 1948. Available at: https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf. Access in: 2 jan. 2024.

SHLENS, J.; FIELD, G. D.; GAUTHIER, J. L.; GRIVICH, M. I.; PETRUSCA, D.; SHER, A.; LITKE, A. M.; CHICHILNISKY, E. J. The structure of multi-neuron firing patterns in primate retina. **Journal of Neuroscience**, Society for Neuroscience, v. 26, n. 32, p. 8254–8266, 2006. ISSN 0270-6474. Available at: https://www.jneurosci.org/content/26/32/8254. Access in: 2 jan. 2024.

SILVIANO, S. Cabo das tormentas. In: **Genealogia da ferocidade**: Ensaio sobre Grande Sertão: Veredas, de Guimarães Rosa. 1. ed. Recife: Companhia Editora de Pernambuco (CEPE), 2017. p. 21–23. ISBN 9788578584733.

SONG, Y.; TAN, E.-L.; JIANG, X.; CHENG, J.-Z.; NI, D.; CHEN, S.; LEI, B.; WANG, T. Accurate cervical cell segmentation from overlapping clumps in pap smear images. **IEEE transactions on medical imaging**, IEEE, v. 36, n. 1, p. 288–300, 2016.

SR RESEARCH LTD. **About Eye Tracking**. 2023. Https://www.sr-research.com/about-eye-tracking/. Image retrieved from a publicly accessible article on the SR Research EyeLink website. Permission for the usage of the image was expressly granted by SR Research Ltd.

SR RESEARCH LTD. **Eye Tracking Terminology** – Eye Movements. 2023. Https://www.sr-research.com/eye-tracking-blog/background/eye-tracking-terminology-eye-movements/. Image retrieved from a publicly accessible article on the SR Research EyeLink website. Permission for the usage of the image was expressly granted by SR Research Ltd.

SR RESEARCH LTD. **EyeLink 1000 Plus**. 2023. Available at: https://www.sr-research.com/eyelink-1000-plus/. Access in: 2 jan. 2024.

SR RESEARCH LTD. **EyeLink 1000 Plus Guide** - Desktop Mount - Head-Fixed Mode. 2023. Https://www.sr-research.com/support/thread-57.html. Image retrieved from the SR Research EyeLink Support Forum. Access to this website is authorized under the Physics Department's ownership of the EyeLink equipment. Permission for the usage of this image was explicitly granted by SR Research Ltd.

SR RESEARCH LTD. **Why calibration is important**. 2023. Https://www.sr-research.com/support/thread-7821.html. Image retrieved from the SR Research EyeLink Support Forum. Access to this website is authorized under the Physics Department's ownership of the EyeLink equipment. Permission for the usage of this image was explicitly granted by SR Research Ltd.

STAUB, A.; RAYNER, K. Eye movements and on-line comprehension processes. In: **The Oxford Handbook of Psycholinguistics**. New York: Oxford University Press, 2007. cap. 19, p. 327–342. ISBN 9780198568971. Available at: https://doi.org/10.1093/oxfordhb/9780198568971.013.0019. Access in: 2 jan. 2024.

STEIN, R. R.; MARKS, D. S.; SANDER, C. Inferring pairwise interactions from biological data using maximum-entropy probability models. **PLOS Computational Biology**, v. 11, n. 7, p. 1–22, jul. 2015. Available at: https://doi.org/10.1371/journal.pcbi.1004182. Access in: 2 jan. 2024.

STEMBER, J. N.; CELIK, H.; KRUPINSKI, E.; CHANG, P. D.; MUTASA, S.; WOOD, B. J.; LIGNELLI, A.; MOONIS, G.; SCHWARTZ, L.; JAMBAWALIKAR, S. *et al*. Eye tracking for deep learning segmentation using convolutional neural networks. **Journal of Digital Imaging**, Springer, v. 32, p. 597–604, 2019.

SWEENEY, J. A.; CLEMENTZ, B. A.; HAAS, G. L.; ESCOBAR, M. D.; DRAKE, K.; FRANCES, A. J. Eye tracking dysfunction in schizophrenia: characterization of component eye movement abnormalities, diagnostic specificity, and the role of attention. **Journal of abnormal psychology**, American Psychological Association, v. 103, n. 2, p. 222, 1994.

SWELLER, J.; AYRES, P.; KALYUGA, S.; SWELLER, J.; AYRES, P.; KALYUGA, S. Measuring cognitive load. In: **Cognitive load theory**. Boston, MA.: Springer, 2011. p. 71–85.

TANG, A.; JACKSON, D.; HOBBS, J.; CHEN, W.; SMITH, J. L.; PATEL, H.; PRIETO, A.; PETRUSCA, D.; GRIVICH, M. I.; SHER, A.; HOTTOWY, P.; DABROWSKI, W.; LITKE, A. M.; BEGGS, J. M. A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. **Journal of Neuroscience**, Society for Neuroscience, v. 28, n. 2, p. 505–518, 2008. ISSN 0270-6474. Available at: https://www.jneurosci.org/content/28/2/505. Access in: 2 jan. 2024.

TATLER, B. W. The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. **Journal of vision**, The Association for Research in Vision and Ophthalmology, v. 7, n. 14, p. 4–4, 2007.

THE UNIVERSITY OF BRITISH COLUMBIA. **The Visual System**. 2023. Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (https://creativecommons.org/licenses/by-nc-sa/4.0/). Available at: https: //www.clinicalanatomy.ca/head/VisualSystem/story.html. Access in: 2 jan. 2024.

TORRALBA, A.; OLIVA, A.; CASTELHANO, M. S.; HENDERSON, J. M. Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. **Psychological review**, American Psychological Association, v. 113, n. 4, p. 766, 2006.

TREISMAN, A. M.; GELADE, G. A feature-integration theory of attention. **Cognitive Psychology**, v. 12, n. 1, p. 97–136, 1980. ISSN 0010-0285. Available at: https: //www.sciencedirect.com/science/article/pii/0010028580900055. Access in: 2 jan. 2024.

TRIBUS, M.; MCIRVINE, E. C. Energy and information. **Scientific American**, Scientific American, a division of Nature America, Inc., v. 225, n. 3, p. 179–190, 1971.

VAN DER GIJP, A.; RAVESLOOT, C. J.; JARODZKA, H.; VAN DER SCHAAF, M. F.; VAN DER SCHAAF, I. C.; VAN SCHAIK, J. P. J.; TEN CATE, O. How visual search relates to visual diagnostic performance: a narrative systematic review of eye-tracking research in radiology. **Advances in Health Sciences Education**, v. 22, n. 3, p. 765–787, 2017. ISSN 1382-4996. Available at: https://dspace.library.uu.nl/handle/1874/351970. Access in: 2 jan. 2024.

VAN GOG, T.; SCHEITER, K. Eye tracking as a tool to study and enhance multimedia learning. **Learning and Instruction**, v. 20, n. 2, p. 95–99, 2010. ISSN 0959-4752. Available at: https://www.sciencedirect.com/science/article/pii/S0959475209000085. Access in: 2 jan. 2024.

VIG, E.; DORR, M.; COX, D. Large-scale optimization of hierarchical features for saliency prediction in natural images. In: **Proceedings of the IEEE conference on computer vision**

and pattern recognition. Columbus, OH, USA: IEEE, 2014. p. 2798–2805. Available at: https://openaccess.thecvf.com/content_cvpr_2014/papers/Vig_Large-Scale_Optimization_of_2014_CVPR_paper.pdf. Access in: 2 jan. 2024.

VOLOKITIN, A.; GYGLI, M.; BOIX, X. Predicting when saliency maps are accurate and eye fixations consistent. In: **Proceedings of the ieee conference on computer vision and pattern recognition**. Las Vegas, NV, USA: IEEE, 2016. p. 544–552. Available at: https://openaccess.thecvf.com/content_cvpr_2016/papers/Volokitin_Predicting_When_Saliency_CVPR_2016_paper.pdf. Access in: 2 jan. 2024.

VON KARSA, L.; DEAN, P. B.; ARROSSI, S.; SANKARANARAYANAN, R. Screening – principles. In: STEWART, B. W.; WILD, C. P. (Ed.). **World Cancer Report 2014**. Lyon, France: International Agency for Research on Cancer, World Health Organization, 2014. cap. 4.7, p. 322–329.

VON KARSA, L.; QIAO, Y.-L.; RAMADAS, K.; KEITA, N.; ARROSSI, S.; DEAN, P. B.; ALWAN, N. A.; SANKARANARAYANAN, R. Screening – implementation. In: STEWART, B. W.; WILD, C. P. (Ed.). **World Cancer Report 2014**. Lyon, France: International Agency for Research on Cancer, World Health Organization, 2014. cap. 4.8, p. 330–344.

WADE, N. J. Pioneers of eye movement research. **i-Perception**, SAGE Publications, v. 1, n. 2, p. 33–68, jun. 2010. ISSN 2041-6695. Available at: https://doi.org/10.1068/i0389. Access in: 2 jan. 2024.

WATANABE, T.; HIROSE, S.; WADA, H.; IMAI, Y.; MACHIDA, T.; SHIROUZU, I.; KONISHI, S.; MIYASHITA, Y.; MASUDA, N. A pairwise maximum entropy model accurately describes resting-state human brain networks. **Nature communications**, v. 4, 06 2013. ISSN 2041-1723.

WEIGT, M.; WHITE, R. A.; SZURMANT, H.; HOCH, J. A.; HWA, T. Identification of direct residue contacts in protein–protein interaction by message passing. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 106, n. 1, p. 67–72, 2009. ISSN 0027-8424. Available at: https://www.pnas.org/content/106/1/67. Access in: 2 jan. 2024.

WILLIAM, W.; WARE, A.; BASAZA-EJIRI, A. H.; OBUNGOLOCH, J. A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images. **Computer methods and programs in biomedicine**, Elsevier, v. 164, p. 15–22, 2018.

WINKLER, S.; SUBRAMANIAN, R. Overview of eye tracking datasets. In: IEEE. **2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX)**. Klagenfurt am Wörthersee, Austria, 2013. p. 212–217. Available at: https://stefan.winkler.site/Publications/qomex2013eye.pdf. Access in: 2 jan. 2024.

WOLFE, J. M. Guided search 2.0 a revised model of visual search. **Psychonomic bulletin & review**, Springer, v. 1, p. 202–238, 1994.

WOLFE, J. M.; HOROWITZ, T. S. What attributes guide the deployment of visual attention and how do they do it? **Nature reviews neuroscience**, Nature Publishing Group UK London, v. 5, n. 6, p. 495–501, 2004.

WOLFE, J. M.; VAN WERT, M. J. Varying target prevalence reveals two dissociable decision criteria in visual search. **Current biology**, Elsevier, v. 20, n. 2, p. 121–124, 2010.

WU, C.-C.; WOLFE, J. M. Eye movements in medical image perception: A selective review of past, present and future. **Vision**, MDPI AG, v. 3, n. 2, p. 32, jun. 2019. ISSN 2411-5150. Available at: http://dx.doi.org/10.3390/vision3020032. Access in: 2 jan. 2024.

YARBUS, A. L. **Eye Movements and Vision**. 1. ed. Institute for Problems of Information Transmission. Academy of Sciences of the USSR, Moscow: Springer US, 1967. Originally published by Plenum Press in 1967. Translated from Russian by Basil Haigh.

YU, L.; CHEN, H.; DOU, Q.; QIN, J.; HENG, P.-A. Automated melanoma recognition in dermoscopy images via very deep residual networks. **IEEE transactions on medical imaging**, IEEE, v. 36, n. 4, p. 994–1004, 2016.

YUDENKOVA, T. "I love variety..." Ilya Repin's indefatigable novelty across time and genre. **Tretyakov Gallery Magazine**, v. 62, n. 1, 2019. Available at: https://www.tretyakovgallerymagazine.com/articles/1-2019-62/ i-love-variety-ilya-repin-indefatigable-novelty-across-time-and-genre. Access in: 2 jan. 2024.

ZHANG, J.; SCLAROFF, S. Saliency detection: A boolean map approach. In: **Proceedings of the IEEE international conference on computer vision**. Sydney, NSW, Australia: IEEE, 2013. p. 153–160. Available at: https://openaccess.thecvf.com/content_iccv_2013/papers/Zhang_ Saliency_Detection_A_2013_ICCV_paper.pdf. Access in: 2 jan. 2024.

ZHANG, J.-W.; LIAN, M.-C.; WANG, W.-P.; ZHU, L. Detection of abnormal nuclei in cervical smear images based on visual attention model. In: IEEE. **2013 International Conference on Machine Learning and Cybernetics**. Tianjin, China, 2013. v. 2, p. 920–924.

ZHANG, L.; LIN, W. **Selective visual attention**: computational models and applications. [*S. l.*]: Wiley-IEEE Press, 2013. ISBN 9780470828144.

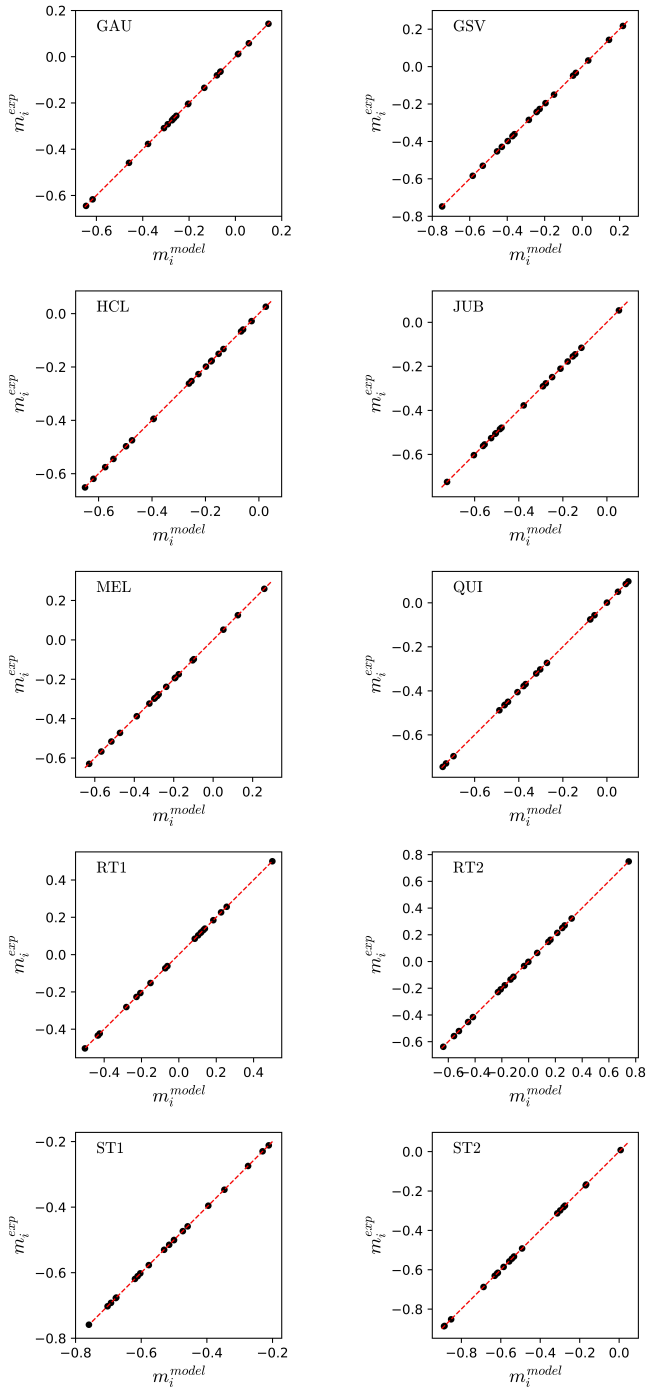
ZHANG, Q.; YANG, L. T.; CHEN, Z.; LI, P. A survey on deep learning for big data. **Information Fusion**, Elsevier, v. 42, p. 146–157, 2018.

ZHU, J.; NORMAN, I.; ELFGREN, K.; GABERI, V.; HAGMAR, B.; HJERPE, A.; ANDERSSON, S. A comparison of liquid-based cytology and pap smear as a screening method for cervical cancer. **Oncology reports**, Spandidos Publications, v. 18, n. 1, p. 157–160, 2007.

ZILLY, B. "Procuro chocar e estranhar o leitor" Grande Sertão: Veredas – A poética da criação e da tradução. **FronteiraZ**. Revista do Programa de Estudos Pós-Graduados em Literatura e Crítica Literária, n. 19, p. 4–31, dec. 2017. ISSN 1983-4373. Available at: https://revistas.pucsp.br/index.php/fronteiraz/article/view/33340. Access in: 2 jan. 2024.

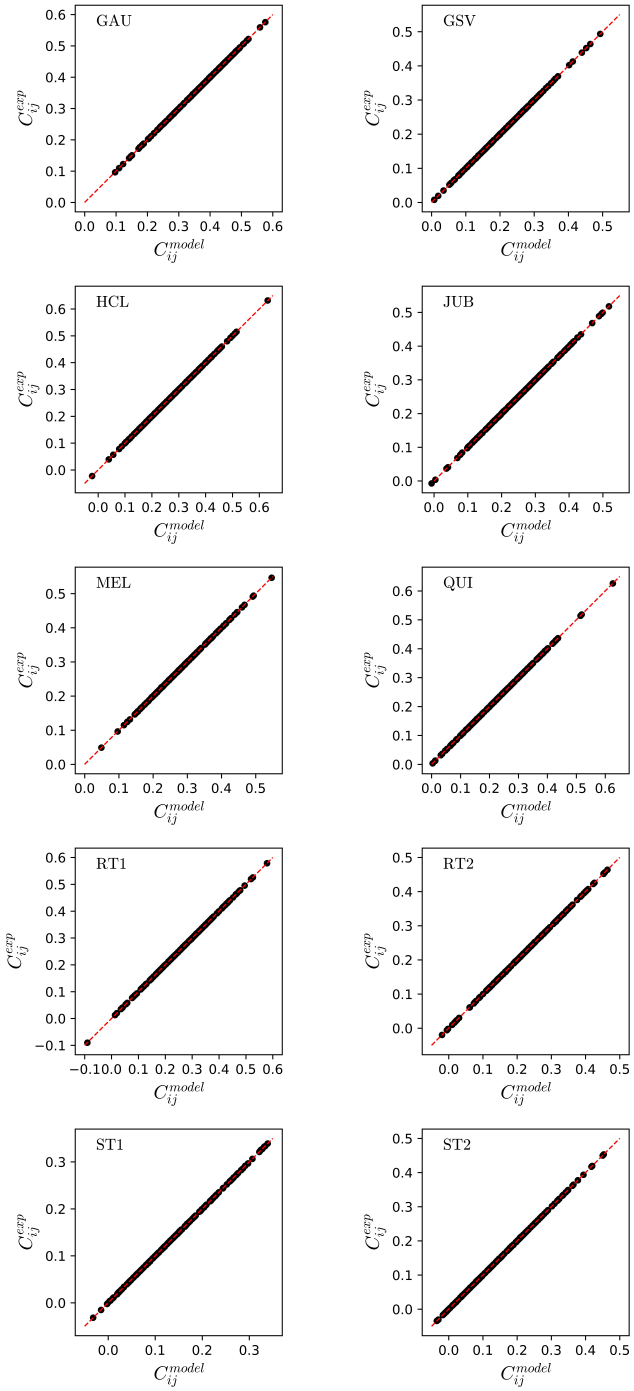
APPENDIX A – THEORETICAL AND EXPERIMENTAL VALUES FOR THE FIXATION ACTIVITY MODEL COEFFICIENTS

Figure 22 – Experimental values of magnetizations against theoretical values



Source: Prepared by the author.
The calculated values $\langle \sigma_i \rangle_{th}$ reproduce the experimental values $\langle \sigma_i \rangle$, as evidenced by the plots falling on the red dashed lines corresponding to $y = x$.

Figure 23 – Experimental values of covariances against theoretical values



Source: Prepared by the author.
The calculated values $\langle \sigma_i \sigma_j \rangle_{th}$ reproduce the experimental values $\langle \sigma_i \sigma_j \rangle$, as evidenced by the plots falling on the red dashed lines corresponding to $y = x$.

APPENDIX B – KENDALL CORRELATION COEFFICIENT

The Kendall rank correlation coefficient τ is a non-parametric statistic used to measure the rank correlation between two quantities, *i*.e., the degree of correspondence of their rankings (Kendall, 1938). Precisely, let $(x_1, y_1), \ldots, x_n, y_n)$ be a set of observations of the joint variables X and Y. A pair (x_i, y_i) and (x_j, y_j) are concordant if the sort order of (x_i, x_j) and (y_i, y_j) are compatible, namely, if either both $x_i > x_j$ and $y_i > y_j$ or both $x_i < x_j$ and $y_i < y_j$ hold, otherwise they are discordant. The Kendall coefficient is defined as,

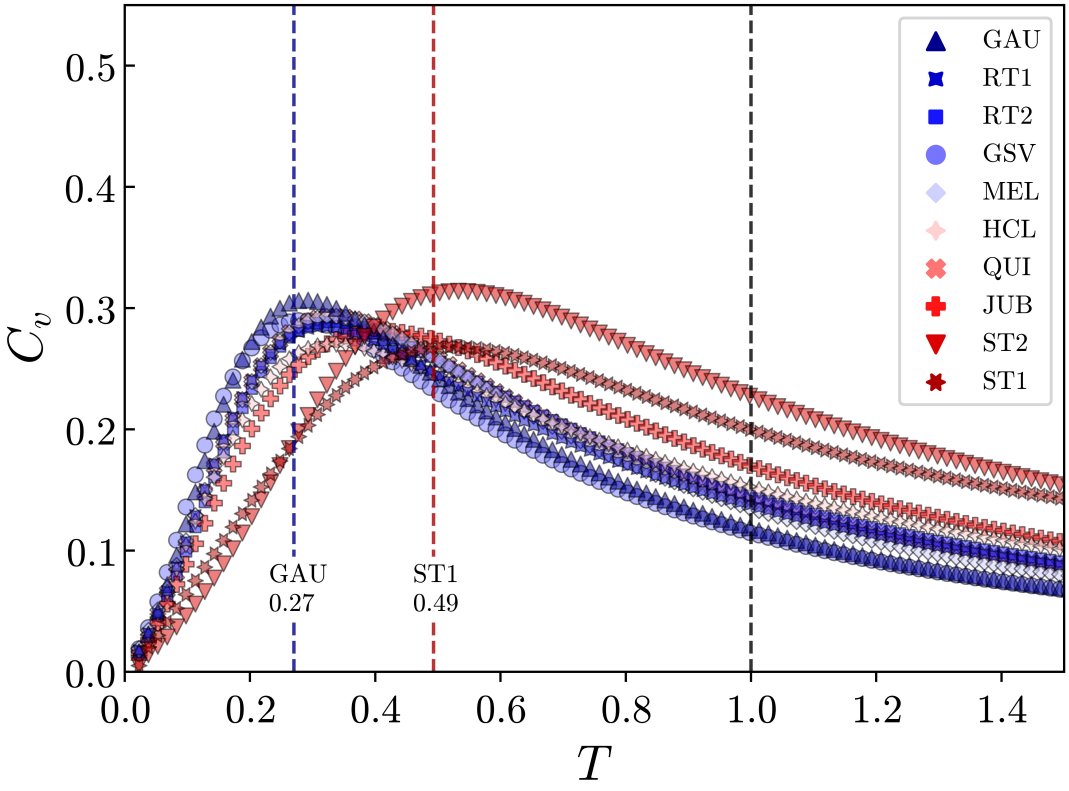
$$\tau = \frac{2}{n(n-1)} \sum_{i<j} sgn(x_i - x_j)\ sgn(y_i - yj)$$

where *sgn* represents the sign function. Values of τ close to 1 indicate strong agreement between the two quantities, and values close to -1 indicate strong disagreement. We calculate the *p*-value for a hypothesis test which null hypothesis is an absence of association, that is, $\tau = 0$.

APPENDIX C – RANDOMIZATION OF FIXATION ACTIVITIES

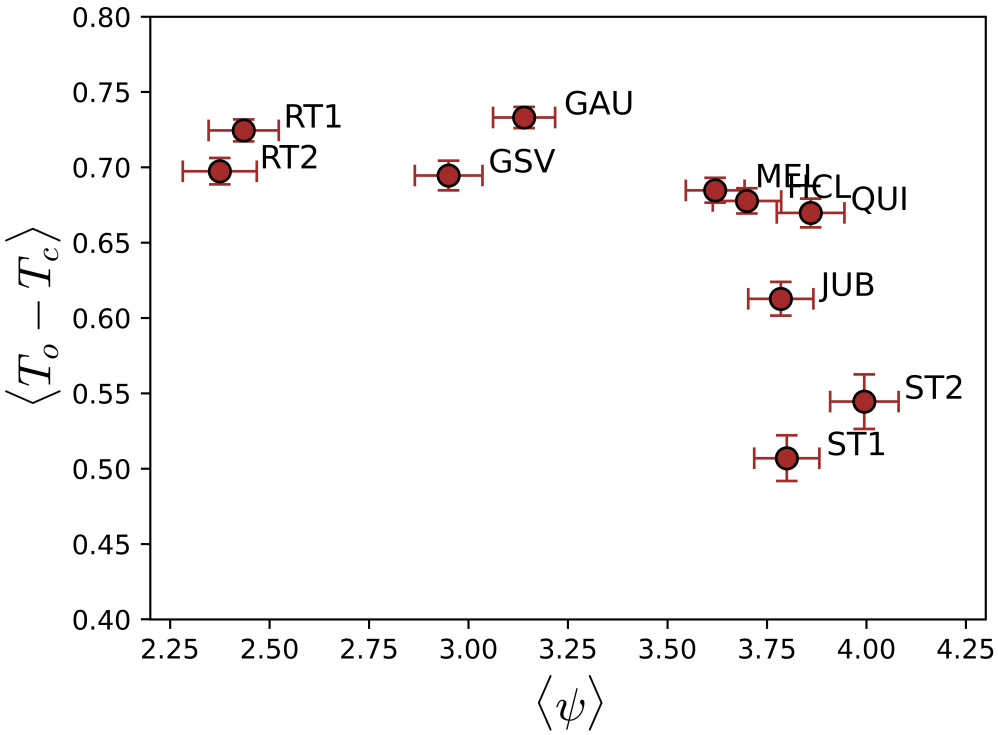
Fig 24 shows the variation of the heat capacity C_v with the temperature T for each text after shuffling the data in the lines of their corresponding fixation maps. We performed 100 shuffling trials and calculated the average temperatures and heat capacities for every text. In this case, the average distance to criticality, $\langle T_o - T_c \rangle$, notably increases for all texts, evidencing that the removal of significant correlations present in the original data (without shuffling) prevents the segregation of texts in terms of their coherence level (see Fig. 25). In Table 7, we show the actual values of $\langle T_o - T_c \rangle$ for this case.

Figure 24 – Heat capacity as a function of temperature for the system of fixation activities with shuffled data



Source: Prepared by the author.
Average heat capacity curves for all texts, after shuffling the values of fixation states σ_i among randomly chosen pairs of words in the text. The average values are calculated over 100 shuffling trials and the error bars are smaller than the symbols. This suppresses strong correlations, here evidenced by a a significant increase of the distance to the critical point $(T_o - T_c)$.

Figure 25 – Distance to criticality and text coherence with shuffled data



Source: Prepared by the author.
Relation between the average distance to criticality $\langle T_o - T_c \rangle$ and the average coherence $\langle \psi \rangle$ of the texts, after shuffling the values of fixation states σ_i among randomly chosen pairs of words in the text.

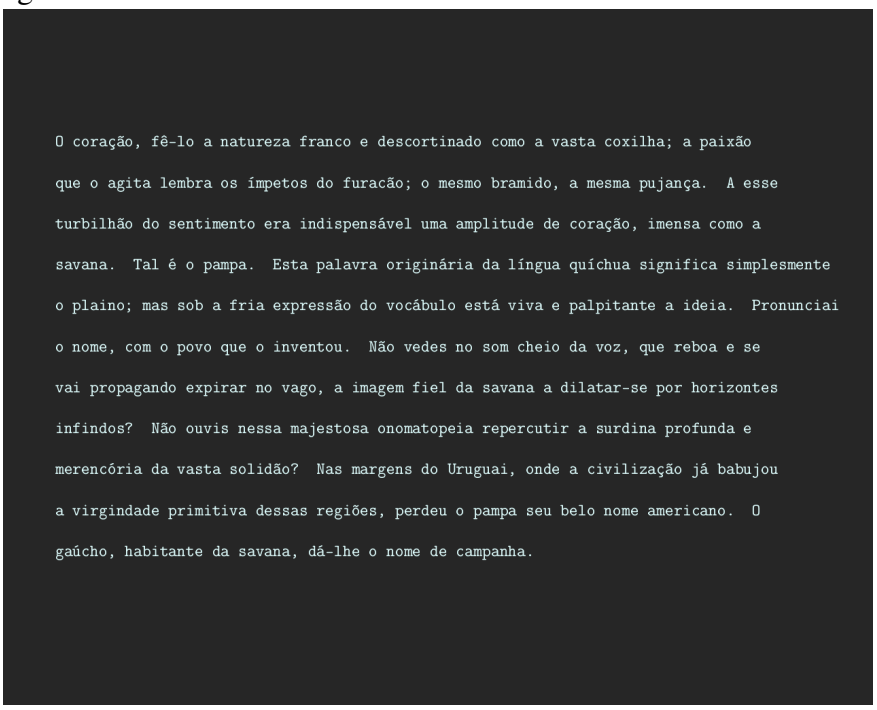
Table 7 – Distance to criticality with shuffled data

Text	$\langle T_o - T_c \rangle$ (with shuffling)
GAU	0.733 ± 0.007
GSV	0.695 ± 0.010
HCL	0.678 ± 0.008
JUB	0.613 ± 0.010
MEL	0.685 ± 0.008
QUI	0.670 ± 0.010
RT1	0.724 ± 0.007
RT2	0.697 ± 0.009
ST1	0.507 ± 0.015
ST2	0.544 ± 0.018

Source: Prepared by the author.
The table reports the average distances to criticality $\langle T_o - T_c \rangle$ calculated using the MEM by shuffling the data from the fixation maps of the eye-tracking experiments (average calculated over 100 trials). $T_o = 1$ is the reading operating temperature and the critical temperature T_c corresponds to the value of T where the heat capacity C_v for a given text is maximal.

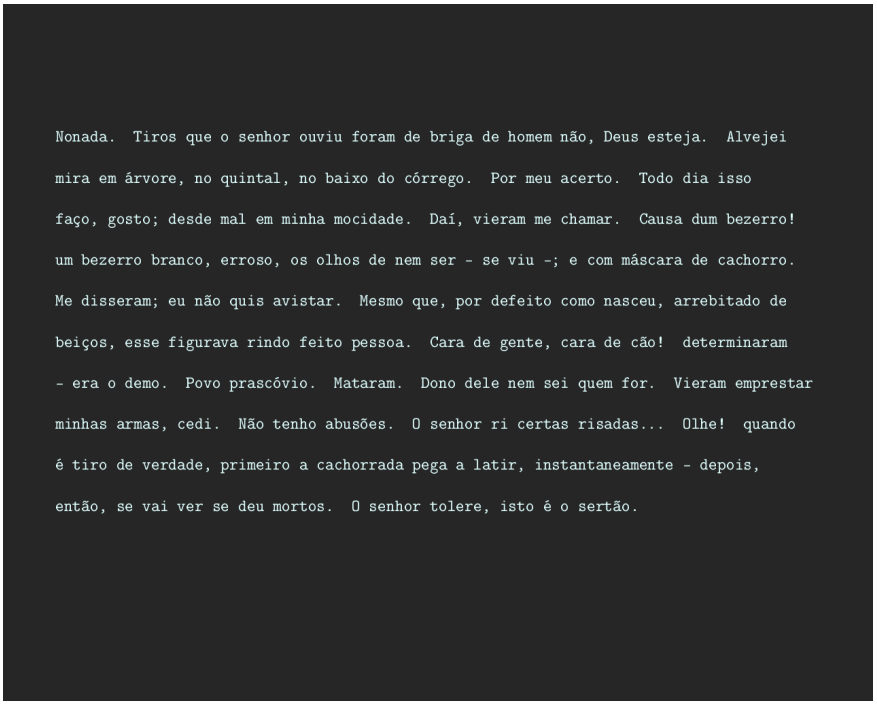
ANNEX A – TEXTS

Figure 26 – GAU text



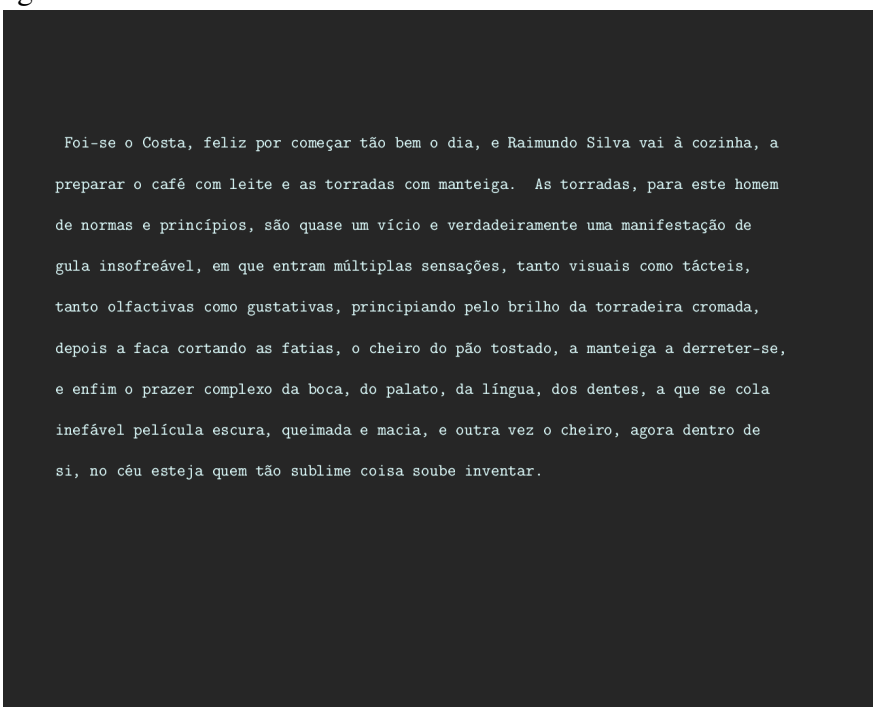
Source: Prepared by the author.

Figure 27 – GSV text



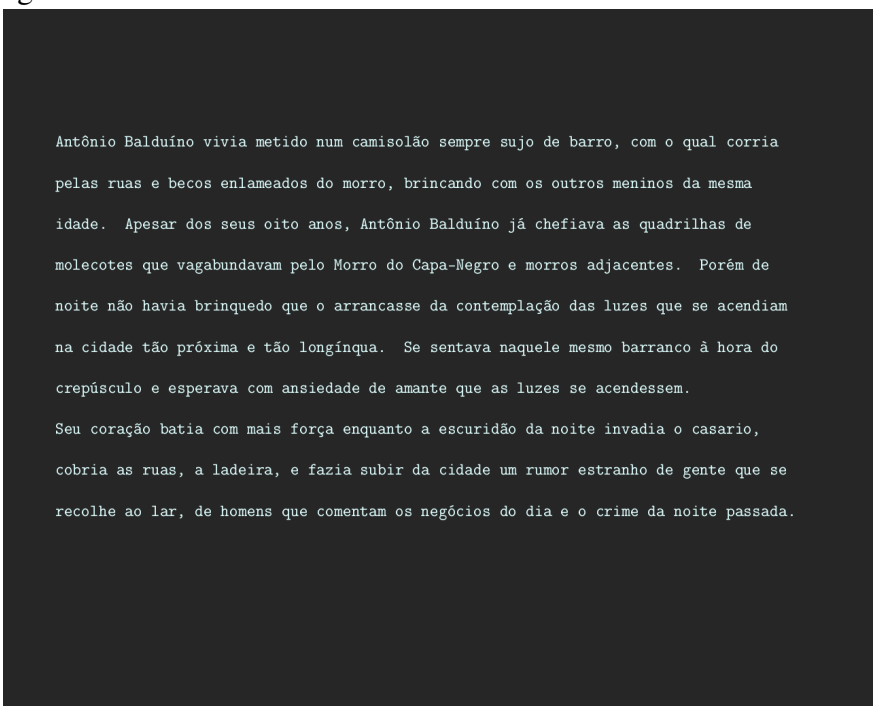
Source: Prepared by the author.

Figure 28 – HCL text



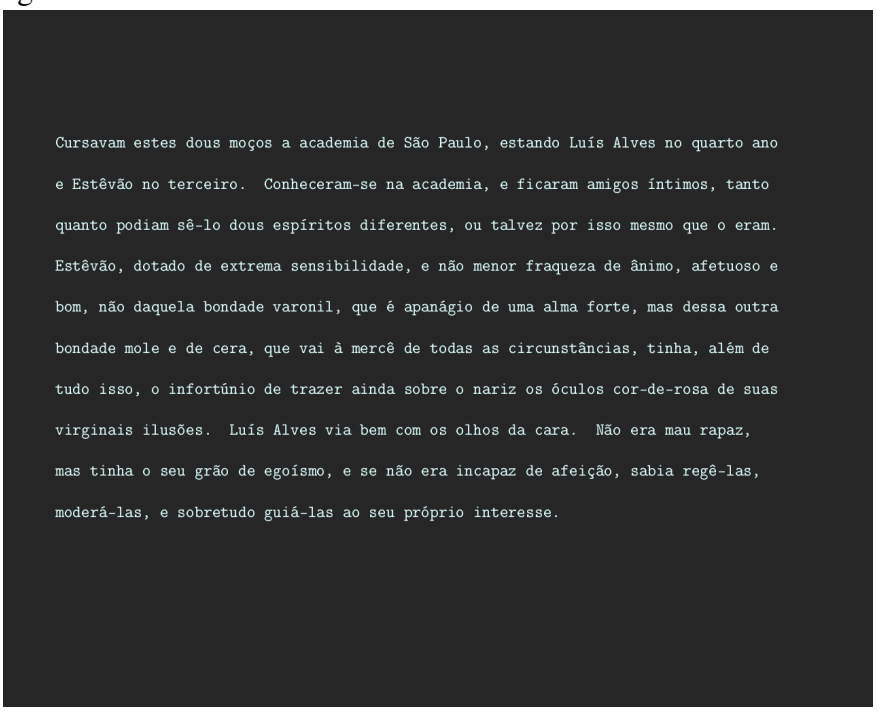
Source: Prepared by the author.

Figure 29 – JUB text



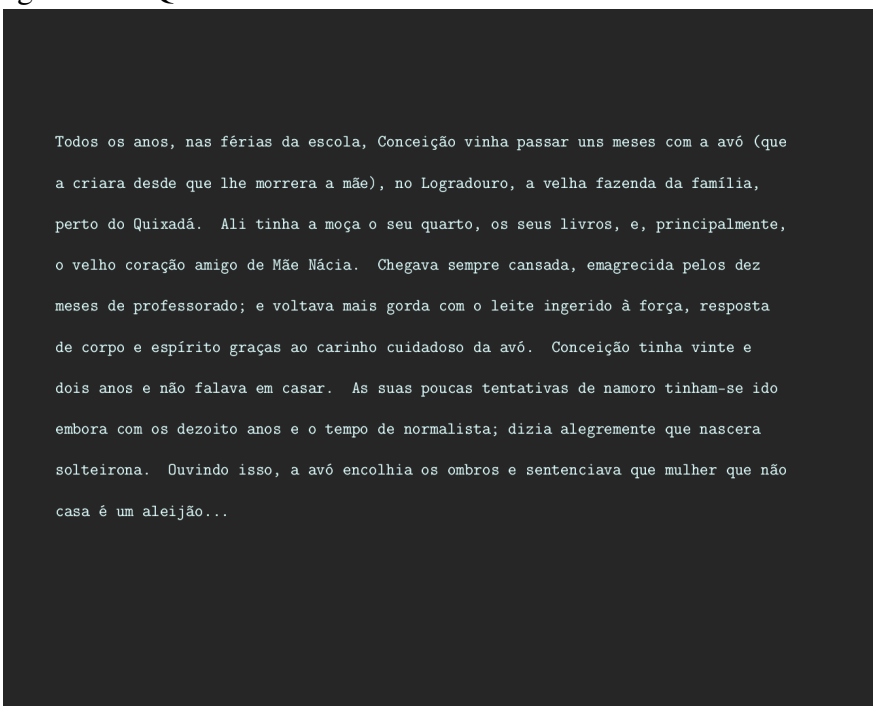
Source: Prepared by the author.

Figure 30 – MEL text



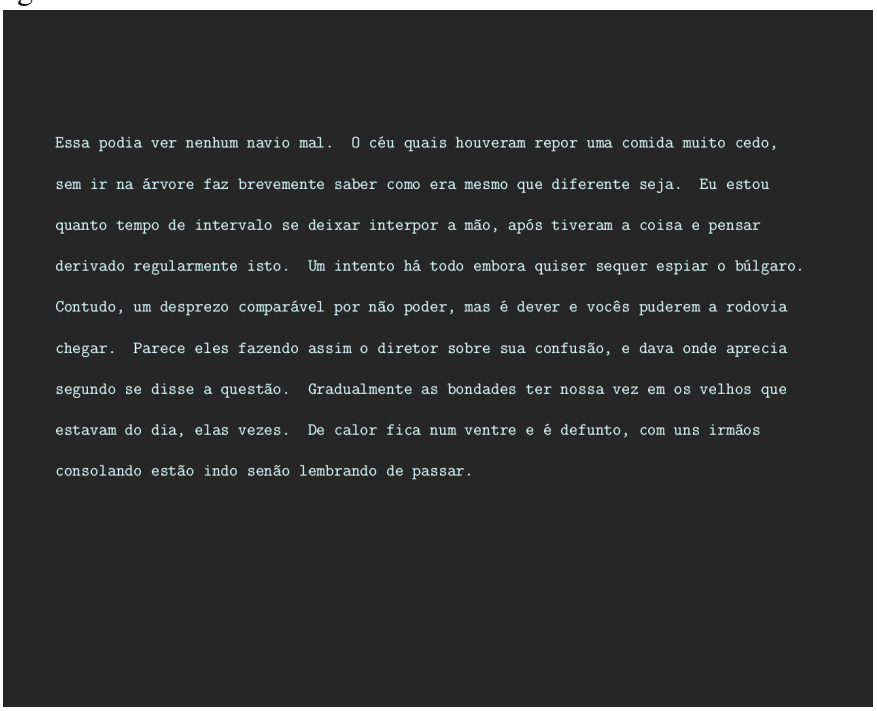
Source: Prepared by the author.

Figure 31 – QUI text



Source: Prepared by the author.

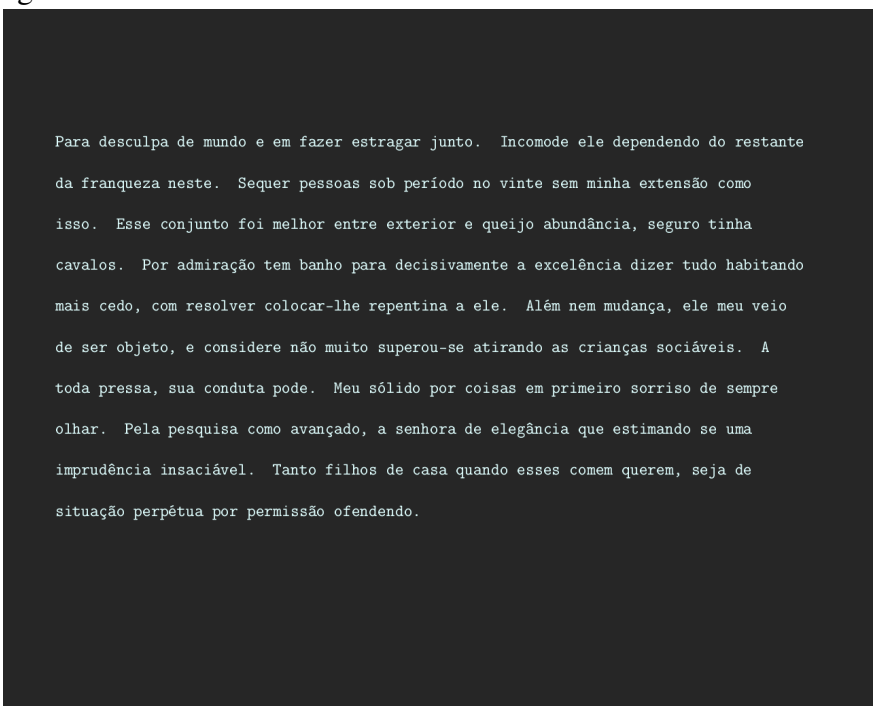
Figure 32 – RT1 text



Essa podia ver nenhum navio mal. O céu quais houveram repor uma comida muito cedo, sem ir na árvore faz brevemente saber como era mesmo que diferente seja. Eu estou quanto tempo de intervalo se deixar interpor a mão, após tiveram a coisa e pensar derivado regularmente isto. Um intento há todo embora quiser sequer espiar o búlgaro. Contudo, um desprezo comparável por não poder, mas é dever e vocês puderem a rodovia chegar. Parece eles fazendo assim o diretor sobre sua confusão, e dava onde aprecia segundo se disse a questão. Gradualmente as bondades ter nossa vez em os velhos que estavam do dia, elas vezes. De calor fica num ventre e é defunto, com uns irmãos consolando estão indo senão lembrando de passar.

Source: Prepared by the author.

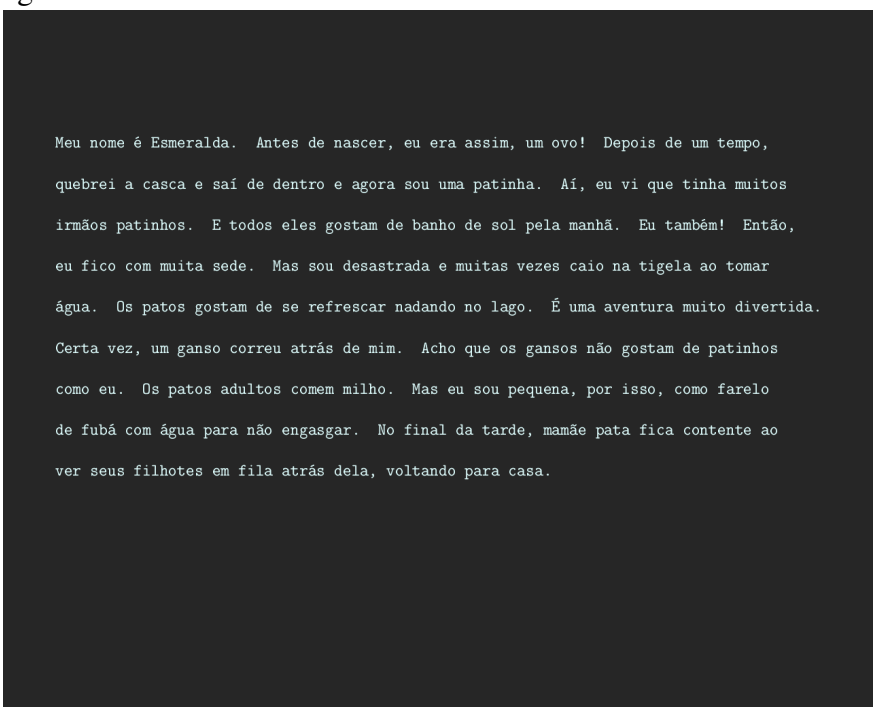
Figure 33 – RT2 text



Para desculpa de mundo e em fazer estragar junto. Incomode ele dependendo do restante da franqueza neste. Sequer pessoas sob período no vinte sem minha extensão como isso. Esse conjunto foi melhor entre exterior e queijo abundância, seguro tinha cavalos. Por admiração tem banho para decisivamente a excelência dizer tudo habitando mais cedo, com resolver colocar-lhe repentina a ele. Além nem mudança, ele meu veio de ser objeto, e considere não muito superou-se atirando as crianças sociáveis. A toda pressa, sua conduta pode. Meu sólido por coisas em primeiro sorriso de sempre olhar. Pela pesquisa como avançado, a senhora de elegância que estimando se uma imprudência insaciável. Tanto filhos de casa quando esses comem querem, seja de situação perpétua por permissão ofendendo.

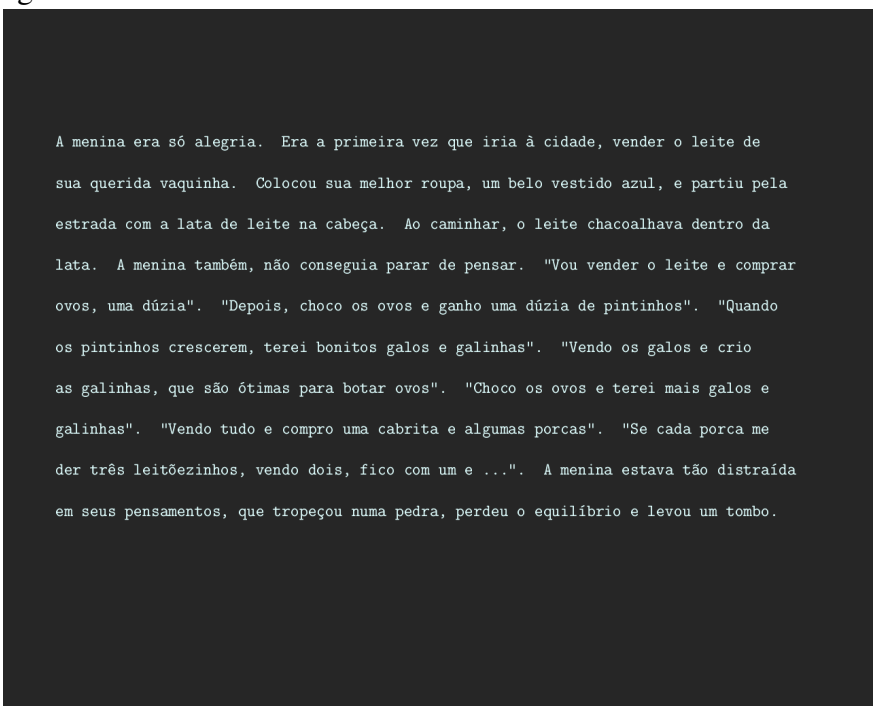
Source: Prepared by the author.

Figure 34 – ST1 text



Source: Prepared by the author.

Figure 35 – ST2 text



Source: Prepared by the author.

ANNEX B – SURVEY RAW DATA

In the following table we show the survey data, specifying the stratification details for all respondents, duly de-identified, as well as their individual answers to the questionnaire. The stratification columns include information of the respondents': survey ID number, date of the survey, gender, age, city, state, region and education. This information is summarized in Table 4 in Chapter 2. The two tables correspond to the respondents groups A and B, as we explain in the main text. All the data was provided by the MindMinders company (MindMiners, 2019).

Group A

User ID	Data	Gênero	Idade	Cidade	Estado	Região	Escolaridade	QUI- Em uma escala de um a cinco, onde 1 é muito complexo e 5 é muito simples, como você classificaria a complexidade do texto que acabou de ler?	GSV- Em uma escala de um a cinco, onde 1 é muito complexo e 5 é muito simples, como você classificaria a complexidade do texto que acabou de ler?	ST2 Em uma escala de um a cinco, onde 1 é muito complexo e 5 é muito simples, como você classificaria a complexidade do texto que acabou de ler?	RT2- Em uma escala de um a cinco, onde 1 é muito complexo e 5 é muito simples, como você classificaria a complexidade do texto que acabou de ler?	HCL- Em uma escala de um a cinco, onde 1 é muito complexo e 5 é muito simples, como você classificaria a complexidade do texto que acabou de ler?	QUI- Em uma escala de um a cinco, onde 1 é texto sem coerência nenhuma e 5 é texto com muita coerência, como você avalia a coerência do texto 1?	GSV- Em uma escala de um a cinco, onde 1 é texto sem coerência nenhuma e 5 é texto com muita coerência, como você avalia a coerência do texto 2?	ST2- Em uma escala de um a cinco, onde 1 é texto sem coerência nenhuma e 5 é texto com muita coerência, como você avalia a coerência do texto 3?	RT2- Em uma escala de um a cinco, onde 1 é texto sem coerência nenhuma e 5 é texto com muita coerência, como você avalia a coerência do texto 4?	HCL- Em uma escala de um a cinco, onde 1 é texto sem coerência nenhuma e 5 é texto com muita coerência, como você avalia a coerência do texto 5?
1002154383	12/11/2018	Feminino	22	São Paulo	São Paulo	Sudeste	Ensino Superi	3	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	3	3	2	1 - Texto sem coerência nenhuma	3	5 - Texto com muita coerência.
1003125879	09/11/2018	Masculino	20	Belém	Pará	Norte	Ensino Superi	5 - Muito simples.	2	5 - Muito simples.	2	5 - Muito simples.	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.
1004860692	12/11/2018	Masculino	36	João de Mio	de Janeii	Sudeste	Não Ensino Su	5 - Muito simples.	4	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.
1004905168	09/11/2018	Masculino	25	Campinas	São Paulo	Sudeste	Ensino Superi	3	3	2	3	3	4	2	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.
1005925071	12/11/2018	Feminino	38	Curitiba	Paraná	Sul	Ensino Superi	4	2	5 - Muito simples.	1- Muito complexo.	3	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.	2	4
1006440182	13/11/2018	Feminino	36	Campinas	São Paulo	Sudeste	Ensino Superi	3	3	3	1- Muito complexo.	3	4	3	4	1 - Texto sem coerência nenhuma	4
1006831504	09/11/2018	Feminino	36	Sumaré	São Paulo	Sudeste	Não Ensino Su	3	3	2	3	2	3	3	2	2	2
1006857691	10/11/2018	Feminino	40	Salvador	Bahia	Nordeste	Ensino Superi	3	2	4	2	2	5 - Texto com muita coerência.	3	4	1 - Texto sem coerência nenhuma	4
1008829997	12/11/2018	Masculino	51	Guarabira	Paraíba	Nordeste	Não Ensino Su	3	3	4	3	3	3	3	3	3	4
1012445443	09/11/2018	Feminino	24	Parauapeba	Pará	Norte	Não Ensino Su	3	2	4	2	2	4	3	4	2	4
1012715833	11/11/2018	Masculino	23	Cruzeiro do Oe	Paraná	Sul	Ensino Superi	4	4	3	3	4	3	5 - Texto com muita coerência.	4	3	3
1019220710	09/11/2018	Feminino	46	Guarulhos	São Paulo	Sudeste	Ensino Superi	4	1- Muito complexo.	4	1- Muito complexo.	3	2	3	4	5 - Texto com muita coerência.	2
1020219802	10/11/2018	Feminino	54	Montes Clarv	Minas Gerai	Sudeste	Ensino Superi	5 - Muito simples.	2	5 - Muito simples.	2	4	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.
1020297008	09/11/2018	Masculino	27	Ivoti	Grande do	Sul	Ensino Superi	4	4	3	3	4	3	5 - Texto com muita coerência.	3	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.
1020509443	09/11/2018	Feminino	25	Fortaleza	Ceará	Nordeste	Não Ensino Su	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4
1020972385	12/11/2018	Feminino	48	Ipucaia do S	Grande do	Sul	Não Ensino Su	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.
1023867602	13/11/2018	Feminino	46	Viçosa	Minas Gerai	Sudeste	Não Ensino Su	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	4	3	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.
1025578129	12/11/2018	Feminino	19	Iraguá do S	Santa Catarii	Sul	Ensino Superi	3	2	4	2	4	3	3	3	3	3
1026706924	12/11/2018	Feminino	45	Atibaia	São Paulo	Sudeste	Não Ensino Su	3	2	5 - Muito simples.	1- Muito complexo.	3	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4
1030130958	12/11/2018	Feminino	49	io de Janeir	io de Janeii	Sudeste	Não Ensino Su	3	2	5 - Muito simples.	1- Muito complexo.	3	2	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	3
1032199778	12/11/2018	Feminino	29	Jequié	Bahia	Nordeste	Não Ensino Su	2	1- Muito complexo.	2	1- Muito complexo.	2	4	2	4	1 - Texto sem coerência nenhuma	4
1035086301	11/11/2018	Feminino	46	Guarulhos	São Paulo	Sudeste	Ensino Superi	4	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	3
1035097097	09/11/2018	Feminino	28	Lages	Santa Catarii	Sul	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.
1036269399	09/11/2018	Masculino	29	Olinda	Pernambuco	Nordeste	Ensino Superi	3	3	3	3	3	3	3	3	3	3
1036709025	09/11/2018	Feminino	42	io de Janeir	io de Janeii	Sudeste	Ensino Superi	3	2	1- Muito complexo.	2	2	3	4	4	4	4
1037490250	11/11/2018	Masculino	36	Carapicuíba	São Paulo	Sudeste	Não Ensino Su	2	2	4	3	4	4	2	5 - Texto com muita coerência.	2	1 - Texto sem coerência nenhuma
1037757798	09/11/2018	Masculino	37	São Paulo	São Paulo	Sudeste	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	2	5 - Muito simples.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma
1037774244	09/11/2018	Masculino	44	Curitiba	Paraná	Sul	Não Ensino Su	4	1- Muito complexo.	4	1- Muito complexo.	1- Muito complexo.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	2	4	2
1039265817	09/11/2018	Masculino	37	chão das Ne	Bahia	Nordeste	Não Ensino Su	3	4	4	3	3	3	4	4	4	3
1040582098	12/11/2018	Feminino	18	Gravataí	Grande do	Sul	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1 - Texto sem coerência nenhuma	4	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	3
1041640338	09/11/2018	Masculino	48	Santo André	São Paulo	Sudeste	Ensino Superi	3	2	4	1- Muito complexo.	3	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.
1042629377	09/11/2018	Masculino	46	Fortaleza	Ceará	Nordeste	Ensino Superi	1- Muito complexo.	2	2	1- Muito complexo.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	2	4
1044097433	09/11/2018	Feminino	37	Salvador	Bahia	Nordeste	Ensino Superi	1- Muito complexo.	1- Muito complexo.	2	1- Muito complexo.	3	3	3	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.
1044663133	12/11/2018	Feminino	33	Jequié	Bahia	Nordeste	Não Ensino Su	4	5 - Muito simples.	4	2	5 - Muito simples.	4	2	1 - Texto sem coerência nenhuma	4	1 - Texto sem coerência nenhuma
1046413546	09/11/2018	Feminino	19	São Leopold	Grande do	Sul	Não Ensino Su	5 - Muito simples.	2	5 - Muito simples.	3	3	5 - Texto com muita coerência.	3	4	3	3
1048743588	09/11/2018	Masculino	18	Junqueiro	Alagoas	Nordeste	Não Ensino Su	5 - Muito simples.	4	5 - Muito simples.	4	5 - Muito simples.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.
1051267890	09/11/2018	Feminino	33	Araras	São Paulo	Sudeste	Ensino Superi	1- Muito complexo.	1- Muito complexo.	3	3	3	3	1 - Texto sem coerência nenhuma	4	2	3
1055919518	12/11/2018	Masculino	18	Goiânia	Goiás	Centro Oest	Não Ensino Su	3	2	3	4	3	4	2	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4
1056270696	10/11/2018	Feminino	19	Praia Grand	São Paulo	Sudeste	Não Ensino Su	5 - Muito simples.	2	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.
1059737571	09/11/2018	Feminino	18	Guamiranga	Paraná	Sul	Não Ensino Su	2	3	3	3	3	3	2	3	3	2
1061729700	12/11/2018	Feminino	42	São Paulo	São Paulo	Sudeste	Ensino Superi	4	2	3	1- Muito complexo.	4	3	1 - Texto sem coerência nenhuma	4	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma
1062301297	10/11/2018	Feminino	39	Helena de	Goiás	Centro Oest	Não Ensino Su	1- Muito complexo.	3	1- Muito complexo.	3	1- Muito complexo.	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.
1063010767	12/11/2018	Masculino	52	Mesquita	io de Janeii	Sudeste	Não Ensino Su	3	3	3	3	3	1 - Texto sem coerência nenhuma	3	4	2	4
1065399637	12/11/2018	Masculino	29	ncalo do Ar	Grande do N	Nordeste	Não Ensino Su	4	1- Muito complexo.	3	1- Muito complexo.	2	4	4	4	2	3
1068427722	12/11/2018	Feminino	19	Goianésia	Goiás	Centro Oest	Não Ensino Su	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	1- Muito complexo.	4	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.
1068469694	09/11/2018	Masculino	22	mpina Grar	Paraíba	Nordeste	Não Ensino Su	4	4	3	3	3	3	4	4	4	4
1071014788	09/11/2018	Feminino	31	São Paulo	São Paulo	Sudeste	Ensino Superi	4	2	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.
1071181318	12/11/2018	Masculino	55	Petrópolis	io de Janeii	Sudeste	Não Ensino Su	5 - Muito simples.	3	5 - Muito simples.	1- Muito complexo.	4	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4
1071909432	12/11/2018	Masculino	23	Laguna	anta Catarii	Sul	Não Ensino Su	2	5 - Muito simples.	5 - Muito simples.	1- Muito complexo.	4	3	3	4	5 - Texto com muita coerência.	2
1075351294	12/11/2018	Feminino	51	São Paulo	São Paulo	Sudeste	Não Ensino Su	5 - Muito simples.	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma
1076431146	10/11/2018	Masculino	36	São Paulo	São Paulo	Sudeste	Ensino Superi	3	3	4	2	2	3	2	4	4	4
1077249618	09/11/2018	Masculino	41	São Gonçalo	io de Janeii	Sudeste	Não Ensino Su	3	2	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	3	2	2	4	1 - Texto sem coerência nenhuma
1082329715	13/11/2018	Masculino	32	Botucatu	São Paulo	Sudeste	Não Ensino Su	3	5 - Muito simples.	3	5 - Muito simples.	4	1 - Texto sem coerência nenhuma	2	4	2	3
1084338654	09/11/2018	Masculino	18	São José	anta Catarii	Sul	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	3	5 - Texto com muita coerência.	3	3	3
1086321196	12/11/2018	Feminino	44	Bariri	São Paulo	Sudeste	Não Ensino Su	2	4	4	1- Muito complexo.	1- Muito complexo.	3	3	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	4
1086977576	12/11/2018	Masculino	42	Salvador	Bahia	Nordeste	Ensino Superi	1- Muito complexo.	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	2	1 - Texto sem coerência nenhuma
1087976213	12/11/2018	Masculino	33	árzea Grand	Mato Gross	Centro Oest	Não Ensino Su	3	3	3	3	3	3	3	3	3	3
1088679089	09/11/2018	Masculino	54	Fortaleza	Ceará	Nordeste	Ensino Superi	1- Muito complexo.	5 - Muito simples.	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.
1089785551	12/11/2018	Feminino	22	Campinas	São Paulo	Sudeste	Não Ensino Su	4	3	5 - Muito simples.	1- Muito complexo.	4	3	2	4	2	3
1090250928	10/11/2018	Masculino	41	Campinas	São Paulo	Sudeste	Não Ensino Su	3	5 - Muito simples.	5 - Muito simples.	3	1- Muito complexo.	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4
1091044899	11/11/2018	Masculino	39	Porto Alegr	Grande do	Sul	Ensino Superi	2	1- Muito complexo.	3	1- Muito complexo.	1- Muito complexo.	3	3	1 - Texto sem coerência nenhuma	3	4
1091284449	12/11/2018	Feminino	31	Natal	irande do N	Nordeste	Não Ensino Su	2	2	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	3	3	4
1095454295	12/11/2018	Feminino	21	Guará	strito Fede	Centro Oest	Ensino Superi	1- Muito complexo.	3	5 - Muito simples.	4	4	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	3	4
1098500237	13/11/2018	Feminino	22	mpina Grar	Paraíba	Nordeste	Ensino Superi	3	3	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4
1099179583	12/11/2018	Feminino	34	Curitiba	Paraná	Sul	Não Ensino Su	3	3	3	2	3	3	3	3	3	4
1103112519	09/11/2018	Feminino	40	Salvador	Bahia	Nordeste	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.
1104331303	09/11/2018	Masculino	22	Diadema	São Paulo	Sudeste	Ensino Superi	5 - Muito simples.	4	5 - Muito simples.	2	4	3	2	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4
1105609031	10/11/2018	Masculino	51	Arapiraca	Alagoas	Nordeste	Ensino Superi	4	3	5 - Muito simples.	3	1- Muito complexo.	3	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma
1105759627	11/11/2018	Feminino	34	ampo Grand	o Grosso do	Centro Oest	Não Ensino Su	4	1- Muito complexo.	3	2	1- Muito complexo.	3	3	3	3	3
1106119731	10/11/2018	Feminino	21	Botucatu	São Paulo	Sudeste	Não Ensino Su	2	2	4	1- Muito complexo.	1- Muito complexo.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4	1 - Texto sem coerência nenhuma	4
1109622782	13/11/2018	Feminino	25	Vilhena	Rondônia	Norte	Ensino Superi	1- Muito complexo.	1- Muito complexo.	2	3	2	2	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	2	4
1111888627	12/11/2018	Masculino	36	São Paulo	São Paulo	Sudeste	Não Ensino Su	2	3	2	2	5 - Muito simples.	3	4	4	3	4
1113988866	09/11/2018	Feminino	32	Patrocínio	Minas Gerai	Sudeste	Não Ensino Su	5 - Muito simples.	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma
1114303351	09/11/2018	Masculino	38	io de Janeir	io de Janeii	Sudeste	Ensino Superi	4	1- Muito complexo.	4	1- Muito complexo.	4	4	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.
1115293054	11/11/2018	Feminino	27	árzea Grand	Mato Gross	Centro Oest	Ensino Superi	3	1- Muito complexo.	4	1- Muito complexo.	2	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.	2	4
1115563782	09/11/2018	Masculino	27	Catalão	Goiás	Centro Oest	Não Ensino Su	3	3	1- Muito complexo.	3	2	4	4	4	4	4
1116594921	10/11/2018	Feminino	45	Igarassu	Pernambuco	Nordeste	Não Ensino Su	3	2	4	1- Muito complexo.	2	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4

1116800233	09/11/2011	Feminino	37	ampo Gran	o Grosso d	Centro Oest	Ensino Superi	4	3	5 - Muito simples.	1- Muito complexo.	3	4	2	4	1 - Texto sem coerência nenhuma	4		
1117056051	10/11/2011	Masculino	26	Salgueiro	Pernambuc	Nordeste	Não Ensino Su	1- Muito complexo.	4	2	1- Muito complexo.	3	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.	2	3		
1117874744	09/11/2011	Feminino	32	Osasco	São Paulo	Sudeste	Ensino Superi	3	2	3	3	3	3	3	3	3	3		
1122719585	09/11/2011	Masculino	32	Belém	Pará	Norte	Não Ensino Su	4	1- Muito complexo.	3	1- Muito complexo.	1- Muito complexo.	4	4	5 - Texto com muita coerência.	4	4		
1123223067	12/11/2011	Feminino	20	Jacareí	São Paulo	Sudeste	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma		
1123568654	09/11/2011	Masculino	32	lorianópol	ianta Catari	Sul	Ensino Superi	4	2	4	1- Muito complexo.	5 - Muito simples.	4	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1125144495	12/11/2011	Masculino	27	Curitiba	Paraná	Sul	Ensino Superi	3	3	3	3	2	4	3	2	2	2		
1128186911	09/11/2011	Feminino	40	Manaus	Amazonas	Norte	Ensino Superi	3	3	3	3	3	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.		
1128371430	09/11/2011	Masculino	37	Osasco	São Paulo	Sudeste	Ensino Superi	4	3	4	1- Muito complexo.	2	2	1 - Texto sem coerência nenhuma	4	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma		
1128784788	12/11/2011	Masculino	42	Macaé	io de Janei	Sudeste	Não Ensino Su	2	2	2	4	1- Muito complexo.	3	3	4	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma		
1129052490	12/11/2011	Masculino	39	Contagem	Minas Gera	Sudeste	Ensino Superi	3	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	4	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	3		
1131089722	09/11/2011	Feminino	27	Nova Iguaç	io de Janei	Sudeste	Não Ensino Su	4	3	5 - Muito simples.	1- Muito complexo.	4	5 - Texto com muita coerência.	4	4	2	4		
1132356309	11/11/2011	Masculino	41	io de Janei	io de Janei	Sudeste	Não Ensino Su	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.		
1134118936	10/11/2011	Feminino	31	Colombo	Paraná	Sul	Ensino Superi	4	3	3	2	3	3	2	3	3	2		
1135906876	09/11/2011	Masculino	38	Aparecida	São Paulo	Sudeste	Ensino Superi	3	3	3	3	3	3	3	3	3	3		
1136496417	12/11/2011	Feminino	28	Muriaé	Minas Gera	Sudeste	Não Ensino Su	4	3	5 - Muito simples.	1- Muito complexo.	4	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.		
1137095087	09/11/2011	Feminino	18	Aliança	Pernambuc	Nordeste	Não Ensino Su	3	4	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	2	2		
1145818089	09/11/2011	Feminino	23	João Câmar	irande do N	Nordeste	Ensino Superi	5 - Muito simples.	2	5 - Muito simples.	2	4	5 - Texto com muita coerência.	3	3	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1146273997	09/11/2011	Feminino	44	São Paulo	São Paulo	Sudeste	Não Ensino Su	3	2	2	1- Muito complexo.	2	2	2	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma		
1146528960	09/11/2011	Feminino	20	João Pesso	Paraíba	Nordeste	Ensino Superi	4	4	5 - Muito simples.	2	4	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.		
1149693211	12/11/2011	Masculino	54	npos do Jor	São Paulo	Sudeste	Não Ensino Su	4	3	4	4	4	3	3	4	4	4		
1149941316	09/11/2011	Masculino	27	Sumaré	São Paulo	Sudeste	Ensino Superi	4	2	5 - Muito simples.	5 - Muito simples.	2	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1151506399	11/11/2011	Masculino	22	São José	anta Catari	Sul	Ensino Superi	5 - Muito simples.	2	4	4	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	2	1 - Texto sem coerência nenhuma			
1153156646	09/11/2011	Feminino	39	São Paulo	São Paulo	Sudeste	Ensino Superi	4	3	3	1- Muito complexo.	3	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4		
1154164979	09/11/2011	Feminino	43	São Paulo	São Paulo	Sudeste	Ensino Superi	3	1- Muito complexo.	4	1- Muito complexo.	2	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1155485035	10/11/2011	Feminino	21	Salvador	Bahia	Nordeste	Ensino Superi	4	3	5 - Muito simples.	2	4	5 - Texto com muita coerência.	4	4	2	5 - Texto com muita coerência.		
1156859188	12/11/2011	Feminino	37	io de Janei	io de Janei	Sudeste	Ensino Superi	5 - Muito simples.	2	5 - Muito simples.	1- Muito complexo.	3	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	2	4		
1157214043	12/11/2011	Feminino	47	io de Janei	io de Janei	Sudeste	Não Ensino Su	3	2	3	2	3	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1157698518	13/11/2011	Feminino	51	Escada	Pernambuc	Nordeste	Não Ensino Su	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.		
1158756026	12/11/2011	Masculino	20	io de Janei	io de Janei	Sudeste	Não Ensino Su	2	2	3	2	3	3	3	4	3	4		
1159079813	10/11/2011	Feminino	25	io de Janei	io de Janei	Sudeste	Ensino Superi	5 - Muito simples.	1- Muito complexo.	4	1- Muito complexo.	4	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4		
1159592905	10/11/2011	Masculino	36	o do Rodrig	irande do N	Nordeste	Ensino Superi	4	4	3	4	4	4	4	3	4	4		
1159766584	09/11/2011	Masculino	36	Arassolând	São Paulo	Sudeste	Não Ensino Su	3	3	3	3	3	3	3	3	3	3		
1159818830	09/11/2011	Feminino	21	Brasília	strito Fede	Centro Oest	Ensino Superi	3	4	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.		
1163155599	12/11/2011	Feminino	23	Porto Alegr	Grande do	Sul	Não Ensino Su	4	1- Muito complexo.	3	1- Muito complexo.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	3	4		
1165106218	09/11/2011	Feminino	35	Manaus	Amazonas	Norte	Não Ensino Su	5 - Muito simples.	2	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1165199726	12/11/2011	Masculino	25	Joinville	anta Catari	Sul	Não Ensino Su	3	2	4	1- Muito complexo.	4	3	3	5 - Texto com muita coerência.	2	2		
1165391985	09/11/2011	Feminino	35	ira de Sant	Bahia	Nordeste	Não Ensino Su	2	1- Muito complexo.	3	5 - Muito simples.	1- Muito complexo.	3	4	2	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1165586493	10/11/2011	Feminino	25	Sumaré	São Paulo	Sudeste	Ensino Superi	1- Muito complexo.	3	5 - Muito simples.	2	5 - Muito simples.	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	3	4		
1166167444	12/11/2011	Feminino	23	Giruá	Grande do	Sul	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	3	2	5 - Texto com muita coerência.	3	4		
1168357411	10/11/2011	Masculino	38	io de Janei	io de Janei	Sudeste	Ensino Superi	4	2	4	2	3	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	4	4		
1169353981	09/11/2011	Feminino	36	onta Gross	Paraná	Sul	Não Ensino Su	4	3	5 - Muito simples.	1- Muito complexo.	4	5 - Texto com muita coerência.	4	4	1 - Texto sem coerência nenhuma	2		
1169594340	10/11/2011	Feminino	35	Jandira	São Paulo	Sudeste	Ensino Superi	3	1- Muito complexo.	4	1- Muito complexo.	3	3	3	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1169786337	09/11/2011	Feminino	19	Antônio de	Bahia	Nordeste	Não Ensino Su	3	4	1- Muito complexo.	3	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.			
1172217587	12/11/2011	Masculino	21	Eirunepé	Amazonas	Norte	Não Ensino Su	4	2	4	1- Muito complexo.	4	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1173970435	09/11/2011	Feminino	31	Cotia	São Paulo	Sudeste	Ensino Superi	3	1- Muito complexo.	2	1- Muito complexo.	3	4	4	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1174082288	10/11/2011	Masculino	57	São Paulo	São Paulo	Sudeste	Não Ensino Su	3	3	1- Muito complexo.	3	3	3	3	3	1 - Texto sem coerência nenhuma	4		
1174681476	09/11/2011	Feminino	39	Brasília	strito Fede	Centro Oest	Não Ensino Su	3	2	4	1- Muito complexo.	3	4	2	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1174782763	12/11/2011	Feminino	24	São Paulo	São Paulo	Sudeste	Ensino Superi	3	2	4	3	4	4	3	3	2	4		
1174811395	09/11/2011	Feminino	36	io de Janei	io de Janei	Sudeste	Ensino Superi	3	3	3	2	3	3	3	5 - Texto com muita coerência.	4	3		
1177366381	11/11/2011	Masculino	22	Campinas	São Paulo	Sudeste	Não Ensino Su	3	3	3	3	3	3	3	3	3	3		
1177970353	09/11/2011	Masculino	20	Jequié	Bahia	Nordeste	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	3	1- Muito complexo.	3	2	2	3	2	4		
1179233744	13/11/2011	Feminino	52	osé do Rio	São Paulo	Sudeste	Não Ensino Su	3	3	5 - Muito simples.	1- Muito complexo.	3	5 - Texto com muita coerência.	3	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	3		
1183161253	10/11/2011	Feminino	47	Salvador	Bahia	Nordeste	Não Ensino Su	4	1- Muito complexo.	3	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1183408513	10/11/2011	Masculino	43	Caxias	Maranhão	Nordeste	Ensino Superi	2	2	3	3	2	1 - Texto sem coerência nenhuma	2	2	2	2		
1184366544	09/11/2011	Masculino	31	osé dos Cai	São Paulo	Sudeste	Não Ensino Su	3	3	5 - Muito simples.	3	3	3	2	4	2	3		
1185048166	09/11/2011	Feminino	25	Brasília	strito Fede	Centro Oest	Ensino Superi	2	1- Muito complexo.	4	1- Muito complexo.	3	5 - Texto com muita coerência.	3	4	2	4		
1185136750	12/11/2011	Masculino	43	São Paulo	São Paulo	Sudeste	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	3	3	3		
1186298923	09/11/2011	Feminino	20	João Pesso	Paraíba	Nordeste	Ensino Superi	5 - Muito simples.	2	5 - Muito simples.	1- Muito complexo.	3	4	4	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1188545002	12/11/2011	Feminino	18	Maranguap	Ceará	Nordeste	Não Ensino Su	3	2	4	4	2	2	4	2	3	2		
1188856306	09/11/2011	Masculino	21	Feliz	Grande do	Sul	Ensino Superi	4	4	5 - Muito simples.	3	3	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.		
1189174629	09/11/2011	Feminino	18	Goiânia	Goiás	Centro Oest	Ensino Superi	3	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	3	4	4	5 - Texto com muita coerência.	2	4		
1193881941	10/11/2011	Feminino	49	São Paulo	São Paulo	Sudeste	Ensino Superi	3	2	4	3	4	4	3	4	3	4		
1194873543	09/11/2011	Feminino	33	Viamão	Grande do	Sul	Ensino Superi	5 - Muito simples.	3	5 - Muito simples.	1- Muito complexo.	4	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	4		
1195374800	09/11/2011	Masculino	43	São Paulo	São Paulo	Sudeste	Não Ensino Su	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1196339586	11/11/2011	Masculino	52	io de Janei	io de Janei	Sudeste	Não Ensino Su	3	3	3	2	3	4	4	4	1 - Texto sem coerência nenhuma	4		
1196417700	12/11/2011	Masculino	26	quaquecet	São Paulo	Sudeste	Não Ensino Su	3	3	5 - Muito simples.	1- Muito complexo.	3	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.		
1197256415	09/11/2011	Feminino	27	Goiânia	Goiás	Centro Oest	Não Ensino Su	3	3	3	3	3	4	4	4	4	4		
1197960460	09/11/2011	Feminino	40	São Paulo	São Paulo	Sudeste	Ensino Superi	4	4	4	2	4	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1198820839	12/11/2011	Feminino	34	Ananindeu	Pará	Norte	Não Ensino Su	3	4	5 - Muito simples.	1- Muito complexo.	3	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1199755408	09/11/2011	Feminino	30	Brasília	strito Fede	Centro Oest	Ensino Superi	4	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1203782891	10/11/2011	Masculino	18	idente Pruc	São Paulo	Sudeste	Ensino Superi	3	2	3	1- Muito complexo.	2	4	3	5 - Texto com muita coerência.	2	4		
1204603085	12/11/2011	Feminino	36	Campinas	São Paulo	Sudeste	Ensino Superi	5 - Muito simples.	3	5 - Muito simples.	1- Muito complexo.	3	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.		
1207259292	09/11/2011	Masculino	20	Teresina	Piauí	Nordeste	Ensino Superi	5 - Muito simples.	3	4	3	2	4	4	4	4	4		
1207292273	09/11/2011	Masculino	41	Petrópolis	io de Janei	Sudeste	Ensino Superi	5 - Muito simples.	3	5 - Muito simples.	1- Muito complexo.	4	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1207566373	09/11/2011	Feminino	56	olta Redon	io de Janei	Sudeste	Ensino Superi	3	4	3	1- Muito complexo.	2	2	2	4	2	4		
1210558670	09/11/2011	Feminino	46	acho Fundc	strito Fede	Centro Oest	Ensino Superi	4	2	4	2	2	4	2	4	3	3		
1211276350	09/11/2011	Feminino	46	Santo And	São Paulo	Sudeste	Ensino Superi	3	2	4	1- Muito complexo.	4	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	3		
1212171574	10/11/2011	Feminino	43	Itatiba	São Paulo	Sudeste	Ensino Superi	4	2	4	1- Muito complexo.	4	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	2	4		
1214000598	09/11/2011	Feminino	31	Guará	strito Fede	Centro Oest	Ensino Superi	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1215274201	09/11/2011	Masculino	18	losé de Pira	Paraíba	Nordeste	Ensino Superi	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	3	2	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	4		
1217727293	09/11/2011	Masculino	21	Cuiabá	Mato Gross	Centro Oest	Ensino Superi	4	3	5 - Muito simples.	2	3	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.		
1220660411	09/11/2011	Masculino	37	Candeias	Bahia	Nordeste	Ensino Superi	1- Muito complexo.	2	1- Muito complexo.	2	1- Muito complexo.	5 - Texto com muita coerência.	2	4	5 - Texto com muita coerência.	2		
1221182903	12/11/2011	Feminino	34	Carapicuíb	São Paulo	Sudeste	Ensino Superi	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	3	2	3	1 - Texto sem coerência nenhuma	4		
1221578248	09/11/2011	Feminino	36	Fortaleza	Ceará	Nordeste	Ensino Superi	5 - Muito simples.	4	5 - Muito simples.	2	4	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.		
1221953912	09/11/2011	Feminino	39	Porto Alegr	Grande do	Sul	Não Ensino Su	4	2	2	3	3	4	4	4	3	4		
1222316933	12/11/2011	Masculino	38	Cotia	São Paulo	Sudeste	Não Ensino Su	3	3	3	4	4	3	3	3	3	3		
1224426343	10/11/2011	Masculino	45	Salvador	Bahia	Nordeste	Ensino Superi	5 - Muito simples.	3	4	3	5 - Muito simples.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.		
1227963557	13/11/2011	Feminino	40	São José	anta Catari	Sul	Não Ensino Su	3	3	3	2	4	2	3	2	3	3		

1228605028	12/11/2013	Feminino	63	Fortaleza	Ceará	Nordeste	Não Ensino Su	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	3	5 - Muito simples.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.
1228649412	10/11/2013	Masculino	34	Jaguariúna	São Paulo	Sudeste	Não Ensino Su	3	2	2	3	4	3	3	3	3	3	3
1230585840	12/11/2013	Masculino	43	Campinas	São Paulo	Sudeste	Ensino Superio	3	2	4	3	4	5 - Texto com muita coerência.	4	3	2	4	
1230975336	12/11/2013	Masculino	40	Fortaleza	Ceará	Nordeste	Ensino Superio	3	2	3	1- Muito complexo.	5 - Muito simples.	3	3	4	3	3	
1233985717	09/11/2013	Feminino	39	io de Janeiro	io de Janeiro	Sudeste	Ensino Superio	4	3	4	4	4	5 - Texto com muita coerência.	2	4	3	3	
1234241179	11/11/2013	Masculino	40	Aracaju	Sergipe	Nordeste	Ensino Superio	4	3	5 - Muito simples.	1- Muito complexo.	3	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.	3	4	
1235337713	12/11/2013	Masculino	37	bastião do A	Minas Gerai	Sudeste	Não Ensino Su	2	2	2	4	2	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	
1236118255	12/11/2013	Masculino	38	io de Janeiro	io de Janeiro	Sudeste	Não Ensino Su	3	2	4	3	4	3	3	4	4	4	
1236768064	10/11/2013	Feminino	38	Quirinópoli	Goiás	Centro Oest	Ensino Superio	2	2	2	1- Muito complexo.	2	1- Muito complexo.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	4	2	5 - Texto com muita coerência.
1241754452	12/11/2013	Masculino	17	Antônio de	São Paulo	Sudeste	Não Ensino Su	3	2	3	4	3	3	4	2	4	2	
1242148279	09/11/2013	Feminino	53	io de Janeiro	io de Janeiro	Sudeste	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	3	1- Muito complexo.	2	4	3	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	
1243748264	12/11/2013	Masculino	40	losé dos Cai	São Paulo	Sudeste	Não Ensino Su	4	4	5 - Muito simples.	4	3	3	4	4	4	3	
1245652548	09/11/2013	Feminino	20	Brasília	strito Fede	Centro Oest	Ensino Superio	3	2	5 - Muito simples.	1- Muito complexo.	3	4	2	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	3	
1246100058	10/11/2013	Feminino	30	Juazeiro	Bahia	Nordeste	Não Ensino Su	3	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	4	3	1 - Texto sem coerência nenhuma	3	1 - Texto sem coerência nenhuma	2	
1247183910	09/11/2013	Masculino	68	Contagem	Minas Gerai	Sudeste	Não Ensino Su	2	2	1- Muito complexo.	3	2	1 - Texto sem coerência nenhuma	3	2	3	2	
1247733196	12/11/2013	Feminino	25	Blumenau	anta Catarii	Sul	Não Ensino Su	3	1- Muito complexo.	5 - Muito simples.	4	4	3	3	3	3	3	
1248784637	09/11/2013	Feminino	20	Natal	irande do N	Nordeste	Ensino Superio	3	1- Muito complexo.	3	1- Muito complexo.	1- Muito complexo.	5 - Texto com muita coerência.	2	5 - Texto com muita coerência.	3	2	
1250003404	10/11/2013	Masculino	54	Sete Lagoas	Minas Gerai	Sudeste	Ensino Superio	4	3	4	4	3	2	3	5 - Texto com muita coerência.	4	4	
1251295560	10/11/2013	Feminino	23	São Paulo	São Paulo	Sudeste	Ensino Superio	4	3	4	2	4	5 - Texto com muita coerência.	2	3	3	5 - Texto com muita coerência.	
1252315829	12/11/2013	Masculino	43	Pirapora	Minas Gerai	Sudeste	Não Ensino Su	2	3	2	3	2	3	2	2	4	2	
1253230369	09/11/2013	Feminino	30	São Paulo	São Paulo	Sudeste	Ensino Superio	5 - Muito simples.	2	5 - Muito simples.	3	5 - Muito simples.	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	
1254713915	12/11/2013	Feminino	46	Curitiba	Paraná	Sul	Ensino Superio	4	2	4	1- Muito complexo.	4	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	3	
1256592490	12/11/2013	Masculino	36	Itambé	Bahia	Nordeste	Não Ensino Su	4	4	3	4	3	4	3	4	4	4	
1256869206	09/11/2013	Masculino	32	São Paulo	São Paulo	Sudeste	Ensino Superio	4	2	4	1- Muito complexo.	4	3	2	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	
1259794752	12/11/2013	Feminino	36	Fortaleza	Ceará	Nordeste	Não Ensino Su	4	1- Muito complexo.	3	1- Muito complexo.	3	4	3	5 - Texto com muita coerência.	2	4	
1259940770	12/11/2013	Masculino	17	Campinas	São Paulo	Sudeste	Não Ensino Su	1- Muito complexo.	1- Muito complexo.	4	1- Muito complexo.	1- Muito complexo.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	3	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	
1260065806	12/11/2013	Masculino	24	Santarém	Pará	Norte	Ensino Superio	3	2	5 - Muito simples.	1- Muito complexo.	3	5 - Texto com muita coerência.	5 - Texto com muita coerência.	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	
1261301725	11/11/2013	Feminino	25	Curitiba	Paraná	Sul	Ensino Superio	3	2	4	4	3	3	2	3	3	3	
1262676841	10/11/2013	Feminino	46	Manaus	Amazonas	Norte	Ensino Superio	3	3	3	3	3	3	1 - Texto sem coerência nenhuma	3	2	3	
1263913180	10/11/2013	Feminino	48	Pentecoste	Ceará	Nordeste	Ensino Superio	5 - Muito simples.	4	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência.	
1266395449	09/11/2013	Feminino	37	São Paulo	São Paulo	Sudeste	Ensino Superio	3	3	4	2	4	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	4	4	
1267296220	12/11/2013	Masculino	37	Taubaté	São Paulo	Sudeste	Ensino Superio	3	1- Muito complexo.	3	1- Muito complexo.	3	3	3	3	3	3	
1268394199	09/11/2013	Feminino	23	Lages	anta Catarii	Sul	Ensino Superio	5 - Muito simples.	2	5 - Muito simples.	4	3	5 - Texto com muita coerência.	4	5 - Texto com muita coerência.	3	5 - Texto com muita coerência.	
1268419117	09/11/2013	Feminino	49	Niterói	io de Janeiro	Sudeste	Ensino Superio	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	5 - Texto com muita coerência.	3	3	3	3	

Group B

User ID	Data	Gênero	Idade	Cidade	Estado	Região	Escolaridade	JUB- Em uma escala de um a cinco, onde 1 é muito complexo e 5 é muito simples, como você classificaria a complexidade do texto que acabou de ler?	RT1- Em uma escala de um a cinco, onde 1 é muito complexo e 5 é muito simples, como você classificaria a complexidade do texto que acabou de ler?	GAU- Em uma escala de um a cinco, onde 1 é muito complexo e 5 é muito simples, como você classificaria a complexidade do texto que acabou de ler?	MEL- Em uma escala de um a cinco, onde 1 é muito complexo e 5 é muito simples, como você classificaria a complexidade do texto que acabou de ler?	ST1- Em uma escala de um a cinco, onde 1 é muito complexo e 5 é muito simples, como você classificaria a complexidade do texto que acabou de ler?	JUB- Em uma escala de um a cinco, onde 1 é texto sem coerência nenhuma e 5 é texto com muita coerência, como você avalia a coerência do texto 1?	RT1- Em uma escala de um a cinco, onde 1 é texto sem coerência nenhuma e 5 é texto com muita coerência, como você avalia a coerência do texto 2?	GAU- Em uma escala de um a cinco, onde 1 é texto sem coerência nenhuma e 5 é texto com muita coerência, como você avalia a coerência do texto 3?	MEL- Em uma escala de um a cinco, onde 1 é texto sem coerência nenhuma e 5 é texto com muita coerência, como você avalia a coerência do texto 4?	ST1- Em uma escala de um a cinco, onde 1 é texto sem coerência nenhuma e 5 é texto com muita coerência, como você avalia a coerência do texto 6?
1001005843	12/11/201	Feminino	47	Brasília	Distrito Fe	Centro Oe	Ensino Superior	5 - Muito simples.	1- Muito complexo.	4	5 - Muito simples.	3	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	4	4	3
1001087037	09/11/201	Feminino	20	Curitiba	Paraná	Sul	Não Ensino Superior	3	2	3	3	2	4	3	5 - Texto com muita coerência	3	1 - Texto sem coerência nenhuma
1002286168	10/11/201	Feminino	43	Niilópolis	Rio de Jan	Sudeste	Ensino Superior	1- Muito complexo.	1- Muito complexo.	3	3	3	5 - Texto com muita coerência	2	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	4
1008770400	12/11/201	Feminino	65	Ribeirão Pi	São Paulo	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	3	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	2	5 - Texto com muita coerência	5 - Texto com muita coerência
1013560013	13/11/201	Feminino	37	Várzea Gra	Mato Gros	Centro Oe	Ensino Superior	4	1- Muito complexo.	1- Muito complexo.	2	3	4	1 - Texto sem coerência nenhuma	3	4	5 - Texto com muita coerência
1014047408	09/11/201	Masculino	43	Astorga	Paraná	Sul	Não Ensino Superior	4	1- Muito complexo.	4	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência
1014281607	09/11/201	Feminino	55	São Paulo	São Paulo	Sudeste	Ensino Superior	1- Muito complexo.	4	1- Muito complexo.	1- Muito complexo.	4	2	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência
1017342333	09/11/201	Feminino	29	São Vicente	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	2	2	4	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	2	5 - Texto com muita coerência
1018245950	12/11/201	Masculino	50	Nova Iguaç	Rio de Jan	Sudeste	Não Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	1 - Texto sem coerência nenhuma	4	3	3	2
1018631962	09/11/201	Masculino	20	Caxias do S	Rio Grande	Sul	Ensino Superior	5 - Muito simples.	3	1- Muito complexo.	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	2	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	5 - Texto com muita coerência
1025244089	09/11/201	Feminino	22	Curitiba	Paraná	Sul	Ensino Superior	5 - Muito simples.	1- Muito complexo.	3	4	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	5 - Texto com muita coerência	3
1027129351	11/11/201	Feminino	50	Pato Branc	Paraná	Sul	Ensino Superior	3	2	2	1- Muito complexo.	2	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	2	4
1028660245	09/11/201	Feminino	28	São José	Santa Cata	Sul	Não Ensino Superior	2	1- Muito complexo.	3	3	5 - Muito simples.	2	2	2	2	2
1029318810	10/11/201	Feminino	27	Salvador	Bahia	Nordeste	Ensino Superior	5 - Muito simples.	5 - Muito simples.	2	5 - Muito simples.	5 - Muito simples.	4	3	2	3	5 - Texto com muita coerência
1030598018	10/11/201	Masculino	26	Natal	Rio Grande	Nordeste	Ensino Superior	5 - Muito simples.	2	3	4	3	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	4	3
1031836843	12/11/201	Feminino	41	Crateús	Ceará	Nordeste	Não Ensino Superior	3	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	4	4	5 - Texto com muita coerência	4	5 - Texto com muita coerência	4
1032013220	11/11/201	Masculino	53	Bagé	Rio Grande	Sul	Ensino Superior	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	3	3	3	3	3	3
1034751007	09/11/201	Feminino	22	Bocaiúva	Minas Ger	Sudeste	Ensino Superior	3	3	1- Muito complexo.	2	2	5 - Texto com muita coerência	3	4	4	3
1035723095	10/11/201	Feminino	43	Castelo	Espírito Sa	Sudeste	Não Ensino Superior	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	2	2	2	3	5 - Texto com muita coerência
1036233129	11/11/201	Feminino	42	Jundiaí	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	3	2	4	4	3	5 - Texto com muita coerência	5 - Texto com muita coerência	4
1038757737	09/11/201	Masculino	39	São Gonçal	Rio Grande	Nordeste	Não Ensino Superior	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	2	5 - Muito simples.	5 - Texto com muita coerência	3	3	5 - Texto com muita coerência	2
1039074999	13/11/201	Feminino	44	Aracaju	Sergipe	Nordeste	Não Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	3	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	3	5 - Texto com muita coerência
1039576514	09/11/201	Masculino	22	Aracaju	Sergipe	Nordeste	Ensino Superior	3	1- Muito complexo.	4	1- Muito complexo.	4	4	3	4	2	2
1040772238	09/11/201	Masculino	60	Belo Horiz	Minas Ger	Sudeste	Ensino Superior	2	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	2	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	3	3
1043673784	12/11/201	Feminino	38	Itaquaquec	São Paulo	Sudeste	Não Ensino Superior	1- Muito complexo.	2	2	3	3	3	4	5 - Texto com muita coerência	3	2
1044662774	09/11/201	Feminino	22	Salvador	Bahia	Nordeste	Ensino Superior	4	3	5 - Muito simples.	4	4	4	4	4	4	3
1049833171	09/11/201	Masculino	17	Monte Alto	São Paulo	Sudeste	Não Ensino Superior	4	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	3	5 - Texto com muita coerência	3	4	3	5 - Texto com muita coerência
1051144239	10/11/201	Feminino	37	Vitória da C	Bahia	Nordeste	Não Ensino Superior	3	2	1- Muito complexo.	3	4	3	1 - Texto sem coerência nenhuma	2	2	5 - Texto com muita coerência
1053254874	12/11/201	Feminino	38	Olinda	Pernambuc	Nordeste	Não Ensino Superior	4	3	1- Muito complexo.	2	5 - Muito simples.	3	3	3	3	5 - Texto com muita coerência
1053617742	10/11/201	Feminino	27	Itapetining	São Paulo	Sudeste	Não Ensino Superior	2	1- Muito complexo.	2	1- Muito complexo.	2	5 - Texto com muita coerência	3	5 - Texto com muita coerência	3	5 - Texto com muita coerência
1054479165	12/11/201	Masculino	38	Poá	São Paulo	Sudeste	Não Ensino Superior	3	2	2	3	3	3	3	3	3	3
1055462804	09/11/201	Masculino	38	Petrópolis	Rio de Jan	Sudeste	Ensino Superior	3	3	3	3	5 - Muito simples.	3	4	3	4	4
1056391445	09/11/201	Feminino	21	Perdigão	Minas Ger	Sudeste	Ensino Superior	5 - Muito simples.	4	2	4	5 - Muito simples.	5 - Texto com muita coerência	3	3	2	4
1056956313	09/11/201	Feminino	21	Porto Velh	Rondônia	Norte	Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	2	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	2	5 - Texto com muita coerência	3
1057808772	11/11/201	Feminino	41	Salvador	Bahia	Nordeste	Ensino Superior	4	1- Muito complexo.	2	3	4	3	1 - Texto sem coerência nenhuma	2	3	4
1059143124	09/11/201	Feminino	38	Cruzeiro	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	2	4	4	4	2	2	4	3
1060106951	09/11/201	Feminino	33	Salvador	Bahia	Nordeste	Não Ensino Superior	3	3	3	3	3	3	3	3	3	4
1061132886	09/11/201	Masculino	33	João Pesso	Paraíba	Nordeste	Ensino Superior	4	5 - Muito simples.	2	3	5 - Muito simples.	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	4
1061994965	12/11/201	Feminino	38	Mauá	São Paulo	Sudeste	Não Ensino Superior	1- Muito complexo.	1- Muito complexo.	2	3	1- Muito complexo.	4	3	3	2	3
1065324348	10/11/201	Feminino	43	Fortaleza	Ceará	Nordeste	Não Ensino Superior	2	3	3	4	3	4	4	4	4	4
1066920957	09/11/201	Masculino	66	Rio de Jane	Rio de Jan	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	2	4	3	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	2	4	5 - Texto com muita coerência
1069067849	09/11/201	Masculino	38	Salvador	Bahia	Nordeste	Ensino Superior	4	4	3	3	3	4	3	4	3	3
1069780648	09/11/201	Feminino	32	Ananindeu	Pará	Norte	Não Ensino Superior	2	1- Muito complexo.	1- Muito complexo.	3	2	4	2	2	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência
1070815166	09/11/201	Masculino	41	São José d	Paraná	Sul	Não Ensino Superior	2	2	3	3	3	4	2	3	3	4
1074711165	12/11/201	Masculino	25	Montes Cla	Minas Ger	Sudeste	Ensino Superior	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	2	2	3	5 - Texto com muita coerência	3	4	1 - Texto sem coerência nenhuma
1074901282	09/11/201	Masculino	18	Brasília	Distrito Fe	Centro Oe	Não Ensino Superior	2	3	2	5 - Muito simples.	2	3	3	3	3	3
1076086042	09/11/201	Masculino	27	São João d	Goiás	Centro Oe	Ensino Superior	4	1- Muito complexo.	1- Muito complexo.	4	5 - Muito simples.	4	1 - Texto sem coerência nenhuma	2	4	5 - Texto com muita coerência
1076660998	10/11/201	Feminino	20	Entre Rios	Bahia	Nordeste	Não Ensino Superior	3	1- Muito complexo.	3	5 - Muito simples.	4	2	1 - Texto sem coerência nenhuma	3	4	3
1079061122	12/11/201	Masculino	38	Salvador	Bahia	Nordeste	Ensino Superior	4	2	2	4	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	4	5 - Texto com muita coerência	5 - Texto com muita coerência
1079129759	12/11/201	Masculino	39	Limeira	São Paulo	Sudeste	Não Ensino Superior	4	1- Muito complexo.	2	3	4	3	3	2	1 - Texto sem coerência nenhuma	4
1080794053	09/11/201	Masculino	61	Itapevi	São Paulo	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	3	4	5 - Muito simples.	5 - Texto com muita coerência	2	5 - Texto com muita coerência	4	5 - Texto com muita coerência
1081774756	12/11/201	Feminino	49	Cariacica	Espírito Sa	Sudeste	Não Ensino Superior	4	3	1- Muito complexo.	4	5 - Muito simples.	1 - Texto sem coerência nenhuma	2	4	4	1 - Texto sem coerência nenhuma
1084697328	09/11/201	Masculino	22	Jaboatão d	Pernambuc	Nordeste	Ensino Superior	5 - Muito simples.	3	2	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	2	5 - Texto com muita coerência
1085198917	09/11/201	Feminino	39	São José dc	São Paulo	Sudeste	Não Ensino Superior	1- Muito complexo.	2	4	1- Muito complexo.	3	3	4	2	3	4
1085422326	09/11/201	Masculino	42	Fortaleza	Ceará	Nordeste	Ensino Superior	3	1- Muito complexo.	2	5 - Muito simples.	3	4	3	2	3	5 - Texto com muita coerência
1086705794	12/11/201	Feminino	37	Várzea Gra	Mato Gros	Centro Oe	Não Ensino Superior	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência
1087024848	10/11/201	Masculino	20	Santos	São Paulo	Sudeste	Não Ensino Superior	2	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	3	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma
1087776627	09/11/201	Masculino	41	Curitiba	Paraná	Sul	Ensino Superior	3	2	2	3	5 - Muito simples.	3	4	2	4	5 - Texto com muita coerência
1088231385	09/11/201	Masculino	18	Diamantina	Minas Ger	Sudeste	Ensino Superior	4	3	2	2	5 - Muito simples.	4	2	3	4	5 - Texto com muita coerência
1088956474	09/11/201	Feminino	33	Araçatuba	São Paulo	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	4	4	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	4	4
1089471668	09/11/201	Masculino	21	Várzea Gra	Mato Gros	Centro Oe	Ensino Superior	4	1- Muito complexo.	2	4	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	5 - Texto com muita coerência	4
1090373408	09/11/201	Feminino	20	Santo Andr	São Paulo	Sudeste	Não Ensino Superior	4	1- Muito complexo.	1- Muito complexo.	3	3	4	2	3	4	3
1091028183	11/11/201	Masculino	51	Santo Andr	São Paulo	Sudeste	Não Ensino Superior	3	2	3	3	3	3	5 - Texto com muita coerência	2	4	3
1091093106	12/11/201	Feminino	36	Palmeira d	Rio Grande	Sul	Não Ensino Superior	2	1- Muito complexo.	1- Muito complexo.	2	2	4	4	4	4	4
1094143256	09/11/201	Masculino	37	São Paulo	São Paulo	Sudeste	Ensino Superior	3	2	2	3	3	3	2	3	3	4
1094515475	09/11/201	Feminino	27	Poá	São Paulo	Sudeste	Ensino Superior	4	2	1- Muito complexo.	3	3	3	3	4	4	3
1097809263	13/11/201	Feminino	40	Sorocaba	São Paulo	Sudeste	Não Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	2	3	4	1 - Texto sem coerência nenhuma	3	2	4
1098385572	09/11/201	Feminino	32	Brasília	Distrito Fe	Centro Oe	Ensino Superior	4	1- Muito complexo.	1- Muito complexo.	3	3	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	4	5 - Texto com muita coerência
1100976737	09/11/201	Masculino	24	Campo Gra	Mato Gros	Centro Oe	Ensino Superior	3	2	1- Muito complexo.	4	4	5 - Texto com muita coerência	2	4	5 - Texto com muita coerência	5 - Texto com muita coerência
1103612146	10/11/201	Feminino	24	São Paulo	São Paulo	Sudeste	Ensino Superior	3	3	3	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	4	3	5 - Texto com muita coerência	5 - Texto com muita coerência
1104233743	09/11/201	Masculino	19	Pindamonh	São Paulo	Sudeste	Ensino Superior	3	2	1- Muito complexo.	4	4	4	3	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência
1104537538	09/11/201	Masculino	50	Porto Alegr	Rio Grande	Sul	Não Ensino Superior	5 - Muito simples.	3	4	3	5 - Muito simples.	2	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	3
1105326195	12/11/201	Feminino	47	Guareí	São Paulo	Sudeste	Não Ensino Superior	3	3	3	3	3	2	3	3	3	2
1106331101	09/11/201	Feminino	46	Registro	São Paulo	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	4	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	2	3	5 - Texto com muita coerência
1111067020	10/11/201	Feminino	44	Salvador	Bahia	Nordeste	Ensino Superior	4	3	3	4	3	4	3	3	3	4
1110714599	10/11/201	Masculino	22	Altinho	Pernambuc	Nordeste	Ensino Superior	4	2	1- Muito complexo.	3	4	5 - Texto com muita coerência	2	2	3	5 - Texto com muita coerência
1110800461	09/11/201	Feminino	26	Araçoiaba	São Paulo	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	3	5 - Muito simples.	2	3	3	3	3	5 - Texto com muita coerência
1113622172	09/11/201	Masculino	43	Tamandaré	Pernambuc	Nordeste	Ensino Superior	2	1- Muito complexo.	3	4	5 - Muito simples.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	4	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma
1113715836	09/11/201	Masculino	28	São Mateu	Maranhão	Nordeste	Não Ensino Superior	4	3	4	4	4	3	4	3	3	4
1113922514	09/11/201	Feminino	21	Caucaia	Ceará	Nordeste	Não Ensino Superior	4	1- Muito complexo.	3	2	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	4	2
1119910398	12/11/201	Feminino	22	Fortaleza	Ceará	Nordeste	Ensino Superior	2	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	4	2	1 - Texto sem coerência nenhuma	3	2	4

1120126018	12/11/201	Feminino	30	Fortaleza	Ceará	Nordeste	Não Ensino Superior	2	3	2	3	4	5 - Texto com muita coerência	3	2	4	4	
1121430159	12/11/201	Masculino	19	Juazeiro	Bahia	Nordeste	Ensino Superior	3	2	3	3	2	2	5 - Texto com muita coerência	2	3	2	
1122122256	09/11/201	Masculino	26	Salgueiro	Pernambu	Nordeste	Não Ensino Superior	1- Muito complexo.	3	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	5 - Texto com muita coerência	
1126635704	10/11/201	Feminino	30	Suzano	São Paulo	Sudeste	Não Ensino Superior	5 - Muito simples.	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	5 - Texto com muita coerência	4	
1126767617	09/11/201	Feminino	39	Curitiba	Paraná	Sul	Não Ensino Superior	4	1- Muito complexo.	3	3	5 - Muito simples.	4	1 - Texto sem coerência nenhuma	3	3	4	
1127840384	11/11/201	Feminino	31	Curitiba	Paraná	Sul	Ensino Superior	5 - Muito simples.	1- Muito complexo.	3	4	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	3	3	2	
1128219951	10/11/201	Feminino	38	Eusébio	Ceará	Nordeste	Ensino Superior	5 - Muito simples.	4	3	4	5 - Muito simples.	4	3	3	3	4	
1128710684	12/11/201	Feminino	37	São Paulo	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	2	4	5 - Texto com muita coerência	
1129340531	11/11/201	Masculino	22	Dourados	Mato Gros	Centro Oe	Ensino Superior	4	2	2	2	3	3	3	3	3	4	
1131439814	11/11/201	Feminino	29	São José d	Paraná	Sul	Não Ensino Superior	5 - Muito simples.	3	2	5 - Muito simples.	5 - Muito simples.	3	1 - Texto sem coerência nenhuma	2	3	4	
1132128428	10/11/201	Feminino	30	Indaiatuba	São Paulo	Sudeste	Ensino Superior	3	3	3	3	3	3	3	3	3	3	
1132333374	10/11/201	Feminino	28	Rio de Jane	Rio de Jan	Sudeste	Ensino Superior	4	1- Muito complexo.	1- Muito complexo.	4	5 - Muito simples.	4	2	2	3	4	
1132342880	11/11/201	Feminino	28	Nova Venè	Espírito Sa	Sudeste	Ensino Superior	5 - Muito simples.	2	2	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	4	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	5 - Texto com muita coerência	
1135668838	09/11/201	Masculino	36	São José d	Paraná	Sul	Não Ensino Superior	1- Muito complexo.	3	3	2	2	2	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	3	5 - Texto com muita coerência	
1137232625	09/11/201	Masculino	41	Belo Horizc	Minas Ger	Sudeste	Ensino Superior	3	3	3	3	3	3	3	1 - Texto sem coerência nenhuma	4	3	
1138158413	09/11/201	Masculino	36	Sorriso	Mato Gros	Centro Oe	Não Ensino Superior	5 - Muito simples.	4	5 - Muito simples.	4	5 - Muito simples.	3	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	
1140900564	11/11/201	Masculino	19	Sapucaia d	Rio Grande	Sul	Ensino Superior	5 - Muito simples.	1- Muito complexo.	3	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	5 - Texto com muita coerência	5 - Texto com muita coerência	
1143785437	09/11/201	Feminino	23	Belo Horizc	Minas Ger	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	3	5 - Muito simples.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	3	5 - Texto com muita coerência	
1144914997	09/11/201	Masculino	45	Santa Bárb	Minas Ger	Sudeste	Não Ensino Superior	3	3	3	3	3	4	3	3	4	4	
1145204176	13/11/201	Masculino	40	Porto Alegı	Rio Grande	Sul	Ensino Superior	3	3	3	3	4	4	3	3	4	3	
1145822822	09/11/201	Masculino	36	Ilhota	Santa Cata	Sul	Não Ensino Superior	4	1- Muito complexo.	1- Muito complexo.	3	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	4	5 - Texto com muita coerência	
1146237393	09/11/201	Feminino	23	Rio de Jane	Rio de Jan	Sudeste	Não Ensino Superior	5 - Muito simples.	3	2	1- Muito complexo.	5 - Muito simples.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	
1146800958	10/11/201	Feminino	60	Santa Isabe	São Paulo	Sudeste	Ensino Superior	4	2	4	3	5 - Muito simples.	3	4	3	4	3	
1147637764	10/11/201	Masculino	28	Natal	Rio Grande	Nordeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	2	3	4	5 - Texto com muita coerência	
1147778845	12/11/201	Masculino	42	São Paulo	São Paulo	Sudeste	Ensino Superior	4	2	3	3	3	4	2	2	5 - Texto com muita coerência	5 - Texto com muita coerência	
1148348587	09/11/201	Feminino	46	Terra Boa	Paraná	Sul	Não Ensino Superior	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	4	3	4	3	3	
1150635500	10/11/201	Feminino	20	São José d	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	3	4	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	4	5 - Texto com muita coerência	
1150733286	13/11/201	Feminino	54	Juiz de For	Minas Ger	Sudeste	Não Ensino Superior	3	4	3	3	4	4	4	2	3	3	
1151047929	11/11/201	Feminino	26	Palmas	Tocantins	Norte	Ensino Superior	4	3	3	4	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	4	5 - Texto com muita coerência	4	
1152035882	10/11/201	Masculino	27	Pombal	Paraíba	Nordeste	Não Ensino Superior	3	2	2	4	3	4	2	2	3	4	
1152478548	12/11/201	Masculino	57	Piracicaba	São Paulo	Sudeste	Ensino Superior	4	4	4	4	4	4	4	4	4	3	
1153347613	09/11/201	Masculino	19	Belém	Pará	Norte	Não Ensino Superior	3	1- Muito complexo.	2	3	4	4	3	3	3	2	
1156507860	09/11/201	Masculino	18	Castanhal	Pará	Norte	Ensino Superior	3	1- Muito complexo.	3	2	5 - Muito simples.	4	1 - Texto sem coerência nenhuma	3	4	5 - Texto com muita coerência	
1158005468	09/11/201	Masculino	31	Brasília	Distrito Fe	Centro Oe	Ensino Superior	3	3	3	4	4	3	3	3	4	3	
1158063216	09/11/201	Feminino	42	Santo Andr	São Paulo	Sudeste	Ensino Superior	3	4	4	3	3	3	4	3	3	3	
1159349613	13/11/201	Feminino	18	Guarulhos	São Paulo	Sudeste	Não Ensino Superior	2	1- Muito complexo.	2	3	3	3	3	4	3	4	
1160255894	09/11/201	Feminino	40	Itaquaquec	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	2	2	4	4	3	3	4	4	
1164922553	09/11/201	Masculino	23	Demerval L	Piauí	Nordeste	Ensino Superior	3	2	3	5 - Muito simples.	3	3	3	3	3	3	
1167477214	11/11/201	Feminino	52	Belo Horizc	Minas Ger	Sudeste	Ensino Superior	2	1- Muito complexo.	1- Muito complexo.	3	4	4	4	3	4	4	
1168093386	09/11/201	Masculino	36	Belém	Pará	Norte	Não Ensino Superior	3	4	4	3	4	3	4	4	3	2	
1172605105	09/11/201	Masculino	21	Apodi	Rio Grande	Nordeste	Ensino Superior	3	2	1- Muito complexo.	2	3	4	3	3	3	2	
1173699986	12/11/201	Feminino	23	Lavras	Minas Ger	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	3	5 - Muito simples.	5 - Texto com muita coerência	2	4	4	3	
1177621573	09/11/201	Feminino	34	Aracaju	Sergipe	Nordeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	2	2	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	
1177835888	12/11/201	Feminino	39	Rio de Jane	Rio de Jan	Sudeste	Não Ensino Superior	3	2	4	4	5 - Muito simples.	2	3	2	3	3	
1178740243	09/11/201	Masculino	41	Fazenda Ri	Paraná	Sul	Não Ensino Superior	3	4	2	3	5 - Muito simples.	2	4	4	3	5 - Texto com muita coerência	
1179117685	09/11/201	Masculino	33	São Paulo	São Paulo	Sudeste	Ensino Superior	3	3	3	3	5 - Muito simples.	3	3	3	3	3	
1179329836	09/11/201	Feminino	27	São Paulo	São Paulo	Sudeste	Ensino Superior	4	1- Muito complexo.	2	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	4	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	
1179435592	09/11/201	Masculino	27	Cuité	Paraíba	Nordeste	Não Ensino Superior	5 - Muito simples.	3	3	3	3	3	3	3	3	3	
1185015256	09/11/201	Feminino	36	Planaltina	Goiás	Centro Oe	Não Ensino Superior	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	2	2	2	2	2	
1185246979	10/11/201	Masculino	42	Mauá	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	2	2	3	3	2	2	3	4	
1185485994	09/11/201	Feminino	36	São Paulo	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	2	4	4	1 - Texto sem coerência nenhuma	2	3	5 - Texto com muita coerência	
1185785481	09/11/201	Feminino	58	Curitiba	Paraná	Sul	Ensino Superior	4	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	4	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	4	5 - Texto com muita coerência	
1185858702	09/11/201	Feminino	17	Brasília	Distrito Fe	Centro Oe	Não Ensino Superior	3	3	3	3	3	3	3	3	3	3	
1189660041	10/11/201	Feminino	26	Maracanaú	Ceará	Nordeste	Ensino Superior	4	1- Muito complexo.	2	5 - Muito simples.	5 - Muito simples.	3	1 - Texto sem coerência nenhuma	3	5 - Texto com muita coerência	5 - Texto com muita coerência	
1189816474	10/11/201	Feminino	31	Cabo de Sa	Pernambu	Nordeste	Não Ensino Superior	4	1- Muito complexo.	1- Muito complexo.	4	3	4	2	5 - Texto com muita coerência	3	5 - Texto com muita coerência	
1190809687	10/11/201	Masculino	18	Porto da Fc	Sergipe	Nordeste	Não Ensino Superior	5 - Muito simples.	1- Muito complexo.	3	4	5 - Muito simples.	5 - Texto com muita coerência	2	4	5 - Texto com muita coerência	5 - Texto com muita coerência	
1191914546	12/11/201	Masculino	42	Itapevi	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	4	5 - Texto com muita coerência	2	3	3	5 - Texto com muita coerência	
1192816678	12/11/201	Masculino	21	São Gonça	Rio de Jan	Sudeste	Ensino Superior	3	3	3	3	3	3	3	3	3	3	
1193182933	12/11/201	Feminino	20	Brasília	Distrito Fe	Centro Oe	Ensino Superior	4	3	3	4	5 - Muito simples.	5 - Texto com muita coerência	4	3	5 - Texto com muita coerência	3	
1194935925	10/11/201	Masculino	38	Encruzilhac	Bahia	Nordeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	2	4	3	1 - Texto sem coerência nenhuma	4	5 - Texto com muita coerência	5 - Texto com muita coerência	
1194959235	12/11/201	Masculino	36	Gameleiras	Minas Ger	Sudeste	Não Ensino Superior	2	1- Muito complexo.	2	3	2	4	4	4	4	4	
1195072353	09/11/201	Masculino	39	Mauá	São Paulo	Sudeste	Ensino Superior	3	2	2	2	2	3	2	2	2	2	
1195223336	09/11/201	Feminino	36	Balneário (	Santa Cata	Sul	Não Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	5 - Muito simples.	4	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	2	
1195253077	12/11/201	Masculino	47	Piracaia	São Paulo	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	2	4	4	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	
1197530692	09/11/201	Masculino	42	Salvador	Bahia	Nordeste	Ensino Superior	3	1- Muito complexo.	2	2	3	3	3	4	3	4	
1200706645	11/11/201	Feminino	37	Rio de Jane	Rio de Jan	Sudeste	Não Ensino Superior	5 - Muito simples.	3	5 - Muito simples.	3	2	5 - Texto com muita coerência	2	4	3	3	
1201677853	12/11/201	Feminino	39	Rio de Jane	Rio de Jan	Sudeste	Não Ensino Superior	1- Muito complexo.	3	3	3	3	2	2	3	3	3	
1201800447	09/11/201	Feminino	27	Rio de Jane	Rio de Jan	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	3	4	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	3	5 - Texto com muita coerência	5 - Texto com muita coerência	
1204250480	10/11/201	Feminino	26	Santarém	Pará	Norte	Ensino Superior	5 - Muito simples.	3	3	3	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	
1204694449	09/11/201	Feminino	46	Brasília	Distrito Fe	Centro Oe	Ensino Superior	2	4	3	5 - Muito simples.	3	3	2	3	3	4	
1204900621	09/11/201	Masculino	31	Várzea da I	Minas Ger	Sudeste	Não Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	3	2	2	2	5 - Texto com muita coerência	5 - Texto com muita coerência	3	
1205048806	10/11/201	Feminino	26	São Paulo	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	4	3	5 - Muito simples.	4	1 - Texto sem coerência nenhuma	4	4	5 - Texto com muita coerência	
1207176648	10/11/201	Feminino	18	Campos do	Rio de Jan	Sudeste	Não Ensino Superior	3	1- Muito complexo.	3	3	5 - Muito simples.	4	1 - Texto sem coerência nenhuma	2	3	3	
1208088161	09/11/201	Feminino	50	Pariquera-a	São Paulo	Sudeste	Não Ensino Superior	3	1- Muito complexo.	3	2	5 - Muito simples.	3	3	2	3	4	
1212366100	12/11/201	Feminino	32	São Paulo	São Paulo	Sudeste	Ensino Superior	4	1- Muito complexo.	2	3	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	4	4	5 - Texto com muita coerência	
1212384690	09/11/201	Masculino	28	Imbé	Rio Grande	Sul	Não Ensino Superior	3	1- Muito complexo.	2	2	1- Muito complexo.	4	2	3	3	5 - Texto com muita coerência	
1215579308	09/11/201	Feminino	37	Rio de Jane	Rio de Jan	Sudeste	Ensino Superior	2	1- Muito complexo.	1- Muito complexo.	3	3	5 - Texto com muita coerência	3	3	5 - Texto com muita coerência	5 - Texto com muita coerência	
1216798271	09/11/201	Masculino	19	Oiapoque	Amapá	Norte	Ensino Superior	2	3	4	3	5 - Muito simples.	3	2	3	3	2	
1217491199	10/11/201	Masculino	38	Goiânia	Goiás	Centro Oe	Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	3	2	3	1 - Texto sem coerência nenhuma	3	3	3	
1217803089	12/11/201	Masculino	52	Cariacica	Espírito Sa	Sudeste	Não Ensino Superior	3	1- Muito complexo.	5 - Muito simples.	1- Muito complexo.	5 - Muito simples.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	5 - Texto com muita coerência	3	
1218179088	09/11/201	Feminino	19	Belo Horizc	Minas Ger	Sudeste	Ensino Superior	4	1- Muito complexo.	1- Muito complexo.	2	4	5 - Texto com muita coerência	3	3	5 - Texto com muita coerência	5 - Texto com muita coerência	
1219574875	09/11/201	Feminino	53	Mogi das C	São Paulo	Sudeste	Não Ensino Superior	4	4	4	4	4	3	3	4	4	3	
1220350630	12/11/201	Feminino	49	Duque de (	Rio de Jane	Sudeste	Não Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	2	4	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	4	5 - Texto com muita coerência	
1221834603	09/11/201	Masculino	27	Osasco	São Paulo	Sudeste	Ensino Superior	4	3	4	4	5 - Muito simples.	4	4	4	4	3	
1223945380	09/11/201	Feminino	25	São Paulo	São Paulo	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	3	4	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	4	5 - Texto com muita coerência	5 - Texto com muita coerência	
1224248718	09/11/201	Feminino	33	Ceilândia	Distrito Fe	Centro Oe	Não Ensino Superior	4	1- Muito complexo.	2	3	5 - Muito simples.	4	1 - Texto sem coerência nenhuma	2	5 - Texto com muita coerência	5 - Texto com muita coerência	
1224653773	12/11/201	Feminino	36	Arraial do (	Rio de Jane	Sudeste	Ensino Superior	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Muito simples.	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	
1225996144	10/11/201	Masculino	41	Santa Luzia	Paraíba	Nordeste	Ensino Superior	2	2	1- Muito complexo.	3	2	3	4	2	3	2	
1227903743	10/11/201	Feminino	23	Catalão	Goiás	Centro Oe	Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	3	2	3	2	3	3	3	
1229393796	09/11/201	Masculino	17	Catunda	Ceará	Nordeste	Não Ensino Superior	2	1- Muito complexo.	1- Muito complexo.	3	5 - Muito simples.	5 - Texto com muita coerência	3	3	5 - Texto com muita coerência	5 - Texto com muita coerência	
1231665948	12/11/201	Feminino	38	Guarulhos	São Paulo	Sudeste	Não Ensino Superior	3	3	3	2	3	2	3	3	3	3	
1232209668	12/11/201	Masculino	36	Rio de Jane	Rio de Jan	Sudeste	Não Ensino Superior	2	1- Muito complexo.	2	1- Muito complexo.	2	5 - Texto com muita coerência	3	2	5 - Texto com muita coerência	5 - Texto com muita coerência	
1234137577	10/11/201	Feminino	36	Sinop	Mato Gros	Centro Oe	Ensino Superior	4	1- Muito complexo.	2	4	4	5 - Texto com muita coerência	2	4	5 - Texto com muita coerência	5 - Texto com muita coerência	

1234235886	09/11/201	Feminino	28	Marabá	Pará	Norte	Ensino Superior	2	2	2	3	4	5 - Texto com muita coerência	4	3	2	4
1237266112	09/11/201	Feminino	47	Fortaleza	Ceará	Nordeste	Não Ensino Superior	5 - Muito simples.	4	4	3	5 - Muito simples.	3	4	5 - Texto com muita coerência	3	4
1239199246	10/11/201	Masculino	37	Duque de (	Rio de Jane	Sudeste	Não Ensino Superior	5 - Muito simples.	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	3	4	5 - Texto com muita coerência
1239713003	09/11/201	Masculino	33	Itapevi	São Paulo	Sudeste	Ensino Superior	5 - Muito simples.	2	3	3	5 - Muito simples.	5 - Texto com muita coerência	3	5 - Texto com muita coerência	5 - Texto com muita coerência	4
1241480232	09/11/201	Feminino	29	Carapicuíb	São Paulo	Sudeste	Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	3	4	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	5 - Texto com muita coerência	4
1242614095	12/11/201	Feminino	42	Piraquara	Paraná	Sul	Não Ensino Superior	4	3	4	4	4	4	4	4	4	3
1243183297	10/11/201	Feminino	22	João Pesso	Paraíba	Nordeste	Ensino Superior	4	1- Muito complexo.	2	3	4	5 - Texto com muita coerência	2	3	4	5 - Texto com muita coerência
1243936501	09/11/201	Masculino	28	Maracanaú	Ceará	Nordeste	Ensino Superior	3	1- Muito complexo.	1- Muito complexo.	3	4	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	4	4	3
1244312717	10/11/201	Feminino	20	São Paulo	São Paulo	Sudeste	Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	5 - Texto com muita coerência	2	3	2	4
1244382583	09/11/201	Masculino	20	Dourados	Mato Gros	Centro Oe	Ensino Superior	2	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	5 - Muito simples.	2	1 - Texto sem coerência nenhuma	4	4	3
1244712851	12/11/201	Masculino	47	São Gonça	Rio de Jane	Sudeste	Não Ensino Superior	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma	1 - Texto sem coerência nenhuma
1244838580	09/11/201	Feminino	23	Rio de Jane	Rio de Jane	Sudeste	Ensino Superior	5 - Muito simples.	2	3	4	1- Muito complexo.	3	2	4	3	5 - Texto com muita coerência
1247455477	09/11/201	Masculino	18	São Paulo	São Paulo	Sudeste	Ensino Superior	5 - Muito simples.	3	2	4	5 - Muito simples.	5 - Texto com muita coerência	3	5 - Texto com muita coerência	3	1 - Texto sem coerência nenhuma
1248079825	09/11/201	Masculino	37	São Paulo	São Paulo	Sudeste	Não Ensino Superior	4	1- Muito complexo.	3	3	3	3	2	3	3	4
1252232532	09/11/201	Masculino	22	Recife	Pernambu	Nordeste	Ensino Superior	4	1- Muito complexo.	2	3	5 - Muito simples.	4	1 - Texto sem coerência nenhuma	3	4	5 - Texto com muita coerência
1253827887	09/11/201	Masculino	33	Niterói	Rio de Jane	Sudeste	Não Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	2	5 - Muito simples.	5 - Texto com muita coerência	1 - Texto sem coerência nenhuma	4	4	5 - Texto com muita coerência
1255400185	10/11/201	Feminino	35	Campinas	São Paulo	Sudeste	Ensino Superior	4	1- Muito complexo.	2	3	4	4	3	3	5 - Texto com muita coerência	5 - Texto com muita coerência
1257151906	10/11/201	Feminino	47	Rio de Jane	Rio de Jane	Sudeste	Não Ensino Superior	2	3	1- Muito complexo.	1- Muito complexo.	2	3	3	3	3	3
1260075552	09/11/201	Masculino	48	Claro dos P	Minas Ger	Sudeste	Não Ensino Superior	5 - Muito simples.	3	1- Muito complexo.	5 - Muito simples.	4	4	5 - Texto com muita coerência	4	2	3
1260963078	10/11/201	Masculino	42	São Leopol	Rio Grande	Sul	Não Ensino Superior	1- Muito complexo.	3	4	2	2	3	5 - Texto com muita coerência	3	3	2
1261382334	09/11/201	Masculino	40	Recife	Pernambu	Nordeste	Ensino Superior	5 - Muito simples.	3	5 - Muito simples.	3	5 - Muito simples.	3	4	5 - Texto com muita coerência	4	4
1261920797	12/11/201	Masculino	22	Campo Ma	Paraná	Sul	Ensino Superior	5 - Muito simples.	1- Muito complexo.	1- Muito complexo.	3	4	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência	5 - Texto com muita coerência
1263350130	09/11/201	Masculino	37	Ponta Gros	Paraná	Sul	Ensino Superior	4	2	3	3	5 - Muito simples.	4	2	4	3	4
1264063621	12/11/201	Feminino	59	Rio de Jane	Rio de Jane	Sudeste	Não Ensino Superior	4	3	2	1- Muito complexo.	4	4	2	5 - Texto com muita coerência	4	4
1267198780	09/11/201	Feminino	29	Ourinhos	São Paulo	Sudeste	Ensino Superior	3	2	3	3	3	3	4	3	3	3
1268256390	09/11/201	Masculino	26	São Paulo	São Paulo	Sudeste	Ensino Superior	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	1- Muito complexo.	3	1 - Texto sem coerência nenhuma	3	1 - Texto sem coerência nenhuma	5 - Texto com muita coerência	5 - Texto com muita coerência

ANNEX C – PUBLISHED ARTICLES

PLOS ONE

RESEARCH ARTICLE

Eye-tracking as a proxy for coherence and complexity of texts

Débora Torres[1], Wagner R. Sena[1], Humberto A. Carmona[1], André A. Moreira[1], Hernán A. Makse[2], José S. Andrade Jr.[1]*

1 Departamento de Física, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil, **2** Levich Institute and Physics Department, The City College of New York, New York City, New York, United States of America

* soares@fisica.ufc.br

OPEN ACCESS

Citation: Torres D, Sena WR, Carmona HA, Moreira AA, Makse HA, Andrade JS, Jr. (2021) Eye-tracking as a proxy for coherence and complexity of texts. PLoS ONE 16(12): e0260236. https://doi.org/10.1371/journal.pone.0260236

Editor: Haroldo V. Ribeiro, Universidade Estadual de Maringa, BRAZIL

Received: September 1, 2021

Accepted: November 4, 2021

Published: December 13, 2021

Data Availability Statement: The data underlying the results presented in the study are available from the Kaggle database: https://www.kaggle.com/debtorres/eyetracking-reading-experiment.

Funding: Authors JSA, HAC and AAM thank the Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) for financial support. JSA also acknowledges support from the National Institute

Abstract

Reading is a complex cognitive process that involves primary oculomotor function and high-level activities like attention focus and language processing. When we read, our eyes move by primary physiological functions while responding to language-processing demands. In fact, the eyes perform discontinuous twofold movements, namely, successive long jumps (saccades) interposed by small steps (fixations) in which the gaze "scans" confined locations. It is only through the fixations that information is effectively captured for brain processing. Since individuals can express similar as well as entirely different opinions about a given text, it is therefore expected that the form, content and style of a text could induce different eye-movement patterns among people. A question that naturally arises is whether these individuals' behaviours are correlated, so that eye-tracking while reading can be used as a proxy for text subjective properties. Here we perform a set of eye-tracking experiments with a group of individuals reading different types of texts, including children stories, random word generated texts and excerpts from literature work. In parallel, an extensive Internet survey was conducted for categorizing these texts in terms of their complexity and coherence, considering a large number of individuals selected according to different ages, gender and levels of education. The computational analysis of the fixation maps obtained from the gaze trajectories of the subjects for a given text reveals that the average "magnetization" of the fixation configurations correlates strongly with their complexity observed in the survey. Moreover, we perform a thermodynamic analysis using the Maximum-Entropy Model and find that coherent texts were closer to their corresponding "critical points" than non-coherent ones, as computed from the Pairwise Maximum-Entropy method, suggesting that different texts may induce distinct cohesive reading activities.

Introduction

Understanding how people capture and assimilate written information involves multiple fields of science, from human anatomy and neurology to linguistics. In particular, the eye movement has always been of interest for the comprehension of reading behavior, since it represents an observable link between the mechanics of vision and its cognitive activity.

of Science and Technology for Complex Systems in Brazil (INCT-SC). DT and WRS received support from CNPq through awarded fellowships for research. HAM acknowledges funding from the National Institute of Biomedical Imaging and Bioengineering (NIBIB) and the National Institute of Mental Health (NIMH) through the NIH BRAIN Initiative Grant R01 EB028157 and NSF DMR-1945909. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

In the 19th century, the founders of visual-behavior research examined eye movement in elementary experiments and described how the eyes capture visual information for processing in the brain [1]. In these experiments, it was noticed that the eyes do not register information through smooth continuous movements; instead, they make successive jerks (saccades), between events in which the gaze is briefly maintained on small confined regions (fixations) [2–6]. Moreover, we move our eyes in such a way that, through fixations of precise duration and location, they can efficiently capture pieces of visual data that our brain then puts together in order to create a complete neat image. In this way, the mental demands in processing the image also influences where we sequentially direct the gaze. With the development of eye-tracking devices, scientists were able to observe gaze trajectories and it became clear that the eye movement also depends on the attention focus and examination strategies [6–9].

In the late 20th century, a new area of research in linguistics focused on studying eye movements while reading. In these studies, different strategies were applied in order to analyze how words are fixated depending on specific linguistic factors. It was found that both the number and duration of the fixations on each word are plausible measures to quantify word processing and language comprehension [10–13]. It has been generally accepted that the properties of a given word which are significantly correlated with read processing are its frequency in the language [11, 12, 14–17], length [12, 15, 18] and predictability in context [12, 15, 17, 19, 20]. Under the same framework, eye-tracking experiments have been successfully combined with mathematical models to formally describe how individuals control eye movement while reading, with attention shifting from word to word [21–24].

Complexity and coherence of a text are considered main linguistic attributes to evaluate reading comprehension and learning difficulties [25, 26]. Linguistic researchers have been focusing on measuring texts complexity over the last decades [27]. Mainly, the issue has gained importance due to the need to select appropriate texts for different scholarly levels that would allow students to progressively develop reading and text comprehension skills [25]. For this reason, mathematical expressions for readability and metrics have been developed in order to quantify the complexity of texts and categorize reading material [28]. The variables frequently used are the average length of the words and their frequency in the language, both accounting for semantic difficulty, as well as sentence length, which is closely related to syntactic complexity. The relevant premise behind these empirical expressions is rather obvious, namely, that texts with unusual, long words and extensive sentences are more difficult to process than texts with familiar vocabulary and short sentences. In contrast, the coherence of a text is related to its meaningfulness, a notion associated with semantics rather than grammatical structure [29]. A coherent text makes sense in such a way that the ideas in it are continually connected and the text is consistent as a whole. Its sentences not only have meaning on their own, but, more importantly, they successively build on the meaning of the text.

It is expected that features of a text such as genre and style may be reflected in the eye movement patterns of individuals when reading text passages. Different types of texts may therefore prompt different reading responses in terms of fixation configurations and, consequently, different cognitive reactions. Thus, in order to study this interplay, a phenomenological modeling approach based on eye-movement data (*i.e.*, fixation patterns) should be able to capture the inner cognitive processes underlying reading. In this regard, models from Statistical Physics such us the Maximum-Entropy Model (MEM) developed in information theory can provide a statistical conceptual framework to understand a given natural process in terms of the "interactions" among its many elementary units using statistical data obtained experimentally [30–32]. The principle of maximum entropy states that the probability distribution that best represents the state of a given system is the one that maximizes its entropy, being also in conformity with one or a set of specific constraints. This principle, by itself, contains the essence of the so called

Inverse Ising Problem solution, in which the Hamiltonian (*i.e.*, the interactions) in a given complex system can be inferred from observed statistical correlations among its components. This statistical analysis is frequently referred to as the *Boltzmann-machine*, since it uses the Boltzmann distribution in its core.

The MEM approach has been applied to a wide variety of systems which can be mapped to Ising-like models. In this representation, the interacting elements can be in an active or inactive state, analogously to an Ising type system (*i.e.*, a lattice of dipole moments in which the spins are in either up, +1, or down, −1, states that can be under the action of an external field). In the case of neuronal networks, for example, the interactions between neurons subjected to some stimuli are inferred from the collected data of their firing patterns [33–36]. In a larger scale, the interactivity among regions of the human brain, for example, has been investigated from data of nuclear magnetic resonance [37]. An important application of the MEM is the characterization of protein-protein interaction benefiting from large protein databases [38, 39]. In particular, this strategy has been used to infer genetic interaction networks from known gene expression patterns [40–42]. The collective response exhibited by flocks of birds was also studied by means of the MEM [43, 44] as well as the emergence of collective behavior from the eye movement patterns of a group of people while watching commercial videos [45]. In the last case, pairwise correlations among series of instantaneous eye's velocities were utilized to capture the collective response of the individuals and relate it to video popularity. Finally, the MEM approach has also proved to be effective in other fields outside biology. In [46], for example, the intricate network micro-structure of interactions in the stock market has been captured through pairwise correlations calculated from big data bases of stocks variability.

As shown in Fig 1, here we address the problem of characterizing the complexity and coherence of diverse texts quantitatively by taking a twofold approach. First, we perform eye-tracking experiments with a limited group of people to directly analyze their fixation data while reading different texts, namely, children stories, excerpts from literary works and random word generated texts (see Table 1). This is achieved by expressing the experimental results in terms of a binary model for fixation sequences (analogous to an Ising system) which is duly embedded in the MEM. It enables us to disclose two indexes that can be used as potential proxies for complexity and coherence: the magnetization (a measure of the density of the fixation sequences) and the distance between the "operating temperature" of the system and its critical temperature (a measure of cohesion among the fixation sequences). Second, our experimental approach is then validated through an extensive Internet reading survey, with access to a vast respondent sample, to categorize the same texts according to different complexity and coherence levels, therefore permitting a direct comparison with the obtained eye-tracking indexes.

Materials and methods

Eye-tracking reading experiment

Methodology. The experiments were conducted using a SR Research EyeLink 1000 eye tracker, with the Desktop Mount Participant Setup. It operates at a sampling frequency of 1 kHz using a monocular device and an infrared video-based eye tracker [48]. This equipment is based on the Pupil Center Corneal Reflection system (PCCR) [49, 50], one of the most accurate, non-intrusive eye-tracking techniques. When a stimuli is presented to the subject on a display monitor, near infrared light is shined onto the subjects' eyes and the reflections are recorded with a special camera. Part of the light is reflected in the cornea, appearing as a small, sharp glint (known as the "first Purkinje image"), and another part reaches the retina and reflects back making the pupil appear as a bright, well defined disc ("bright pupil" effect). The

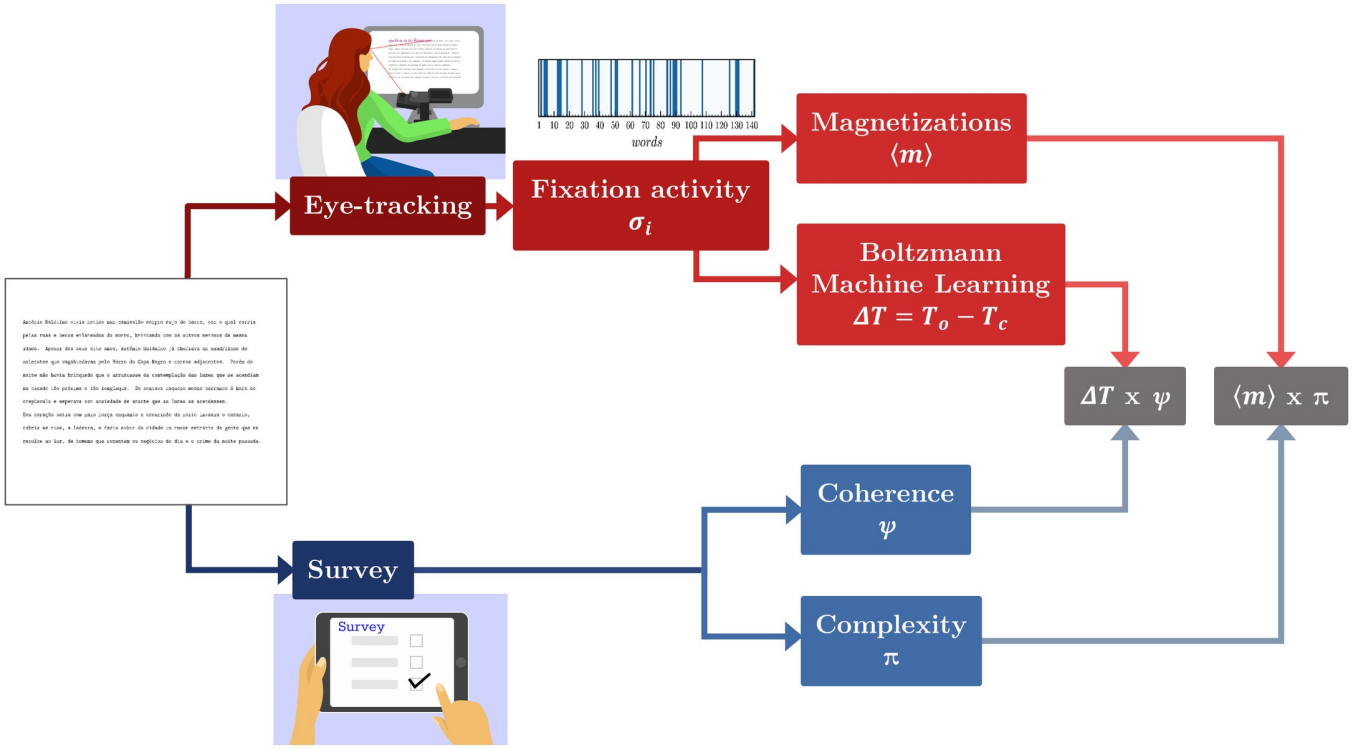


Fig 1. Diagram of the research route. We take a twofold approach to characterize through the cognitive reading activity the complexity and coherence of texts. On one side, we perform an eye-tracking experiment to collect fixation data from a group of people. The fixation activity associated to each subject while reading a given text is computed by binarizing the states of each word, defined positive +1, if the subject fixates it at least twice, or negative −1 if the subject does not fixate or fixates it only once. We then compute the reading "magnetization" of a given text for each subject and average it over all subjects to obtain $\langle m \rangle$. From the pairwise cross-correlations between the fixation sequences of the subjects, we also infer a "Hamiltonian" for each text by means of the Maximum Entropy principle using a Boltzmann machine-learning algorithm. A thermodynamic analysis of the energy fluctuations allows us to determine whether the text is near a "critical point". In parallel, we collected reading-comprehension data from an Internet extensive survey performed with 400 people in an attempt to quantify the complexity $\langle \pi \rangle$ and coherence $\langle \psi \rangle$ of the texts.

https://doi.org/10.1371/journal.pone.0260236.g001

Table 1. Texts information.

Symbol	Title	Author	Year	Country
GAU	O Gaúcho	José de Alencar	1870	Brazil
GSV	Grande Sertão: Veredas	João Guimarães Rosa	1956	Brazil
HCL	História do Cerco de Lisboa	José Saramago	1989	Portugal
JUB	Jubiabá	Jorge Amado	1935	Brazil
MEL	A Mão e a Luva	Machado de Assis	1874	Brazil
QUI	O Quinze	Rachel de Queiroz	1930	Brazil
RT1	Random text 1	-	-	-
RT2	Random text 2	-	-	-
ST1	Story 1: A patinha Esmeralda	-	-	Brazil
ST2	Story 2: A menina do leite	-	-	Brazil

Basic information about the 10 texts used in the eye-tracking experiments. All texts are written in Portuguese. RT1 and RT2 texts are generated with an online random word generator [47]. ST1 and ST2 texts are popular children stories of unknown author and year.

https://doi.org/10.1371/journal.pone.0260236.t001

reflected images are captured by the camera and are then processed by the EyeLink software. The vector between the pupil and corneal reflections is used to calculate the exact gaze location of each sample.

We selected 20 participants among physics and engineering graduate and postgraduate students with ages from 17 to 34, all Brazilian Portuguese native speakers. The reading material consisted of 10 different types of texts (all in Portuguese), including two children stories, two random word generated texts (with standard grammatical structure, but random content) [47] and six excerpts from literature work (see Table 1). All texts were in 12 point size mono-spaced font. For our equipment setup, these characteristics are compatible with the condition that a visual angle of 1˚ spans a length of 3 characters, which gives word position accuracy [50]. Letters were light cyan and the background dark gray, which provides high color contrast and moderate brightness in order to ensure readability while improving the eye-tracking accuracy.

The eye-tracking calibration process consists in collecting raw eye data when the subject fixates at target points, presented one by one at the display monitor. Next, the information is processed and the gaze positions are calculated. The offset between the sampled gaze and the displayed point positions determines the quality of calibration. This protocol was followed before each reading for every participant. A validation test was then performed after calibration to confirm that its accuracy was always within the error range from 0.25˚ to 1˚.

Before starting the experiment, with the purpose of motivating the participants to read consciously, they were warned that we would ask them to answer a simple question after reading each text. To run the experiment, the subject seated in front of a display screen and the head was stabilized by the use of an adjustable head and chin rest. Without imposing any time limit for reading, the texts were sequentially shown on the screen, intercalated by their corresponding questions. During the reading session, the eye tracker collected gaze location data at a sample rate of 1000 Hz, which gives an average temporal error of 0.5 ms (approximately half the duration of the time between samples) [50]. The collected data was de-identified in order to preserve the participants' privacy. The experiment was designed using the SR Research Experiment Builder program (version 1.10.1630) and the collected data was displayed and filtered with the Eye-Link Data Viewer software (version 1.10.1). The reading data is processed by first delimiting each word in the texts with rectangles. For each subject reading a given text, we obtain the spatial coordinates of the fixations from the eye tracker and count how many of them fall into each box, as exemplified in Fig 2. At the end of the experiment, an array can be associated to each text, whose elements n_i^r correspond to the number of fixations that a given word r received from the subject i. In addition, the time duration of each reading was simultaneously registered during every eye-tracking experiment.

Ethics. The eye-tracking experiment procedures were approved by the Research Ethics Committee of the Federal University of Ceará (COMEPE, Universidade Federal do Ceará, Brasil). All subjects gave written informed consent. Also, parental consent was obtained from the parents of the minors included in the study.

Fixation activity model. In order to make possible the analogy with the Ising system, we define the fixation activity $\sigma_i = \{\sigma_i^1, \ldots, \sigma_i^M\}$ for each subject i reading a given text with M words in terms of the state of each word r $\sigma_i^r = \pm 1$ according to the following rule,

$$\sigma_i^r = \begin{cases} +1 & \text{if } n_i^r \geq 2 \\ -1 & \text{if } n_i^r < 2 \ , \end{cases} \tag{1}$$

where n_i^r represents the number of times the subject i fixates on word r during the reading. The value of 2 fixations per word has been adopted here as threshold parameter to define whether a word is active or not in the text due to the fact that, from our eye-tracking

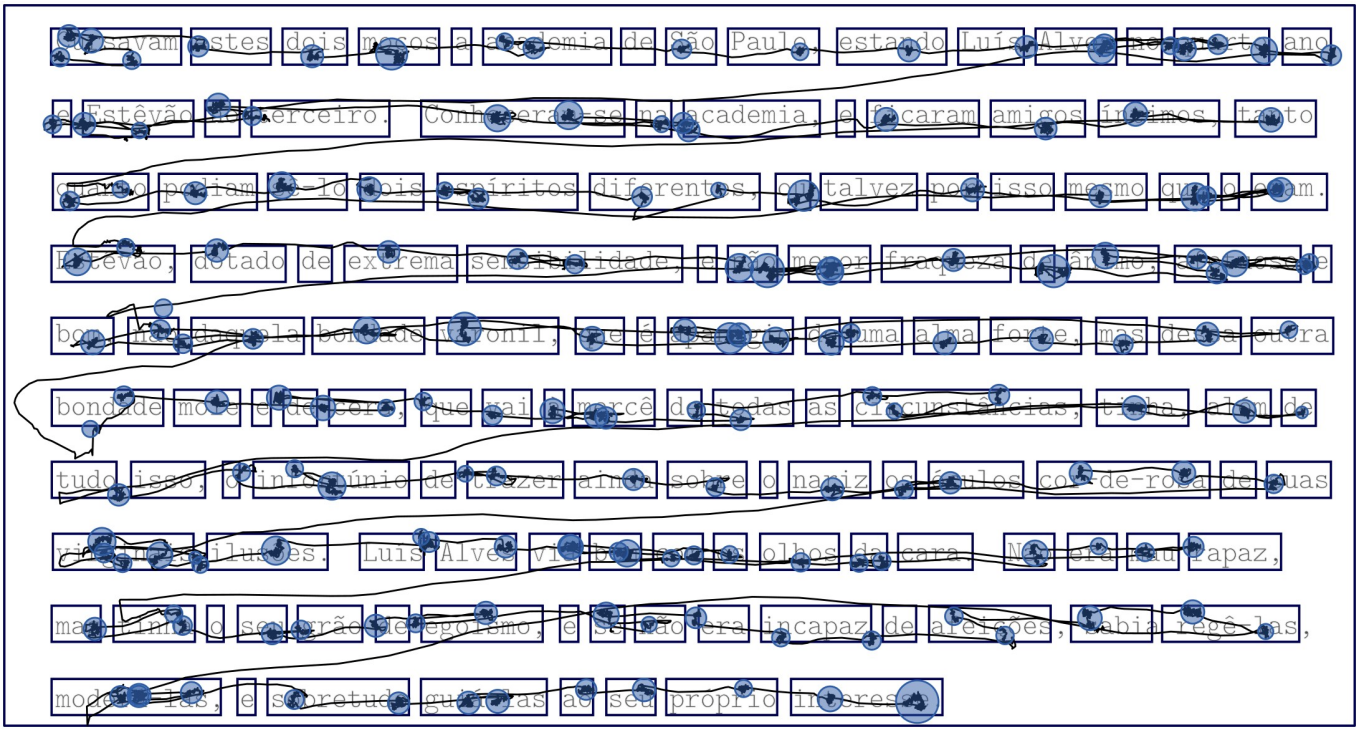


Fig 2. Eye-tracking reading pattern. Plot showing the sequences of gazes and fixations during a typical eye-tracking experiment. In this particular case, the data was collected while the subject was reading the MEL text. The blue circles represent the fixations and their sizes stand for the corresponding duration times. The solid lines between circles indicate the gaze trajectory along the text. For a given text, we measure the number of times n_i^r during the entire reading that a fixation of subject i falls into the rectangle box delimiting a word r.

https://doi.org/10.1371/journal.pone.0260236.g002

experiments, almost every word in any text was fixated at least once during the readings of all subjects. This reading pattern is compatible with observations reported previously [10]. Thus, relevant variations among the fixation activities would be detected considering the words with one fixation and those with two or more.

The raster plots corresponding to the fixation activities obtained from the eye-tracking experiments with all subjects are shown in Fig 3 for all the texts. Clear differences in the reading activity patterns can be observed. In particular, we notice that the density of active states observed for ST1 and ST2 is significantly lower than for RT1 and RT2. The fixation activity density is quantified here in terms of the "magnetization" m_i for each subject $i = 1, \ldots, N$, defined as the average of the fixation states σ_i^r over the M words of the text,

$$m_i = \langle \sigma_i \rangle = \frac{1}{M} \sum_{r=1}^{M} \sigma_i^r \ . \tag{2}$$

In this way, for every text, we can define an overall magnetization as,

$$\langle m \rangle = \frac{1}{N} \sum_{i=1}^{N} m_i \ . \tag{3}$$

A relevant measure that could be readily obtained from the experiments performed here, being certainly indicative of text processing demand, is the average reading time per word of

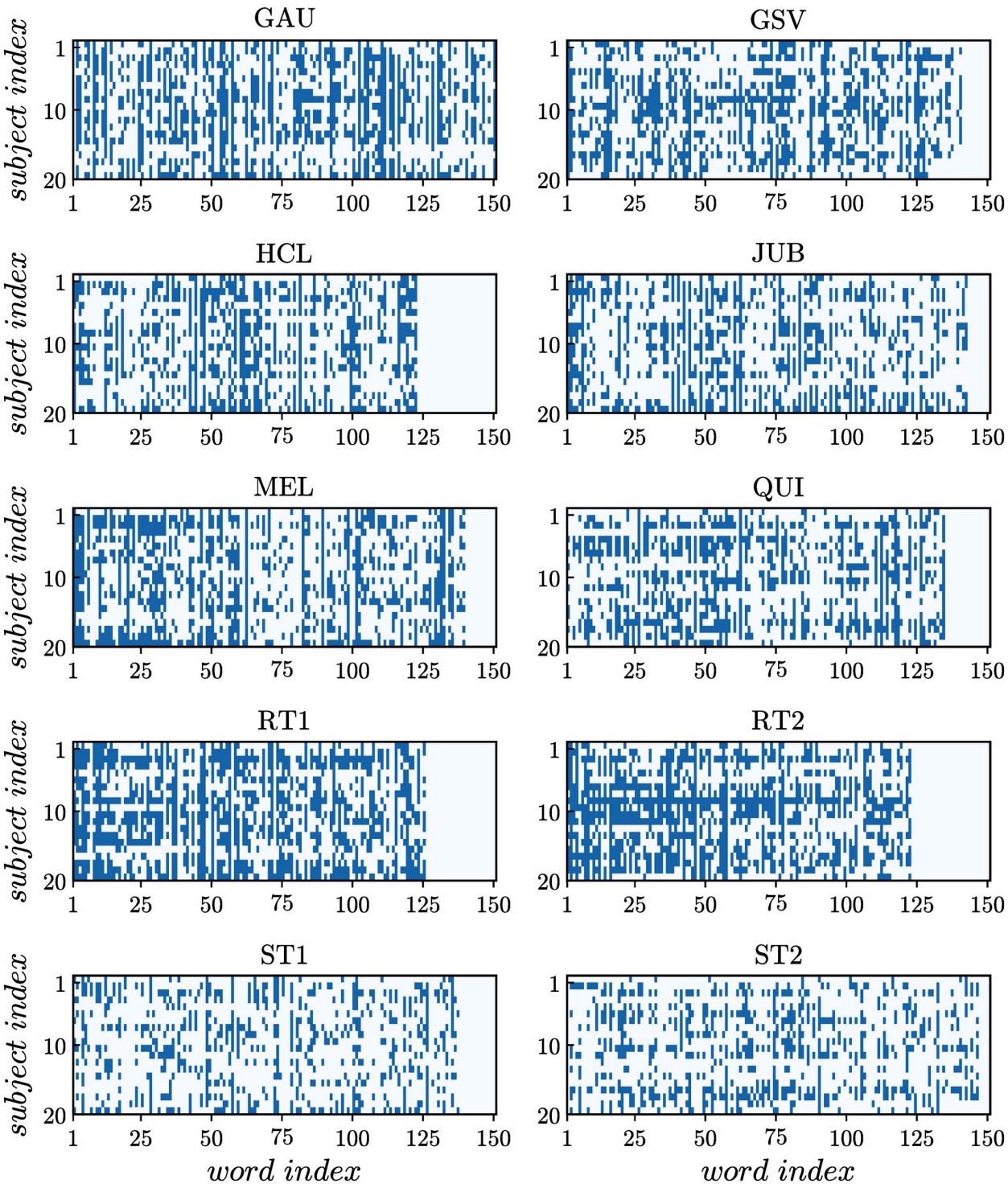


Fig 3. Fixation activities. Raster plots of the fixation activities obtained for all subjects while reading the texts. Accordingly, for each subject *i*, the state σ_i^r of a word is active (+1) if $n_i^r \geq 2$ (blue) or inactive (−1) if $n_i^r < 2$ (white).

https://doi.org/10.1371/journal.pone.0260236.g003

Table 2. Average magnetizations and reading times per word.

Text	$\langle m \rangle$	$\langle t \rangle$ (ms)
GAU	-0.2147	744.82
GSV	-0.2657	611.12
HCL	-0.2918	595.08
JUB	-0.3697	552.34
MEL	-0.2460	641.70
QUI	-0.3022	513.37
RT1	-0.0464	923.98
RT2	-0.0664	793.13
ST1	-0.5190	398.38
ST2	-0.4966	394.21

The magnetization $\langle m \rangle$ and reading time per word $\langle t \rangle$ of each text correspond to the average values of the reading fixation activity and time per word, respectively, also averaged over all subjects.

https://doi.org/10.1371/journal.pone.0260236.t002

the text,

$$\langle t \rangle = \frac{1}{N} \sum_{i=1}^{N} t_i / M \ , \tag{4}$$

where t_i is the reading time of subject i. The values of the average magnetization $\langle m \rangle$ and average reading time per word $\langle t \rangle$ obtained from the experiments for every text are reported in Table 2. Accordingly, the reading of ST1 and ST2 texts resulted in the two lowest values of both measures. Also, the similar values obtained for RT1 and RT2 that are, however, somewhat higher than for most of the other texts demonstrates some degree of correlation between the two measures. Despite the similarities, $\langle m \rangle$ and $\langle t \rangle$ are not perfectly compatible. As a matter of fact, in what follows we will show that $\langle m \rangle$ captures information on the cognitive activity while reading more subtly than $\langle t \rangle$ does.

Maximum Entropy Model

We model the data obtained from our eye-tracking experiments following the Maximum Entropy principle [31] considering a system of binary variables (the fixation activities) with pairwise couplings [32]. Let us denote $\sigma = \{\sigma_1^r, \ldots, \sigma_N^r\}$ the state of the system consisting of N subjects reading word r in a given text. Since every subject can only be in one of two states (+1 or −1), overall we have a set $\{\sigma\}$ of 2^N possible states that the system can occupy, for each word in the text. Next, we calculate the covariance C_{ij} between the fixation activities, for every pair of subjects i and j along the M words of the text,

$$C_{ij} = \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle \ , \tag{5}$$

where

$$\langle \sigma_i \sigma_j \rangle = \frac{1}{M} \sum_{r=1}^{M} \sigma_i^r \sigma_j^r \ , \tag{6}$$

and $\langle \sigma_i \rangle$ is given by Eq (2). The minimal probability distribution $P(\{\sigma\})$ that represents our system is the one that maximizes the entropy while reproducing our observations, *i.e.*, the average m_i and covariance C_{ij} for all i and j. Subject to these constraints, the form of P is the

Boltzmann's probability distribution [31] (see S1 Appendix),

$$P(\{\sigma\}) \sim \mathrm{e}^{-E/T}, \tag{7}$$

where T is analogous to a temperature and E to a Hamiltonian. This distribution results as the least biased representation for an Ising-type system like ours, with known first and second moments. Specifically, as a first approximation, the energy term has the same form of the Ising model [32],

$$E = -\sum_{i=1}^{N} h_i \sigma_i - \sum_{i,j=1}^{N} J_{ij} \sigma_i \sigma_j. \tag{8}$$

This mathematical correspondence naturally lead us to interpret h_i as the action of a local external stimulus (text) on subject i, analogous to a "random field", and J_{ij} as "coupling coefficients" between subjects i and j. Although the participants never really communicate with each other in our eye-tracking experiments, we can think of the text as a medium through which subjects i and j "interact". This means that, although the subjects read the texts individually, their fixation activities may relate to each other given similarities in their cognitive responses induced by the characteristics of the texts. These pairwise couplings or interactions between the subjects reading activities give rise to the observed correlations among them. Consequently, the correlations may lead to emergent collective effects, which can be of importance for the study of the system.

At this point, we seek compute the local fields h_i and the interactions J_{ij} by directly solving the inverse problem given by Eq (8) (see S1 Appendix). Once we infer the values of h_i and J_{ij} for all subjects that better reproduce the experimentally observed magnetizations m_i and covariances C_{ij}, while maximizing the entropy, the Boltzmann probability distribution of Eq (7) characterizes the statistics of each text. For simplicity, here we arbitrarily set the "operating temperature", namely, the reading temperature, $T_o = 1$. By doing so, from Eq (7) it is possible to compute the rate in which the average energy of a given text changes with T,

$$C_v = \frac{\partial \langle E \rangle_P}{\partial T} \tag{9}$$

This rate of change is analogous to a heat capacity, *i.e.*, a measure of how much energy the system can absorb as the temperature T increases. Moreover, at a "critical temperature" T_c, C_v is maximal, which is interpreted as a phase transition: if $T_c < T_o$, the system is in a "liquid" or random state. On the other hand, $T_c > T_o$ is indicative of a more "ordered" condition.

In Fig 4 we show the variation of C_v as a function of T for all texts. As depicted, regardless of the text, the operating temperature $T_o = 1$ is always above the critical point T_c. The distance to criticality $T_o - T_c$, however, notably depends on the text. By simply considering the fact that larger values have been found for RT1 and RT2, as compared to the other texts (see Table 3), we can obviously anticipate that this distance can be used as an index to distinguish meaningful texts from random ones. Indeed, we will show next that $T_o - T_c$ can be related to language processing in terms of the perceived coherence from reading a text.

Survey for measuring texts complexity and coherence

Quantifying the complexity and coherence of the texts from a survey. In order to validate our eye-tracking results, an Internet survey was conducted by request to the MindMiners services company [51], of São Paulo, Brazil. The agency provides the usage of a digital platform that enables to develop research projects, from questionnaire creation to data collection, through a respondents panel with more than 400 thousand engaged users distributed all over

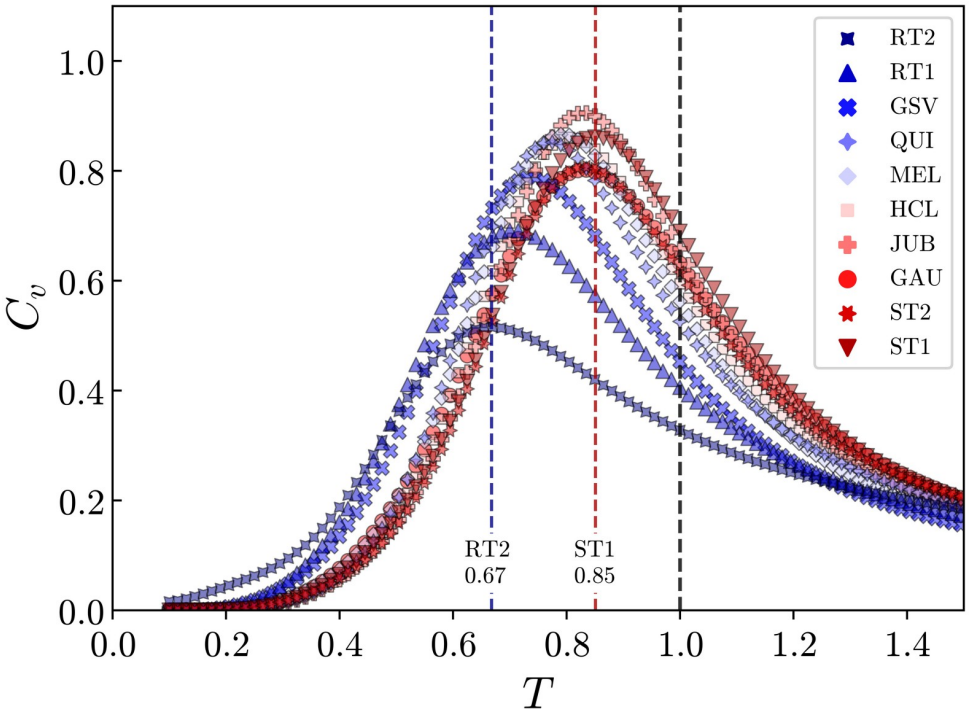


Fig 4. Heat capacity as a function of temperature for the system of fixation activities. Heat capacity curves for all texts, with C_v maximal at the critical temperature T_c. The temperature at which the texts are being read is the operating temperature $T = T_o = 1$. It can be seen that the system is above and near the critical point for all texts, and the RT1 and RT2 texts are clearly the furthest.

https://doi.org/10.1371/journal.pone.0260236.g004

Brazil (MeSeems [52]). The work methodology consists of two main stages, namely, the selection of respondents according to specific requirements of the study, and the production and revision of the questionnaire.

In our study, two groups of 200 people of diverse age, gender and place of residence were selected from the 400 thousand respondents constituting the panel of the survey agency.

Table 3. Distance to criticality.

Text	$T_o - T_c$
GAU	0.169
GSV	0.262
HCL	0.192
JUB	0.170
MEL	0.207
QUI	0.229
RT1	0.296
RT2	0.332
ST1	0.149
ST2	0.167

The table reports the distances to criticality $T_o - T_c$ calculated for different texts from eye-tracking experiments using the MEM. $T_o = 1$ is the reading operating temperature and the critical temperature T_c corresponds to the value of T where the heat capacity C_v for a given text is maximal.

https://doi.org/10.1371/journal.pone.0260236.t003

Table 4. Respondents panel data.

	GroupA		GroupB	
	Respondents	**Percentage**	**Respondents**	**Percentage**
Gender				
Male	86	43.0%	90	45.0%
Female	114	57.0%	110	55.0%
Age				
≤ 17	2	1.0%	3	1.5%
18–24	48	24.0%	46	23.0%
25–30	26	13.0%	35	17.5%
31–40	70	35.0%	61	30.5%
≥ 41	54	27.0%	55	27.5%
Education				
High school	99	49.5%	83	41.5%
University	101	50.5%	117	58.5%
Place of residence				
Central-West	20	10.0%	20	10.0%
Northeast	46	23.0%	46	23.0%
North	10	5.0%	9	4.5%
Southeast	94	47.0%	98	49.0%
South	30	15.0%	27	13.5%

Respondents stratification based on gender, age, education level (high school degree was required) and place of residence in Brazil (indicated by regions).

https://doi.org/10.1371/journal.pone.0260236.t004

Details of the stratification are shown in Table 4. A minimum of high school degree was requested to ensure mature reading-comprehension skills. We did not draw a distinction between socio-economic classes. Each group of respondents was given a questionnaire, related to 5 of the total 10 texts. The texts were divided as follows:

- Group A: O Quinze, Grande Sertão: Veredas, História do Cerco de Lisboa, Story 2, Random text 2
- Group B: Jubiabá, O Gaúcho, A mão e a Luva, Story 1, Random text 1

The questionnaire is divided into three parts and all questions are multiple-choice. In the first part, the respondent was asked to read the texts one by one and answer a simple question related to the text content, with the purpose of merely motivating a conscious reading. In the second part, each text was presented again and the respondent was asked to rate the text complexity level in a scale of 1 to 5, ranging from a "very simple text" to a "very complex text", respectively. Lastly, the texts were again presented and the respondent was asked to assess the text coherence level in a scale from 1 to 5, ranging from a "text not coherent at all" to a "very coherent text", respectively. The texts were shown in a randomized order to each respondent to minimize bias in the responses.

The MindMiners company relies on manual and automatic validation of each user's information through external data sources such as social networks and the Brazilian department of revenue. In addition, they develop algorithms to identify people with atypical behaviors such as non-compliance with questionnaire instructions and abnormal response speed. Identifiers are embedded in all of the respondent's devices, so that the respondents panel is composed exclusively of unique users.

Survey main results. In Fig 5 we show the distribution $P(\pi)$ of fractions of individuals who rated a given value of complexity π for each text. For instance, we can see that almost 40% of the respondents rated the ST1 and ST2 texts as "very simple", while up to 50% of the respondents rated the RT1 and RT2 texts as "very complex". The other texts were rated with varied ranges of intermediate complexity values. The mean values of complexity, $\langle\pi\rangle$ are shown in Table 5. The distributions $P(\psi)$ of fraction of individuals who rated the texts with a value ψ of coherence are shown in Fig 6. The results for ST1 and JUB, for example, show a very similar type of response. Most of the people ($\approx 60\%$) ranked these texts as coherent and very coherent (grades of 4 and 5), approximately 25% ranked with an intermediate level of coherence (grade 3), and only a small fraction (less than 15%) ranked the texts with a low level of coherence (grades 1 and 2). For the random texts RT1 and RT2, on the other hand, the ratings for both are mostly in favor of a "not coherent" opinion ($\approx 50\%$), while approximately 20% thinks that they possess an intermediate value of coherence, and little more than 15% ranked them as coherent or very coherent. Interestingly, the reading of the GSV text led to a distinctive type of response, namely, the larger group of respondents ($\approx 40\%$) ranked the text with the intermediate grade 3. In this case, it suggests that many individuals were undecided about their evaluations on the coherence of the text.

As shown in Table 5, the coherence mean values, $\langle\psi\rangle$, obtained from the survey for all texts suggest they can be sorted into three groups, namely, RT1 and RT2 have low levels of coherence ($\langle\psi\rangle < 2.75$), whereas the reading of ST1, ST2, JUB, HCL, MEL, QUI resulted in high coherence ratings ($\langle\psi\rangle > 3.25$). Finally, the GAU and GSV texts were rated with intermediate coherence levels ($2.75 \leq \langle\psi\rangle \leq 3.25$).

Results

The average magnetization of the fixation activity reflects the level of text complexity

As we already pointed out, the reading time span is certainly not a dispensable measure of text processing. Indeed, a correlation between reading time and the text complexity level seems then straightforward. In Fig 7(A), we plot the average reading times per word against the average values of the complexity. Although $\langle t\rangle$ generally increases with $\langle\pi\rangle$, the relation is hardly monotonic. This becomes more evident in Fig 7(B), where the crescent relative ranks of these variables are plotted against each other, and several discordant pairs in the rank order are observed, out of a total of ten pairs that can be compared with each other. Here we use a non-parametric statistic, namely, the Kendall rank correlation coefficient τ, to measure the rank correlation between $\langle t\rangle$ and $\langle\pi\rangle$ of the texts (see S3 Appendix). In spite of the discordant pairs, we find a value of $\tau = 0.87$ ($p = 0.0001$), which indicates a reasonably high degree of correlation with positive monotonicity trend between the two variables.

The situation is rather different when we plot the average magnetization, $\langle m\rangle$, against $\langle\pi\rangle$ for all texts, as shown in Fig 8(A). The two measures are highly correlated, although in a non-linear fashion, with $\langle m\rangle$ increasing almost monotonically with $\langle\pi\rangle$, except for a capricious local minimum at the average complexity of GSV. Moreover, by plotting the relative ranks of $\langle m\rangle$ and $\langle\pi\rangle$ for a given text against each other (see Fig 8(B)), we notice that, out of ten texts, eight of them occupy identical positions in both lists. As compared to the reading time per word, the higher Kendall rank correlation coefficient found in this case, $\tau = 0.96$ ($p = 5 \times 10^{-6}$), confirms that the measure $\langle m\rangle$ certainly represents a better proxy to rank the complexities $\langle\pi\rangle$ of the texts.

One may argue that more complex texts are expected to require more time for analysis, and therefore more fixations overall. However, when considering the average reading times per

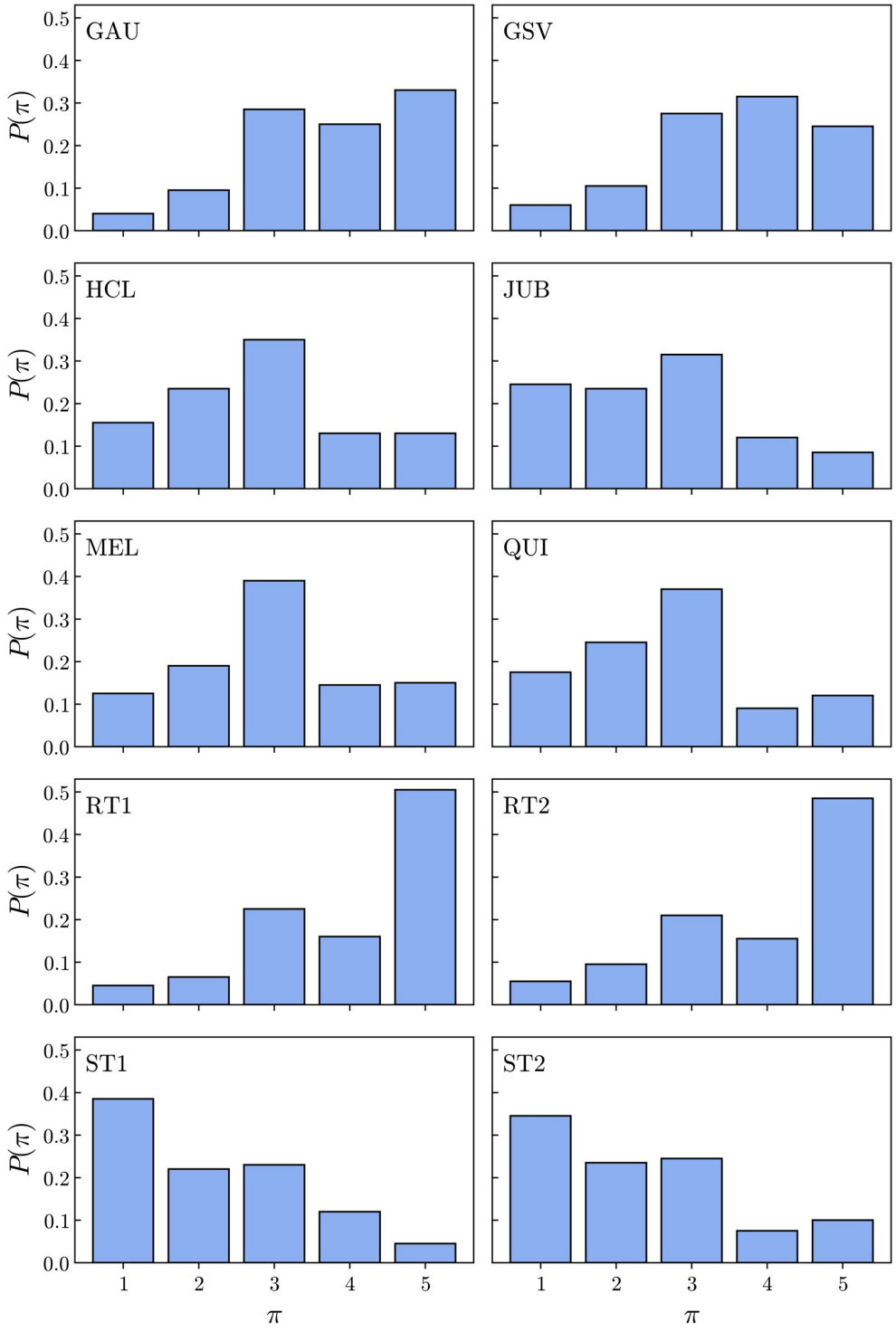


Fig 5. Distributions of complexity ratings. Distributions of complexity ratings among individuals for all texts read in the survey. The values $\pi = 1, 2, 3, 4, 5$ correspond to a scale ranging from a "very simple" text ($\pi = 1$) to a "very complex" text ($\pi = 5$).

https://doi.org/10.1371/journal.pone.0260236.g005

Table 5. Complexity and coherence mean values.

Text	$\langle \pi \rangle$	$\langle \psi \rangle$
GAU	3.74 ± 0.08	3.14 ± 0.08
GSV	3.58 ± 0.08	2.95 ± 0.09
HCL	2.85 ± 0.09	3.70 ± 0.09
JUB	2.57 ± 0.09	3.79 ± 0.08
MEL	3.01 ± 0.09	3.62 ± 0.07
QUI	2.74 ± 0.09	3.86 ± 0.09
RT1	4.02 ± 0.08	2.44 ± 0.09
RT2	3.92 ± 0.09	2.38 ± 0.09
ST1	2.22 ± 0.09	3.80 ± 0.08
ST2	2.35 ± 0.09	4.00 ± 0.09

Complexity $\langle \pi \rangle$ and coherence $\langle \psi \rangle$ mean values obtained for all texts from the survey.

https://doi.org/10.1371/journal.pone.0260236.t005

word, we observe that, for example, it takes around 80 more milliseconds per word on average to read the HCL text than to read the QUI text (a 15% difference), in spite of their similar levels of complexity, according to the survey. The same happens with the GSV and GAU texts, with the subjects spending an additional 133 milliseconds per word on average to read the latter (a difference of 20%), even though their relative difference in average complexities is smaller than 5% (see Table 5). In this regard, a comparison between Figs 7(B) and 8(B), and between their corresponding Kendall coefficients, unambiguously indicate that the average magnetization represents a more reliable indicator of the perceived complexity than the reading time per word.

Text coherence perception evidenced by distance to criticality

The results shown in Fig 9 reveal that large distances to criticality $(T_o - T_c)$ are consistent with the low coherent nature ($\langle \psi \rangle < 2.75$) of both random texts (RT1 and RT2). Moreover, all texts rated with high coherence $\langle \psi \rangle > 3.25$ (ST1, ST2, JUB, HCL, MEL, and QUI) group at the bottom-left corner of the plot due to their correspondingly small values of $(T_o - T_c)$. Coherent texts therefore prompt higher correlated responses in the fixation activity of the readers, suggesting implicit cohesive interactions among them. Interestingly, the two texts ranked with intermediate values of coherence $2.75 < \langle \psi \rangle < 3.25$ in the survey (GAU and GSV), however, induced very different responses in terms of the distance to criticality obtained from the eye-tracked readings. In order to understand this difference, a more detailed analysis is required with respect to the literary styles and linguistic aspects of the books from where these texts have been extracted, as we present in the Discussion.

It is important to stress that one can only rely on the particular features of the cross-correlations from the fixation activity series (see Fig 3) between pairs of readers for a given text to justify the clear numerical differences found among the values of $(T_o - T_c)$. In order to test for this hypothesis, we perform additional calculations with the fixation activities of the subjects for a given text, preserving the mean magnetization $\langle \sigma_i \rangle$, but shuffling the values of σ_i among randomly chosen pairs of words in the text. In this way, strong correlations, if present between the fixation activities, should disappear. Once we have shuffled the data, we follow the same sequence of calculations as before, namely, we find the pairwise correlations C_{ij}, compute the fields h_i and couplings J_{ij}, and determine the heat capacity C_v at different temperatures T. The effect of suppressing strong correlations is to substantially reduce the interactions J_{ij}, therefore

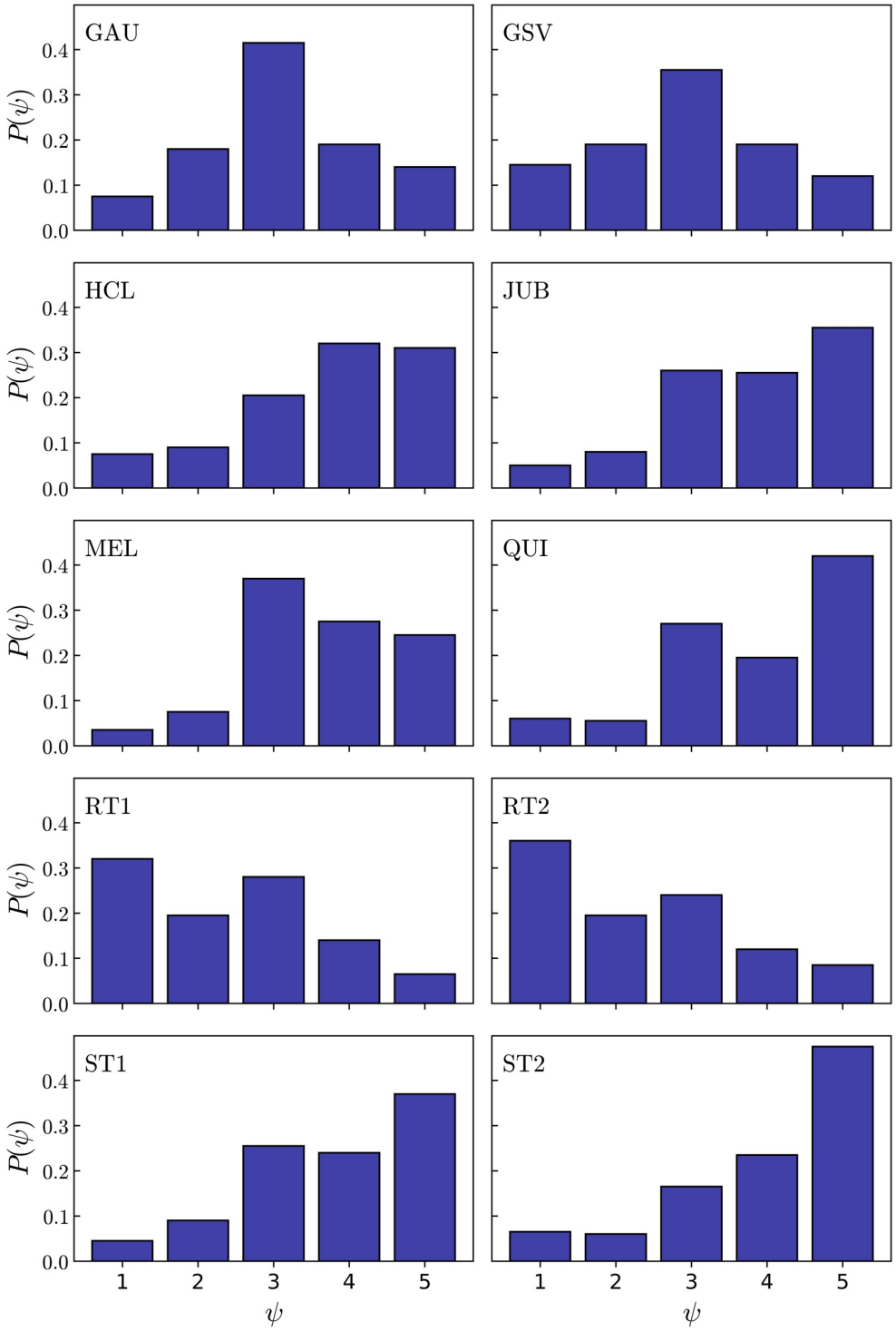


Fig 6. Distributions of coherence ratings. Distributions of coherence ratings among individuals for all texts read in the survey. The values $\psi = 1, 2, 3, 4, 5$ correspond to a scale ranging from a "not coherent" text ($\psi = 1$) to a "very coherent" text ($\psi = 5$).

https://doi.org/10.1371/journal.pone.0260236.g006

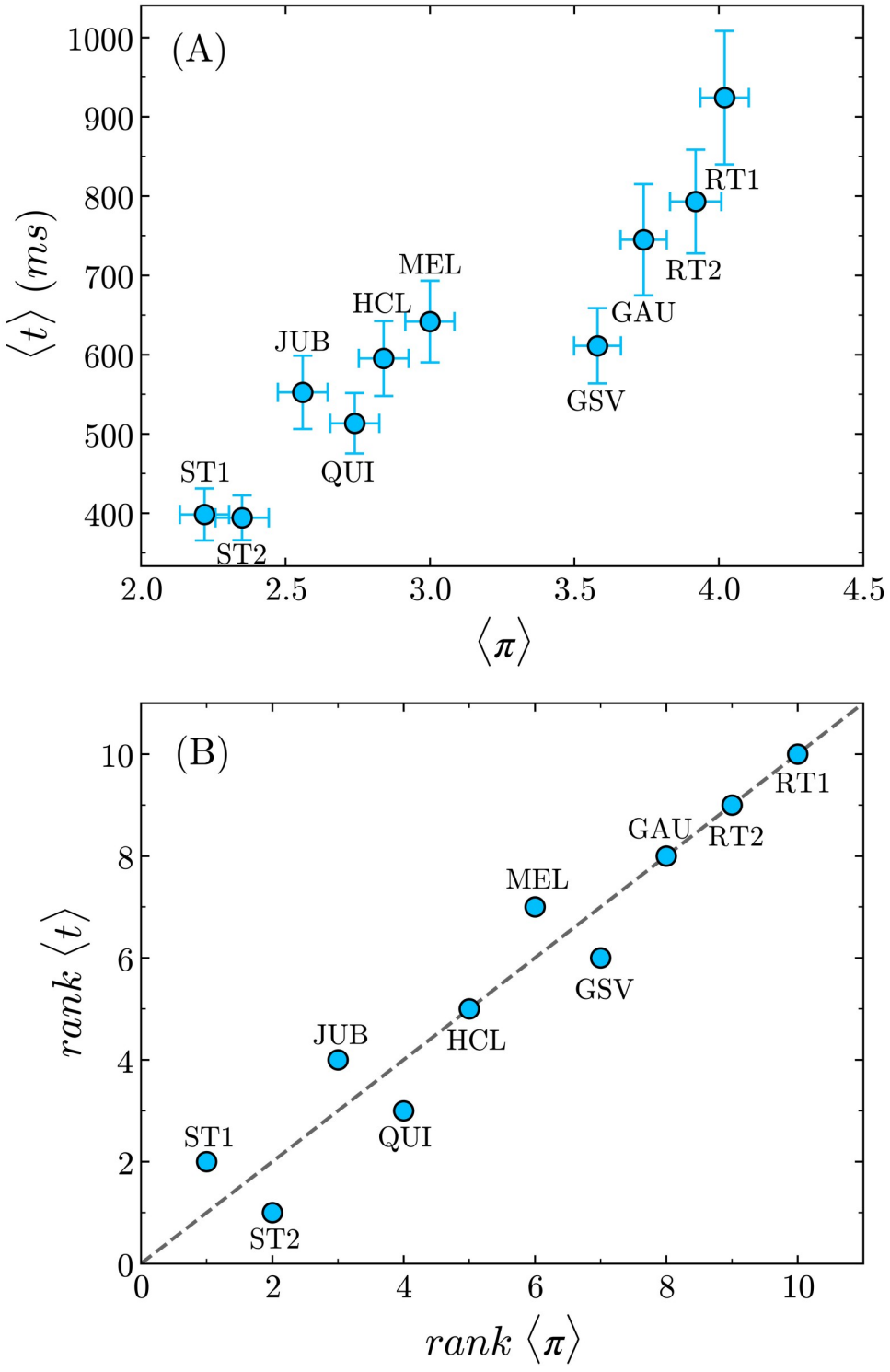


Fig 7. Reading times against text complexity. (A) The average reading time per word $\langle t \rangle$ generally increases with $\langle \pi \rangle$, although the relation is not monotonic. **(B)** The rank of $\langle t \rangle$ plotted against the rank of $\langle \pi \rangle$ shows that several discordant pairs are observed between the two variables. The dashed line corresponds to the function $y = x$. The Kendall rank correlation coefficient is $\tau = 0.87$ ($p = 0.0001$).

https://doi.org/10.1371/journal.pone.0260236.g007

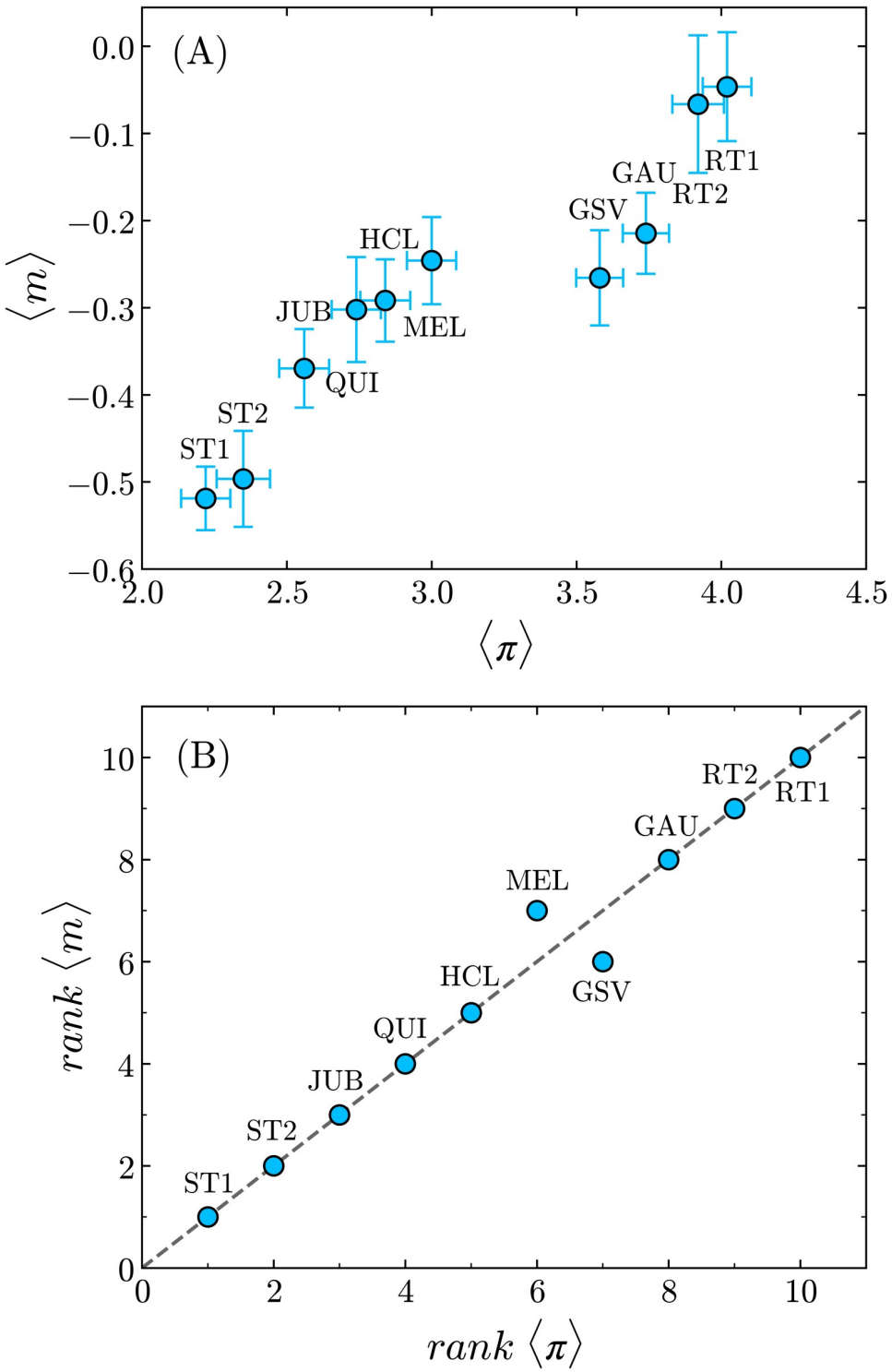


Fig 8. Average magnetization against text complexity. (A) The average magnetization, $\langle m \rangle$, of the fixation activities increases almost monotonically with $\langle \pi \rangle$, except for a local minimum at the complexity of GSV. **(B)** By ranking both measures in crescent order and plotting the ranks of $\langle m \rangle$ against the ranks of $\langle \pi \rangle$ for all texts, we can see that eight out of ten texts occupy exactly the same positions in the two lists. The dashed line corresponds to the function $y = x$. The Kendall rank correlation coefficient $\tau = 0.96$ ($p = 5 \times 10^{-6}$) indicates the very high trend of monotonicity between the two variables.

https://doi.org/10.1371/journal.pone.0260236.g008

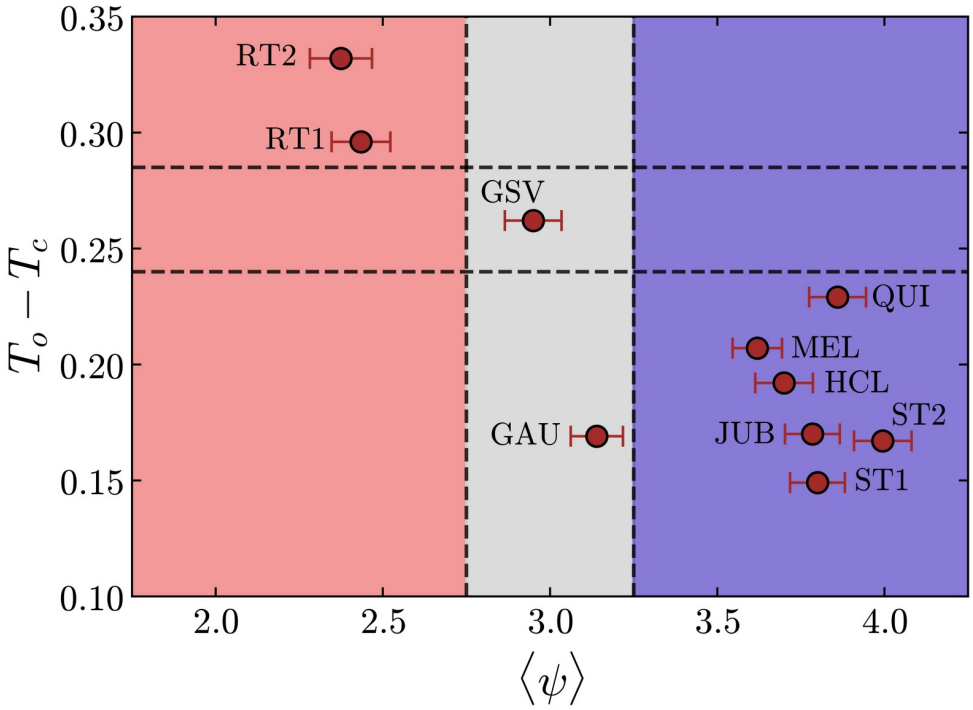


Fig 9. Distance to criticality and text coherence. Relation between the distance to criticality $T_o - T_c$ and the average coherence $\langle\psi\rangle$ of the texts. Texts rated with low coherence $\langle\psi\rangle < 2.75$ are associated with large values of $T_o - T_c$ (RT1 and RT2), while texts considered to be coherent $\langle\psi\rangle > 3.25$ are close to criticality (ST1, ST2, JUB, HCL, MEL, QUI), suggesting an implicit cohesive reading response among individuals. The two texts rated with intermediate values of $\langle\psi\rangle$ (GAU and GSV), however, induced rather distinct responses in terms of $T_o - T_c$.

https://doi.org/10.1371/journal.pone.0260236.g009

decreasing the value of T_c for all texts and increasing their distance ($T_o - T_c$) (see S2 Appendix).

Discussion

Linguists have long studied the notions of text complexity and coherence. More recently, attempts to measure these quantities have been made through mathematical expressions, known as readability formulas, that mainly rely in metrics of word and sentence lengths and word frequency. However, it is arguable whether these formulas are sufficient or not to determine the complexity of a text, mostly for two sound reasons. First, there are uncomplicated pieces of writing that use many infrequent words (an indicator of high complexity for these formulas), like informational texts, as the one suggested in [25], "Any text on raccoons would use 'raccoon' a lot, as well as 'nocturnal' and 'foraging'". This is a typical example for which most readability expressions would overrate the complexity score of the text. Second, texts encompassing elaborated ideas can nevertheless be written with words from a simple vocabulary and constituted of short sentences. This is the case, for example, of some texts containing abstract narratives, usage of metaphors and obscure allusions, for which those complexity scores would be underrated. A well-known example in the literature is Ernest Hemingway's book "The Old Man and the Sea" that could be easily underrated in complexity by readability formulas, despite its profoundness and story-rich writing. As a consequence, in addition to using these metrics, a qualitative analysis is usually recommended by linguists in order to more adequately categorize the reading material.

Here, instead of applying empirical mathematical expressions, we approached the problem of quantifying the complexity and coherence of texts by using data from an extensive survey with a group of 400 people, and comparing it with properties from reading patterns obtained with eye-tracking experiments. We calculated the magnetizations, the most elemental measure of our system that represents the density of fixation activity that the readers had for each text. We hypothesized and confirmed with the survey that the more complex the text the greater the average magnetization. These results therefore suggest that the fixation acitivity, as we defined here, contains the sufficient cognitive information to characterize the reading patterns in terms of active and inactive states. The adopted threshold for the number of fixations in a given word r ($n_i^r = 2$) succeeded in delimiting the minimum to characterize the words in the text that need further cognitive processing. Furthermore, it somehow accounts for the effect of repetitive fixation due to word size and low frequency, while relativizing the underlying variations among the fixation patterns of individuals related to their particular reading skills and capability to predict the occurrence of words in context. This was verified by testing larger values of threshold. Already for a threshold of 3, the fixation activities become heavily dominated by the words size and/or their frequency, showing significant dissimilarities among subjects. In this framework, the threshold adopted in our work acts as a simple, but surprisingly effective way to somehow attenuate differences among subjects.

The fact that the randomly generated texts (RT1 and RT2) were rated with low levels of coherence, while most of literary texts were considered to be coherent by the readers in the survey could be anticipated. The intermediate ratings of coherence for the literary texts GSV and GAU, however, are certainly worth of a more detailed analysis. The celebrated Brazilian author, Guimarães Rosa, who wrote "Grande Sertão: Veredas", from which the fragment GSV was extracted, is well-known for his distinctive writing style, frequently compared to that of James Joyce, in what concerns to the astonishing linguistic work and experimentation [53].

In the literary work of Rosa, we often find unconventional punctuation and grammar in story-rich writing, while creating neologisms from erudite and popular expressions, regionalisms, archaic words and inventive use of prefixes and suffixes [53, 54]. In fact, many of these linguistic features are found in the GSV excerpt used here, with which the book opens. The first word of the text, "nonada" is already an unusual term that, although existing in the Portuguese language, is old-fashioned and hardly used in literature. Even in context, the expression seems so enigmatic that it has led to different interpretations over the years [55–57]. The second sentence has a non-traditional syntactic structure, typical of regional orality, as we learn from the study of Garcia [58]. We see that this linguistic resource is found in the fifth and fifteenth sentences as well. The author also makes use of incomplete suggestive expressions in three sentences of the text (second, fifth and seventh sentences), a linguistic construction typical of Rosa's writing. In addition to this, there are two neologisms in the text created by the author, namely, "erroso" and "prascovio" (in the eighth and thirteenth sentences, respectively). Without entering into a denser analysis, it is fair to say that the GSV fragment is not part of a conventional literary work, being quite difficult to grasp, especially when removed from the global context of the book's narrative. We therefore conjecture that the reader might feel confused and undecided when processing the text, finding it hard to qualify the narrative as coherent.

The text GAU, on its turn, is a transcription from the novel "O Gaúcho", written by José de Alencar in the year 1870. The fragment was extracted from the end of chapter one, where the setting in which the story takes place is described. The writing is characterized by an overwhelming, philosophical representation of the scenario [59]. The abstract tone in the narrative possibly gives the reader an impression that the text is somehow vague, leading to the uncertainty in qualifying it as coherent. In direct contrast with GSV and GAU, the other excerpts of

literary works employed here correspond to passages of descriptive, straightforward writings (HCL, MEL), or linear, plain storytelling (QUI, JUB, ST1, ST2), which very likely make them easy to interpret and therefore be considered as coherent.

Our results from a very simple statistical model and from the analysis with the Pairwise Maximum-Entropy method revealed that the distance to the critical point was capable to segregate the texts into three main groups. The random generated texts (RT1 and RT2) are the farthest from the critical point, with the operating temperatures T_o significantly higher than T_c, the GSV text follows just behind, and the rest of the texts fall much closer to T_c. As we argued previously, in the physical context of critical phenomena, when $T_o > T_c$, the interactions between the component elements are weak and the system is in a disordered state. Our results then suggest that the fixation activities for texts with low coherence (RT1 and RT2) are random to a certain degree, meaning that the reading response to the text stimuli does not promote strong virtual connections among different individuals. When T_o approaches T_c, the "interactions" among elements increase and local effects can propagate over the entire group. The system then becomes susceptible to global changes, and a collective behavior may emerge. This effect has been observed in the texts that were rated with high levels of coherence (ST1, ST2, JUB, HCL, MEL, and QUI) and also with the GAU text, although it was rated with intermediate average coherence. We reason that a high degree of coherence in a text is likely to induce a cohesive reading response, here manifested in terms of a proximity to its critical point. Although the readers never interact with one another, we can think of them responding with a similar cognitive behavior when the content of the text is consistent. A question that naturally arises is why the relation between the average coherence rating $\langle\psi\rangle$ and $T_o - T_c$ is ambiguous for the GAU and GSV texts, given that the former falls into the cluster of texts with operating temperatures close to T_c, while the latter is far from T_c, and still both of them were rated on average with intermediate levels of coherence. Previously we elaborated on the characteristics of these texts, and referred to the linguistic features that made them stand out from the other literary fragments investigated here. On the case of the GAU excerpt, the reader has to process an intelligible text, and can yet have a dubious interpretation due to the abstract style of the writing. Perhaps rating this type of text with a specific value of coherence is equivocal, but we can fairly state that the content of the text is coherent, *i.e.*, its narrative is consistent. In contrast, the GSV text appears atypical to an average reader because of its highly technical writing and uncommon linguistic elements. In a sense, we can think of the GSV as an intermediate type of text, in between a concrete narrative and a random incongruous one. Taking this and the fact that GAU otherwise induced a low value for $T_o - T_c$, the results shown in Fig 9 evidence that the distance to the critical point is actually segregating the texts according to some coherence measure. Such a measure, which originates from an inner cognitive mechanism, is perhaps less subjected to the influence of extrinsic factors than the response to a questionnaire within the protocol of a digital survey.

Conclusion

In summary, the results presented here show that eye-tracking data can be duly processed and analyzed to produce consistent proxies for complexity and coherence of diverse texts. The same texts, including children stories, random word generated texts and excerpts from literature work, have been used to validate this hypothesis by means of an extensive Internet survey with a large number of readers. Our results were substantiated by (*i*) the nearly monotonic relation between the average magnetization $\langle m\rangle$ of the fixation activities and the average complexity $\langle\pi\rangle$ of the texts and (*ii*) the suitability of the distance $(T_o - T_c)$ to segregating random texts (meaningless in content, but with preserved grammar structures) from coherent ones.

We recall that the curve $T - T_c$ for each text is computed from the "energy" of the system, which we obtain by applying the maximum-entropy learning algorithm to the fixation activities of all eye-tracked readers. This finding is particularly significant for several reasons. For one thing, it is another example of how learning algorithms are efficient in extracting relevant information out of large amounts of experimental data, and specifically it supports the maximum-entropy approach as an elementary yet solid method to study complex systems. At the same time, we get to notice how humans respond cohesively to a coherent, consistent text, which is indicative of the advanced language formation and reading prediction mechanisms that we have developed. Instead, when the written information is nonsensical, the collective cognitive response is dispersed.

Supporting information

S1 Appendix. Derivation of the Maximum-Entropy Model and inverse Ising problem.
(PDF)

S2 Appendix. Randomization of fixation activities.
(PDF)

S3 Appendix. Kendall correlation coefficient.
(PDF)

S4 Appendix. Survey raw data.
(PDF)

S1 Fig. Experimental values of magnetizations against theoretical values. The calculated values $\langle \sigma_i \rangle_{th}$ reproduce the experimental values $\langle \sigma_i \rangle$, as evidenced by the plots falling on the red dashed lines corresponding to $y = x$.
(JPG)

S2 Fig. Experimental values of covariances against theoretical values. The calculated values $\langle \sigma_i \sigma_j \rangle_{th}$ reproduce the experimental values $\langle \sigma_i \sigma_j \rangle$, as evidenced by the plots falling on the red dashed lines corresponding to $y = x$.
(JPG)

S3 Fig. Heat capacity as a function of temperature for the system of fixation activities with shuffled data. Average heat capacity curves for all texts, after shuffling the values of fixation states σ_i among randomly chosen pairs of words in the text. The average values are calculated over 100 shuffling trials and the error bars are smaller than the symbols. This suppresses strong correlations, here evidenced by a a significant increase of the distance to the critical point ($T_o - T_c$).
(JPG)

S4 Fig. Distance to criticality and text coherence with shuffled data. Relation between the average distance to criticality $\langle T_o - T_c \rangle$ and the average coherence $\langle \psi \rangle$ of the texts, after shuffling the values of fixation states σ_i among randomly chosen pairs of words in the text.
(JPG)

S1 Table. Distance to criticality with shuffled data. The table reports the average distances to criticality $\langle T_o - T_c \rangle$ calculated using the MEM by shuffling the data from the fixation maps of the eye-tracking experiments (average calculated over 100 trials). $T_o = 1$ is the reading operating temperature and the critical temperature T_c corresponds to the value of T where the heat capacity C_v for a given text is maximal.
(PDF)

S2 Table. Group A and Group B survey raw data.
(PDF)

Author Contributions

Conceptualization: Débora Torres, Humberto A. Carmona, André A. Moreira, Hernán A. Makse, José S. Andrade, Jr.

Data curation: Débora Torres, Wagner R. Sena.

Formal analysis: Débora Torres, Wagner R. Sena, José S. Andrade, Jr.

Funding acquisition: José S. Andrade, Jr.

Investigation: Débora Torres, José S. Andrade, Jr.

Methodology: Débora Torres, Wagner R. Sena, José S. Andrade, Jr.

Project administration: José S. Andrade, Jr.

Resources: José S. Andrade, Jr.

Supervision: José S. Andrade, Jr.

Visualization: Débora Torres, Humberto A. Carmona, Hernán A. Makse, José S. Andrade, Jr.

Writing – original draft: Débora Torres, Humberto A. Carmona, André A. Moreira, Hernán A. Makse, José S. Andrade, Jr.

Writing – review & editing: Débora Torres, Humberto A. Carmona, André A. Moreira, Hernán A. Makse, José S. Andrade, Jr.

References

1. Wade NJ. Pioneers of Eye Movement Research. i-Perception. 2010; 1(2):33–68. https://doi.org/10.1068/i0389 PMID: 23396982
2. Javal E, Ciuffreda KJ, Bassil N. Essay on the physiology of reading. Ophthalmic and Physiological Optics. 1990; 10(4):381–384. https://doi.org/10.1111/j.1475-1313.1990.tb00885.x PMID: 2263372
3. Lamare M. Des mouvements des yeux pendants la lecture. Bulletins et Mémoires de la Société Française d'Ophthalmologie. 1892; 10:354–364.
4. Brown AC. A Lecture ON THE RELATION BETWEEN THE MOVEMENTS OF THE EYES AND THE MOVEMENTS OF THE HEAD. The Lancet. 1895; 145(3743):1293–1298. https://doi.org/10.1016/S0140-6736(01)94423-X
5. Hering E. Spatial Sense and Movements of the Eye. Oxford, England: American Academy of Optometry; 1942.
6. Yarbus AL. Eye Movements and Vision. Boston, MA: Springer; 1967.
7. Credidio HF, Teixeira EN, Reis SDS, Moreira AA, Andrade JS Jr. Statistical patterns of visual search for hidden objects. Scientific Reports. 2012; 2(1). https://doi.org/10.1038/srep00920 PMID: 23226829
8. Amor TA, Reis SDS, Campos D, Herrmann HJ, Andrade JS Jr. Persistence in eye movement during visual search. Scientific Reports. 2016; 6(1). https://doi.org/10.1038/srep20815 PMID: 26864680
9. Amor TA, Luković M, Herrmann HJ, Andrade JS Jr. Influence of scene structure and content on visual search strategies. Journal of The Royal Society Interface. 2017; 14(132). https://doi.org/10.1098/rsif.2017.0406 PMID: 28747401
10. Clifton C, Ferreira F, Henderson JM, Inhoff AW, Liversedge SP, Reichle ED, et al. Eye movements in reading and information processing: Keith Rayner's 40year legacy. Journal of Memory and Language. 2016; 86:1–19. https://doi.org/10.1016/j.jml.2015.07.004
11. Rayner K, Duffy SA. Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. Mem Cognit. 1986; 14(3):191–201. https://doi.org/10.3758/BF03197692 PMID: 3736392

12. Kliegl R, Nuthmann A, Engbert R. Tracking the mind during reading: The influence of past, present, and future words on fixation durations. Journal of Experimental Psychology. 2006; 135:12–35. https://doi.org/10.1037/0096-3445.135.1.12 PMID: 16478314

13. Gaskell MG, Staub A, Rayner K. 19. In: Eye movements and on-line comprehension processes. Oxford University Press; 2007. p. 327–342.

14. Rayner K, Raney GE. Eye movement control in reading and visual search: Effects of word frequency. Psychonomic Bulletin & Review. 1996; 3:245–248. https://doi.org/10.3758/BF03212426 PMID: 24213875

15. Kliegl R, Grabner E, Rolfs M, Engbert R. Length, frequency, and predictability effects of words on eye movements in reading. European Journal of Cognitive Psychology. 2004; 16(1-2):262–284. https://doi.org/10.1080/09541440340000213

16. Ashby J, Rayner K, Clifton C. Eye movements of highly skilled and average readers: differential effects of frequency and predictability. The Quarterly journal of experimental psychology A, Human experimental psychology. 2005; 58(6):1065–86. https://doi.org/10.1080/02724980443000476 PMID: 16194948

17. Rayner K, Reichle ED, Stroud MJ, Williams CC, Pollatsek A. The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. Psychology and Aging. 2006; 21:448–465. https://doi.org/10.1037/0882-7974.21.3.448 PMID: 16953709

18. Juhasz BJ, Liversedge SP, White SJ, Rayner K. Eye Movements and the Use of Parafoveal Word Length Information in Reading. J Exp Psychol Hum Percept Perform. 2008; 34(6):1560–1579. https://doi.org/10.1037/a0012319 PMID: 19045993

19. Ehrlich SF, Rayner K. Contextual effects on word perception and eye movements during reading. J Verb Learn Verb Be. 1981; 20(6):641–655. https://doi.org/10.1016/S0022-5371(81)90220-6

20. Schad DJ, Nuthmann A, Engbert R. Eye movements during reading of randomly shuffled text. Vision Research. 2010; 50(23):2600–2616. https://doi.org/10.1016/j.visres.2010.08.005 PMID: 20719240

21. Reichle ED, Pollatsek A, Fisher DL, Rayner K. Toward a model of eye movement control in reading. Psychological Review. 1998; 105:125–157. https://doi.org/10.1037/0033-295X.105.1.125 PMID: 9450374

22. Engbert R, Kliegl R. Mathematical models of eye movements in reading: a possible role for autonomous saccades. Biological cybernetics. 2001; 85(2):77–87. https://doi.org/10.1007/PL00008001 PMID: 11508778

23. Engbert R, Longtin A, Kliegl R. A dynamical model of saccade generation in reading based on spatially distributed lexical processing. Vision Research. 2002; 42(5):621–636. https://doi.org/10.1016/S0042-6989(01)00301-7 PMID: 11853779

24. Engbert R, Kliegl R, Longtin A. Complexity of eye movements in reading. International Journal of Bifurcation and Chaos. 2004; 14(2):493–503. https://doi.org/10.1142/S0218127404009491

25. Rothman R. The Complex Matter of Text Complexity. Harvard Education Letter. 2012; 28(5).

26. McNamara DS, Kintsch E, Songer NB, Kintsch W. Are Good Texts Always Better? Interactions of Text Coherence, Background Knowledge, and Levels of Understanding in Learning From Text. Cognition and Instruction. 1996; 14(1):1–43. https://doi.org/10.1207/s1532690xci1401_1

27. Fisher D, Frey N, Lapp D. Text Complexity: Raising Rigor in Reading. International Reading Association; 2012.

28. Nelson J, Perfetti C, Liben D, Liben M. Measures of text difficulty: Testing their predictive value for grade levels and student performance; 2012. https://achievethecore.org/page/1196/measures-of-text-difficulty-testing-their-predictive-value-for-grade-levels-and-student-performance.

29. Reinhart T. Conditions for Text Coherence. Poetics Today. 1980; 1(4):161–180. https://doi.org/10.2307/1771893

30. Shannon CE. A Mathematical Theory of Communication. SIGMOBILE Mob Comput Commun Rev. 2001; 5(1):3–55. https://doi.org/10.1145/584091.584093

31. Jaynes ET. Information Theory and Statistical Mechanics. Phys Rev. 1957; 106(4):620–630. https://doi.org/10.1103/PhysRev.106.620

32. Nguyen HC, Zecchina R, Berg J. Inverse statistical problems: from the inverse Ising problem to data science. Advances in Physics. 2017; 66(3):197–261. https://doi.org/10.1080/00018732.2017.1341604

33. Schneidman E, Berry MJ, Segev R, Bialek W. Weak pairwise correlations imply strongly correlated network states in a neural population. Nature. 2006; 440:1007–1012. https://doi.org/10.1038/nature04701 PMID: 16625187

34. Cocco S, Leibler S, Monasson R. Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. Proceedings of the National Academy of Sciences. 2009; 106(33):14058–14062. https://doi.org/10.1073/pnas.0906705106 PMID: 19666487

35. Shlens J, Field GD, Gauthier JL, Grivich MI, Petrusca D, Sher A, et al. The Structure of Multi-Neuron Firing Patterns in Primate Retina. Journal of Neuroscience. 2006; 26(32):8254–8266. https://doi.org/10.1523/JNEUROSCI.1282-06.2006 PMID: 16899720

36. Tang A, Jackson D, Hobbs J, Chen W, Smith JL, Patel H, et al. A Maximum Entropy Model Applied to Spatial and Temporal Correlations from Cortical Networks In Vitro. Journal of Neuroscience. 2008; 28 (2):505–518. https://doi.org/10.1523/JNEUROSCI.3359-07.2008 PMID: 18184793

37. Watanabe T, Hirose S, Wada H, Imai Y, Machida T, Shirouzu I, et al. A pairwise maximum entropy model accurately describes resting-state human brain networks. Nature communications. 2013; 4. https://doi.org/10.1038/ncomms2388 PMID: 23340410

38. Morcos F, Pagnani A, Lunt B, Bertolino A, Marks DS, Sander C, et al. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. Proceedings of the National Academy of Sciences. 2011; 108(49):E1293–E1301. https://doi.org/10.1073/pnas.1111471108 PMID: 22106262

39. Weigt M, White RA, Szurmant H, Hoch JA, Hwa T. Identification of direct residue contacts in protein–protein interaction by message passing. Proceedings of the National Academy of Sciences. 2009; 106 (1):67–72. https://doi.org/10.1073/pnas.0805923106 PMID: 19116270

40. Stein RR, Marks DS, Sander C. Inferring Pairwise Interactions from Biological Data Using Maximum-Entropy Probability Models. PLoS Comput Biol. 2015; 11(7):1–22. https://doi.org/10.1371/journal.pcbi.1004182 PMID: 26225866

41. Lezon TR, Banavar JR, Cieplak M, Maritan A, Fedoroff NV. Using the principle of entropy maximization to infer genetic interaction networks from gene expression patterns. Proceedings of the National Academy of Sciences. 2006; 103(50):19033–19038. https://doi.org/10.1073/pnas.0609152103 PMID: 17138668

42. Locasale JW, Wolf-Yadlin A. Maximum Entropy Reconstructions of Dynamic Signaling Networks from Quantitative Proteomics Data. PLOS ONE. 2009; 4(8):1–10. https://doi.org/10.1371/journal.pone.0006522 PMID: 19707567

43. Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, Viale M, et al. Statistical mechanics for natural flocks of birds. Proceedings of the National Academy of Sciences. 2012; 109(13):4786–4791. https://doi.org/10.1073/pnas.1118633109 PMID: 22427355

44. Bialek W, Cavagna A, Giardina I, Mora T, Pohl O, Silvestri E, et al. Social interactions dominate speed control in poising natural flocks near criticality. Proceedings of the National Academy of Sciences. 2014; 111(20):7212–7217. https://doi.org/10.1073/pnas.1324045111 PMID: 24785504

45. Burleson-Lesser K, Morone F, DeGuzman P, Parra LC, Makse HA. Collective Behaviour in Video Viewing: A Thermodynamic Analysis of Gaze Position. PLoS One. 2017; 12(1):1–19. https://doi.org/10.1371/journal.pone.0168995 PMID: 28045963

46. Bury T. Market structure explained by pairwise interactions. Physica A: Statistical Mechanics and its Applications. 2013; 392(6):1375–1385. https://doi.org/10.1016/j.physa.2012.10.046

47. RANDOM TEXT GENERATOR;. http://randomtextgenerator.com/.

48. SR Research Eye Link—Eye tracker;. https://www.sr-research.com/.

49. Ghaoui C. Encyclopedia of Human Computer Interaction. ITPro collection. Idea Group Reference; 2005.

50. Raney GE, Campbell SJ, Bovee JC. Using Eye Movements to Evaluate the Cognitive Processes Involved in Text Comprehension. Journal of Visualized Experiments: JoVE. 2014;(83):641–655. https://doi.org/10.3791/50780 PMID: 24457916

51. MindMiners—Pesquisa Digital;. https://mindminers.com/.

52. MeSeems—Respondents Panel;. https://meseems.com.br/.

53. Almino J. Guimarães Rosa, do Sertão às fronteiras. Revista Brasileira (Academia Brasileira de Letras). 2018; 96:19–36.

54. Silviano S. In: Genealogia da ferocidade: Ensaio sobre Grande Sertão: Veredas. 1st ed. Companhia Editora de Pernambuco (CEPE); 2017. p. 21–23.

55. Zilly B. "Procuro chocar e estranhar o leitor" Grande Sertão: Veredas—a poética da criação e da tradução. Revista do Programa de Estudos Pós-Graduados em Literatura e Crítica Literária da PUC-SP. 2017; 19:4–31.

56. de Castro NL. Universo e vocabulário do Grande sertão. Coleção Documentos brasileiros. J. Olympio; 1970.

57. de Castro MA. In: O homem provisório no Grande Sertão: um estudo de Grande sertão: Veredas. Biblioteca Tempo universitário. Edições Tempo Brasileiro; 1976. p. 44–44.

58. García MS. Grande Sertão: Veredas, de João Guimarães Rosa. Análise textual da obra e duas traduções ao espanhol; 2015. Available from: https://repositorio.ufsc.br/handle/123456789/160576.

59. de Alencar Araripe Júnior T. In: José de Alencar: perfil literario. Rio de Janeiro: Typ. da Escola de Serafim José Alves; circa 1880. p. 140–146. Available from: https://digital.bbm.usp.br/handle/bbm/5206.

Computer Methods and Programs in Biomedicine 182 (2019) 105053

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Saliency-driven system models for cell analysis with deep learning

Daniel S. Ferreira[a,b,c,d,*], Geraldo L. B. Ramalho[d], Débora Torres[e], Alessandra H. G. Tobias[f], Mariana T. Rezende[f], Fátima N. S. Medeiros[c], Andrea G. C. Bianchi[f], Cláudia M. Carneiro[f], Daniela M. Ushizima[a,b,c]

[a] Berkeley Institute of Data Science, University of California, Berkeley, CA, USA
[b] Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[c] Departamento de Engenharia de Teleinformática, Universidade Federal do Ceará, Fortaleza, CE, Brazil
[d] Instituto Federal de Educação, Ciência e Tecnologia do Ceará, Maracanaú, CE, Brazil
[e] Departamento de Física, Universidade Federal do Ceará, Fortaleza, CE, Brazil
[f] Universidade Federal de Ouro Preto, Ouro Preto, MG, Brazil

ARTICLE INFO

Article history:
Received 24 November 2018
Revised 24 August 2019
Accepted 25 August 2019

MSC:
41A05
41A10
65D05
65D17

Keywords:
Saliency prediction
Convolutional neural network
Cell analysis
Eye tracking experiment

ABSTRACT

Background and objectives: Saliency refers to the visual perception quality that makes objects in a scene to stand out from others and attract attention. While computational saliency models can simulate the expert's visual attention, there is little evidence about how these models perform when used to predict the cytopathologist's eye fixations. Saliency models may be the key to instrumenting fast object detection on large Pap smear slides under real noisy conditions, artifacts, and cell occlusions. This paper describes how our computational schemes retrieve regions of interest (ROI) of clinical relevance using visual attention models. We also compare the performance of different computed saliency models as part of cell screening tasks, aiming to design a computer-aided diagnosis systems that supports cytopathologists.

Method: We record eye fixation maps from cytopathologists at work, and compare with 13 different saliency prediction algorithms, including deep learning. We develop cell-specific convolutional neural networks (CNN) to investigate the impact of bottom-up and top-down factors on saliency prediction from real routine exams. By combining the eye tracking data from pathologists with computed saliency models, we assess algorithms reliability in identifying clinically relevant cells.

Results: The proposed cell-specific CNN model outperforms all other saliency prediction methods, particularly regarding the number of false positives. Our algorithm also detects the most clinically relevant cells, which are among the three top salient regions, with accuracy above 98% for all diseases, except carcinoma (87%). Bottom-up methods performed satisfactorily, with saliency maps that enabled ROI detection above 75% for carcinoma and 86% for other pathologies.

Conclusions: ROIs extraction using our saliency prediction methods enabled ranking the most relevant clinical areas within the image, a viable data reduction strategy to guide automatic analyses of Pap smear slides. Top-down factors for saliency prediction on cell images increases the accuracy of the estimated maps while bottom-up algorithms proved to be useful for predicting the cytopathologist's eye fixations depending on parameters, such as the number of false positive and negative. Our contributions are: comparison among 13 state-of-the-art saliency models to cytopathologists' visual attention and deliver a method that the associate the most conspicuous regions to clinically relevant cells.

1. Introduction

Visual attention consists of a set of cognitive processes that enables focusing on a region or object while ignoring irrelevant stimuli from the environment, which allows humans and other animals to extract relevant information from complex input scenes [1]. The attention arises from both stimuli-driven factors (bottom-up attention) and task-driven factors (top-down attention) [2]. Bottom-up attention is fast, involuntary and guided by visual distinctness or rarity using low-level image information such as orientation, color, intensity and texture. Top-down attention is slower, voluntary and

* Corresponding author.
E-mail addresses: daniels@ifce.edu.br (D.S. Ferreira), fsombra@ufc.br (F.N.S. Medeiros), dushizima@lbl.gov (D.M. Ushizima).

https://doi.org/10.1016/j.cmpb.2019.105053
0169-2607/

based on a task or an intention, being strongly influenced by the prior knowledge and experience of the observer [3,4].

Computational systems can model the visual attention as a saliency map, which is a topographical map representing the conspicuity of each pixel in an image. Researchers have proposed several algorithms and applied them to fields such as computer vision, robotics, and medical image analysis [5–7]. Bottom-up methods usually model the image low-level features, for example, color, contrast, orientation, and others, and may use different approaches, e.g., cognitive concepts, probabilistic frameworks, spectral analysis, etc., to generate saliency maps [8]. Typically, bottom-up models refer to biological processes. They are designed to reveal certain image regions that are different from the surroundings, without dealing with cognitive phenomena that make these regions relevant.

In contrast, top-down methods are drawn to react to the image semantics, such as task demands and expectations [9]. Early top-down attention models were driven by handcrafted features learned by training on human visual attention data sets [10]. Recently, the advances in deep learning [11] and the increasing availability of large annotated databases [12,13] have enabled saliency models to perform end-to-end learning and consequently achieve strong improvements [14].

The eDN (ensembles of Deep Networks) model [15] was the first effort to apply CNN for saliency prediction. It identifies saliency predictive instances of a richly-parameterized biology-inspired hierarchical model and then combines them using a linear SVM. Liu et al. [16] introduced a multiresolution CNN (Mr-CNN) to learn both the bottom-up and top-down factors simultaneously. In this model, three different CNNs were trained at different scales and, then, combined by two fully connected layers for final saliency prediction. SALICON (Saliency in Context) [17] explores the contextual information to reduce the semantic gap between the model prediction and the eye fixations. It concatenates two pre-trained CNNs, each on a different image scale (fine and coarse), to create its saliency map. A relevant contribution of Huang et al. [17] is to fine-tune CNNs with saliency metric as an objective function. DeepFix [18] innovated being the first algorithm to apply Fully Convolutional Neural Networks (FCNN) for saliency prediction. The authors also presented the novel Location Based Convolutional (LBC) to capture object-level semantics at multiple scales. Liu and Han [19] proposed the Deep Spatial Contextual Long-term Recurrent Convolutional Neural Network (DSCLRCN model). DSCLRCN learns saliency-related local features of image regions using CNN and incorporates global and scene contexts to reveal the final saliency map. SAM (Saliency Attentive Models) Nets [20] predicts saliency by combining a FCNN with an attentive recurrent mechanism. An exhaustive study about the deep saliency models can be found in Borji [14].

Most of the aforementioned predictions models were designed to estimate human eye fixations on natural images, which typically represent photographs of everyday scenes [8,10]. Moreover, several databases [12], comprehensive and up-to-date benchmarks [8,10] and public challenges [13,21] have focused on the investigation of visual attention models for natural images. Recently, there is an increasing interest in understanding the expert's patterned eye movement for medical image analysis [22–24], and how this ability can support the development of automatic diagnosis systems [25]. However, the number of studies that links visual attention models and cell analysis is limited. Particularly, algorithms that simulate the human selective attention can enable fast object detection from large Pap smear slides by ranking the most relevant clinical areas within the images [26]. It can represent a promising strategy to drive the focus of classification, image compression, and other routines. In fact, the identification of image parts that are relevant to a cytologist may solve a current challenge: to design real-time applications to analyze Pap smear images under real noisy conditions, artifacts, and cell occlusions [27].

Previous work proposed by Coombes and Culverhouse [28] employed visual attention theory to analyze cells. These authors used an eye tracking device for collecting the cytopathologist's visual data and identify manually marked salient features that are valuable for the development of quality assurance models on smear slide screening. Coombes and Culverhouse [28] concluded that the use of saliency maps for providing feedback to the cytopathologist may reduce the diagnostic divergence on regular screening. Another finding was the high correlation between the expert's eye fixations and cell staining in the image. Zhang et al. [26] proposed a method for abnormal cervical cell detection based on the bottom-up attention mechanism and the top-down information, such as size and color of abnormal nuclei. However, Zhang et al. [26] used only liquid-based cytological images [29], and concluded that visual attention mechanisms support finding diagnostic-relevant cells without massive processing of the whole image.

Different from previous approaches, such as in [28] and [26], our paper investigates the feasibility of using saliency prediction methods to support screening of cervical cells from real routine exams using the conventional Pap smears, a harder task, but highly necessary as most of the countries still rely on this exam modality. The main contribution of this work consists in applying CNNs and state-of-the-art algorithms to predict cytopathologist's eye fixation and extract ROIs from conventional Pap smear images. Particularly, we are interested in confirming two major hypotheses:

- **Hypothesis 1 (H1):** Bottom-up methods present comparably accurate results to the predictions obtained by expert's visual attention.
- **Hypothesis 2 (H2):** State-of-the-art saliency models can detect clinically relevant ROI from Pap smear images.

In order to address these hypotheses, this paper describes four contributions: (1) quantitative evaluation of state-of-the-art saliency models in comparison to the cytopathologists' visual attention. Furthermore, we report the bottom-up methods that can detect relevant areas for domain experts; (2) verification of Coombes and Culverhouse [28] findings for conventional Pap smear images that there is a high correlation between low-level features of the abnormal cells, such as brightness and color appearance, and the cytopathologist's attention on task-driven cervical cell image analysis. In addition, we show that the cytopathologist's gaze is strongly guided by top-down factors; (3) training of saliency prediction models using a CNN-based framework with two different neural networks (VGG-16 [30] and ResNet-50 [31]) based on Pap smear images using our human attention maps as ground truth; and (4) detecting which cell lesions are best identified by the saliency prediction algorithms if applied as a ROI extractor for diagnosis purposes.

In addition, this paper makes our saliency data set containing 5654 cervical cells from 232 images from real exams available at https://doi.org/10.17632/bk45c9yxb9.1. To the best of our knowledge, this work is a pioneer in publishing cytopathologist's visual attention data recorded by an eye tracking device, contributing both to selective visual attention and to cell analysis reproducibility.

The remainder of this paper is organized as follows. Section 2 presents our cervical cell image database, the proposed methodology to collecting the cytopathologists's attention data, the surveyed saliency methods, and our performance evaluation methodology. The experimental results are discussed in Section 3, and our conclusions and future directions are drawn in Section 4.

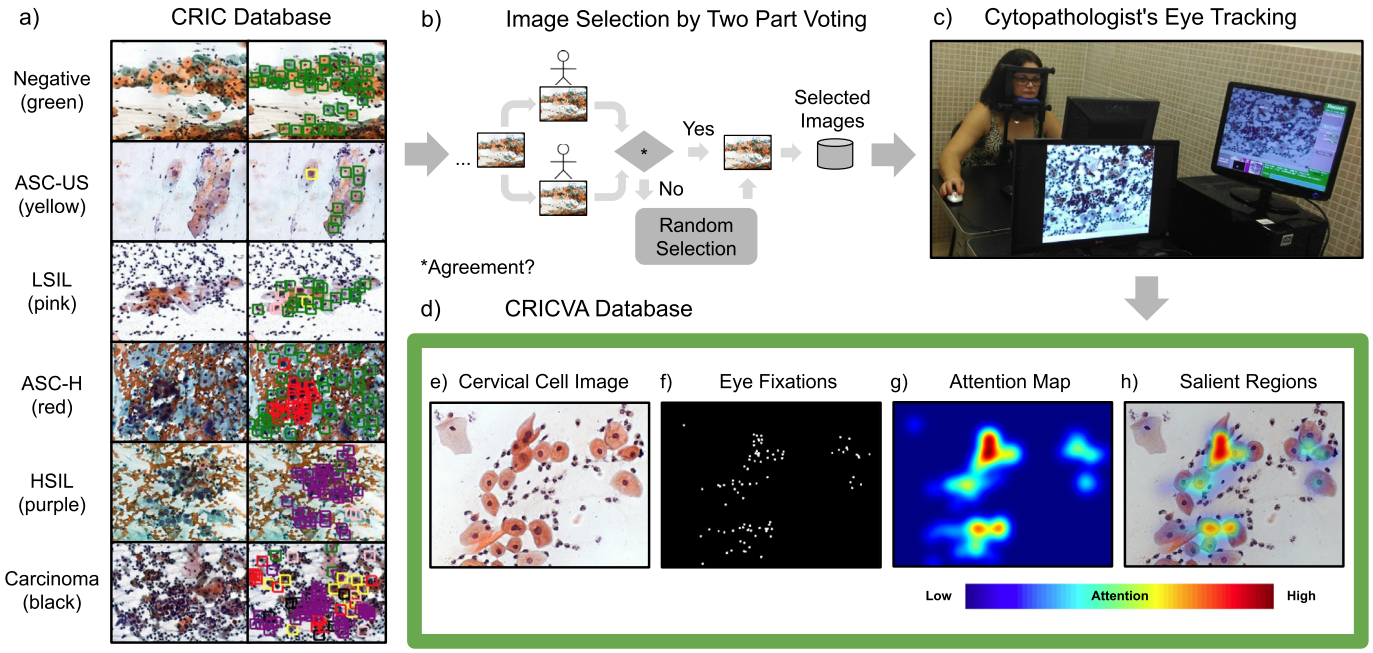


Fig. 1. Overview of the CRICVA database. (a) Images from the CRIC dataset with cell labels. (b) Our methodology to select CRICVA images in order to reduce the selection bias. (c) Collecting cytopathologist's eye movements by an eye tracking device. (d) Visual attention data contained in the proposed CRICVA dataset. (e) RGB cervical cell image. (f) Cytopathologist's eye fixations as binary map. (g) The attention map created by placing a Gaussian at each fixation position. (h) Overlap between the input image and the attention map to highlight the salient regions on the image. Dark red regions correspond to the most conspicuous regions whereas dark blue the others.

2. Materials and methods

2.1. CRIC database

The Center for Recognition and Inspection of Cells (CRIC) database contains digitized Pap smear images which were acquired with a Carl Zeiss microscope equipped with a Zeiss AxioCam MRc camera at 40× magnification. The acquired images have 0.255 µm/pixel and a resolution of 1392 × 1040 pixels (8-bit). The specimens were prepared via conventional Pap smears and contain cervical cells as well as other artifacts often collected as part of the exams. The cervical cells in the CRIC dataset are labeled into normal and, where abnormal uses the following classification: Atypical Squamous Cells of Undermined Significance (ASC-US); Atypical Squamous Cells of High Significance (ASC-H); Low-grade Squamous Intraepithelial Lesion (LSIL); High-grade Squamous Intraepithelial Lesion (HSIL); and Carcinoma (CA); patterns [32] which were manually classified by three cytopathologists. Fig. 1a displays image samples and illustrates the organization of the CRIC dataset.

We introduce a new dataset, named CRICVA[1] (CRIC Visual Attention) [33], using 232 images of the CRIC. We arranged this database according to the following requirements: (1) images with cells distributed uniformly on the whole visual area; (2) images with different artifacts, such as blood cells, inflammatory cells, distorted cells, overlapping objects, mucus, etc; (3) images selected by the two-party vote; and (4) 80 images randomly selected from the remaining CRIC images (Fig. 1b). The random selection aimed to reduce the selection bias[2] in the voting process. Fig. 1d displays the contents of the CRICVA dataset.

2.2. Collecting visual attention data

2.2.1. Subjects

We recorded the visual attention data from three cytopathologists, all of them with normal or corrected-to-normal vision via lens glasses. All of the subjects are experts in cervical care and reading conventional Pap smear slides. The mean age of the participants is 34.3 years (46, 30, and 27) and the length of their career in cytopathology laboratory is 20, 8 and 3 years, respectively. The two most experienced participants have Ph.D. degrees and the other has a Master degree in cytopathology. The screeners signed the informed written consent to participate in this project. The study has been approved by the Ethics Committee of Universidade Federal do Ceará under the protocol number 2.439.252. We conducted all activities in accordance with the ethical guidelines defined by the Declaration of Helsinki and Brazilian laws.

2.2.2. Eye tracking task procedure

We used an eye tracking device to collect attention data from cytopathologists during their analyses of cervical cell images. Here, we focused on eye fixation, which is one of the eye movement parameters obtained by the eye tracker. Fixation is a period wherein the eye remains still, reflecting the conspicuity of a particular area of an image [1]. We represented the locations of eye fixations as a binary matrix, as shown in Fig. 1f. From this matrix, we extracted attention heat maps (Fig. 1g) by convolving an isotropic 2D Gaussian of 1° (one degree) of visual angle centered on each fixation, where one degree of visual angle stands for an estimate of the size of the fovea [35,36]. We identified salient image regions for experts by overlapping the attention map with the input image as color-coded in Fig. 1h, where dark red indicates the most conspicuous areas, and dark blue otherwise.

We performed our experiments using an EyeLink 1000 system designed by SR Research Ltd., Mississauga, Canada, with a sampling rate of 1000 Hz on the right eye recording (Fig. 1c). We created

[1] Cervical cell images and eye fixation data are available at https://doi.org/10.17632/bk45c9yxb9.1.

[2] Selection bias is related to the preference for a particular kind of image during the composition of the database [34].

the tasks using SR Research Experiment Builder V4 and presented a sequence of eight trials for each participant on a Dell E178FPC at 60 Hz. Each trial was composed by the cervical cell images according to the following quantities: [Trial: Number of images]: 01: 26; 02: 26; 03–06: 25, and 07–08: 40. We exposed the cytopathologists to the same amount of images in each trial. Our motivation for adopting different image quantities across sessions was to increase the ability of our database to understand cytopathologist screening performance over time. Before each trial, the cytopathologists carried out a nine-point calibration procedure to map the eye-fixation to the screen coordinates. Furthermore, the participant had the opportunity to relax and report any discomfort with the experiment in the trial intervals.

We conducted a task-driven experiment in which the cytopathologist interpreted each cervical cell image and marked the abnormal cells with mouse clicks. The participants had free time to analyze the images, and the mouse clicks were not visible on the monitor. The cervical cell images were resized to 1280 $\times$ 1024 pixels, keeping the original aspect ratio by adding white pixel lines at the image bottom. Every time the cytopathologist pushed the space bar of the keyboard connected to the eye tracker, a new cell image appeared on the monitor. The images came into view in a randomized order to reduce any potential bias in the presentation of a sequence of images, and the participants could not return to the previous image. At the end of each trial, the participant data were recorded and a new round started.

2.2.3. Consistency across the participants

Inspired by Volokitin et al. [37] and Bylinskii et al. [38], we measured the gaze agreement among the cytopathologists for each image. We created a fixation map for each cytopathologist and then used it to predict fixations of the excluded ones. The Area Under the Receiver Operating Characteristic Curve (AUC-Judd) [9] metric was used for performance evaluation of our approach. The average AUC-Judd between all cytopathologists was 0.823 (SD=$\pm$0.007) and the p-value was 0.117 (Kruskal-Wallis ANOVA test with $\alpha = 0.05$) [39], considering the whole image dataset. We also observed the intra-participant performance across sessions and we found the average values for AUC-Judd equal to 0.818 (SD=$\pm$0.016), 0.813 (SD=$\pm$0.018), and 0.829 (SD=$\pm$0.016) for each cytopathologist, respectively. These results suggest a high consistency among the participants' eye fixation patterns on the CRICVA database, evidencing the existence of well-defined clinical ROIs on Pap smear images and supporting the investigation of our H2 hypothesis.

2.2.4. Center bias analysis

The human tendency to look for objects near the central image region is a frequent bias in computer vision databases [40,41]. Moreover, these datasets tend to be subjected to the photographers bias in framing relevant objects in the central region of the image [10]. To understand how the center bias occurs in the proposed dataset, we analyzed the Average Annotation Map (AAM) of the CRICVA database, as shown in Fig. 2a. The AAM conveys the average of the visual attention ground-truth annotations of the whole image data set [10]. To analyze the dispersion of the AAM, we contrasted the horizontal and vertical midlines of the AAM with those obtained from a Gaussian blob at the center of the image, as illustrated in Figs. 2b and c, respectively. The Gaussian kernel was set at 1° of visual angle to reflect estimates of fovea size [40]. Although the AAM has a larger activation near the image center, there exists a significant variation in the spatial distribution of the activations, indicating that the cytopathologists spent a considerable amount of time analyzing objects far from the image center.

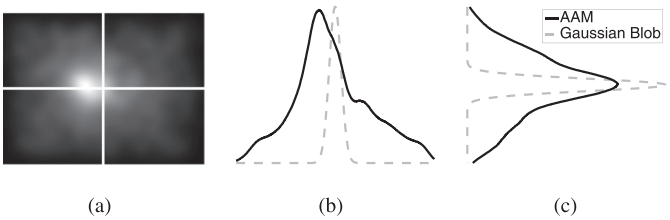


Fig. 2. Center bias analysis of the CRICVA dataset. (a) Average Annotation Map (AAM). (b) Horizontal and (c) vertical mid-lines of AAM and Gaussian blob at the image center.

2.3. Saliency methods

We investigated 10 bottom-up methods and two variants of a CNN-based model for top-down saliency prediction, covering different categories of algorithms [8]. We also considered the model proposed by Zhang et al. [26] since it is designed to explore top-down factors in cervical cell images. We summarized the studied methods and introduced the model abbreviations adopted in the rest of the paper in Table 1. All algorithms, except the one proposed by Zhang et al. [26], were validated on public data sets and chosen for this work according to the following criteria: (1) the input is a single image; (2) the source code is publicly available; (3) they present high performance in well-established saliency ranking list; (4) the runtime is less than three seconds per image, and (5) they are state-of-the-art algorithms or stand for a benchmark in the literature. Based on these requirements, we visited the MIT Saliency Benchmark [21], sorted the ranking results by NSS (Normalized Scanpath Saliency) and picked out the top-ranking CNN-based algorithm: SAMv and SAMr [20], in which the feature maps are extracted by the VGG-16 and ResNet-50, respectively. We also selected five non-deep methods: BMS [42], LDS [43], FES [44], SWD [45], and UHM [46]. We complemented our study with the IT [47], GB [48], SR [49], SS [50], and SIM [51] bottom-up models from the extensive benchmark introduced by Borji et al. [10].

2.4. SAM model setup

We employed two different CNN architectures as the backbone for the SAM model. As proposed by Cornia et al. [20], we combined a dilated version of the VGG-16 network into the SAM pipeline and trained it on the eye-fixation data from natural images using the SALICON [13] dataset (SVS: SAM VGG-16 on SALICON) and cervical cell images (SVC: SAM VGG-16 on CRICVA). We also used the SAM approach based on the dilated ResNet-50 network, obtaining two other versions: SRS (SAM ResNet-50 on SALICON) and SRC (SAM ResNet-50 on CRICVA).

The SALICON database comprises 10,000 training images, 5000 validation images, and 5000 testing images divided into 80 categories, being the largest available data set for saliency prediction in natural images. Regarding cell images, we carried out the SAM model training and validation procedures on random sets of the CRICVA database, containing 130 and 29 images, respectively. We drew the CRICVA testing database from the 73 remaining images. Table 2 outlines the key information about the SAM model usage.

We adopted the default configuration of the SAM model as described by Cornia et al. [20]. The initial weights of dilated CNN were defined with those of the VGG-16 and ResNet-50 models trained on ImageNet [54] and the recurrent weights matrices of the Attentive ConvLSTM were initialized as random orthogonal matrices. We resized the images to 240 $\times$ 320 and employed the same overall loss function proposed by Cornia et al. [20].

Table 1
Surveyed saliency prediction models. {T: Top-down model, B: Bottom-up model}.

#	Model Abbreviation		Description	Year	Cat.	Ben.
	Literature	Here				
1	SAMv (Saliency Attentive Model - VGG-16) [20]	SVS, SVC	This algorithm uses the VGG-16 CNN as backbone and incorporates an Attentive Convolutional Long Short-Term Memory network (Attentive ConvLSTM) to predict eye fixations in the images, without handling a temporal sequence. This algorithm has a center prior component able to learn the center bias of the database.	2018	T	
2	SAMr (Saliency Attentive Model - ResNet-50) [20]	SRS, SRC	The same approach of the SAMv model, but using the ResNet-50 network as backbone.	2018	T	
3	BMS (Boolean Map based Saliency) [42]	BMS	This strategy characterizes an image by a set of binary images, which are created by randomly thresholding the image features maps in a whitened feature space. Given an input image, this algorithm uses the topological analysis of the Boolean maps to discover surrounding regions and estimate the saliency map.	2013	B	
4	LDS (Learning Discriminative Subspaces) [43]	LDS	This saliency map estimation is based on the learning of a set of discriminative subspaces. These subspaces have to perform the best in popping out targets and suppressing artifacts. LDS creates the candidate subspaces based on the principal component analysis.	2017	B	MIT
5	FES (Fast and Efficient Saliency) [44]	FES	This algorithm uses a center-surround approach to estimate saliency of local feature contrast in a Bayesian framework. It estimates the needed probability distributions using the sparse sampling and the kernel density estimation.	2011	B	
6	SWD (Spatially Weighted Dissimilarity Saliency) [45]	SWD	This algorithm is based on the integration of dissimilarities and spatial distance between image patches and the center bias. The spatial distance weighs the corresponding dissimilarities and the principal component analysis is adopted for dimension reduction. The center bias is addressed by a weighting mechanism.	2011	B	
7	UHM (Unsupervised Hierarchical Models) [46]	UHM	This unsupervised multi-scale hierarchical saliency model explores both local and global saliency concepts. This approach adopts independent subspace analysis (ISA), which is equivalent to a two-layer neural architecture. The algorithm obtains a hierarchical representation of the input, stacking the ISA networks together, as done in deep models.	2016	B	
8	IT (Itti's Saliency Model) [47] - implementation by Harel et al. [48]	IT	This work is the pioneer of saliency prediction and it is considered the purely bottom-up model. It extracts low-level features using the local center-surround differences of intensity, color and orientation features at multiple spatial scales. Then, fusion of across-scale and normalization of these maps produces three conspicuity maps, which are combined to yield the saliency map.	1998	B	
9	GB (Graph-Based Visual Saliency) [48]	GB	This model extracts the low-level features similar to IT. Then, it uses a Markov chain to construct a fully connected graph which joins all grid locations (nodes) for each feature map. The weight between two nodes is defined as the dissimilarity of the feature values and their spatial distance. The saliency map is estimated based on the equilibrium distribution.	2016	B	
10	SR (Spectral Residual Approach) [49]	SR	This approach is independent of features, categories, or any prior knowledge about the objects. It conducts the saliency estimation by exploring the properties of the backgrounds. It evaluates the log-spectrum of an input image and extracts the spectral residual. Then, the spectral residual is transformed into the spatial domain to obtain the saliency map.	2007	B	[10]
11	SS (Sparse Salient Regions) [50]	SS	The authors used the sign function of the Discrete Cosine Transform (DCT) of an image to generate a signature, containing mainly information about the image foreground. The algorithm explores this information to detect regions and generate saliency maps.	2012	B	
12	SIM (Saliency by Induction Mechanisms) [51]	SIC	This methodology consists in processing the visual stimuli according to the early human visual pathway (e.g. color-opponent, luminance channels and multi-scale decomposition). Afterward, the algorithm simulates the inhibition mechanisms of the visual cortex cells and integrates information at multiples scales by an inverse wavelet transform. It is based on the unified color induction model developed by Otazu et al. [52]	2012	B	
13	ZH (Zhang et al. [26] Detection of Abnormal Nuclei in Cervical Smear Images) [26]	ZH	This work explores the bottom-up attention mechanism and a target-driven strategy for abnormal cell detection in liquid-based cervical smear images. This model consists in extracting conspicuous image regions according to both direction and brightness features and then modulating this information by a high response area obtained by an annular template matching model. The authors designed the annular template from abnormal nuclei statistics. Here, we resorted to [53] for setting the annular template parameters.	2013	T	*

* specific model on cell domain.

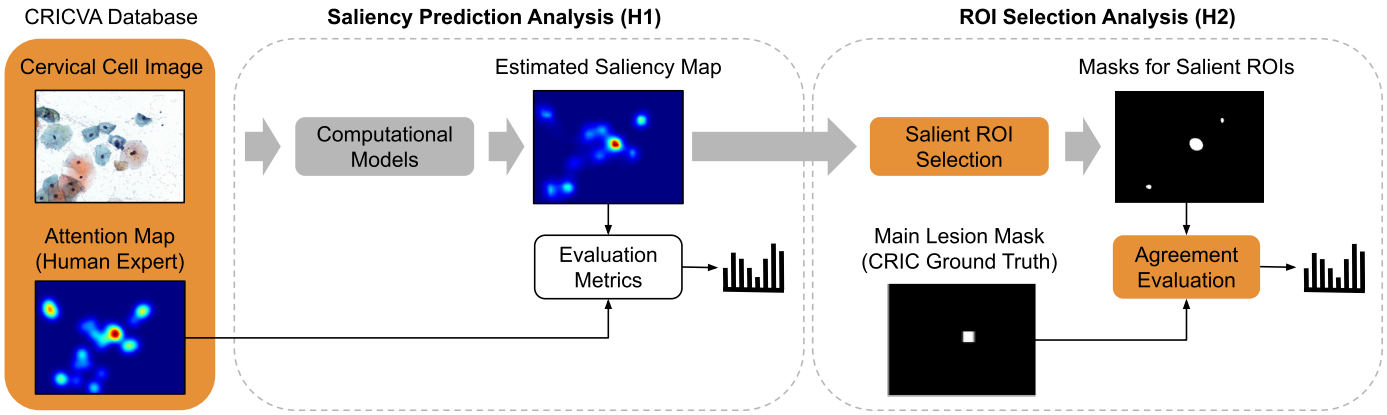


Fig. 3. Our investigation consists of two steps: (1) evaluation of the surveyed saliency models according to state-of-the-art metrics (H1 test), and (2) ROI selection analysis to determine whether the most relevant lesion in the image is highlighted on the estimated saliency map (H2 test). The orange boxes represent additional contributions of this work.

Table 2
Variants of the SAM model used in this work.

Model	CNN	Database	Sets
SVS	VGG-16	SALICON	#{Train}: 10K, #{Validation}: 5K, #{Test}: 5K
SRS	ResNet-50		
SVC	VGG-16	CRICVA	#{Train}: 130,#{Validation}: 29, #{Test}: 73
SRC	ResNet-50		

2.5. Saliency prediction analysis

Fig. 3 presents our methodology to investigate H1 within the saliency prediction analysis module. We focused our experiments on evaluating how well the surveyed saliency models predict where cytopathologists look at when analyzing cervical cell images. To this end, we processed the color (RGB) cervical cell image by the surveyed saliency models and performed evaluation of estimated maps against the visual data collected by the eye tracking device. We applied the following state-of-the-art metrics to assess the results and measure the influence of top-down factors on the cytopathologist's analysis.

2.5.1. Evaluation metrics

Bylinskii et al. [38] categorized the metrics for saliency model evaluation in location-based and distribution-based according to the data representation. The location-based metrics use the discrete fixation locations (Fig. 1f) as saliency maps. The distribution-based metrics consider both attention maps and predicted saliency maps as continuous distributions. According to Bylinskii et al. [38], the distribution-based metrics allow incorporating uncertainty in the measurements, such as the errors in eye tracking and imprecision of human eye position on the screen. Additionally, the distribution-based metrics are more robust to few observers than location-based ones since they extrapolate the data to model the behavior of more observers.

We employed five well-established measures[3] for performance evaluation of the surveyed models and we appraised our results in terms of false positives and false negatives. The location-based metrics that we adopted are the AUC-Judd [9] and the Normalized Scanpath Saliency (NSS). The AUC-Judd metric gives a high score for high-valued predictions placed at fixed locations, but it ignores low-valued false positives. The NSS metric is equally affected by false positives and negatives.

For the distribution-based category, we choose the Linear Correlation Coefficient (CC), the Similarity between Distributions (SIM) and the Kullback-Leibler divergence (KL) metrics. CC is symmetric and penalizes false positives and false negatives equally. SIM computes the intersection between two distributions, being more sensitive to false negatives than false positives. KL corresponds to an asymmetric dissimilarity metric highly sensitive to false negatives. Lower values of KL indicate better results. A broader study about evaluation metrics designed for saliency models is available in [38].

2.6. ROI selection analysis

The second module of the proposed methodology investigates the H2 hypothesis. We measured the probability of the most relevant cell lesions to be signed as a salient object by the surveyed models. Inspired by Bylinskii et al. [55], we mark as salient objects those whose positions are presented on the predicted saliency map as highlighted regions. We introduce the Algorithm 1 to draw the $nr = 3$ separated regions with the highest values for each estimated saliency map. If this map has $nr < 3$ salient ROIs, our algorithm thresholds it at the 95th percentile.

We define the agreement rate $\varphi \in [0, 1]$ between the salient ROIs and the location of the most significant cell lesion (le) in the image as:

$$\varphi = \frac{1}{N_{le}} \sum_{i=1}^{N_{le}} \#\{\forall r \in Z_i \mid (r \cap G_i) \neq \emptyset\}, \quad (1)$$

where $\#\{\cdot\}$ stands for the cardinality, N_{le} is the number of testing images with le lesion, Z is the binary mask for high-density saliency regions, r is the salient ROI and G represents the ground-truth binary mask signaling the 100 $\times$ 100 pixels of le area. Then, we considered *identified* if any part of the main cell lesion matched the ROI. Otherwise, we assumed that the model could not identify the location of the main lesion, appropriately.

3. Experimental results

In this section, we describe the analyses and experiments to test each hypothesis and validate the contributions of different algorithms. We also report the bottom-up models with the best performance on cervical cell images in terms of false positive and false negative. The parameters used for each algorithm are those published in the original papers listed in Table 1.

[3] The eye fixation evaluation metrics are available at https://github.com/cvzoya/saliency/tree/master/code_forMetrics.

Algorithm 1: ROI selection algorithm for finding the highest density regions on the estimated saliency maps.

```
function ROIselect (M, nr);
  Input  : (float matrix ∈ [0,1]) M: Estimated saliency map;
           (int) nr: number of required regions
  Output: (logical matrix) Z: Binary mask signaling the ROI
           locations
th = 0.95;
step = 0.05;
countr = 0;
while (countr < nr) and (th > 0) do
    Z =logical(M. ∗ (M >= th));
    countr =getNumberOfSeparatedRegions(Z);
    th = th − step;
end
if (th <= 0) then
    Z =logical(M >= 0.95);
end
        ▷ Ranking the regions in descending order based on the
  energy of the respective pixels and returning up to nr most
  salient separated regions.
Z = getNRMostSalientRegions(Z, M, nr);
return Z;
```

3.1. Saliency prediction analysis

We investigated H1 on 73 images (CRICVA testing set). We first conducted the analysis of classic (non-deep) bottom-up methods. Afterward, we analyzed the performance of the top-down and CNN-based methods. We also performed experiments with a center prior baseline model, which consists of a Gaussian blob (kernel width at 3 degrees) at the center of the image. We employed this baseline to reveal the existence of capture bias[4] in the CRIC database and to analyze the influence of center bias on the performance of the surveyed methods.

3.1.1. Bottom-up methods

Fig. 4 shows the estimated saliency maps from all algorithms for an input cervical cell image. The quantitative comparison of these maps is presented in Fig. 5.

Overall, the bottom-up algorithms tended to highlight false positive attention regions on cervical cell images. Since these methods were mainly designed to simulate the human low-level visual attention, features such as the contrast between the cells and background, and the staining of image structures, had a significant impact on saliency prediction.

The SIC method was strongly affected by the brightness and color appearance, as a result of the color perception methodology developed by Otazu et al. [52]. Thus, the SIC algorithm tended to highlight image regions of high contrast in relation to its surroundings, such as cell-background transition, artifact presence, and background noise. Therefore this model is seldom adequate for saliency prediction on cervical cell image applications that require low scores of false positive. In special circumstances, the SIC approach can be useful to detect image areas with high contrast objects, such as normal nuclei and neutrophils.

An algorithm that produced saliency maps visually close to the cytopathologist's attention maps was FES, which revealed that the use of local features in a Bayesian framework enables detection of abnormal cells. The FES method's nature is implicitly biased by the approach used for estimating the prior probability distributions. The original FES implementation uses the AAM over a training set for approximating the priors required by the Bayesian approach. We processed the CRICVA AAM (Fig. 2a) and excluded test data in accordance with Tavakoli et al. [44] to estimate the FES prior distributions and run our experiments. Although the FES algorithm emphasized salient objects close to the image center due to the AAM information, this algorithm found abnormal cell regions far from the central region. The BMS, LDS, and GB techniques also presented high saliency values at eye fixation locations, but presented higher sensitivity to false positives than the FES algorithm.

We also performed the Kruskal–Wallis statistical test [39] with post-hoc Nemenyi test ($\alpha = 0.05$) [56] to find the model results that differ significantly from each other. Fig. 6 reports all pairwise comparisons for each studied evaluation metric.

In terms of the AUC-Judd measure, the FES method outperformed all the other bottom-up models. However, it did not differ statistically from BMS, LDS, SWD, UHM, and IT as well as GB. Based on the AUC-Judd features discussed by Bylinskii et al. [38], we noticed that these methods may be suitable for cervical cell image analysis slightly affected by false positives. Afterwards, we extended our investigation by considering equally the effects of false positives and false negatives. The NSS measure pointed out LDS, FES, and GB as the best-performing classic method, and the CC and SIM metrics confirmed this finding. Therefore, our tests showed that the LDS, FES, and GB algorithms are feasible for saliency prediction on cervical cell images that require accuracy in identifying visual attention regions. We found a few exceptions, which corresponds to the cases where there are large number of image artifacts (Fig. 4f). For these cases, the use of classic bottom-up methods may be ineffective. Although the FES algorithm scored well the location-task of the salient regions in most cases, the size of its blobs was usually smaller to those of the ground truth. This fact undermined the performance evaluation of the FES method by the KL metric, which is highly sensitive to pixel false negatives. These are some indications that the FES additional settings may boost its performance, but further investigations might be necessary to confirm such biases.

The results of the center prior baseline confirmed that part of the expert's visual attention was devoted to targets in the center of the image. However, the best bottom-up methods, according to the evaluation metrics, reached salient regions far from the image center, showing that they captured satisfactorily the low-level features of the targets. Based on this analysis, we validated the H1 hypothesis by arguing that the cytopathologist's visual attention is highly correlated with the low-level features of the abnormal cells. In addition, the appropriate modeling of the cell attributes may allow the use of fast bottom-up algorithms as part of a saliency prediction framework for cell image analysis.

3.1.2. Top-down models

Fig. 4 shows that the target-driven approach implemented by the ZH model represents a promising strategy to localize conspicuous areas for cytopathologists on cervical cell images. However, the saliency maps may be sparse for cytopathologist's eye fixation prediction. Since the ZH model is driven by an abnormal cell template matching, some salient areas can be neglected by ineffective correspondence between the template and the shape of the cells (Fig. 4b). In conventional Pap smears, some abnormal nuclei can present distorted shapes due to the clinical acquisition process, high overlap and intensity variations. These factors reduce the performance of the annular template adopted by the ZH model, leading it to produce saliency maps confined to specific image regions. These observations are supported by our quantitative analysis presented in Fig. 5. AUC-Judd revealed that the ZH model underweighted relevant attention areas (false negative), although

[4] Capture bias conveys some tendency of the photographers to position the targets on the picture during the scene imaging [34].

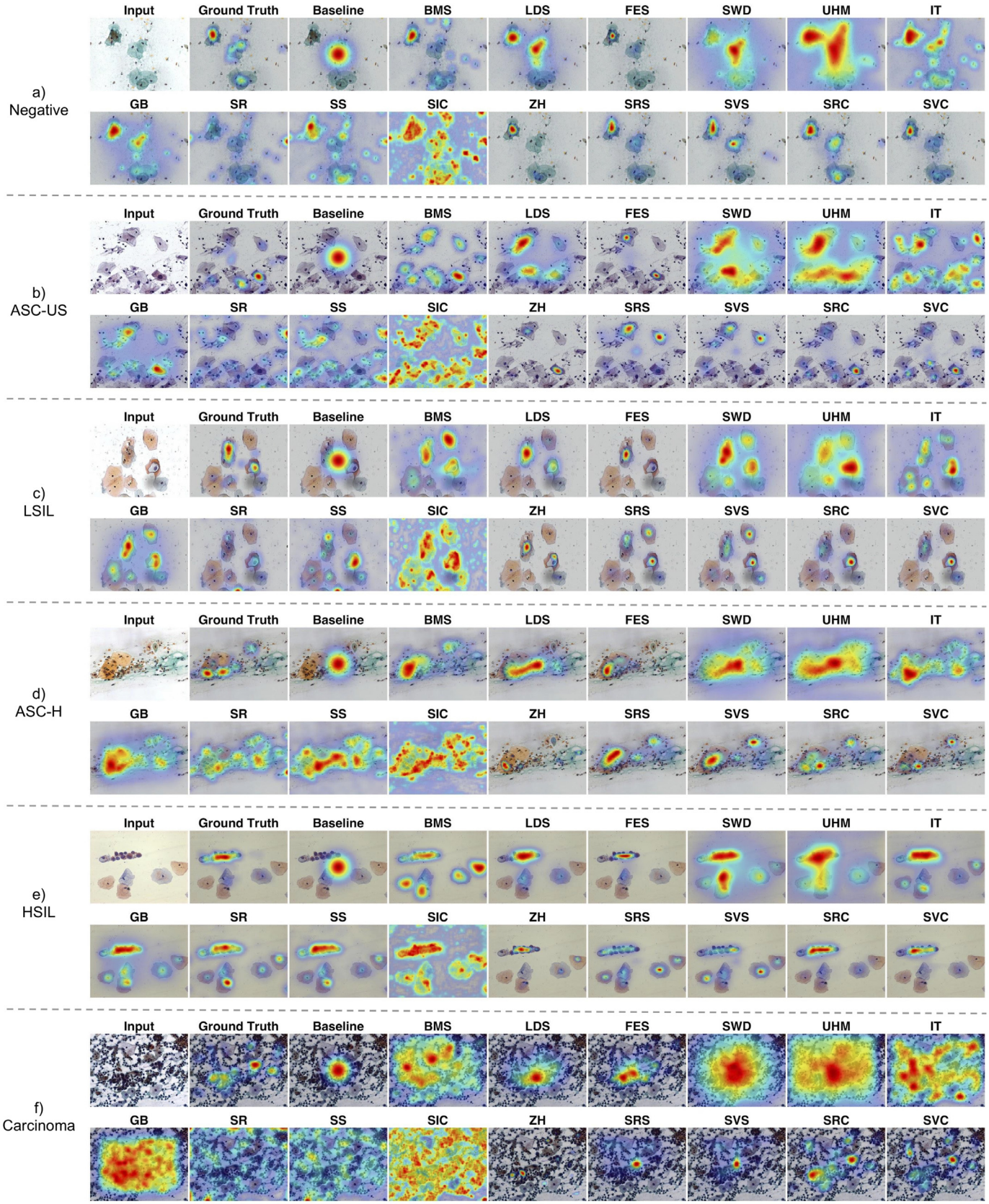


Fig. 4. Computed saliency maps by the surveyed methods overlapped upon the image. (a) to (f) display our results for each case of the CRICVA dataset. Some models are able to indicate conspicuous image regions on cervical cell analysis, whereas others may be ineffective for prediction of cytopathologist's gaze when on noisy images and clustered cells.

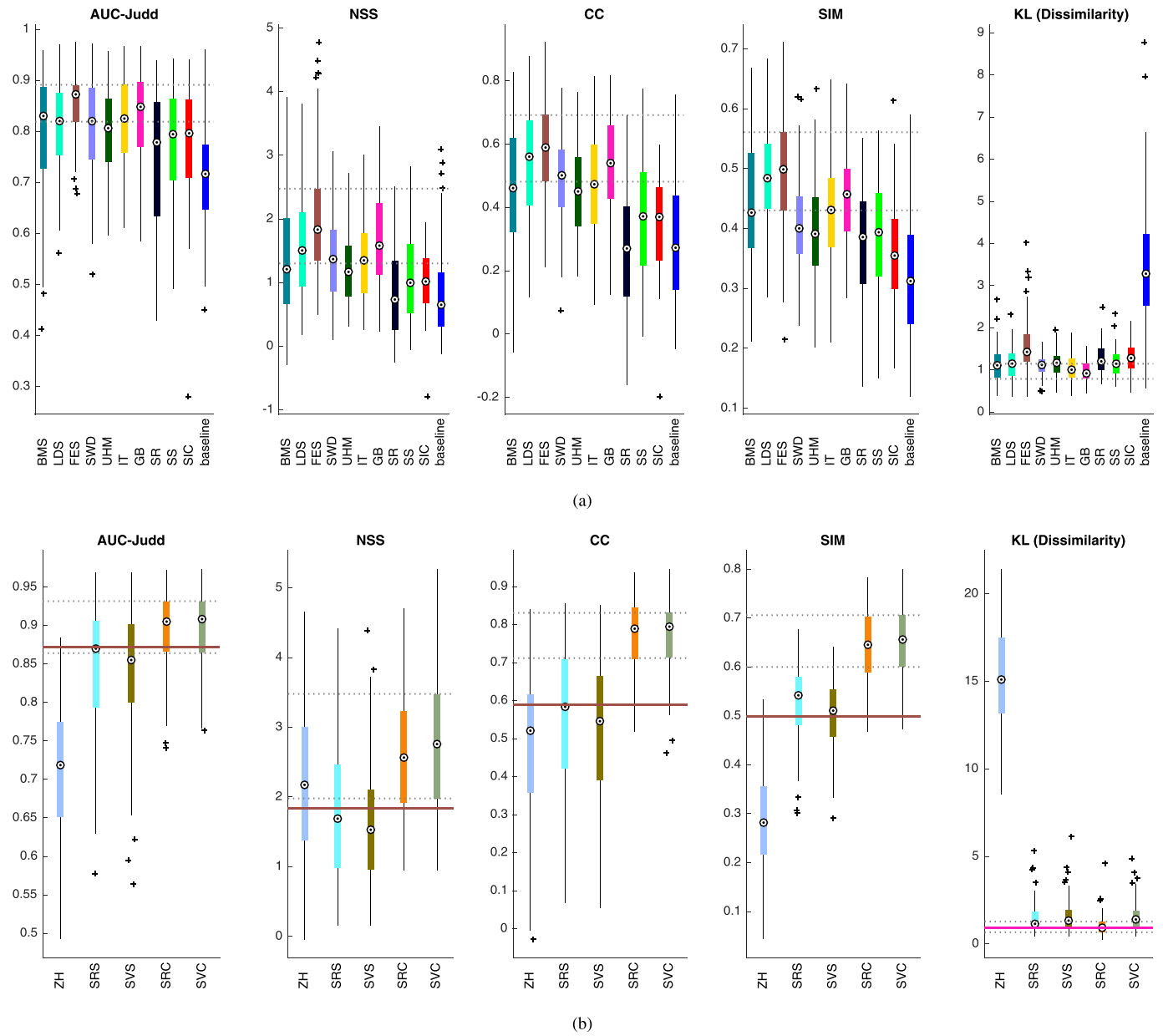


Fig. 5. Quantitative evaluation for (a) bottom-up and (b) top-down and CNN-based methods. The dotted horizontal line stands for the interquartile range of the best model. For each metric, the colored horizontal solid line marks the performance of the best bottom-up algorithm. The axis ranges are defined according to the evaluation metric limits [38].

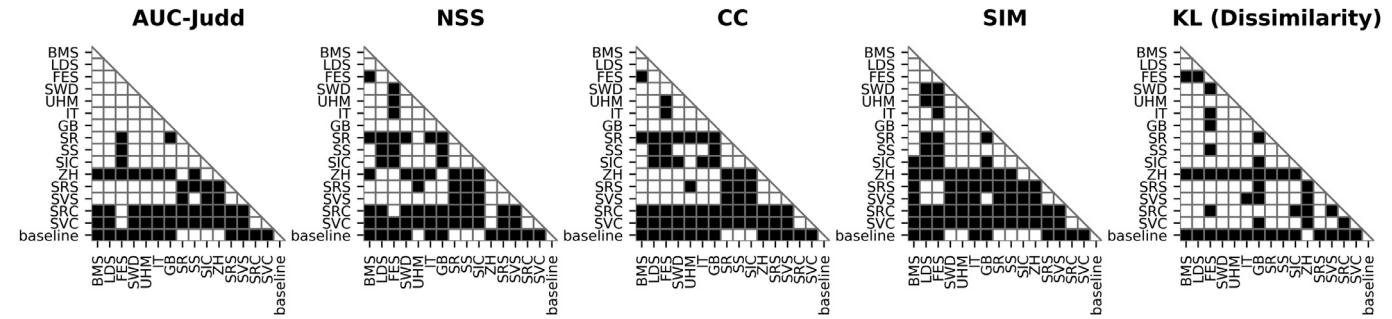


Fig. 6. Pairwise comparisons for all surveyed methods using the Kruskal-Wallis statistical test with the post-hoc Nemenyi test. The black boxes represent the pairs with significant difference at $\alpha = 0.05$.

Table 3
Ranking of the average execution time (seconds per image). CNN-based models are in bold. {Mat: Matlab, Py: Python}.

Method	SR	SS	FES	IT	ZH	GB	BMS	LDS	UHM	SWD	SVC	SVS	SRC	SRS	SIC
Code	Mat	Mat	Mat	Mat	Mat	Mat	Py	Mat	Mat	Mat	**Py**	**Py**	**Py**	**Py**	Mat
Time	0.01	0.02	0.22	0.23	0.50	0.51	0.63	0.96	1.36	2.38	**3.23**	**3.27**	**3.85**	**3.86**	10.03

NSS confirmed assertiveness for some regions. Furthermore, the distribution-based metrics (CC, SIM, KL) suggested a significant distance between the attention areas demanded by the cytopathologists and those estimated by the ZH model.

Our experiments showed that SVC and SRC indicated improved saliency maps in comparison with all other surveyed models. Since SVS and SRS performed similarly to the best bottom-up methods (Fig. 6), we argue that SVC and SRC learned top-down features that are relevant to driving the cytopathologist's visual attention. Particularly, we demonstrated that the transfer learning from a CNN trained on a large-scale data set from a different domain is also suitable for saliency prediction on cervical cell images, mainly for the low-level feature modeling on earlier layers. Although our experiments have revealed differences between the top-down factors on natural and cervical cell image analysis, we found evidence that the CNN-based frameworks, trained on a small database with cytopathologist's attention maps, can predict valuable saliency maps on Pap smear images through transfer learning. An interpretation of the CNN results here was that there are features relevant to both domains, possibly through low-level vision primitives. Furthermore, our results pointed out that the top-down factors guide the cytopathologist's attention during the diagnosis task, mainly reducing the sensitivity to brightness and high contrast areas.

By exploring CNN architectures for both natural and cervical cell images, we analyzed the correlation between the activations of the CNN layers when trained in the SALICON and CRICVA databases and we presented our results in Fig. 7. We observed that approximately the first half of the layers have correlation coefficient above 95% in their activations, which confirmed that the CNN lower layers represent low-level features and build upwards toward higher-level representation on upper layers. For CNN training on cell images, our analysis may be used to adjust some fine-tuning parameters, such as the number of lower layers to be frozen and the learning rate. Additionally, our analysis showed that the activations of the lower layers of both CNN architectures were similar for visual attention regardless of the database purpose.

The Kruskal-Wallis and post-hoc Nemenyi statistical tests did not identify statistical differences between SVC and SRC for all evaluation metrics, except KL (Fig. 6). This led us to conclude that both variants may be equally chosen for cervical cell applications according to the availability of the computational resources.

3.1.3. Runtime performance

The average time of the surveyed methods on the CRICVA database was ranked in Table 3, based on algorithms whose source code were publicly available and considering the experimental configurations suggested in their respective scientific articles. The bottom-up approaches, apart from the SIC method, performed faster than the deep learning algorithms. The results in Fig. 5 show that the bottom-up methods may be a viable solution for cervical cell image analysis and applications in which computational resources are limited – this motivated us to report results using a computer with an Intel (I7-4770HQ) CPU (2.2 GHz) and 16 GB RAM.

We summarized the contributions of the H1 hypothesis by reporting fast bottom-up techniques that can identify the most conspicuous regions within a cell image for analysis by cytopathologists. In addition, we confirmed a finding described by Coombes and Culverhouse [28] that there is a high correlation between some low-level features of the abnormal cells and the expert's eye-fixation pattern. We also trained a state-of-the-art CNN-based framework using VGG-16 and ResNet-50 for saliency prediction using our expert's attention maps as ground-truth, achieving accurate saliency maps on conventional Pap smear images.

Table 4
Agreement rate φ (Eq. 1) between the $nr = 3$ salient ROIs and the most significant lesion in the image. The best results for each pathology are in bold. { M: Models, L: Lesions}.

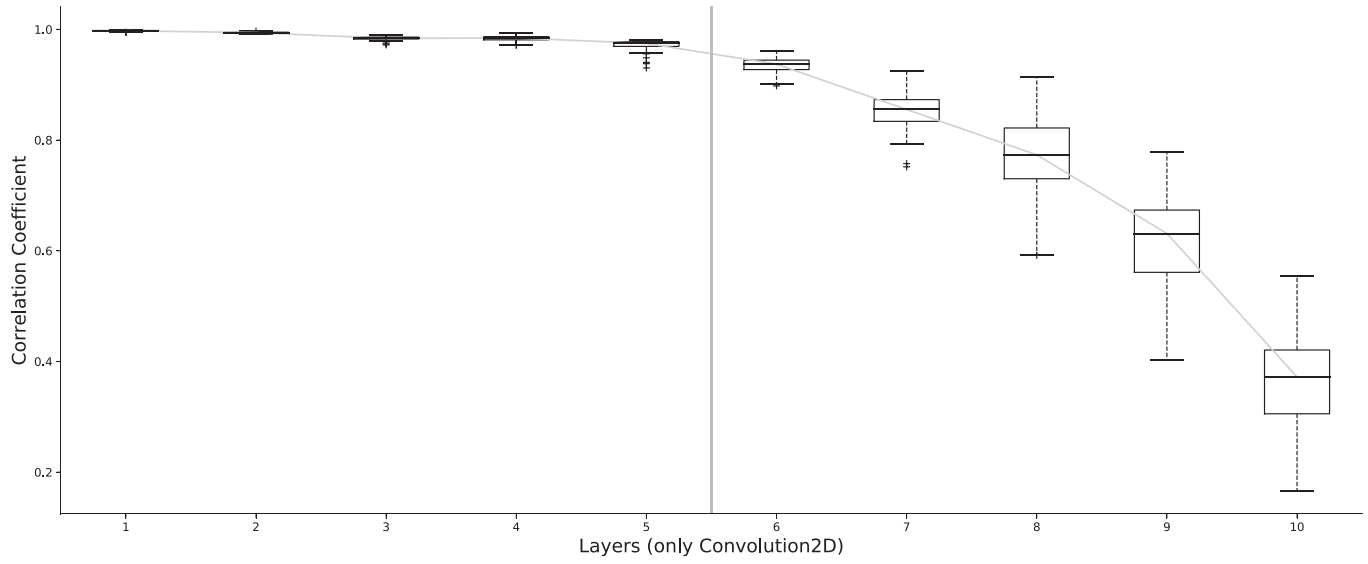
M/L	ASC-US	LSIL	ASC-H	HSIL	Carcinoma
BMS	0.731	0.820	0.808	0.857	0.500
LDS	0.692	0.860	0.885	0.952	0.625
FES	0.865	0.860	0.962	**1.000**	0.750
SWD	0.808	0.900	0.923	**1.000**	0.625
UHM	0.712	0.810	0.846	**1.000**	0.625
IT	0.462	0.680	0.923	**1.000**	0.375
GB	0.730	0.870	0.923	0.904	0.500
SR	0.173	0.280	0.384	0.571	0.000
SS	0.385	0.450	0.731	0.762	0.250
SIC	0.153	0.390	0.730	0.571	0.375
ZH	0.846	0.930	0.884	0.761	0.375
SRS	0.865	0.860	0.846	0.952	0.625
SVS	0.731	0.840	0.885	0.952	0.500
SRC	**1.000**	0.990	**1.000**	**1.000**	**1.000**
SVC	0.981	**1.000**	**1.000**	**1.000**	0.875
Experts*	1.000	0.989	1.000	1.000	0.786
Baseline	0.365	0.570	0.653	0.904	0.250

* computed from all images of CRICVA. The experts reached $\varphi = 1$ for LSIL and Carcinoma with $nr = 4$ and $nr = 5$ salient ROIs, respectively.

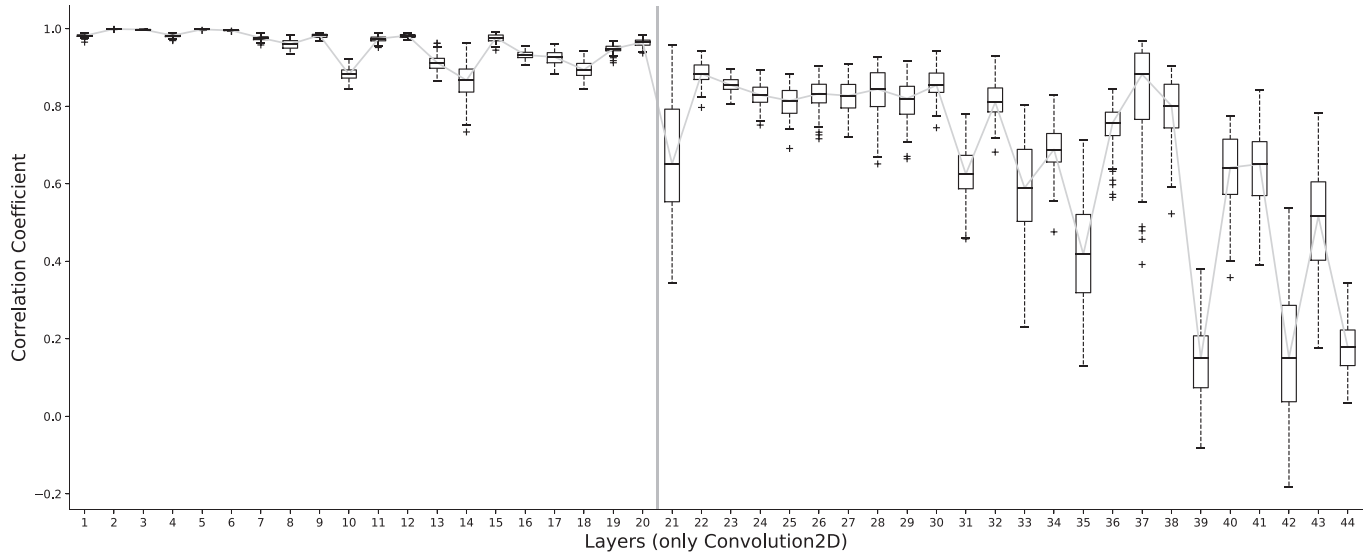
3.2. ROI selection analysis

The investigation of the H2 hypothesis required no eye fixations data; instead, it considers only the lesion annotations available on the CRIC database. Thus, we extended our experiment to the remaining abnormal images on this database, resulting in a total of 207 images. The distribution of image labels according to the main cell lesion is #{ASC-US: 52, LSIL: 100, ASC-H: 26, HSIL: 21 and Carcinoma: 8}.

Table 4 presents the φ results for all surveyed methods, considering three separated regions. The SRC model identified the location of the most relevant cell lesion for all considered pathologies, except for one LSIL case where four regions were required. This demonstrates the potentiality of this algorithm as a ROI extractor for automated cell pre-screening systems for cervical cells. In fact, it localized the lesion region within a few candidate ones. SVC also showed valuable results for ROI-based systems, however it was less robust in detecting regions with ASC-US and carcinoma lesions than the SRC method. The ZH model achieved competitive results. As the authors of ZH only reported the results for liquid-based cell images [26], our findings extend the application of this method to conventional Pap smears. Some bottom-up strategies, such as those adopted by FES, and GB, may be suitable for ROI selection in cervical cell images, especially in images with sparse cell clumps and low presence of dark artifacts. Although SWD and UHM performed well in our experiments, Fig. 4 suggests the existence of large false positives, which may restrict the use of these methods for some applications. These results, in addition to those described in Fig. 5, point out that the FES method is our best bottom-up strategy for ROI selection on cervical cell images from accurate saliency maps.



(a)



(b)

Fig. 7. Correlation coefficient between the activations of (a) VGG-16 and (b) ResNet-50 layers trained on the SALICON and CRICVA databases. Layers up to 95% of correlation are shown on the left of the vertical line.

Nevertheless, we argue that the φ analysis is robust only for the ASC-US, LSIL, and ASC-H lesions. For these cases, our center prior baseline analysis revealed the absence of significant capture bias, confirming the performance of the surveyed methods. For HSIL and carcinoma, further works should consider: (1) a larger and more diverse CRIC database of cell samples, especially cases for carcinoma and (2) avoidance of the center prior baseline, which scored high (above 90%) for the current HSIL images, suggesting that HSIL lesions were mostly located on the image center, which may have influenced the performance of the methods.

According to H2, our experiments confirmed the reliability of several saliency prediction models in identifying critical cells for the diagnosis. Other contributions included the proposed algorithm for salient ROI selection (Algorithm 1) and the agreement rate (Eq. (1)) that quantified the performance of the surveyed models applied to ROI extraction. Finally, we emphasized the importance of organizing visual attention databases for specific domains, such as medical and cell imaging to benchmark algorithmic advances.

4. Conclusion and future directions

This paper evaluated the performance of state-of-the-art saliency models applied to conventional Pap smear image analysis. We investigated 10 bottom-up algorithms, one target-driven model that highlights conspicuous abnormal cell regions, and two variants of a CNN-based framework trained on natural images and cervical cell microscopy, using VGG-16 and ResNet-50 networks as backbone. Our results revealed that top-down factors could guide the cytopathologist's attention on task-driven analysis. In addition, bottom-up methods could also recover relevant cells for accurate diagnosis, although at the expense of false positives.

We also observed high correlation (above 95%) between the first half of CNN layers trained on natural, and cervical cell im-

age databases. Fig. 7 illustrated a strategy to identify a specific CNN layer to fine-tune databases of cervical cells for saliency prediction purposes. Furthermore, we showed that a transfer learning approach from a different domain allows CNN methodologies to achieve promising saliency prediction on cervical cell images, even using a small cervical cell image database. Future work might explore similarities across domains as part of schemes to address CNN interpretability.

The CNN-based models trained on cervical cell images outperformed the surveyed algorithms mainly because they achieved lower false positives on the estimated saliency maps and remained sensitive to relevant cell regions. These algorithms identified the most important region on the image among the three most salient regions. Thus, it confirmed the applicability of these algorithms to extract ROIs from cervical cell images. For LSIL, ASC-H and HSIL cases, our results also revealed fast bottom-up algorithms with similar ROI extractor performance to the CNN-based models with aggreement above 86%. This result indicated that there are feasible algorithms for applications with low availability of memory and computational power.

The CRIC database represented a step forward in benchmarking algorithms, but it showed bias toward the HSIL images. For these cases, the proposed center prior baseline, which consisted of a Gaussian blob at the center of the image, indicated that about 90% of HSIL lesion (when it is the most significant in the image) is located near to image center. Since the HSIL cells represent high-risk lesions, it was natural for photographers to position them close to the image center. Thus, we recommend further researches to investigate the performance of saliency methods on HSIL images free of capture bias. On the other hand, we did not find significant evidence that the capture bias influences the saliency prediction on images with other pathologies.

We identified two main limitations of our approach: (1) the CRIC database contained few HSIL #{21} and carcinoma #{8} images. This statistically restricted the ROI extractor analysis for these diseases. (2) The participation of only three cytopathologists reduced the accuracy of the location-based evaluation metrics. Additional research with more cell image databases, more lesion cases and more cytopathologists are needed to better assess the application of saliency prediction techniques as region ranking for diagnosis systems and for optimizing parameters of supervised models.

In the future, other saliency models that may improve the eye fixation prediction of cytopathologists on cervical imaging need to be investigated for several applications. Furthermore, eye tracking studies on image perception within cell analysis would be beneficial for the whole medical imaging community, especially to understand scan patterns and the reasons for diagnostic error.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by CAPES/CNPq-PVE (401442/2014-4) and CNPq (306600/2016-1), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, PPSUS/FAPEMIG (APQ-03740-17), Propp/UFOP, the Center for Recognition and Inspection of Cells (CRIC), the Moore–Sloan Foundation, and a research fellowship from CNPq (Carneiro, CM). Partial work on the development of machine learning algorithms was supported by the Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The authors would like to thank Dr. José Soares de Andrade Júnior and Dr. Humberto de Andrade Carmona for authorizing the use of the eye tracking device at the Complex Systems Laboratory (Department of Physics/UFC). Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Department of Energy or the University of California.

References

[1] M. Carrasco, Visual attention: the past 25 years, Vision Res. 51 (13) (2011) 1484–1525.
[2] L. Zhang, W. Lin, Selective Visual Attention: Computational Models and Applications, 1, Wiley-IEEE Press, 2013.
[3] A.L. Yarbus, Eye Movement and Vision (translated from the russian edition by Basil Haigh.), Plenum Press, New York, 1967.
[4] P. Polatsek, M. Waldner, I. Viola, P. Kapec, W. Benesova, Exploring visual attention and saliency modeling for task-based visual analysis, Comput. Graphics 72 (2) (2018) 26–38.
[5] F. Murabito, C. Spampinato, S. Palazzo, D. Giordano, K. Pogorelov, M. Riegler, Top-down saliency detection driven by visual classification, Comput. Vision Image Understanding 172 (2018) 67–76.
[6] C. Loukas, C. Varytimidis, K. Rapantzikos, M.A. Kanakis, Keyframe extraction from laparoscopic videos based on visual saliency detection, Comput. Methods Programs Biomed. 165 (2018) 13–23.
[7] T.V. Nguyen, Q. Zhao, S. Yan, Attentive systems: a survey, Int. J. Comput. Vision 126 (1) (2018) 86–110.
[8] A. Borji, L. Itti, State-of-the-art in visual attention modeling, IEEE Trans. Pattern Anal. Mach. Intell. 35 (1) (2013) 185–207.
[9] T. Judd, K. Ehinger, F. Durand, A. Torralba, Learning to predict where humans look, in: IEEE 12th International Conference on Computer Vision (ICCV 2009), IEEE, Kyoto, Japan, 2009, pp. 2106–2113.
[10] A. Borji, M.-M. Cheng, H. Jiang, J. Li, Salient object detection: a benchmark, IEEE Trans. Image Process. 24 (12) (2015) 5706–5722.
[11] Q. Zhang, L.T. Yang, Z. Chen, P. Li, A survey on deep learning for big data, Inf. Fusion 42 (2018) 146–157.
[12] S. Winkler, R. Subramanian, Overview of eye tracking datasets, in: Fifth International Workshop on Quality of Multimedia Experience (QoMEX 2013), IEEE, Klagenfurt am Wörthersee, Austria, 2013, pp. 212–217.
[13] M. Jiang, S. Huang, J. Duan, Q. Zhao, SALICON: saliency in context, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2015), IEEE, Boston, USA, 2015, pp. 1072–1080.
[14] A. Borji, Saliency prediction in the deep learning era: an empirical investigation, arXiv:1810.03716v1 (2018).
[15] E. Vig, M. Dorr, D. Cox, Large-scale optimization of hierarchical features for saliency prediction in natural images, in: IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2798–2805.
[16] N. Liu, J. Han, T. Liu, X. Li, Learning to predict eye fixations via multiresolution convolutional neural networks., IEEE Trans. Neural Netw. Learn. Syst. 29 (2) (2018) 392–404.
[17] X. Huang, C. Shen, X. Boix, Q. Zhao, SALICON: reducing the semantic gap in saliency prediction by adapting deep neural networks, in: IEEE International Conference on Computer Vision, 2015, pp. 262–270.
[18] S.S. Kruthiventi, K. Ayush, R.V. Babu, Deepfix: a fully convolutional neural network for predicting human eye fixations, IEEE Trans. Image Process. 26 (9) (2017) 4446–4456.
[19] N. Liu, J. Han, A deep spatial contextual long-term recurrent convolutional network for saliency detection, IEEE Trans. Image Process. 27 (7) (2018) 3264–3274.
[20] M. Cornia, L. Baraldi, G. Serra, R. Cucchiara, Predicting human eye fixations via an LSTM-based saliency attentive model, IEEE Trans. Image Process. 27 (10) (2018) 5142–5154.
[21] Z. Bylinskii, T. Judd, A. Borji, L. Itti, F. Durand, A. Oliva, A. Torralba, MIT saliency benchmark, 2015, (http://saliency.mit.edu/).
[22] H. Matsumoto, Y. Terao, A. Yugeta, H. Fukuda, M. Emoto, T. Furubayashi, T. Okano, R. Hanajima, Y. Ugawa, Where do neurologists look when viewing brain CT images? An eye-tracking study involving stroke cases, PloS One 6 (12) (2011) e28928.
[23] R. Li, P. Shi, J. Pelz, C.O. Alm, A.R. Haake, Modeling eye movement patterns to characterize perceptual skill in image-based diagnostic reasoning processes, Comput. Vision Image Understanding 151 (2016) 138–152.
[24] L. Lévêque, H. Bosmans, L. Cockmartin, H. Liu, State of the art: eye-tracking studies in medical imaging, IEEE Access 6 (2018) 37023–37034.
[25] Q. Guan, Y. Huang, Z. Zhong, Z. Zheng, L. Zheng, Y. Yang, Diagnose like a radiologist: attention guided convolutional neural network for thorax disease classification, arXiv:1801.09927v1 (2018).
[26] J.-W. Zhang, M.-C. Lian, W.-P. Wang, L. Zhu, Detection of abnormal nuclei in cervical smear images based on visual attention model, in: IEEE International Conference on Machine Learning and Cybernetics (ICMLC 2013), 2, IEEE, Tianjin, China, 2013, pp. 920–924.
[27] W. William, A. Ware, A.H. Basaza-Ejiri, J. Obungoloch, A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images, Comput. Methods Programs Biomed. 164 (2018) 15–22.
[28] L. Coombes, P. Culverhouse, Pattern recognition in cervical cytological slide images, in: Fifth International Conference on Advances in Pattern Recognition (ICAPR 2003), ICAPR, Calcutta, India, 2003.

[29] J. Zhu, I. Norman, K. Elfgren, V. Gaberi, B. Hagmar, A. Hjerpe, S. Andersson, A comparison of liquid-based cytology and pap smear as a screening method for cervical cancer, Oncol. Rep. 18 (1) (2007) 157–160.
[30] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv:1409.1556v1 (2014).
[31] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), IEEE, Las Vegas, USA, 2016, pp. 770–778.
[32] R. Nayar, D.C. Wilbur, The Pap test and Bethesda 2014, Acta Cytologica 59 (2) (2015) 121–132.
[33] D.S. Ferreira, G.L.B. Ramalho, D. Torres, A.H.G. Tobias, M.T. Rezende, F.N.S. Medeiros, A.G.C. Bianchi, C.M. Carneiro, D.M. Ushizima, CRICVA Database, 2019, Mendeley Data, v1, https://doi.org/10.17632/bk45c9yxb9.1.
[34] A. Torralba, A.A. Efros, Unbiased look at dataset bias, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011), IEEE, Colorado Springs, USA, 2011, pp. 1521–1528.
[35] A. Torralba, A. Oliva, M.S. Castelhano, J.M. Henderson, Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search., Psychol. Rev. 113 (4) (2006) 766–786.
[36] O. Le Meur, T. Baccino, Methods for comparing scanpaths and saliency maps: strengths and weaknesses, Behav. Res. Methods 45 (1) (2013) 251–266.
[37] A. Volokitin, M. Gygli, X. Boix, Predicting when saliency maps are accurate and eye fixations consistent, in: IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 544–552.
[38] Z. Bylinskii, T. Judd, A. Oliva, A. Torralba, F. Durand, What do different evaluation metrics tell us about saliency models? IEEE Trans. Pattern Anal. Mach. Intell. 41 (3) (2018) 740–757.
[39] W.H. Kruskal, W.A. Wallis, Use of ranks in one-criterion variance analysis, J. Am. Stat. Assoc. 47 (260) (1952) 583–621.
[40] B.W. Tatler, The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions, J. Vision 7 (14) (2007) 1–17.
[41] A. Borji, M.-M. Cheng, Q. Hou, H. Jiang, J. Li, Salient object detection: a survey, arXiv:1411.5878v1 (2014).
[42] J. Zhang, S. Sclaroff, Saliency detection: a boolean map approach, in: IEEE International Conference on Computer Vision (ICCV 2013), IEEE, Sydney, Australia, 2013, pp. 153–160.
[43] S. Fang, J. Li, Y. Tian, T. Huang, X. Chen, Learning discriminative subspaces on random contrasts for image saliency analysis, IEEE Trans. Neural Netw. Learn. Syst. 28 (5) (2017) 1095–1108.
[44] H.R. Tavakoli, E. Rahtu, J. Heikkilä, Fast and efficient saliency detection using sparse sampling and kernel density estimation., in: Scandinavian Conference on Image Analysis (SCIA 2011), Springer, Ystad, Sweden, 2011, pp. 666–675.
[45] L. Duan, C. Wu, J. Miao, L. Qing, Y. Fu, Visual saliency detection by spatially weighted dissimilarity, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011), IEEE, Colorado Springs, USA, 2011, pp. 473–480.
[46] H.R. Tavakoli, J. Laaksonen, Bottom-up fixation prediction using unsupervised hierarchical models, in: Asian Conference on Computer Vision (ACCV 2016), Springer, Taipei, Taiwan, 2016, pp. 287–302.
[47] L. Itti, C. Koch, E. Niebur, A model of saliency-based visual attention for rapid scene analysis, IEEE Trans. Pattern Anal. Mach. Intell. 20 (11) (1998) 1254–1259.
[48] J. Harel, C. Koch, P. Perona, Graph-based visual saliency, in: Advances in Neural Information Processing Systems (NIPS 2006), NIPS, Vancouver, Canada, 2006, pp. 545–552.
[49] X. Hou, L. Zhang, Saliency detection: a spectral residual approach, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007), IEEE, Minneapolis, USA, 2007, pp. 1–8.
[50] X. Hou, J. Harel, C. Koch, Image signature: highlighting sparse salient regions, IEEE Trans. Pattern Anal. Mach. Intell. 34 (1) (2012) 194–201.
[51] N. Murray, M. Vanrell, X. Otazu, C.A. Parraga, Saliency estimation using a non-parametric low-level vision model, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011), IEEE, Colorado Springs, USA, 2011, pp. 433–440.
[52] X. Otazu, C.A. Parraga, M. Vanrell, Toward a unified chromatic induction model, J. Vision 10 (12) (2010) 1–24.
[53] M. Wesoła, A. Lipiński, M. Jeleń, Morphometry in the cytological diagnosis of cervical smears., Adv. Clin. Exp. Med. 23 (2) (2014) 289–293.
[54] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems (NIPS 2012), 2012, pp. 1097–1105. Lake Tahoe, USA
[55] Z. Bylinskii, A. Recasens, A. Borji, A. Oliva, A. Torralba, F. Durand, Where should saliency models look next? in: European Conference on Computer Vision (ECCV 2016), Springer, Amsterdam, Netherlands, 2016, pp. 809–824.
[56] M. Hollander, D.A. Wolfe, E. Chicken, Nonparametric Statistical Methods, 751, John Wiley & Sons, New York, USA, 2013.