Analyzing and Predicting Gaze Behavior of People with Visual Impairments

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München

vorgelegt von

Jesse W. Grootjen

M.Sc. Artificial Intelligence

München, 8th July 2025

Erstgutachter: Prof. Dr. Albrecht Schmidt
Zweitgutachter: Prof. Dr. Hans Gellersen
Drittgutachter: Prof. Dr. Stefan van der Stigchel

Tag der mündlichen Prüfung: 29.10.2025

Abstract

Visual impairments, including cataracts, age-related macular degeneration, and glaucoma, affect millions globally, often progressing gradually and unnoticed until a significant part of vision is lost. These visual impairments lead to irreversible vision loss if left untreated. At the same time, treatment is available to slow the progression of these conditions, underscoring the importance of early detection and intervention. This dissertation explores how gaze behavior changes in response to cataracts, age-related macular degeneration, and glaucoma, and evaluates the feasibility of using these changes for early detection through gaze-based machine learning models. Framed within the broader goals of preventive healthcare, this work introduces a comprehensive framework for building scalable, low-burden, and privacy-preserving diagnostic systems grounded in eye tracking.

To achieve this, the dissertation addresses three core research objectives. First, it investigates how simulated cataracts, glaucoma, and age-related macular degeneration affect gaze patterns during visual tasks in high-fidelity virtual environments. Itutilizes perceptually validated simulations, refined through structured interviews with individuals directly affecteds. This work captures nuanced behavioral adaptations, such as altered fixation rates, reduced saccadic amplitudes, and increased task durations. Notably, the early stages of cataract and glaucoma often resulted in measurable gaze changes before participants became consciously aware of any visual degradation, highlighting gaze-based detection as a promising avenue for early screening.

Second, building on the finding that visual impairments induce measurable changes in eye movements, the thesis investigates how these changes can be reliably detected using machine learning. A central challenge in this effort is the presence of blink-related artifacts and missing data, which can distort gaze signals and compromise model performance. We conduct a structured literature review and identify inconsistencies in how such artifacts are handled across existing studies. In response, we introduce and validate a standardized preprocessing pipeline using 11 public datasets. Our empirical evaluation demonstrates that trimming blink-related artifacts up to 70 ms before and 118 ms after a blink with empirically derived thresholds and applying interpolation methods such as linear or cubic splines significantly enhances model reliability, particularly in sequence models such as long short-term memory neural networks. Using this pipeline, we allow for more consistent and reproducible gaze data handling, which is crucial for training robust predictive models.

Third, my dissertation evaluates the application of gaze-based models for real-time detection and monitoring of central field loss. Using an existing dataset with simulated gaze-contingent scotomas, we compare traditional feature-based classifiers (support vector machines and random forests) with deep learning models (long short-term memory neural networks) trained on raw gaze sequences. Both approaches achieved high accuracy and precision in identifying central field loss conditions. Importantly, we introduce a novel cost-benefit framework that considers both model performance and computational latency. This analysis

identifies a 1600 ms gaze window as the optimal balance between accuracy and responsiveness, enabling deployment on consumer-grade devices such as smartphones or tablets.

After this, we integrate the findings across the three chapters to demonstrate that gaze behavior can reveal subtle and condition-specific signs of visual impairment before conscious awareness. Simulations of cataracts, glaucoma, and age-related macular degeneration in virtual reality revealed measurable changes in gaze behavior, such as increased fixation rates and altered saccades, that aligned with clinical symptomatology from the literature. By validating simulation realism through interviews, the studies aim to stay close to ecological fidelity. Further, a standardized preprocessing pipeline for blink-induced gaps improved signal quality and model reliability. Finally, long-short-term memory neural networks proved capable of detecting early-stage central field loss from short gaze windows, balancing accuracy and efficiency. Together, these results position gaze-based monitoring as a promising tool for scalable, preventive screening of vision disorders.

In conclusion, this dissertation makes three key contributions: (1) a detailed characterization of gaze behavior under different types and severities of visual impairment, (2) a standardized and validated preprocessing pipeline to handle blink-induced artifacts in gaze data, and (3) the demonstration of gaze-based predictive models for early detection of central field loss, optimized for real-time, low-latency applications. These findings form the foundation for future gaze-based diagnostic systems that integrate seamlessly into daily interactions, enabling scalable and privacy-aware screening tools that support timely intervention and reduce the long-term burden of undiagnosed vision loss.

Zusammenfassung

Sehbeeinträchtigungen wie Katarakte, altersbedingte Makuladegeneration und Glaukome betreffen weltweit Millionen von Menschen und schreiten oft schleichend und unbemerkt voran, bis ein beträchtlicher Teil des Sehvermögens verloren gegangen ist. Unbehandelt können diese Seheinschränkungen zu irreversiblem Sehverlust führen. Mittlerweile stehen Therapien zur Verfügung, die das Fortschreiten dieser Erkrankungen verlangsamen können, was die Relevanz einer frühzeitigen Diagnose und Intervention unterstreicht. Diese Dissertation untersucht, wie sich das Blickverhalten als Reaktion auf Katarakte, altersbedingte Makuladegeneration und Glaukom verändert, und evaluiert das Potenzial der Nutzung dieser Veränderungen zur Früherkennung durch blickbasierte maschinelle Lernmodelle. Es wird ein umfassendes Framework für den Bau skalierbarer, belastungsarmer und datenschutzfreundlicher Diagnosesysteme auf der Grundlage von Eye-Tracking vorgestellt, das thematisch in die umfassenderen Ziele der Gesundheitsvorsorge eingebettet ist.

Um dies zu erreichen, werden in dieser Dissertation drei zentrale Forschungsziele verfolgt. Zunächst wird untersucht, wie simulierte Katarakte, Glaukome und altersbedingte Makuladegeneration das Blickverhalten bei Sehaufgaben in einer realitätsnahen virtuellen Umgebung beeinflussen. Dabei werden wahrnehmungsbasierte Simulationen verwendet, die durch strukturierte Interviews mit Betroffenen verfeinert wurden. Diese Arbeit deckt subtile Verhaltensanpassungen auf, wie z. B. veränderte Fixationsraten, reduzierte sakkadische Amplituden und eine erhöhte Aufgabendauer. Bemerkenswert ist, dass die frühen Stadien von Katarakt und Glaukom oft zu messbaren Blickveränderungen führten, bevor sich die Teilnehmer einer Sehverschlechterung bewusstwurden, was die blickbasierte Früherkennung als einen vielversprechenden Ansatz hervorhebt.

Zweitens wird, basierend auf der Erkenntnis, dass Sehbehinderungen messbare Veränderungen der Augenbewegungen hervorrufen, untersucht, wie diese Veränderungen mithilfe von maschinellem Lernen zuverlässig erkannt werden können. Eine wesentliche Herausforderung dabei sind Artefakte, die durch Blinzeln und fehlende Daten entstehen, wodurch Blicksignale verzerrt und die Leistung des Modells beeinträchtigt werden kann. Wir führen eine strukturierte Literaturrecherche durch und identifizieren Inkonsistenzen im Umgang mit solchen Artefakten in bestehenden Studien. Als Antwort darauf führen wir eine standardisierte Aufbereitungspipeline ein und validieren sie anhand von 11 öffentlichen Datensätzen. Unsere empirische Auswertung zeigt, dass das Abschneiden von blinzelbezogenen Artefakten bis zu 70 ms vor und 118 ms nach einem Blinzeln anhand empirisch abgeleiteter Schwellenwerte und die Anwendung von Interpolationsmethoden wie linearen oder kubischen Splines die Modellzuverlässigkeit deutlich erhöht, insbesondere bei Sequenzmodellen wie neuronalen Netzen mit Long Short-Term Memory. Mit dieser Pipeline ermöglichen wir eine konsistentere und reproduzierbarere Handhabung der Blickdaten, was für das Training robuster Vorhersagemodelle entscheidend ist.

Drittens evaluiert meine Dissertation die Anwendung von blickbasierten Modellen für die

Echtzeit-Erkennung und Überwachung des Verlusts des zentralen Gesichtsfeldes. Anhand eines bestehenden Datensatzes mit simulierten blickabhängigen Skotomen vergleichen wir traditionelle merkmalsbasierte Klassifikatoren (Support Vector Machines und Random Forests) mit Deep-Learning-Modellen (neuronale Netze mit Long Short-Term Memory), die auf rohen Blicksequenzen trainiert wurden. Beide Ansätze erreichten eine hohe Korrektklassifikationsrate und Relevanz bei der Identifizierung zentraler Gesichtsfeldausfälle. Insbesondere führen wir ein neues Kosten-Nutzen-Framework ein, das die Modellleistung und die Berechnungslatenz berücksichtigt. Die Analyse hat gezeigt, dass ein Blickfenster von 1600 ms das optimale Gleichgewicht zwischen Genauigkeit und Reaktionsschnelligkeit darstellt und damit den Einsatz auf Verbrauchergeräten wie Smartphones oder Tablets ermöglicht. Anschließend werden die Ergebnisse aus den drei Kapiteln zusammengeführt, um zu zeigen, dass das Blickverhalten subtile und krankheitsspezifische Anzeichen einer Sehbehinderung aufzeigen kann, bevor diese bewusst wahrgenommen wird. Simulationen von Katarakten, Glaukomen und altersbedingter Makuladegeneration in der virtuellen Realität ergaben messbare Veränderungen im Blickverhalten, wie erhöhte Fixationsraten und veränderte Sakkaden, die mit der klinischen Symptomatik aus der Fachliteratur übereinstimmten. Durch Überprüfung der Realitätstreue der Simulation anhand von Befragungen zielen die Studien darauf ab, die Realität genau abzubilden. Darüber hinaus werden die Signalqualität und die Zuverlässigkeit des Modells durch eine standardisierte Aufbereitung der Lücken, die durch Blinzeln verursacht werden, verbessert. Schlussendlich hat sich gezeigt, dass neuronale Netze mit Long Short-Term Memory in der Lage sind, den Verlust des zentralen Gesichtsfeldes im Frühstadium aus kurzen Blickfenstern zu erkennen und dabei ein Gleichgewicht zwischen Genauigkeit und Effizienz zu erzielen. Diese Ergebnisse machen Blickerfassung zu einem vielversprechenden Instrument für ein skalierbares, präventives Screening von Sehstörungen.

Diese Dissertation leistet daher drei wichtige Forschungsbeiträge: (1) eine detaillierte Charakterisierung des Blickverhaltens bei verschiedenen Arten und Schweregraden von Sehbeeinträchtigungen, (2) eine standardisierte und validierte Pipeline zur Aufbereitung von Blickdaten, um durch Blinzeln verursachte Artefakte zu behandeln, und (3) die Demonstration von blickbasierten Vorhersagemodellen für die frühzeitige Erkennung des Verlusts des zentralen Gesichtsfelds, die für Echtzeitanwendungen mit geringer Latenzzeit optimiert sind. Diese Forschungsergebnisse bilden die Grundlage für künftige blickbasierte Diagnosesysteme, die sich nahtlos in den Alltag integrieren lassen und skalierbare und datenschutzfreundliche Screening-Tools ermöglichen. Diese Tools können das rechtzeitige Eingreifen unterstützen und die langfristige Belastung durch nicht diagnostizierte Sehbeeinträchtigungen verringern.

Acknowledgements

The past years have been a challenging yet rewarding journey. Throughout this time, I have had the privilege of meeting many inspiring and wonderful people who have shaped me into the person I am today. First, I would like to extend my deepest gratitude to my supervisor, **Albrecht Schmidt**. You have created an intellectually stimulating environment where curiosity thrives, and I am grateful for your encouragement to pursue ambitious research questions and explore new directions. I always felt trusted and supported throughout my PhD journey, and I deeply appreciate being welcomed into the vibrant research community you have built. I am also profoundly grateful to **Sven Mayer**, whose day-to-day guidance and technical expertise shaped every aspect of this dissertation. Your patience through countless iterations of eye-tracking experiments, your insightful feedback on manuscripts, and your collaborative spirit made you not just a supervisor but a true research partner. Thank you for the countless hours of discussion, for pushing me to think more critically about methodology, and for the occasional coffee break when we both needed to step back. I would also like to thank the members of my dissertation committee. To **Hans Gellersen**, your expertise in gaze interaction and eye-tracking systems is unparalleled, and having your perspective on this work means a great deal to me. And to **Stefan van der Stigchel**, I am deeply grateful for your expert view on this research. Your background in visual cognition and eye-tracking methodology brought an essential perspective, and knowing that this work met your standards is truly meaningful.

I also owe my gratitude to **Pascal Knierim** for showing me the ropes as a new PhD student and introducing me to many of the systems in the lab, including the coffee management. To **Francesco Chiossi, Steeven Villa**, and **Luke Haliburton**, for being my PhD twin-bothers and beating the bureaucracy together. To **Bettina Eska**, for dealing with the many hours of open windows, especially during the winter. To **Thomas Kosch**, thank you for the amazing keyboard sounds and other music coming from your office at all hours of the day, I am still surprised and amazed how you can keep this up. **Matthias Hoppe**, your unrelenting enthusiasm for VR made even me embrace it, and thanks you for being a well-seasoned Nerf battle buddy. **Fiona Draxler**, for teaching me to teach during the first semester with practical web programming. To **Lewis Chuang** and **Tonja Machulla**, for hiring me in the first place. **Linda Hirsch**, for all the great studies to show me stuff beyond quantitative research, and your great enthusiasm for carnival and museums. **Henrike Weingärtner** and **Yannick Weiss**, for all the Christmas cookie baking, Pokémon Go adventures, and movie nights. **Sebastian Feger**, for being the best motorcycle gang member. **Robin Welsch**, for teaching me all about GLMMs and Bayesian statistics, even during conferences. **Jakob Karolus**, for your interest in eye tracking and the biggest laugh. **Florian Lang** for your amazing knowledge of German bureaucracy, amazing t-shirt collection, and choice of food. **Jan Leusmann**, for your music choice, amazing robots, and navigating the teaching of PML. **Maximiliane Windl**, for bringing the unstoppable energy to the lab. **Julian Rasch**, for your amazing mustache. **Sophia Sakel**, for your love for plants and the planticipants. **Verena Winterhalter**, for all the tea breaks and the amazing baked

goods. **Teodora Mitrevska**, for sending back all those memes. **Rifat Amin**, for the best biryani I ever had. **Thomas Weber**, for having more supplies than our actual lab, and for being my IUI co-lecturer and partner in making students understand intelligent user interfaces. **Kathrin Schnizer, Clara Sayffaerth**, and **Philipp Thalhammer**, who apparently watched me struggle through a PhD and thought, "That looks fun, I should do that too." You all somehow decided to pursue PhDs in the same lab. I'm not sure if I was a good advertisement or a terrible warning, but I'm glad you're here nonetheless. **Sarah Prange** and **Lukas Mecke**, for all your help with Uni2Work and LSF, keeping up with all the bombardment of questions and answering all of them. **Dennis Dietz**, for showing how the OptiTrack works. **Carl Oechsner**, for the amazing house parties. Thank you, **Heinrich Hussmann**, for founding the Media Informatics group and always being a warm, smiling presence in the halls. You are missed.

Beyond my formal advisors, I am deeply grateful to several senior colleagues who generously shared their wisdom and guidance with me throughout my PhD, especially **Florian Alt**, **Bastian Pfleging**, **Paweł Woźniak**, and **Jasmin Niess**.

To all the many HCI people from other labs, it is always a blast to meet up at conferences and events. I'm looking forward to the next one. Especially **Christopher Katins, Evropi Stefanidi, Nadine Wagener, Alexandra Sipatchin, Franziska Prummer, Mikołaj Woźniak**, and many others. To the many amazing students I had the pleasure of supervising along the way. I learned a great deal from each and every one of you.

Anja Mebus, thank you for being patient with me when all the different German documents did not make sense to me. **Rainer Fink**, we seriously could not do this without you. Thank you for ensuring the lab does not fall apart and for lending out all the necessary equipment for every project. **Christa** and **Franziska**, thank you all for helping with many little things along the way.

To my friends, you kept me grounded and reminded me that life exists beyond research papers. Thank you for the countless nights at Dicky Woodstock, dancing terribly at Pump up the 90's, the delicious meals, and all the moments of joy and laughter that made these years so much richer. You are wonderful.

Finally, to my dear family, thank you for your love and support over the years!

To my **Mom** and **Henk**, thank you for your unwavering support throughout these years. Your interest in my research meant more than you know.

To my dear **Dad** and **Marianne**, thank you for your endless curiosity about my research and for always asking thoughtful questions. Dad, you were the best worst critic. Your tough questions pushed me to think more clearly, even when I didn't want to hear them. And thank you both for the wonderful distractions: the delicious dinners, the fun outings, and the moments that reminded me there's life beyond the PhD.

To my sisters **Lionne** and **Barbara**, thank you for your love and support. Lionne, what a gift to have been on parallel PhD journeys! Sharing this experience with you made all the difference—you understood the highs and lows like no one else could. Barbara, I'm incredibly lucky to have you as my sister. Thank you for your unwavering support, your joy, and for all the experiences we share together. I'm so grateful for you both. And finally to **Mia**, who arrived just in time for this milestone, welcome! We've apparently set an unreasonably high bar for academic achievement in the family. My apologies in advance.

Thank you. Danke. Bedankt.

TABLE OF CONTENTS

1 INTRODUCTION

Seeing clearly and responding effectively to visual information is fundamental to nearly all aspects of everyday life. Vision enables navigation, social interaction, reading, and countless daily tasks often taken for granted until impairment occurs. Yet, visual impairments remain among the most prevalent and underdiagnosed health conditions globally [279]. As societies age and the global population continues to grow, the incidence of vision-related disorders is expected to rise significantly [175], placing new demands on public health infrastructures. In response, many healthcare systems increasingly turn to technological innovations for early screening, behavioral monitoring, and individualized risk prediction [118, 255]. Early detection and preventive strategies for vision loss have become an urgent priority within this paradigm shift. This comes as no surprise, as current technologies for detecting vision impairments are not scalable (and are therefore expensive) and place the burden on the patient. Gaze data is a great choice for use as an early detection modality as it is passive, unobtrusive, and rich in behavioral indicators. Unlike retinal imaging, it does not require specialized equipment or trained personnel. Moreover, privacy can be preserved by avoiding collecting raw video or biometric identifiers, a major concern in digital health applications [241]. These characteristics make eye tracking attractive for proactive screening and long-term monitoring across various health domains. In this dissertation, we investigate how eye-tracking and machine learning can be combined to detect early signs of cataracts, age-related macular degeneration, and glaucoma by analyzing gaze behavior. This approach offers scalable and low-burden screening alternatives that align with a broader shift toward preventive healthcare systems.

Preventive healthcare has emerged as a prevalent paradigm in modern medicine. Traditionally, healthcare systems focused on diagnosing and treating disease after symptoms arise. However, late-stage disease progression's economic and societal costs have motivated a global pivot toward early detection and prevention [168, 184]. Countries such as Japan, the United States, and members of the European Union have expanded national health policies to include screening programs for non-communicable diseases, which account for more than 70% of global mortality [76, 250]. Advances in digital health technologies, wearable sensors, and AI-based decision-making tools have accelerated this transition. These tools have demonstrated considerable promise in cardiovascular care [258], cancer risk prediction [63], and metabolic disease monitoring [270]. However, despite its considerable global burden, vision loss has been comparatively overlooked in preventive care frameworks. The World Health Organization reports that at least 1 billion people are affected by vision impairments that could have been prevented or have yet to be addressed [278]. Many of these impairments, such as cataracts or early-stage age-related macular degeneration (AMD), are either treatable or manageable if detected early. This highlights the need to develop proactive, continuous, and accessible screening mechanisms for ocular health.

In this dissertation, we contribute to addressing that gap by introducing gaze-based screening as a viable method for early detection of cataracts, age-related macular degeneration, and glaucoma. These conditions can all lead to gradual, often unnoticed vision decline. For example, central field loss (CFL) typically arises from AMD [75], while peripheral vision loss is commonly associated with glaucoma [178], and cataracts cause generalized blurring and contrast loss [133]. Early stages of these impairments are frequently asymptomatic or misattributed to normal aging [71], resulting in delayed diagnosis and lost opportunities for intervention. Individuals often adapt to progressive vision changes without conscious awareness, developing compensatory behaviors such as altered fixation strategies or increased reliance on peripheral cues, like the preferred retinal loci for CFL [222]. However, gaze behavior—including fixation stability, saccadic latency, and exploration patterns—is known to change systematically in response to visual field degradation [51, 136]. This makes eye tracking a compelling candidate for non-invasive, behaviorally-grounded vision screening, particularly in community or home-based contexts where access to specialized diagnostics may be limited.

To investigate the feasibility of gaze-based visual impairment detection, this dissertation adopts a multidisciplinary approach that spans simulations, data collection, machine learning, and user studies. A central method in this work is using high-fidelity simulated impairments, such as gaze-contingent CFL, to approximate vision loss in normally-sighted participants. This method has gained traction in vision science and accessibility research for its ability to generate reproducible conditions [89, 132, 232]. Simulated impairments offer several advantages: they enable controlled comparisons between healthy and impaired conditions, allow for rapid iteration in experimental designs, and provide training datasets that are otherwise difficult to collect from patient populations due to clinical constraints or diversity in disease presentation.

Throughout this dissertation, we leverage eye tracking within controlled simulation environments to model visual behavior under various impairments, specifically focusing on cataracts, glaucoma, and AMD. In these experiments, we analyze changes in oculomotor parameters such as fixation dispersion, saccade trajectories, and pupil dynamics. For example, we explore how AMD causes destabilized fixations, slower target acquisition, and increased microsaccadic activity. These subtle yet quantifiable behavioral signals provide the foundation for building predictive models. We then use this data to train and evaluate machine learning models capable of distinguishing between impaired and unimpaired conditions.

To enable scalability and lower the barrier to early detection, we aim to leverage machine learning models to evaluate individuals' gaze behavior for patterns that may indicate the presence of visual impairments. To evaluate the feasibility, we have implemented and compared several machine learning algorithms, including Support Vector Machines (SVMs), Random Forests (RFs), and Long Short-Term Memory (LSTM) networks. Each of these models offers distinct advantages. SVMs are well-suited for high-dimensional classification problems and have been used extensively in eye-tracking research [171]. RFs provide robustness, feature interpretability, and built-in mechanisms for dealing with missing or noisy data, which is especially valuable in clinical contexts [54]. LSTMs are particularly effective for modeling time-series data, making them ideal for analyzing gaze sequences with temporal dependencies. Their ability to learn from sequential patterns allows for more sophisticated modeling of gaze behavior, especially in prolonged interactions or dynamic tasks.

One challenge that is still not addressed is the presence of gaps in eye-tracking data (e.g., through blinks). LSTM and other neural network models are unable to handle missing data, and these gaps in data will not only interrupt the data stream but can also introduce misleading signals if not properly accounted for. Although commercial eye trackers like EyeLink and Tobii offer built-in blink detection, the methods used vary in reliability. These methods are not always reported in academic studies that we found in our work. Moreover, post-processing strategies for dealing with blink-related missing data are inconsistently applied across the literature. We systematically review blink detection practices and evaluate several infilling strategies (e.g., linear interpolation, cubic splines) to mitigate their impact on machine learning performance. Our findings demonstrate that certain preprocessing pipelines can significantly reduce the error introduced by blink artifacts, thereby increasing model robustness.

The last important contribution of this dissertation is introducing a cost-benefit framework that accounts for model accuracy and runtime performance. The trade-off between inference time and predictive reliability is critical in clinical screening applications, particularly in resource-constrained environments or real-time systems. Our evaluations show that while LSTMs achieve high accuracy with longer gaze windows, their computational cost can be prohibitive for embedded systems or continuous monitoring. We identify optimal window sizes (e.g., 1600 ms) that balance predictive accuracy and latency, enabling practical deployments without compromising diagnostic value. This cost-aware perspective aligns with broader trends in applied AI, where deployment constraints are considered alongside model performance.

In summary, this dissertation contributes to the advancement of gaze-based health diagnostics by demonstrating that there are differences in eye movements, even before the user becomes consciously aware of the visual impairment, introducing a gaze processing pipe line and by demonstrating that central field loss, a leading cause of vision impairment worldwide, can be detected from natural eye movement behavior using machine learning. It bridges the gap between technical feasibility and healthcare relevance, supporting the integration of vision screening into the evolving landscape of preventive, personalized, and accessible medicine. Ultimately, the work aims to inform the design of future diagnostic systems that are clinically effective, socially responsible, ethically grounded, and widely deployable.

1.1 Research Questions

This dissertation investigates how gaze behavior and eye-tracking data can be leveraged to detect early signs of visual impairments through a combination of simulation, machine learning, and robust data processing. The overarching aim is to explore the feasibility of scalable diagnostic methods that align with the goals of preventive healthcare.

To achieve this, the research is structured around three main themes: (1) the design and use of simulated visual impairments in virtual reality (VR) to understand their effects on gaze patterns, (2) the role of blink artifacts and missing data in gaze-based systems and their impact on reliability, and (3) the development of machine learning models for real-time CFL detection. Table 1.1 provides an overview of the research questions (RQs) that guided the work throughout this thesis.

Table 1.1: Overview of the research questions addressed in this dissertation.

RQ	Research Question	Chapter
RQ1	Can gaze patterns in VR simulations reveal signs of visual impairments?	3
RQ2	How should we deal with gaps in eye-tracking data for machine learning models?	4
RQ3	Is real-time detection of central field loss possible using eye-tracking data?	5

The first question (**RQ1**) focuses on how simulations of vision loss in virtual reality can reveal changes in gaze behavior and facilitate early-stage detection of impairments. This includes identifying the specific characteristics in eye movements associated with different types and severities of simulated vision loss and determining the threshold at which gaze-based indicators become reliably detectable even in the absence of conscious awareness.

The second question (**RQ2**) investigates how blink-related artifacts affect data quality and system performance in gaze-based applications. Given the variability in blink detection and handling strategies across studies, this question explores current HCI practices and the generalizability of findings across different hardware and experimental contexts. Furthermore, the impact of artifact mitigation techniques on deep learning performance is examined in detail.

The final question (**RQ3**) addresses the feasibility of real-time CFL detection using continuous eye-tracking data. Here, we evaluate the performance of traditional machine learning approaches on engineered features and investigate whether deep learning models can improve predictive accuracy when trained on raw gaze sequences. We also examine the trade-offs between accuracy and computational cost to identify optimal configurations for real-time deployment.

1.2 Research Context

This thesis is built on work conducted at the Chair for Human-Centered Ubiquitous Media and the Media Informatics Group at LMU Munich over approximately five and a half years. My PhD was primarily supervised by Prof. Albrecht Schmidt, the head of the research group.

Throughout this time, I collaborated with researchers and project partners across multiple institutions to conduct studies, co-organize events, and co-author papers. These collaborations are highlighted in Table 7.1 at the end of the thesis.

Integrated Diagnosis and e-Assistance for people with visual impairments (IDeA) The first part of this thesis was conducted as part of IDeA, which was funded by the BMBF (grant no. 16SV8102) and the work of many involved project partners. Blickshift GmbH, NMY Mixed-Reality Communication GmbH, Talkingeyes&More GmbH, and the University of Tübingen. Within IDeA, we investigated the use of eye-tracking data to detect visual impairments and to develop assistive technologies for people with visual impairments. The work conducted within IDeA resulted in the following publications [79, 143].

Munich Center for Machine Learning (MCML): The final portion of this thesis was conducted as part of the Munich Center for Machine Learning (MCML) at LMU Munich. The MCML connects fifty principal investigators and over 100 PhD students from LMU Munich and TUM researching AI-related topics. Within project *C5: Humane AI*, we investigate human-computer interaction and other human-centered AI topics. The work conducted with MCML resulted in several papers [67, 80, 81, 82, 83, 84, 85, 86, 273].

Due to the collaborative nature of research, every publication was a joint effort with fellow researchers and students. As such, I use the scientific term “we” throughout the thesis.

1.3 Thesis Outline

This dissertation comprises seven chapters, the bibliography, appendix, and the enumerating lists. This work contains the results and evaluations of a total of seven empirical studies between fall 2019 and early 2025. This work is structured as follows:

Chapter 1: Introduction contains the description and motivation of this dissertation, an overview of the research questions, the research context, and this outline.

Chapter 2: Related Work comprises the related work to put preventive healthcare through eye tracking central for visual impairments. Moreover, this chapter also contains the specific related work to connect the three research questions with its specific related work.

Chapter 3: Gaze Patterns in Visual Impairments contains three studies to investigate how simulated visual impairments influence gaze behavior during realistic visual search tasks within virtual environments. We designed the simulations based on semi-structured interviews with individuals experiencing these conditions. After which we systematically explored how different levels of visual impairment affect eye movements and visual search performance. This chapter addresses **RQ1**.

Chapter 4: Standardizing investigates how we can go from gaze data to machine learning models. We first investigate how literature has dealt with blink detection and the handeling of gaps in eye tracking data. We then evaluate the impact of these different preprocessing techniques on the error they introduce compared to the ground truth. Finally, we looked at how blink artifacts affect the performance of a machine learning model, and how different mitigation strategies influence the accuracy of this model. In this chapter, we address **RQ2**.

Chapter 5: Applications adresses the challenge on how eye-tracking data can be used detect CFL in real-time. We first investigate the performance of traditional machine learning models on engineered features, and then we investigate whether deep learning models can improve predictive accuracy when trained on raw gaze sequences. We also examine the trade-offs between accuracy and computational cost to identify optimal configurations for real-time deployment. This chapter answers **RQ3**.

Chapter 6: Discussion reflects on the findings from the previous chapters in relation to the overarching research questions. It synthesizes how different types and severities of simulated visual impairments manifest in gaze behavior, evaluates the effectiveness of various preprocessing strategies, and considers the practical implications of deploying gaze-based models for early detection. The chapter highlights how measurable gaze changes often precede user awareness, underlining their diagnostic value. It also situates these insights within the context of related work, emphasizing how this research advances both methodological and applied contributions in gaze-based health monitoring.

Chapter 7: Conclusion summarizes the core contributions of this dissertation and outlines its broader relevance for preventive healthcare. It presents a framework for gaze-based diagnostics, spanning simulation, preprocessing, and classification. The chapter concludes by articulating a vision for future health technologies where eye tracking is embedded in everyday devices to support continuous, privacy-respecting, and accessible screening. It positions the dissertation as a foundation for future work in gaze-based sensing and highlights opportunities for expanding this research to broader domains of behavioral health monitoring.

2 RELATED WORK

This chapter synthesizes prior work across four key areas relevant to this dissertation: (1) the emerging role of preventive healthcare and the need to integrate vision into this paradigm, (2) visual impairment simulations and their effects on behavior, (3) machine learning approaches for detecting visual impairments from gaze data, and (4) the role of blink artifacts and their handling in gaze-based systems. Together, these areas highlight the growing potential of low-burden, gaze-based methods for the early detection of central field loss, supporting broader goals in preventive health and personalized care.

2.1 Preventive Healthcare

Preventive healthcare has emerged as a cornerstone of modern health policy, marking a shift from reactive treatment to proactive risk reduction on a global scale. International commitments over the past decade, including United Nations declarations and the WHO Global Action Plan on Non-Communicable Diseases (NCDs), reflect a consensus that focusing on prevention is crucial for sustainable health systems [184]. NCDs now account for over 70% of global deaths [76], and a significant portion of this burden is avoidable through early interventions and lifestyle modifications [16, 278]. In response, major healthcare systems have expanded preventive services in policy and practice. For example, the United States increased access to routine screenings under national health reforms by requiring insurance coverage for evidence-based preventive services without cost-sharing [157]. Similarly, the European Union prioritized reducing chronic disease rates via prevention to curb the 70–80% of health expenditures consumed by managing such conditions [225], and Japan legally mandates annual health check-ups for adults aged 40–74 to facilitate early disease detection [177]. These efforts have already improved early diagnosis and reduced disease complications, demonstrating long-term cost savings for society [183].

In parallel, there is growing recognition that preventive care should encompass vision health, which has often been underemphasized in public health strategies. Yet visual impairments impose substantial societal costs, with many cases that are preventable or manageable with timely care. The WHO's World Report on Vision noted that of 2.2 billion people with vision impairment worldwide, at least 1 billion have conditions that were preventable or remain unaddressed [278]. Integrating routine eye screenings into health systems can therefore yield significant benefits. Early detection of ocular conditions enables interventions that preserve sight and reduce downstream disability costs. CFL, often due to macular degeneration or diabetic macular edema, exemplifies where prevention can be transformative. Age-related macular degeneration (AMD), a leading cause of CFL in older adults, already affects roughly one in ten people over age 50 [205] and incurs high treatment expenses, often tens of thousands of dollars per case [206]. However, detecting AMD progression early allows timely therapies that slow vision loss and preserve visual function [70]. Similar approaches, such as regular diabetic retinopathy screening, have proven effective in preventing vision loss [278]. By broadening the preventive paradigm to include vision health, particularly through low-burden screening for early signs of central vision loss, health systems can improve outcomes, maintain quality of life, and reduce the long-term societal costs associated with visual impairment.

2.2 Simulating Visual Impairments to Understand Behavior

To support this early detection, understanding the changes that happen in eye movement behavior and perception that may be an indicator is crucial. Thus, understanding how visual impairments affect perception and behavior has been a focus of human vision research. Advances in simulation technology now enable controlled studies of impairments such as AMD, glaucoma, and cataracts, helping researchers replicate and explore their real-world impact.

2.2.1 Simulating Central Field Loss and AMD

AMD is a leading cause of CFL, progressively impairing fine vision and forcing reliance on eccentric fixation [136]. Eye movements in AMD patients show increased fixation instability, longer saccadic reaction times, and altered patterns during reading and visual search [37, 51, 87]. These changes result in shorter forward saccades, more regressions, and reduced reading speeds [39]. Importantly, oculomotor adaptation strategies differ depending on whether the vision loss is central (e.g., AMD) or peripheral (e.g., glaucoma), as shown by Burton et al. [38].

To study these behavioral effects under controlled conditions, researchers developed various methods to simulate AMD. Early approaches included affixing materials to computer screens [274], while more recent efforts employ contact lenses [6], computer-based visual filters [107, 163], and head-mounted displays [4, 89]. Gaze-contingent simulations allow real-time adaptation of the scotoma to user gaze, thereby more accurately reflecting the experience of individuals with CFL [232].

In addition to studying task performance, these simulations have proven valuable for research in accessibility and user-centered design. Simulation environments can be tailored to evaluate specific perceptual challenges users face, providing empirical evidence for changes in design, interface, or architectural planning. Such systems have been used to study navigation efficiency, decision-making, and visual search in both controlled environments and semi-realistic tasks, making them broadly relevant to both HCI and rehabilitation research.

These simulations are not just tools for academic inquiry. They have been used to assess accessibility in healthcare settings [213], evaluate the impact of design on usability, and even foster empathy among designers and medical students [91]. Empirical studies using simulations show that reading speed declines with increasing scotoma size [274], and face recognition becomes significantly slower and less accurate [98, 163]. The alignment between simulation outcomes and patient performance [198] supports their validity as proxies for real impairments.

2.2.2 Simulating Peripheral Vision Loss and Glaucoma

Glaucoma causes progressive damage to the optic nerve, usually beginning with peripheral vision loss and eventually leading to tunnel vision or blindness [202]. This peripheral degradation impairs scene exploration, spatial awareness, and mobility. Simulations of glaucomatous vision, including VR-based overlays [244] and impairment goggles [148], help evaluate how field loss influences task performance and perception.

These simulations have proven valuable in multiple contexts. For example, Lam et al. [141] found that glaucoma patients had longer search times and higher error rates when navigating virtual environments. Similarly, Sippel et al. [234] showed that binocular field loss slowed visual search in a supermarket task. VR simulations reveal that search and scanning inefficiencies [236, 248], difficulty with object identification in cluttered scenes [178], and impaired driving [146] are all common consequences of glaucoma. Such findings are consistent with those observed in clinical populations, suggesting that simulation-based evaluation can anticipate real-world impairments.

Furthermore, simulation platforms may assist in screening strategies and rehabilitation planning. By offering controlled exposure to daily challenges, such tools can personalize assessments or inform adaptive technologies tailored to an individual's specific visual field loss. Future directions for glaucoma simulation include integration with eye tracking and real-time feedback mechanisms, improving diagnostic realism and usability testing.

2.2.3 Simulating Cataracts

Cataracts, characterized by lens opacities, obstruct light transmission and scatter incoming light (straylight), reducing contrast and sharpness [261]. Their visual impact depends on opacity type (nuclear, cortical, subcapsular) and pupil size [166, 167]. These impairments are associated with prolonged fixations, delayed saccades, and decreased search efficiency [253, 268].

Cataract simulations include fixed goggles [277], contact lenses [6], and virtual environments with parameterized visual distortions [132, 148]. While widely used to assess visibility and empathy, surprisingly few studies examine their impact on eye movements. One exception is Wan et al. [268], who found that cataracts prolong fixation durations and reduce recognition performance. Simulating cataract severity in dynamic, task-based contexts remains an open area that this dissertation seeks to address.

In summary, simulations targeting visual impairment vision are often static or lack the temporal fidelity needed for high-resolution behavioral studies. As VR systems continue to develop, there is a growing opportunity to model visual impairments effects dynamically, incorporating progression over time and user behavior. We have taken these opportunities to incorporate progression over time into one of our user studies.

2.3 Gaze and Machine Learning for Vision Loss Detection

The rise of eye tracking and artificial intelligence has made it possible to explore visual impairments from a computational perspective. In particular, machine learning techniques are increasingly employed to detect vision-related conditions based on distinctive gaze behavior.

2.3.1 Eye Movement Characteristics in AMD

Such distinctive gaze behavior can be observed in individuals with AMD. They develop Preferred Retinal Loci (PRLs) to compensate for central vision loss, resulting in unstable fixations [51, 222]. These fixations span wider areas and lack the precision of foveal vision [207]. Saccades also become non-foveating, exhibiting lower velocities, longer durations, and curved trajectories [210, 275].

These patterns affect everyday activities: reading rates drop [214], search becomes inefficient [131], and face recognition suffers [224]. Despite these measurable effects, AMD often remains undetected for years [71], showing the need for unobtrusive early detection.

Importantly, these gaze behaviors are not only indicators of visual deficits but can also offer insight into compensatory mechanisms and potential interventions. Adaptive technologies and training programs could benefit from understanding how gaze behavior evolves with vision loss. The challenge remains to extract robust and generalizable features across users, tasks, and environments which allow the detection of visual impairments.

2.3.2 Machine Learning from Eye Movements

One method that has been used to successfully detect visual and cognitive impairments from gaze data, is machine learning. Rello and Ballesteros [209] trained SVMs to detect dyslexia from reading behavior, while Chen et al. [42] used gaze deviations to detect strabismus. In glaucoma research, Smith et al. [237] applied Gaussian Mixture Models to evaluate reading in patients.

Deep learning offers further promise. LSTM networks capture temporal dependencies in time-series data and outperform traditional models in gaze-based predictions [33, 93, 242].

Such models learn rich representations directly from raw input, reducing the need for hand-crafted features. Applications range from mental health diagnostics [57] to driver attention tracking [264].

However, many ML models require substantial, clean data, often impractical in everyday scenarios and raises privacy concerns.

2.3.3 Cost and Privacy Considerations

Extensive gaze collection raises privacy concerns. Gaze patterns can reveal identity, gender, age [30, 172], sexual preferences [139], and even hormonal cycles [18]. Tools like PrivacEye [241] attempt to anonymize gaze data, but may also strip away features critical for detecting conditions like CFL. This motivates our investigation into cost-efficient models using minimal, privacy-preserving gaze data.

In this dissertation, we propose to optimize gaze-based detection by reducing input window sizes and computing time. Following prior work [280], we model the trade-off between accuracy and runtime. By applying both traditional (SVM, RF) and deep learning (LSTM) models, we benchmark detection performance and latency, advancing gaze-based diagnostics toward real-world applicability.

In addition, we evaluate how much data is truly necessary to detect early signs of impairment. Prior work often assumes long observation periods are required. However, our findings challenge this assumption and show that robust models can work effectively with short windows, paving the way for real-time monitoring and feedback systems.

2.4 Blink Artifacts and Their Role in Gaze-Based Systems

Blinks introduce challenges in eye tracking by creating periods of signal loss or distortion. Accurately detecting and compensating for these artifacts is critical for both diagnostic and interactive gaze-based applications.

2.4.1 Blink Dynamics and Influencing Factors

Blinks are rapid eyelid closures that protect the eye and maintain tear-film stability [65]. Blink rate and duration vary with environmental conditions [190], age [243], cognitive load [259], and fatigue [185]. In high-load tasks, blink frequency tends to decrease, while duration may increase.

These variations complicate their use in interactive systems or medical applications. For instance, blinks have been linked to attention [144], mental workload [34], and neurological disorders [23]. However, they also produce missing gaze samples and transient artifacts, potentially distorting gaze-based inferences if not properly handled.

2.4.2 Blink Detection Approaches

Modern eye trackers include built-in blink detectors, such as the EyeLink 1000[1] (SR Research Ltd., Ottawa, ON, Canada) and BeGaze[2] (SensoMotoric Instruments GmbH., Toltow, Germany) systems. These rely on pupil size or occlusion thresholds to flag blinks, yet they cannot distinguish tracking loss from true eyelid closure [101]. Consequently, fixations adjacent to blinks are often noisy as they are interrupted with blinks.

2.4.3 Impact on Predictive Models and Interactive Systems

Blinks interrupt gaze sequences used in predictive models like LSTM networks. If untreated, these gaps degrade performance [83]. Studies infill blink gaps using local averaging [12] or extrapolation [33], but practices vary widely.

In interactive systems, blinks have been used as intentional input (e.g., menu navigation [156]), but also complicate dwell-based selection [123]. As gaze interfaces grow, understanding and mitigating blink interference becomes increasingly important. Our approach explicitly quantifies blink effects and selects preprocessing techniques based on empirical performance.

Summary

Prior research has demonstrated the utility of simulating vision loss to study behavior and accessibility. Gaze behavior changes can be leveraged to develop predictive models. However, challenges related to data quality, privacy, and computational cost remain. This dissertation builds on prior work by systematically evaluating the accuracy and cost trade-offs of different modeling strategies, while accounting for blink artifacts and varying amounts of gaze data. The following chapters describe our methods and results in addressing these challenges.

[1] `https://www.sr-research.com/eyelink-1000-plus/`, accessed April 20, 2026

[2] `https://www.dpg.unipd.it/sites/dpg.unipd.it/files/BeGaze2.pdf`, accessed April 20, 2026

3 GAZE PATTERNS IN VISUAL IMPAIRMENTS

Preventive healthcare emphasizes early detection and timely intervention to mitigate or delay the onset and progression of diseases. This can significantly improve individual health outcomes and reducing healthcare burdens. Visual impairments represent an important target for preventive health strategies, given their profound impact on quality of life and the substantial global prevalence.

Early stages of visual impairments often go unnoticed due to their subtle and gradual progression. For example, cataracts, currently affecting over 65 million individuals globally, progresses slowly, making early identification challenging [278]. Similarly, glaucoma and AMD frequently remain undetected until significant and irreversible vision loss has occurred. These conditions significantly affect daily activities, independence, and overall quality of life, highlighting the need for preventive measures that enable earlier diagnosis and more effective management.

Simulation techniques using augmented and virtual reality (AR/VR) offer valuable tools for preventive healthcare research. By accurately representing visual impairments, simulations provide crucial insights into their early manifestations and progression. Previous research utilizing simulations has effectively demonstrated how visual impairments affect task performance and gaze behaviors, serving as essential platforms for understanding impairment impacts and developing preventive strategies [105, 132, 133]. However, they often user simulations that do not accurately represent how the world is perceived by those affected, are often only visualizing one static severity that does not allow for the comparison of normal eye movements with those during simulation-affected trials.

In this chapter, we investigate how different levels of simulated visual impairments influence gaze behaviors during realistic visual search tasks within virtual environments. Our approach involves detailed visual impairment pipelines for cataracts, glaucoma, and AMD, meticulously designed based on semi-structured interviews with individuals experiencing these conditions. These simulations feature adjustable parameters, enabling precise replication of impairment severity and progression. Across multiple controlled user studies (N=101 total), we systematically explore how different levels of visual impairment affect eye movements and visual search performance.

Our findings demonstrate that moderate-to-advanced stages of simulated impairments notably alter gaze patterns and task performance. Conversely, early-stage impairments often remain undetected by participants and show minimal immediate impact, emphasizing the critical importance of objective measurement techniques capable of identifying subtle early-stage changes. Specifically, we highlight how conventional visual impairment simulations often inadequately represent patient-reported experiences, particularly with conditions like AMD. By refining simulation realism, we facilitate earlier detection and better preventive strategies, ultimately contributing to enhanced patient outcomes and reduced healthcare costs associated with advanced visual impairment management.

In this chapter, we present the work of the first three included publications [84, 85, 86], which outlines a vision for advancing preventive healthcare by leveraging realistic virtual reality simulations of visual impairments to identify early-stage diagnostic indicators through detailed analyses of eye movement patterns and task performance[1]. Overall, this chapter addresses **RQ1**.

RQ 1: *Can gaze patterns in VR simulations reveal signs of visual impairments?*

This chapter is based on the following publications.

Grootjen, Jesse W. et al. 'The Effects of AMD Severities on Eye Movements in Virtual Environments.' In: *Mensch und Computer 2025*. New York, NY, USA location = Chemnitz, Germany: Association for Computing Machinery, August 1, 2025. DOI: 10.1145/3743049.3748561

Grootjen, Jesse W. et al. 'Understanding the Impact of Glaucoma Severities Through Virtual Reality Simulations.' In: *Proceedings of Mensch und Computer 2025*. MuĆ25. Chemnitz, Germany: Association for Computing Machinery, August 31, 2025. DOI: 10.1145/10.1145/3743049.3743082

Grootjen, Jesse W. et al. 'Effects of Cataracts Severities on Eye Movements and Task Performance During a Visual Search Task Through Virtual Reality.' In: *Mobile and Ubiquitous Multimedia 2025*. New York, NY, USA location = Enna, Italy: Association for Computing Machinery, December 1, 2025. DOI: 10.1145/3771882.3771884

[1]The following sections are copied with minor changes from the publications: subsection 3.1.1 until subsection 3.1.4, subsection 3.2.1 until subsection 3.2.4, and subsection 3.3.1 until subsection 3.3.9

3.1 Understanding the Impact of Age-Related Macular Degeneration

In this section, we aim to understand how AMD is subjectively perceived by those living with the condition and evaluate whether existing visual simulations reflect this perceptual reality. To do so, we conducted semi-structured interviews with individuals diagnosed with AMD to assess the limitations of current simulation techniques. Building on these insights, we then developed a simulation pipeline that attempts to more accurately reproduce the effects of central-field loss in a gaze-contingent manner. Finally, we evaluated the impact of these visual changes on task performance and gaze behavior in a virtual reality environment. Through this progression—from perception to simulation to behavioral evaluation—we seek to bridge the gap between clinical symptoms and the actual visual experiences of people with AMD.

3.1.1 Structured Interviews

From the simulations mentioned in the related work, we found that related work mostly visualized the central-field loss from AMD as a dark circle with a gradient around the edges when looking at images. We held semi-structured interviews with individuals who have AMD to verify the current quality of visualizations. The interviews all consisted of two parts: demographic questions and, following this, questions about how they perceive the world. All interviews were conducted over video call or in person.

Procedure Prior to the interview, participants were provided with consent forms and a detailed explanation of the procedure, giving them ample time to review the information. At the start of each session, participants were invited to ask any questions regarding the consent form and the process. The interview began with a set of demographic questions relevant to the study, followed by questions about their perception of the world. Examples of these guiding questions include: “How does glaucoma affect your daily life? What challenges do you face”? and “Do specific lighting conditions influence your perception, and if so, how”?

Participants In total, we interviewed six individuals who currently have AMD. One participant was excluded as they reported having other visual impairments as well, which are more influential to their vision. For the remaining five participants, P1 (female, 86 years old) reported $3-5\%$ residual vision in the left eye and no remaining vision in the right eye. P2 (female, 54 years old) reported 3% residual vision in her left eye; in the right eye, she receives just light. P3 (male, 24 years old) reported 3% vision remaining in both eyes (Reported Juvenile Macular Degeneration, with similar characteristics to AMD). P4 (male, 63 years old), reported 5% residual vision in both eyes. Finally, P5 (female, 76 years old) reported 3% residual vision in her right eye and 10% residual vision in her left. It is important to note that each of the participants reported having normal vision earlier on in life.

Results During the interviews, we found that all but one of the participants reported that they did not experience AMD as it usually is visualized, i.e., a dark spot with fading around the edges. Four of the interviewed participants reported a similar perception as those with normal vision experience in the blind spot. "I see a road continue as normal, but I cannot see if there is a bike standing in front" (P3); they mentioned furthermore that the "black spots used for simulations not realistic, the brain infills". Similar reports were made by P2: "While standing in the kitchen, I was not able to see the keys; it was just empty". similar reports were made for P4 and P5. However, P1 reported her perception to be more of a heavy fog, of a grey color, obscuring her vision (If participant statements were not made in English, quotes have been translated by a native speaker with full professional proficiency in English).

3.1.2 AMD Simulation

In the following, we describe the implementation of our effects pipeline using two parameters: central-field loss and fading. It is important to note that we assume that the visualization is completely round. This is oversimplified as there are indefinite variations.

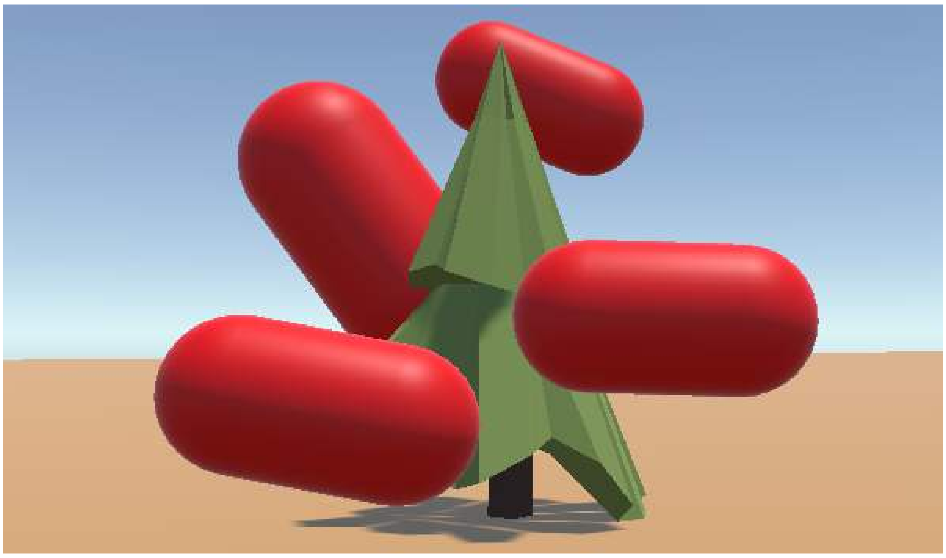


Figure 3.1: A low-poly landscape without any of the AMD visualizations applied.

Central-Field Loss We simulate the central-field loss using a scotoma with a spherical shape of size S (diameter of the sphere) at a fixed distance (d) from each camera while being gaze-contingent. SC is the degrees of central vision to be obstructed by our AMD simulation. Using the following formula, we determine the size of the sphere: $S = 2 \times d \times \tan(SC/2)$.

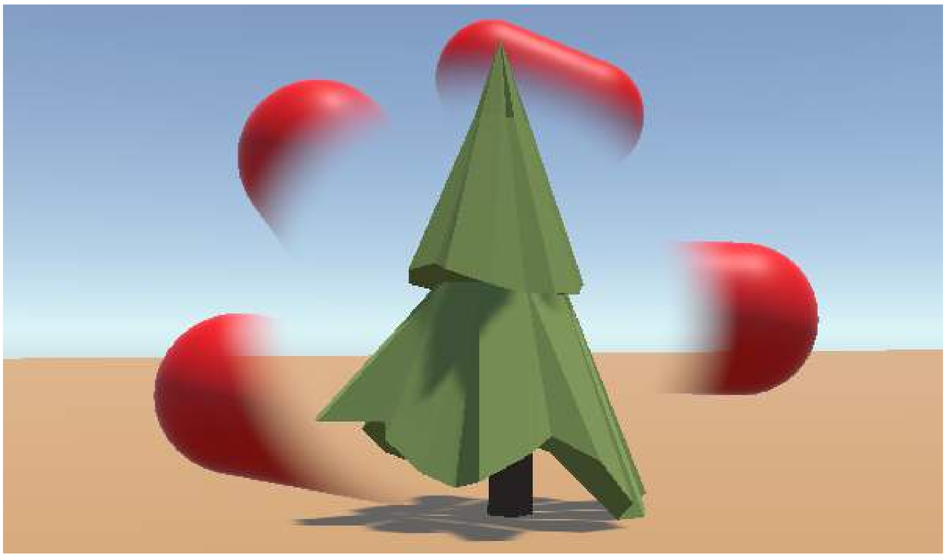


Figure 3.2: A low-poly landscape with the central-field loss and fading visualization applied. In this image, the red capsules have a custom material that hides these items in the central field if they are obscured by our visualizations. This simulation is gaze-contingent.

Fading We implemented the fading on the edges using an adaptation of the `smoothstep` function[2]. The `smoothstep` function uses three parameters, the lower edge (e_0), the upper edge (e_1), and the source variable for interpolation (x). If x is less than or equal to e_0, it outputs 0; if x is greater than or equal to e_1, it outputs 1; and if x is between e_0 and e_1, it interpolates between e_0 (start edge) and e_1 (end edge) according to the Hermite interpolation. The source variable for interpolation (x) is the normalized distance of the objects' texture fragment concerning the sphere's central position in the screen space. Finally, we use the output of this `smoothstep` function as alpha values of the objects' texture to control the objects' visualization.

$$f(x) = \begin{cases} 0 & \text{if } x \leq e_0 \\ 3(\frac{x-e_0}{e_1-e_0})^2 - 2(\frac{x-e_0}{e_1-e_0})^3 & \text{if } e_0 < x < e_1 \\ 1 & \text{if } x \geq e_1 \end{cases}$$

By applying the combination of the central field loss and the fading as a custom shader in Unity, we were able to simulate the loss of central vision where objects disappear when they are obscured by the sphere while leaving the rest of the unity scene not affected, see Figure 3.2. We chose to have our fading dynamically adjust to the size of our CFL, namely being 20% of the CFL, if our AMD simulation is $30°$, our fading ends at $30°$, and the CFL encompasses 80% of the $30°$, specifically $24°$.

[2]see, `http://www.thebookofshaders.com/`, last accessed: April 20, 2026

3.1.3 User Study

In the following section, we will discuss the user study performed with 17 participants, the virtual environment and hardware used for this user study, and the analysis. We ensured that luminance in all rooms was consistent.

We chose to simulate eight conditions of AMD, all Latin Square balanced, of no visualization, $4.3°$, $8.6°$, $12.9°$, $17.2°$, $21.5°$, $25.8°$ and $30°$. These were chosen to keep the interval between conditions equal. Furthermore, we decided not to go beyond $30°$ to keep the visualization lower than what has been reported in previous work to go unnoticed [71] and to stay inside of the near periphery [245].

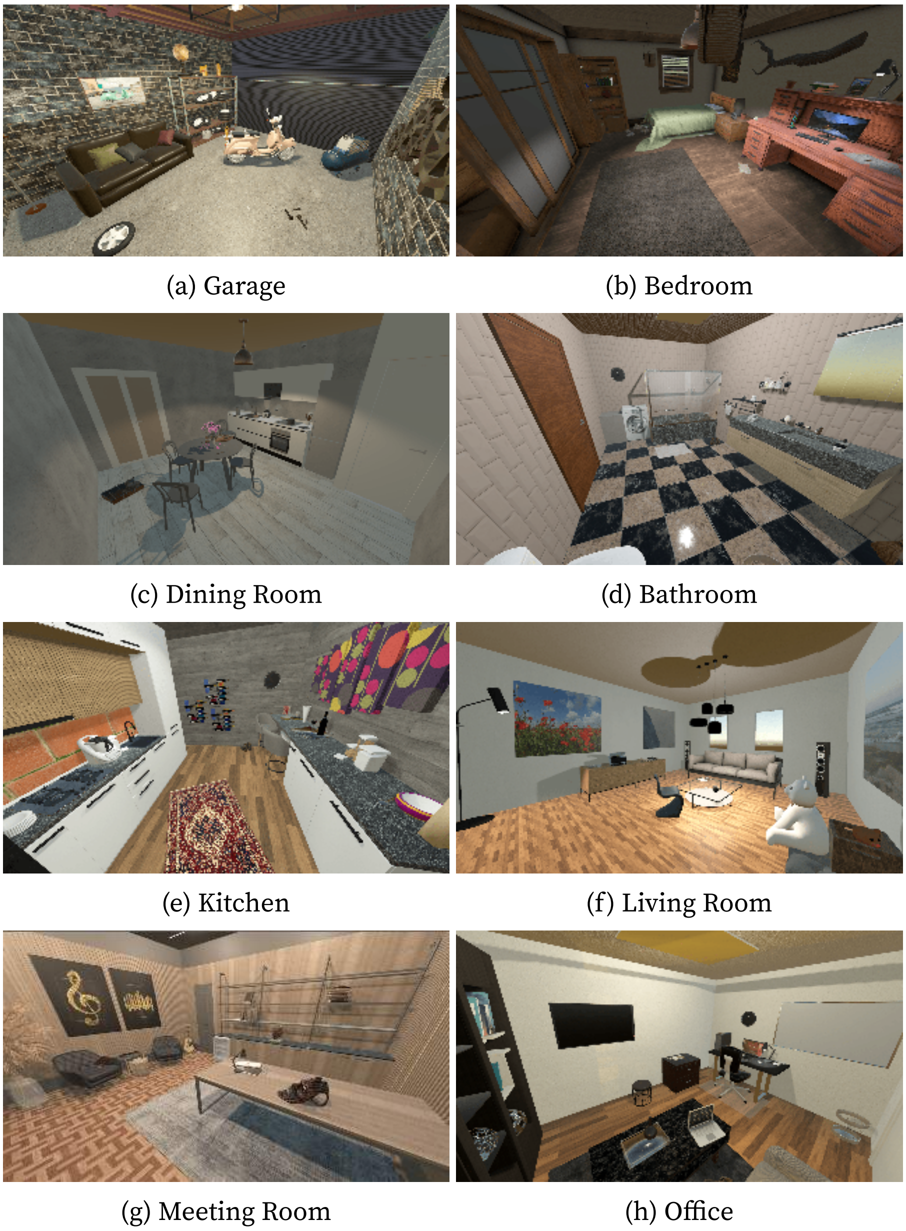
Apparatus

We selected a visual search task to assess the impact of the AMD simulation on participants' behavior and performance. Visual search tasks have been employed widely to investigate visual impairments both in patients with visual impairments [52, 198] as well as in simulations of visual impairments [56, 113, 137].

Pollmann et al. [198] showed that participants with AMD are significantly less deficient in search tasks compared to healthy participants in realistic scenes (such as rooms in a home) due to the possibility of contextual cueing. This shows that displaying realistic scenes is necessary to assess the deficiencies of visual impairment in search tasks adequately. Therefore, we opted to show realistic environments of different rooms in the virtual environment. Similar to David et al. [56], we built eight virtual environments to represent different rooms in a house; see Figure 3.3. For the search items, we used items that could be found in a household, namely, a medicine bottle, screwdriver, calculator, flashlight, glass, hairdryer, tape dispenser, kettle, eraser, sticky notes, phone, highlighter, apple, duck, hole puncher, pencil sharpener, first-aid kit. All rooms and items were chosen to preserve contextual cueing.

We built the virtual environments in Unity3D and deployed it to a desktop with a Ryzan 9 7900X3D, 64GB RAM, an NVIDIA RTX 4080 Super graphics card, and a connected HTC Vive Pro Eye with a built-in Tobii eye tracker. The built-in Tobii eye tracker was accessed using the Vive SRanipal SDK at a sampling frequency of 120 Hz.

To facilitate our analysis, we converted the head directional 3D vectors, eye-in-world directional 3D vectors, and directional 3D vectors into 2D Fick angles using the Fick-gimbal method [92]. This conversion involves two rotations, one over the vertical axis and another over the nested horizontal axis, effectively characterizing the position of each vector. These resulting 2D Fick angles for eye and head directions were the basis for our following analysis. For the analysis, we focused on the eye-tracking data for each trial. Given the relatively short duration of our trials, averaging approximately 4.64 ± 3.41 seconds across all trials, we conducted our eye tracking analysis at the trial level rather than examining individual behaviors within each trial.



(a) Garage (b) Bedroom (c) Dining Room (d) Bathroom (e) Kitchen (f) Living Room (g) Meeting Room (h) Office

Figure 3.3: All rooms used in the user study.

Procedure

The study coordinator first welcomed the participants, and the coordinator then introduced and explained the study procedure. After signing the consent form, we asked them to answer a few demographic questions. Finally, before entering VR, the study coordinator explained the hardware and controls relevant to the study to the participants.

After putting on the VR headset, the study coordinator navigated the participant to the integrated eye-tracking calibration. After this, each participant did 4 practice trials without

Figure 3.4: Items participants were tasked with searching for, listed from left to right and top to bottom, are eraser, sticky notes, glass, medicine bottle, tape dispenser, hole puncher, duck, apple, first aid kit, screwdriver, pencil sharpener, smartphone, calculator, hairdryer, highlighter, kettle, and, flashlight.

any simulation to get familiar with the task before starting the experiment. After which, the experiment was started. For each trial, we present the item the participant was tasked with finding on a gray background. Once the participant confirms, the item spawns at a random location in the room. These spawning locations were randomly assigned throughout the room, and potential collisions with other items were checked. If there was a collision, a different spawning location was selected. This was then repeated for 15 items in each room. We balanced the different severities using Latin square balancing for 8 conditions. Varying from no simulated central scotoma to 30 degrees of central scotoma in equal increments.

We applied a drift correction after each room, following the procedure from Sipatchin et al. [233]. This drift correction was designed to compensate for eye tracking data quality decay in VR due to the participant's movement, which is known to induce drift into the precision of the eye tracker [47]. These drifts can influence the gaze-contingent simulation of the peripheral loss.

Participants

We recruited 17 participants. However, we had to exclude four participants (one for technical issues, one was too exhausted to participate in the study properly, and two did not follow the instructions.). The remaining 13 participants had a mean age of 23.2 ($SD = 2.2$), 8 identifying as female and 5 as male. All participants had normal vision.

3.1.4 Results

In this section, we first present the results of our preprocessing. After which, we present the findings of how our simulation affects eye movements and task performance while being affected by different severities of our AMD simulation.

Peprocessing

We calculated fixations and saccades using pymovements [129]. We leveraged ID-T algorithm [217] with specific fixations thresholds set at a minimum fixation duration of 83 ms and a maximum dispersion of $1.8°$, in line with prior work [22, 260]. This allows us to calculate the following fixation-based metrics: total fixation duration, average fixation duration, fixation count, and the time from the start of the trial to the last fixation in the trial. We used the microsaccade algorithm [62] to extract the saccades. This algorithm allows us to calculate the saccadic amplitude and saccade frequency. The saccadic amplitude was calculated as the angle in degrees between the saccade onset and offset, while the saccadic frequency was determined by dividing the number of saccades within a trial by the trial duration in minutes. To calculate the average pupil dilation for each trial, we computed the mean of the left and right pupil sizes and subsequently calculated the standard deviation to facilitate our analysis at the trial level. Finally, for computing the Index of Pupillary Activity (IPA), we employed the implementation by Duchowski et al. [59]. Thus, we utilized discrete wavelet transforms to analyze pupil diameter signals, specifically employing a two-level discrete wavelet transform. We normalized the wavelet coefficients to ensure a uniform analysis and identified key peaks in the signal to mark significant changes in pupil diameter. We then applied a universal threshold to filter out noise.

In total, we have 1560 trials across all conditions and participants (120 per participant). We filtered out trials where we found excessive head movement as we were interested in gaze. Excessive head movements are trials where the participant moved their head more than 15 cm higher or lower than their average overall. We also looked at if participants found the object, i.e., looked at the object. For this, we took the maximum angular diameter of the object, i.e., if the object was not perfectly square, we took the largest measurement of the height, width, and length. Using the following formula: $\delta = \arctan\left(\frac{d}{2D}\right)$, we calculated the maximum angular diameter of each sample δ. We then compared the absolute angle of the gaze to the center of the object during the trial and the angle that the object takes up in the virtual environment. If, for the last ten samples in the trial, more than 50% of the absolute gaze angle is lower than 180% of δ, the trial was kept. Otherwise, the trials were not included in the remainder of the analysis. Lastly, we excluded trials that were longer than 20 seconds. After all filtering, we were left with 1356 trials. 178 trials remained without central-field loss simulation, 182 trials with $4.3°$ of central-field loss simulation, 178 trials with $8.6°$, 172 trials with $12.9°$, 168 trials with $17.2°$, 169 trials with $21.5°$, 158 trials with $25.8°$ and 151 trials with $30°$ of central-field loss simulation remaining.

To investigate if there is an effect in the number of trials remaining after the filtering, we used

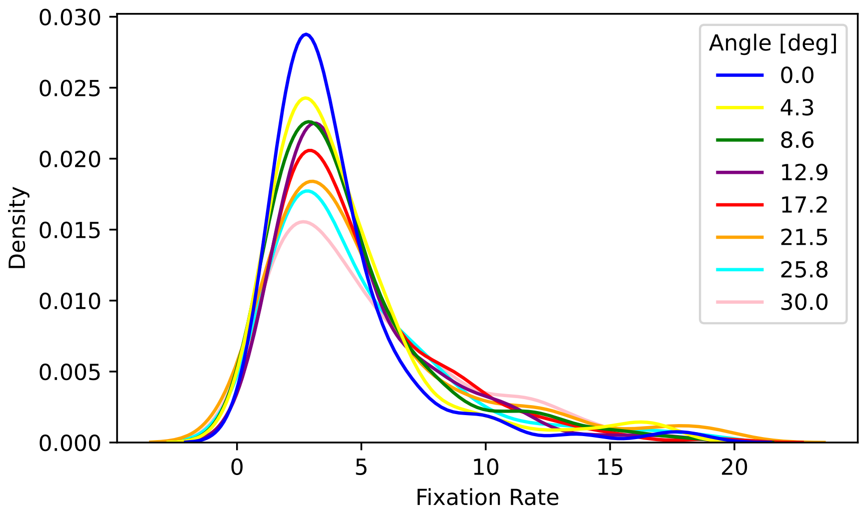


Figure 3.5: Kernel Density Estimate of Time per Trial for each of the different central-field loss simulations.

a Bayes repeated measures ANOVA using JASP with default prior scales, which we performed to estimate the likelihood of an effect. This resulted in a $BF_{10} = 0.169$, thereby providing strong evidence for the absence of an effect.

General Linear Mixed Model Results

We use generalized linear mixed-effect models (GLMMs) to analyze the data from our complex experiment design as it allows us to account for random effects due to individual variability, rooms, objects, spawning locations, or learning effects and is particularly suited for non-normal data and our repeated measures design [41, 229]. For each dependent variable, we include a brief description of each part, including their observed distribution (i.e., corresponding distribution family). These distribution families are selected based on comparing all possible distributions against the data and selecting the best-fitted one. For each dependent variable, our formula was $DV \sim condition + (1|Room) + (1|Object) + (1|Location) + (1|TrialNr) + (1|Pid)$ where DV is our dependent variable. We conducted our analysis in R with lme4 [21].

Our model for time per trial as a function of central-field loss condition exhibits substantial explanatory power, with a conditional $R^2 = 0.70$ and a marginal $R^2 = 0.56$ related to the fixed effects using an inverse Gaussian family with an identity link. The intercept of the model, representing no central-field loss simulation, was estimated at 1.37 ($t = 13.48$, $p < 0.001$). We found that participants took longer to complete their tasks with a higher area of their central vision affected by the central-field loss (see Figure 3.5). The time required for each trial was largest in the case of the $30°$ central-field loss, as confirmed by the GLMM analysis ($t = 2.72$, $p = .007$). We further found significant differences for $25.8°$ ($t = 2.14$, $p = .033$), $21.5°$ ($t = 3.13$, $p = .002$), $17.2°$ ($t = 2.35$, $p = .019$) and $12.9°$ ($t = 1.97$, $p = .049$). We found

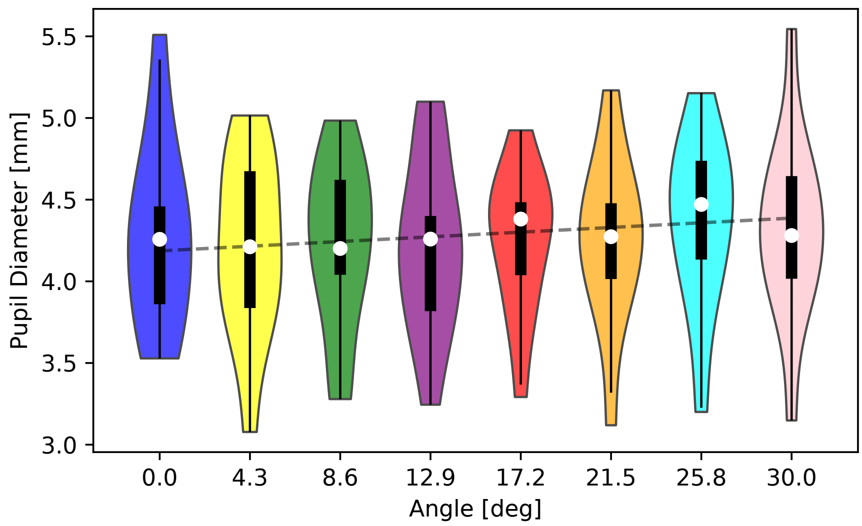


Figure 3.6: Pupil diameter for each of the different central-field loss simulations.

no significant differences in time per trial for the two lowest central-field loss simulations, namely $8.6°$ and $4.3°$.

Pupillometry

Index of Pupillary Activity We did not find significant differences for the index of pupillary activity with a $R^2 = 0.04$ and a marginal $R^2 < 0.01$ related to the fixed effects using an inverse Gaussian family with an identity link. The intercept of the model, representing no central-field loss simulation, was estimated at 0.66 ($t = 16.29$, $p < .001$). We found no significant differences in the index of pupillary activity for all conditions of central-field loss simulations, $30°$ ($t = -0.66$, $p = .509$), $25.8°$, ($t = -0.88$, $p = .379$) $21.5°$ ($t = -1.22$, $p = .223$), $17.2°$ ($t = -1.51$, $p = .131$), $12.9°$ ($t = -2.11$, $p = .055$), $8.6°$ ($t = -0.30$, $p = .767$), and $4.3°$ of central-field loss ($t = -0.15$, $p = .879$).

Pupil Diameter Our model for pupil diameter as a function of central-field loss condition exhibits substantial explanatory power, with a conditional $R^2 = 0.84$ and a marginal $R^2 = 0.54$ related to the fixed effects using a Gamma family with an identity link. The intercept of the model, representing no central-field loss simulation, was estimated at 4.39 ($t = 15.58$, $p < .001$). We found that participants had a larger pupil diameter with a higher area of their central vision affected by the central-field loss (see, Figure 3.6). The largest pupil diameter for each trial was in the case of the $30°$ central-field loss, as confirmed by the GLMM analysis ($t = 3.01$, $p = .003$), we further found significant differences for $25.8°$ ($t = 2.72$, $p = .007$), $21.5°$ ($t = 2.36$, $p = .018$), $17.2°$ ($t = 2.20$, $p = .028$). We found no significant differences in the pupil diameter for three conditions of central-field loss simulations, namely $12.9°$ ($t = 1.63$, $p = .104$), $8.6°$ ($t = 0.71$, $p = .481$) and $4.3°$ ($t = 0.64$, $p = .525$).

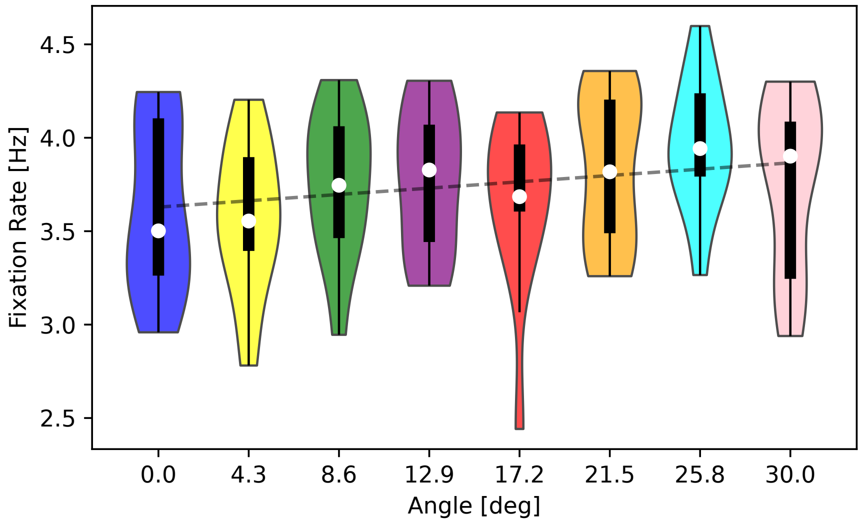


Figure 3.7: Fixation rate for each of the different central-field loss simulations.

Fixations

Fixation Rate Our model for fixations rate as a function of central-field loss condition exhibits substantial explanatory power, with a conditional $R^2 = .69$ and a marginal $R^2 = .44$ related to the fixed effects using a Gamma family with an identity link. The intercept of the model, representing no central-field loss simulation, was estimated at 25.93 ($t = 19.78$, $p < .001$). We found that participants made more fixations with a higher area of their central vision affected by the central-field loss (see, Figure 3.7). The highest fixation rate for each trial was in the case of the $25.8°$ central-field loss, as confirmed by the GLMM analysis ($t = 4.52$, $p < .001$), we further found significant differences for $30°$ ($t = 3.04$, $p = .002$), $17.2°$ ($t = 2.95$, $p = .003$), $12.9°$ ($t = 3.03$, $p = .003$). We found no significant differences in fixation rate for three conditions of central-field loss simulations, namely $21.5°$ ($t = 1.75$, $p = .080$), $17.2°$ ($t = 1.24$, $p = .216$) and $4.3°$ ($t = 0.45$, $p = .650$).

Mean Fixation Duration We did not find significant differences for mean fixation duration with a $R^2 = 0.08$ and a marginal $R^2 < 0.01$ related to the fixed effects using an inverse Gaussian family with an identity link. The intercept of the model, representing no central-field loss simulation, was estimated at 0.16 ($t = 18.77$, $p < .001$). We found no significance between conditions. We found no significant differences in mean fixation duration for all conditions of central-field loss simulations, $30°$ ($t = -1.03$, $p = .301$), $25.8°$, ($t = 0.70$, $p = .485$) $21.5°$ ($t = -1.28$, $p = .200$), $17.2°$ ($t = 0.66$, $p = .508$), $12.9°$ ($t = -0.54$, $p = .591$), $8.6°$ ($t = -0.97$, $p = .335$), and $4.3°$ of central-field loss ($t = -1.10$, $p = .272$).

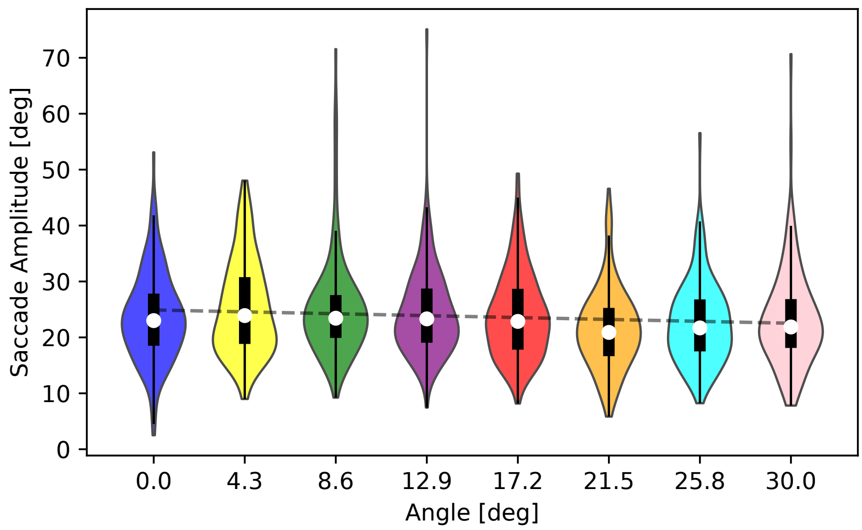


Figure 3.8: Saccade amplitude for each of the different central-field loss simulations.

Saccades

Saccade Amplitude Our model for saccade amplitude as a function of central-field loss condition exhibits substantial explanatory power, with a conditional $R^2 = 0.79$ and a marginal $R^2 = 0.54$ related to the fixed effects using a Gamma family with an identity link. The intercept of the model, representing no central-field loss simulation, was estimated at 25.93 ($t = 13.55$, $p < 0.001$). We found that participants took smaller saccades with a higher area of their central vision affected by the central-field loss (see, Figure 3.8). The smallest saccades for each trial was in the case of the $21.5°$ central-field loss, as confirmed by the GLMM analysis ($t = -4.89$, $p < 0.001$), we further found significant differences for $30°$ ($t = -3.22$, $p = .001$), $25.8°$ ($t = 2.64$, $p = .008$), $17.2°$ ($t = 2.25$, $p = .025$). We found no significant differences in saccade amplitude for the three lowest central-field loss simulations, namely $12.9°$ ($t = -1.83$, $p = .067$), $8.6°$ ($t = 0.02$, $p = .982$) and $4.3°$ ($t = 0.19$, $p = .852$).

Saccade Frequency Our model for saccade frequency as a function of central-field loss condition exhibits substantial explanatory power, with a marginal $R^2 = 0.34$ related to the fixed effects using a Gaussian family with an identity link. The intercept of the model, representing no central-field loss simulation, was estimated at 2.60 ($t = 32.08$, $p < .001$). We found that participants made more saccades with a higher area of their central vision affected by the central-field loss (see, Figure 3.9). The highest saccade frequency for each trial was in the case of the $30°$ central-field loss, as confirmed by the GLMM analysis ($t = -2.31$, $p = .021$). However, this was also the only significant difference we found. We found no significant differences in saccade frequency for the remaining central-field loss simulations, namely, $25.8°$, ($t = -0.80$, $p = .424$) $21.5°$ ($t = 1.42$, $p = .155$), $17.2°$ ($t = -0.78$, $p = .433$), $12.9°$ ($t = 1.62$, $p = .106$), $8.6°$ ($t = 0.68$, $p = .497$), and $4.3°$ of central-field loss ($t = -0.92$,

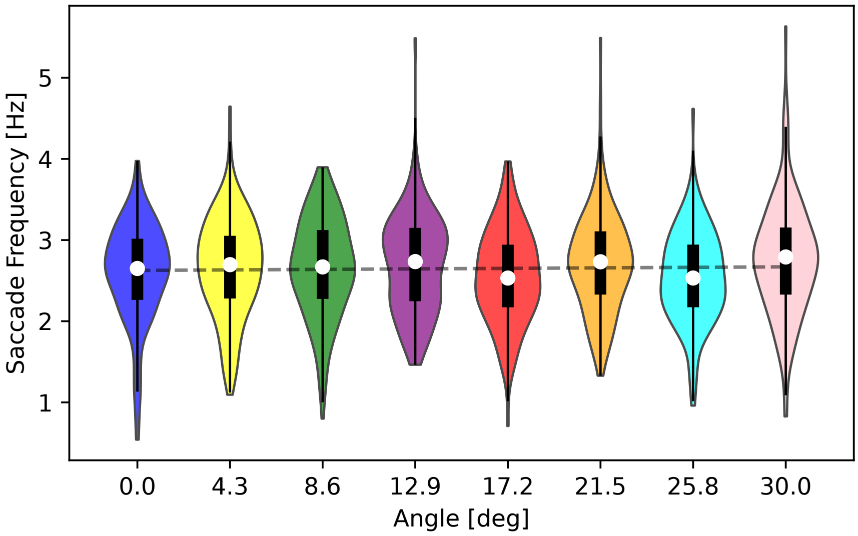


Figure 3.9: Saccade frequency for each of the different central-field loss simulations.

$p = .357$).

3.2 Understanding the Impact of Glaucoma Severities

In this section, we investigate how people with glaucoma perceive their visual environment and whether these experiences are accurately captured by current visualization methods. To this end, we conducted semi-structured interviews with individuals diagnosed with glaucoma, analyzing their subjective reports of vision changes. Drawing from their accounts, we implemented a gaze-contingent simulation of peripheral field loss that aims to more closely reflect lived experience. Finally, we evaluated the behavioral and physiological impact of varying glaucoma severities through a virtual reality study involving a visual search task. By grounding simulation design in real-world perception and quantifying its influence on attention and gaze behavior, we contribute a more realistic and evidence-based model of glaucomatous vision loss.

3.2.1 Structured Interviews

From the simulations mentioned in the related work, we found that peripheral vision loss from glaucoma is mostly visualized with dark visualizations obscuring the periphery with a gradient towards the middle of images. We held semi-structured interviews with individuals who experience glaucoma to verify the current quality of the visualizations. The interviews all followed the same structure, namely, demographic questions and, following this, questions about how they perceive the world. We conducted all interviews over video call or in person.

Procedure

Before the interview, we provided participants with consent forms and a detailed description of the procedure to ensure they had sufficient time to review the materials. At the beginning of each interview, we gave participants the opportunity to ask any questions regarding the consent form and the procedure. The interview commenced with a series of demographic questions pertinent to the study. Once the demographic information was collected, the interview proceeded with questions about their perception of the world (see Supplemental Materials). Examples of guiding questions are: "How does glaucoma impact your everyday life? What are the challenges"? and "Do specific lighting conditions have influence on your perception? And if can you describe how/what changes"?

Participants

In total, we interviewed four individuals who currently have glaucoma. One participant was excluded as they reported having other visual impairments as well, which are more influential to their vision. For the remaining three participants, P1 (female, 68 years old) reported no effect in her central vision and 20–30% remaining peripheral vision. P2 (male, 52 years old) reported 40% remaining vision. P3 (female, 74 years old) reported 30% remaining vision. It is important to note that each of the participants reported having normal vision earlier on in life, which allows them to comment on the loss of vision.

Results

During the interviews, we found that all participants did not experience glaucoma as it usually is visualized, i.e., a dark periphery with a gradient fading towards their central vision. All of our interviewees reported similar experiences. P1 said "The first sign that something was off was when I noticed I was bumping into things more often". P2: "I just noticed things were missing; I wouldn't see them until I basically tripped over them". Furthermore, P3 states, "I can see the road, but I cannot see the bike coming". (If participant statements were not made in English, quotes have been translated by a native speaker with full professional proficiency in English). When we showed them how glaucoma is often visualized, they mentioned the following. P2 stated, "The periphery of my vision isn't just dark; I can't see what's there, but it's not completely black either. It is more like things fade away into nothingness". During the interviews, we talked about and showed different visualizations of the effect, i.e., the visualization used by Stock et al. [244] and Lewis et al. [148].

3.2.2 Glaucoma Simulation

Next, we describe the implementation of our effects pipeline for simulating glaucoma as described in the interviews. For our effects pipeline, we use two parameters, namely, the peripheral field loss and fading around the periphery. It is important to note that we assume the periphery to be obscured evenly and in a circular shape. This is oversimplified as there are indefinite variations of how the peripheral loss exactly manifests.

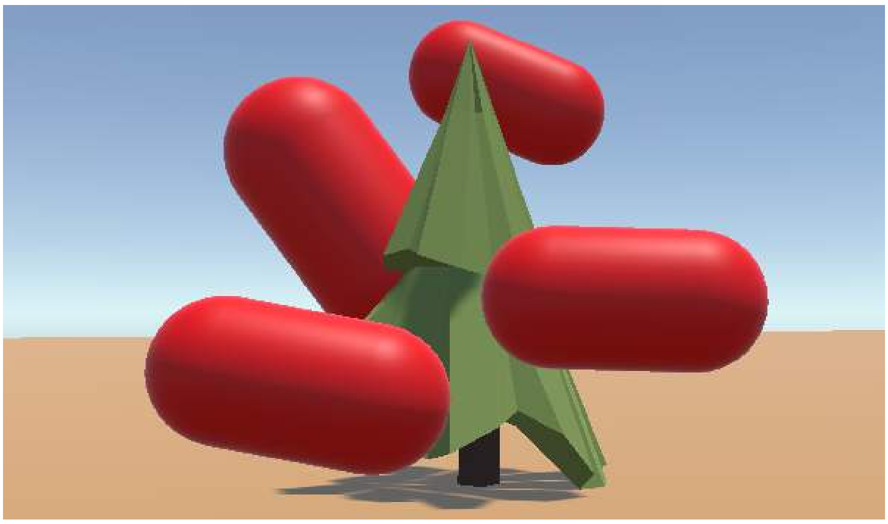


Figure 3.10: A low-poly landscape without any of the Glaucoma visualizations applied.

Peripheral Field Loss

We simulate the peripheral vision loss using a spherical shape of size S (diameter of the sphere) at a fixed distance (d) from the camera for each eye in VR while being gaze-contingent,

$NC_{degrees}$ is the degrees of central vision to be not obstructed by our glaucoma simulation. We determine the size of the sphere using the following the angular diameter formula: $S = 2 \times d \times \tan(NC/2)$. This allows us to manipulate the area of the vision that is not obstructed by our glaucoma simulation.

Fading

$$\texttt{smoothstep}(x) = \begin{cases} 1 & \text{if } x \leq e_0 \\ 1 - \left(3\left(\frac{x-e_0}{e_1-e_0}\right)^2 - 2\left(\frac{x-e_0}{e_1-e_0}\right)^3\right) & \text{if } e_0 < x < e_1 \\ 0 & \text{if } x \geq e_1 \end{cases}$$

To prevent the peripheral field loss from resulting in a hard cutoff, we implemented a fading effect on the inside of the sphere determined by the peripheral field loss as described above. Our fading implementation is an adaptation and modification of the `smoothstep` function (see, `http://www.thebookofshaders.com/` last accessed: April 20, 2026). The modified `smoothstep` function requires three parameters: the lower edge (e_0), the upper edge (e_1), and the source variable for interpolation (x). If x is less than or equal to e_0, it outputs 1; if x is greater than or equal to e_1, it outputs 0; and if x is between e_0 and e_1, it interpolates between e_0 (start edge) and e_1 (end edge) according to the Hermite interpolation. The source variable for interpolation (x) is the normalized distance of the objects' texture fragment concerning the sphere's central position in the screen space. Finally, we use the output of this modified `smoothstep` function as alpha values of the objects' texture to control the objects' visualization.

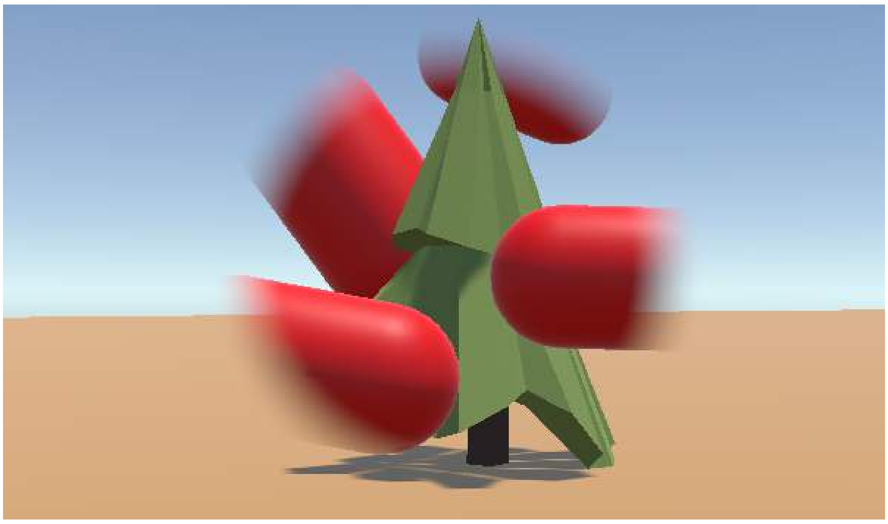


Figure 3.11: A low-poly landscape with the peripheral-field loss and fading visualization applied. In this image, the red capsules have a custom material that hides these items in the peripheral field if they are obscured by our visualizations. This simulation is gaze-contingent.

By applying this combination of the peripheral field loss and fading as a custom shader effect in Unity, we were able to simulate the loss of peripheral vision where objects treated with a custom material disappear when they are not obstructed by the sphere. This, while leaving the rest of the Unity scene unaffected, see Figure 3.11. We chose to have the fading dynamically adjust to the size of the sphere, namely, being 20% of the sphere. If our sphere is $40°$, our fading starts at $40°$ with items treated with the material being completely invisible and ends at $32°$ where items would be completely visible.

3.2.3 User Study

In the following section, we will discuss the user study performed with 26 participants, the virtual environment and hardware used for this user study, and the analysis. We ensured that luminance in all rooms was consistent.

We chose to simulate eight conditions of glaucoma, all Latin Square balanced, $100°$, $90°$, $80°$, $70°$, $60°$, $50°$, $40°$, $30°$ of remaining vision. These were chosen to keep the interval between conditions equal. Furthermore, we decided to not have our simulation affect more than $30°$ of remaining vision to stay outside of the near periphery [245] and on the basis of the reported vision loss of our interviewees, see subsection 3.2.1. Our starting visualization of $100°$ is larger than what a single display shows using the HTC Vive Pro Eye headset, as such there is no visualization during this condition, and will be our control condition.

Apparatus

We use a visual search task to evaluate the effects of the glaucoma simulation on participants' eye movement behavior and task performance. Visual search tasks are commonly used to investigate the impact of visual impairments, both in patients with actual impairments [52, 198] and in simulations of visual impairments [56, 113, 137]. Pollmann et al. [198] demonstrated that patients with visual impairments performed significantly better in search tasks within realistic environments, such as rooms in a home, compared to those in non-realistic scenes, due to the availability of contextual cues. This finding highlight the importance of using realistic environments to accurately assess the deficiencies caused by visual impairments during search tasks. Accordingly, we implemented realistic virtual environments. Following the approach of David et al. [56], we designed eight high-fidelity virtual environments representing different rooms in a household, see Figure 3.3. The search items included common household objects such as a medicine bottle, screwdriver, calculator, and flashlight, among others, see Figure 3.12. The rooms and items were carefully selected to preserve contextual cueing throughout the task.

The virtual environments were developed using Unity and deployed on a desktop equipped with a Ryzen 9 7900X3D processor, 64GB of RAM, and an NVIDIA RTX 4080 Super graphics card. The setup included an HTC Vive Pro Eye headset with an integrated Tobii eye tracker. The Tobii eye tracker, operating at a 120 Hz sampling frequency, was accessed via the Vive SRanipal SDK. For our analysis, head direction, eye-in-world direction, and other 3D directional vectors were converted into 2D Fick angles using the Fick-gimbal method [92]. These 2D Fick

Figure 3.12: Items participants were tasked with searching for, listed from left to right and top to bottom, are eraser, sticky notes, glass, medicine bottle, tape dispenser, hole puncher, duck, apple, first aid kit, screwdriver, pencil sharpener, smartphone, calculator, hairdryer, highlighter, kettle, and, flashlight.

angles for head and eye directions served as the foundation for our subsequent analysis. We concentrated on eye-tracking data for each trial, and given the relatively short trial durations ($M = 4.95$, $SD = 3.63$ seconds), our analysis was performed at the trial level rather than on individual behaviors within each trial.

Procedure

The study began with the coordinator welcoming participants and providing an overview of the study procedure. After participants had signed the consent form, we asked them to fill out a brief demographic questionnaire. Before entering VR, the coordinator explained the necessary hardware and controls to ensure participants were comfortable with the setup. Once participants put on the VR headset, they were guided through eye-tracking calibration. Following this, each participant completed four practice trials without any visual impairment simulation to familiarize themselves with the task.

During the experiment, an item to be located was presented on a gray background at the start of each trial. After the participant confirmed, the item was randomly placed in the virtual room, with adjustments made to avoid collisions with other objects. Participants completed a total of 15 trials per room. We balanced eight different conditions, ranging from no simulated peripheral vision loss to 30 degrees of remaining central vision in equal increments, using a latin square design. After each room, we applied a drift correction following the method of Sipatchin et al. [233], compensating for the known issue of eye-tracker drift caused by participant movement in VR [47]. This correction ensured the accuracy of gaze-contingent simulations of peripheral vision loss throughout the study.

Participants

We recruited 27 participants for the evaluation. However, we had to exclude three participants (one for technical issues and two who did not follow the experimenters' instructions). The remaining 24 participants had a mean age of 23.7 ($SD = 2.6$), 13 identifying as female and 11 as male. Two participants had corrected to normal vision using contacts during the experiment, all others had normal vision.

3.2.4 Results

In this section, we first outline the results of our preprocessing steps. We then present the findings on how varying severities of our glaucoma simulation influenced participants' eye movements and task performance.

Preprocessing

We analyzed fixations and saccades using pymovements [129]. Specifically, we employed its implementation of the ID-T algorithm [217], using a minimum fixation duration threshold of 83 ms and a maximum dispersion threshold of $1.8°$, based on previous studies [22, 260]. This enabled the extraction of key fixation metrics, including total fixation duration, average fixation duration, fixation count, and the time between the start of a trial and the last fixation. For saccades, we utilized pymovements' microsaccade detection algorithm [62], which allowed us to compute saccadic amplitude (the angular distance between saccade onset and offset) and saccade frequency (the number of saccades per trial normalized by trial duration). Pupil dilation was averaged across both eyes, and standard deviations were calculated to facilitate trial-level analysis. To compute the Index of Pupillary Activity (IPA), we followed the method of Duchowski et al. [59].

Across all conditions and participants, 2880 trials were collected (120 per participant). Trials with excessive head movement, defined as head displacement greater than 15 cm above or below the participant's average, were excluded, as our focus was on gaze behavior. To determine whether participants visually engaged with the target object, we calculated the maximum angular diameter using the formula $\delta = \arctan\left(\frac{d}{2D}\right)$, where d is the object size, and D is the distance to the object. We compared the absolute gaze angle to the object's center during the trial, and if over 50% of the final ten gaze samples fell within 180% of δ, the trial was retained. Additionally, trials exceeding 20 seconds in duration were discarded. After applying these filters, 2438 trials remained for analysis. 320 trials in the control condition, 319 with $90°$, 325 with $80°$, 310 with $70°$, 298 with $60°$, 303 with $50°$, 294 with $40°$, and 270 with $30°$ of vision remaining. To examine if the filtering process introduced bias in trial retention across conditions, a Bayesian repeated measures ANOVA was conducted using JASP with default prior scales. The analysis yielded a Bayes Factor of $BF_{10} = 0.031$, providing strong evidence against any condition-related effect on trial retention.

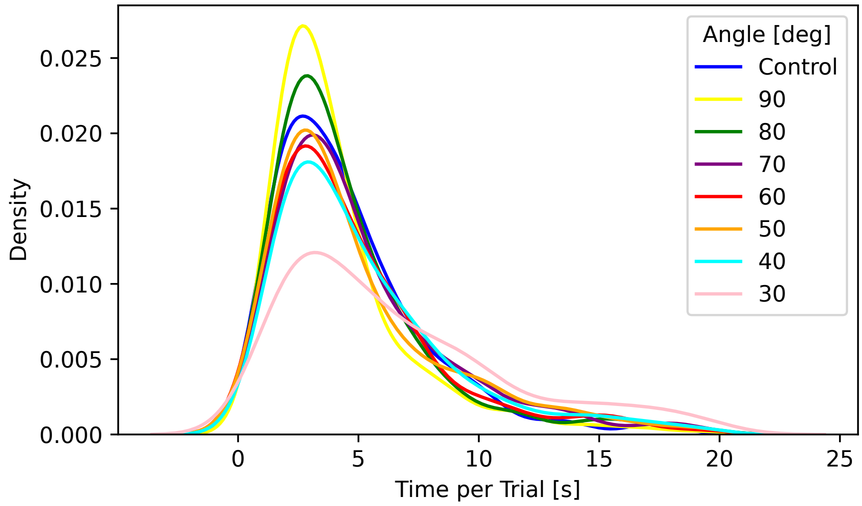


Figure 3.13: Kernel Density Estimate of Time per Trial for each of the different peripheral vision loss simulations.

General Linear Mixed Model Results

To analyze the data from our complex experimental design, we employed generalized linear mixed-effect models (GLMMs), which are well-suited for accounting for random effects due to individual differences, room variability, object placement, spawning locations, and potential learning effects. This approach is particularly effective for handling non-normal data and repeated measures designs [41, 229]. The general formula used for each dependent variable was $DV \sim condition + (1|Room) + (1|Object) + (1|Location) + (1|TrialNr) + (1|Pid)$, where DV represents the dependent variable. We conducted the analysis in R using the lme4 package [21]. We report conditional R^2 as R^2_c and marginal R^2 as R^2_m.

Time Per Trial

Our model for time per trial as a function of peripheral vision loss condition exhibits substantial explanatory power, with a conditional $R^2 = 0.80$ and a marginal $R^2 = 0.50$ related to the fixed effects using an inverse Gaussian family with an identity link. The intercept of the model, representing no peripheral vision loss simulation, was estimated at 4.63 ($t = 13.06$, p $< .001$). We found that participants took longer to complete their tasks with a higher area of their peripheral vision affected by the peripheral vision loss (see, Figure 3.13). The time required for each trial was largest in the case of the $30°$ of vision remaining, as confirmed by the GLMM analysis ($t = 7.93$, p $< .001$), we further found significant differences for $40°$ ($t = 3.13$, p $= 0.002$), $50°$ ($t = 2.38$, p $= 0.018$), and $70°$ ($t = 2.06$, p $= 0.040$). We found no significant differences in time per trial for the three lowest peripheral vision loss simulations, namely $60°$ ($t = 1.47$, p $= 0.141$), $80°$ ($t = 1.00$, p $= 0.316$), and $90°$ ($t = -0.40$, p $= 0.690$) of vision remaining.

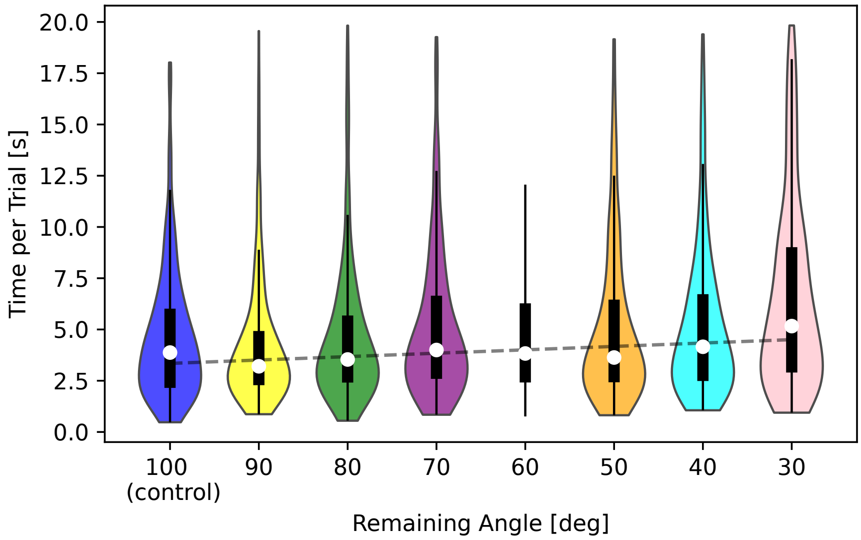


Figure 3.14: Time per Trial for each of the different peripheral vision loss simulations.

Fixations

Fixation Rate We did not find significant differences for fixation rate with a $R^2 = 0.59$ and a marginal $R^2 = 0.18$ related to the fixed effects using a Gamma family with an identity link. The intercept of the model, representing no peripheral vision loss simulation, was estimated at 3.67 ($t = 23.72$, p $< .001$). We found no significance between conditions. 30° ($t = 0.21$, p $= 0.836$), 40° ($t = 1.75$, p $= 0.079$), 50° ($t = 0.10$, p $= 0.918$), 60° ($t = 0.45$, p $= 0.653$), 70° ($t = -0.09$, p $= 0.926$), 80° ($t = 0.05$, p $= 0.963$), and 90° ($t = 0.79$, p $= 0.428$) of vision remaining.

Mean Fixation Duration We did not find significant differences for mean fixation duration with a $R^2 = 0.09$ and a marginal $R^2 < 0.01$ related to the fixed effects using an inverse Gaussian family with an identity link. The intercept of the model, representing no peripheral vision loss simulation, was estimated at 0.16 ($t = 23.03$, p $< .001$). We found no significance between conditions. 30° ($t = 0.25$, p $= 0.804$), 40° ($t = -0.08$, p $= 0.936$), 50° ($t = -0.77$, p $= 0.441$), 60° ($t = 1.35$, p $= 0.176$), 70° ($t = 0.97$, p $= 0.332$), 80° ($t = -0.95$, p $= 0.343$), and 90° ($t = -0.60$, p $= 0.548$) of vision remaining.

Saccades

Saccade Amplitude Our model for saccade amplitude as a function of peripheral vision loss condition exhibits substantial explanatory power, with a conditional $R^2 = 0.92$ and a marginal $R^2 = 0.42$ related to the fixed effects using a Gamma family with an identity link. The intercept of the model, representing no peripheral vision loss simulation, was estimated at 25.74 ($t = 16.24$, p $< .001$). We found that participants made smaller saccades complete

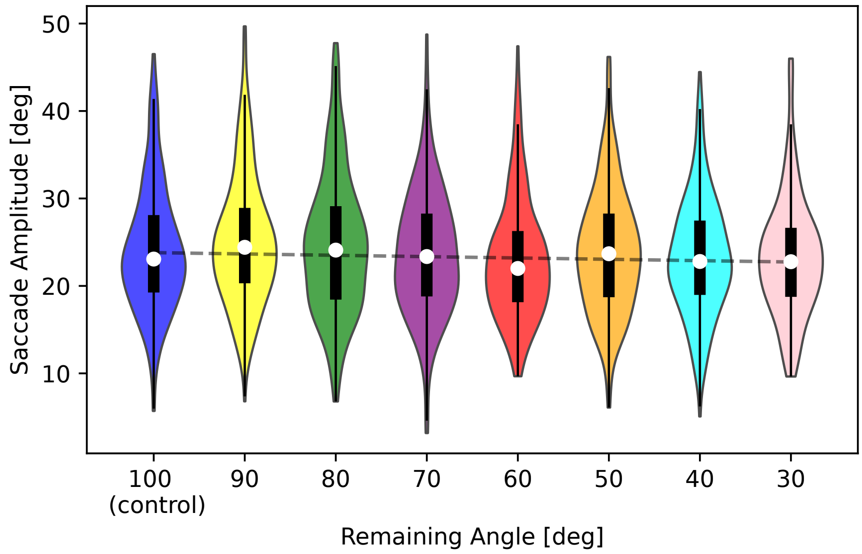


Figure 3.15: Saccade amplitude for each of the different peripheral vision loss simulations.

their tasks with a higher area of their peripheral vision affected by the peripheral vision loss (see, Figure 3.15). The time required for each trial was largest in the case of the $30°$ of vision remaining, as confirmed by the GLMM analysis ($t = -3.09$, p $= 0.002$), we further found significant differences for $40°$ ($t = -2.51$, p $= 0.012$), $60°$ ($t = -2.97$, p $= 0.003$). We found no significant differences in saccade amplitude for four peripheral vision loss simulations, namely $50°$ ($t = -0.76$, p $= 0.450$), $70°$ ($t = -1.33$, p $= 0.182$), $80°$ ($t = 0.45$, p $= 0.652$), and $90°$ ($t = 1.08$, p $= 0.279$) of vision remaining.

Saccade Frequency We did not find significant differences for saccade frequency with a $R^2 = 0.40$ and a marginal $R^2 = 0.17$ related to the fixed effects using a Gamma family with an identity link. The intercept of the model, representing no peripheral vision loss simulation, was estimated at 2.71 ($t = 23.20$, p $< .001$). We found no significance between conditions. $30°$ (-0.07, p $= 0.940$), $40°$ ($t = 0.87$, p $= 0.383$), $50°$ ($t = 0.15$, p $= 0.878$), $60°$ ($t = 0.37$, p $= 0.712$), $70°$ ($t = -0.22$, p $= 0.827$), $80°$ ($t = -0.93$, p $= 0.353$), and $90°$ ($t = -0.94$, p $= 0.348$) of vision remaining.

Pupillometry

Index of Pupilar Activity Our model for the index of pupillary activity as a function of peripheral vision loss condition exhibits moderate explanatory power, with a conditional $R^2 = 0.19$ and a marginal $R^2 = 0.03$ related to the fixed effects using a Gamma family with an identity link. The intercept of the model, representing no peripheral vision loss simulation, was estimated at 0.65 ($t = 17.38$, p $< .001$). We found that participants had a lower index of pupillary activity while completing their tasks with a higher area of their

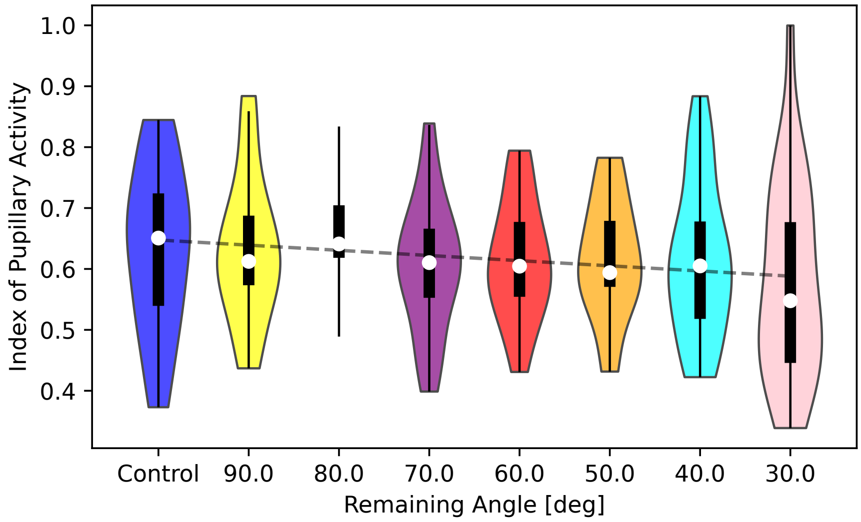


Figure 3.16: Index of pupillary activity for each of the different peripheral vision loss simulations.

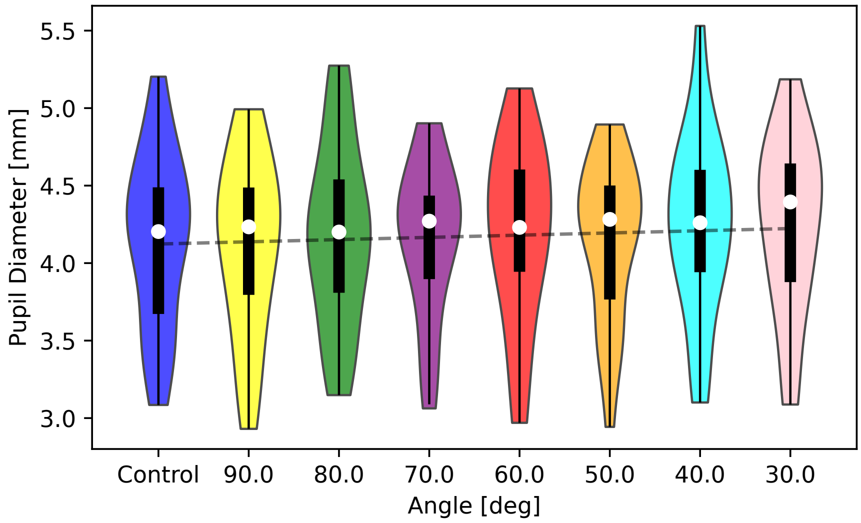


Figure 3.17: Pupil size for each of the different peripheral vision loss simulations.

peripheral vision affected by the peripheral vision loss (see, Figure 3.16). The pupil size during each trial was largest in the case of the $30°$ of vision remaining, as confirmed by the GLMM analysis ($t = -3.32$, p $< .001$), we further found significant differences for $40°$ ($t = 2.39$, p $= 0.017$). We found no significant differences in saccade amplitude for four peripheral vision loss simulations, namely $50°$ ($t = -0.99$, p $= 0.320$), $60°$ ($t = -1.00$, p $= 0.315$), $70°$ ($t = -1.35$, p $= 0.177$), $80°$ ($t = 0.99$, p $= 0.322$), and $90°$ ($t = -0.45$, p $= 0.654$) of vision remaining.

Pupil Size Our model for pupil size as a function of peripheral vision loss condition exhibits substantial explanatory power, with a conditional $R^2 = 0.90$ and a marginal $R^2 = 0.32$ related to the fixed effects using a Gamma family with an identity link. The intercept of the model, representing no peripheral vision loss simulation, was estimated at 4.37 ($t = 20.22$, p $< .001$). We found that participants had larger pupil sizes while completing their tasks with a higher area of their peripheral vision affected by the peripheral vision loss (see, Figure 3.17). The pupil size during each trial was largest in the case of the $30°$ of vision remaining, as confirmed by the GLMM analysis ($t = 4.31$, p $< .001$), we further found significant differences for $40°$ ($t = 3.59$, p $< .001$), $50°$ ($t = -2.97$, p $= 0.003$) and $60°$ ($t = 2.08$, p $= 0.037$). We found no significant differences in saccade amplitude for four peripheral vision loss simulations, namely $70°$ ($t = -0.86$, p $= 0.390$), $80°$ ($t = 0.14$, p $= 0.892$), and $90°$ ($t = -1.56$, p $= 0.119$) of vision remaining.

3.3 Effects of Cataracts Severities on Eye Movements

Cataracts gradually impair vision, yet their subtle early-stage effects are often overlooked by both individuals and standard diagnostics. In this chapter, we investigate how different severities of simulated cataract progression influence eye movements and task performance in virtual environments. Building on a validated simulation pipeline informed by patient interviews, we examine both subtle monocular and pronounced binocular impairments over varying durations. Through controlled visual search tasks, we assess behavioral and physiological changes, aiming to understand when impairments become noticeable and how eye tracking might serve as a foundation for early, gaze-based cataract detection systems.

3.3.1 Cataracts Simulation

For our simulation, we re-implemented the effects pipeline by Krösl et al. [133] in Unity using the five parameters: reduced VA, contrast loss, color shift, loss of peripheral vision, and sensitivity to light.

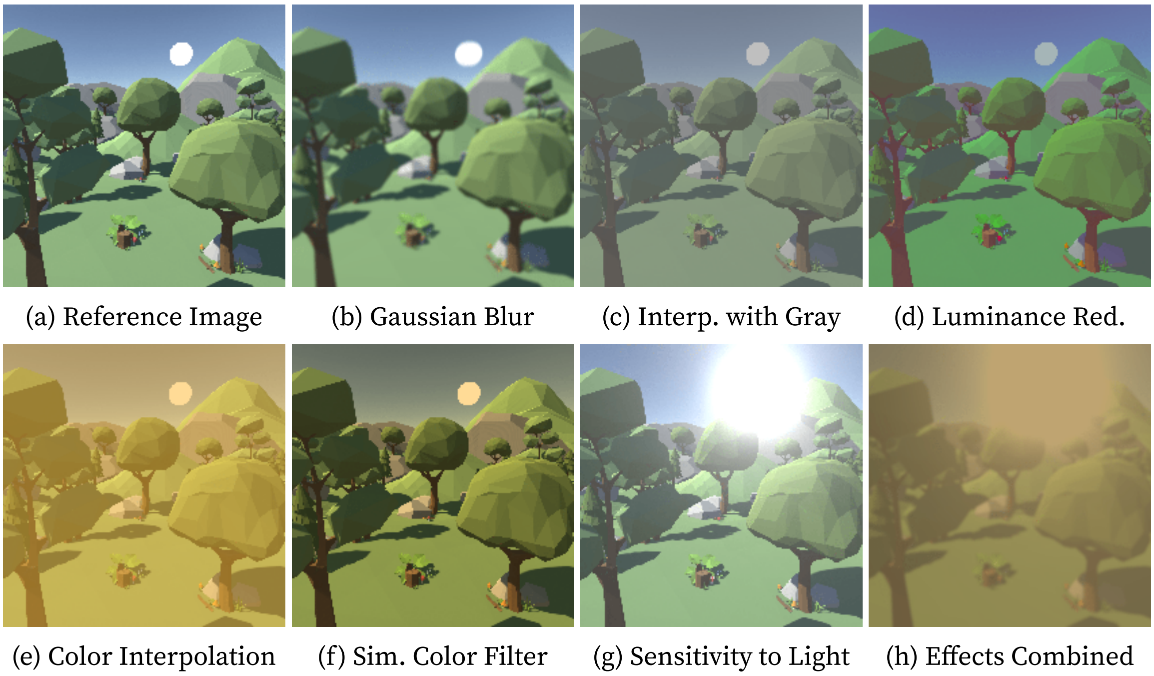


Figure 3.18: An example image of a low-poly landscape with the implemented effects applied. **(a)** No visualizations applied. **(b)** Only the loss of visual acuity filter is applied [133].

3.3.2 Contrast Loss

The work from Krösl et al. [133] already implemented a reduced contrast using gray interpolation. Additionally, we re-implemented the compression of the luminance value from

Krösl et al. [132]; allowing for more control in simulating the mechanism of the human eye since it is implemented in the *CEILAB* color space (also referred to as *L*a*b**). This is because adjusting RGB color channels linearly does not produce linear changes in contrast [132]. To reduce the contrast by compressing the luminance, the RGB value of each pixel needs to be transformed into the *CEILAB* color space first. After this, we can apply the following formula $L = L_1 \times l + 50 \times (1 - l)$ where L represents the new luminance value, L_1 the initial luminance, and $l \in [0, 1], \quad l \in \mathbb{R}$ as a scaling factor. To preserve the average lightness $50 \times (1 - l)$ is added [132]. We visualize this in Figure 3.18d.

3.3.3 Color Shift

While Krösl et al. [133] implemented a color filter, we additionally re-implemented the interpolation by Krösl et al. [132]. We implemented this option as it seems to be more representative of online visualizations used to illustrate cataracts. Hence, we implemented both to be evaluated in interviews later.

3.3.4 Sensitivity to Light

As Krösl et al. [133] used Unreal Engine, we could not use the same bloom post-processing. Instead, we used the Bloom effect provided by Unity as it is a similar post-processing effect. This processing effect allows us to manipulate the light's intensity, threshold, and diffusion. These allow us to simulate the effect of images blurring, with bright lights proving especially problematic due to the intense glare they produce [132]. In Figure 3.18g, we visualize this part of our effects pipeline.

3.3.5 Loss of Peripheral Vision

Lastly, cataracts result in lens clouding. While for nuclear cataracts, this clouding spans the entire lens uniformly, subcapsular cataracts are often visualized as a dark shadow in the center of the lens, and cortical cataracts are often visualized as a dark shadow in the periphery of the lens. We simulate this loss of peripheral vision using a gaze-contingent stencil shader implementation that hides items with a specific material. This results in hidden items outside the area (i.e., in the peripheral). We chose this approach as we found during pilot testing that participants always perceived these dark shadows simulations that other work had used [132, 133], even when these changes were minimal. This is the case as current hardware cannot keep up with human perception regarding refresh rate and latency, thus making it unsuitable for our study. Figure 3.20 visualized the loss of peripheral vision from our pipeline.

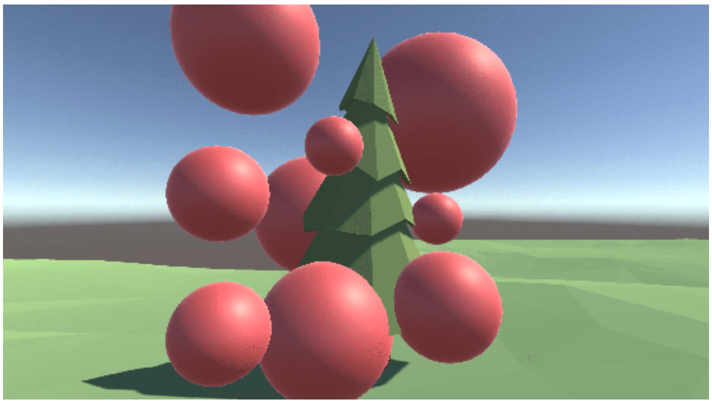


Figure 3.19: A low-poly landscape without any of the cataract parameters applied (Normal Vision).

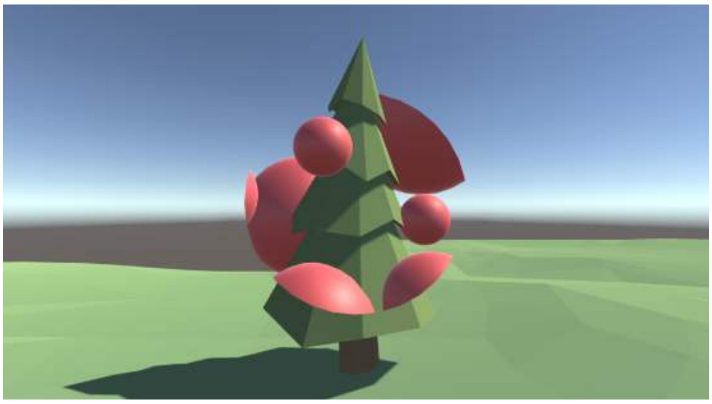


Figure 3.20: A low-poly landscape with the peripheral loss visualization applied. In this image, the red balls have a special invisible material that hides these items in the periphery. This simulation is gaze-contingent.

3.3.6 Cataracts Simulation Validation

In an effort to validate our implementation of the different cataract parameters and the progression of cataracts, we held semi-structured interviews with individuals who either currently have cataracts in one eye or recently had cataract surgery. The interviews all consisted of two parts: demographic questions and, following this, an iterative parameter adjustment. All were done over video call or in person on a normal screen.

Procedure

We prepared a sample scene with a low-poly landscape for the iterative parameter adjustment. In this scene, we implemented parameters for VA reduction, contrast loss, color shift, and sensitivity to light. When two approaches were available, we implemented both, all according to subsection 3.3.1. Before the interview, we sent out the consent forms and a copy of the procedure so the participants had ample time to read through them before participating in the interview. One interview took place in person, and the remaining was over video call. At the start of the interview, we asked if the participant had any questions about the consent form procedure. After this, we started off with a series of demographic questions relevant to the study at hand. After the demographic questions, we came to the iterative parameter adjustment. During the iterative parameter adjustment, we inquired about the different implementations for contrast loss and color shift, which are more capable of recreating their visual perception. We did not represent the loss of peripheral or central vision using the stencil shader approach; instead, we applied a vignette or inverted vignette effect, as the stencil shader effect is not "visible" and therefore, participants were unable to indicate the size accurately. During the iterative parameter adjustment, we used the sliders to adjust the parameters to recreate the participants' visual impression of their cataracts. First, one by one, after which all parameters are combined, as some affect each other.

Participants

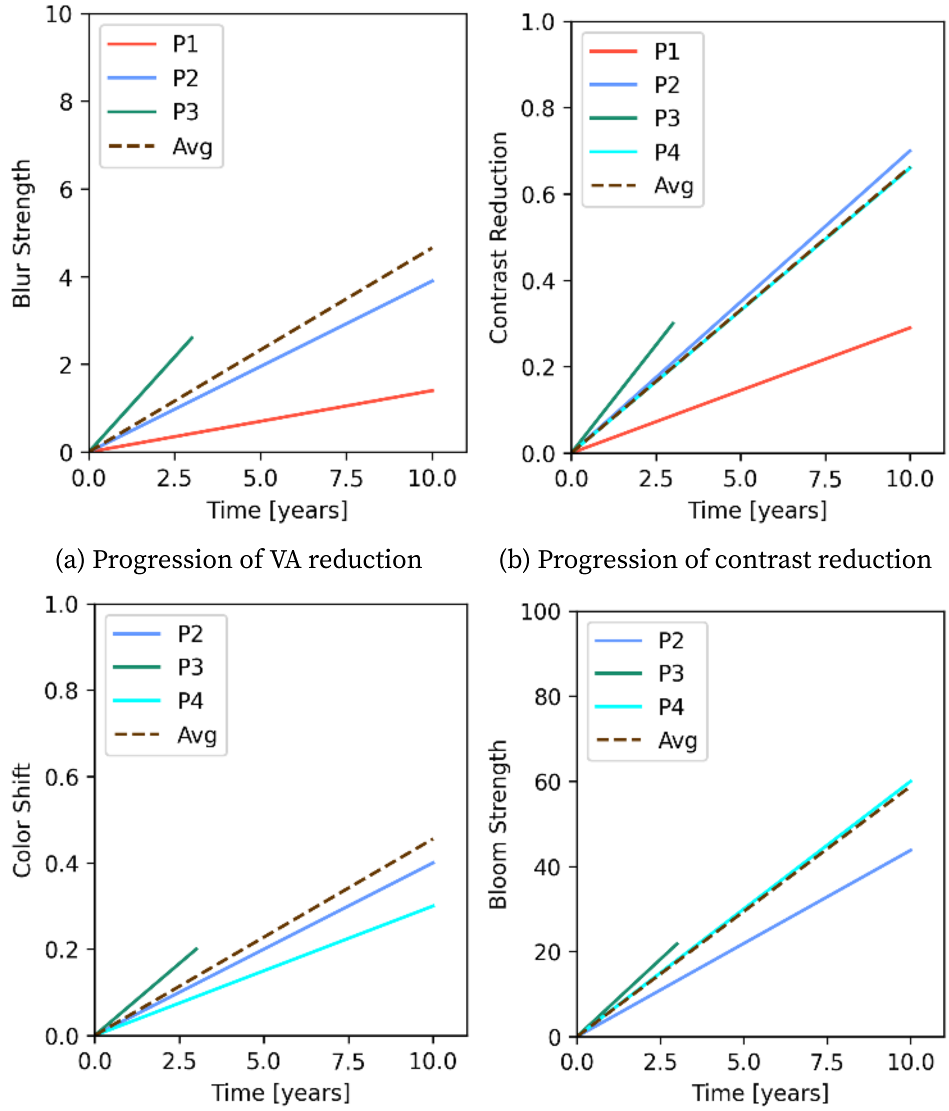
In total, we interviewed six individuals who either currently have cataracts in one eye or recently have had cataract surgery. Two participants were excluded as one reported to have far progressed age-related macular degeneration, which is more influential to their vision, and one participant was excluded as they reported the progression of cataracts to happen within the span of one month, which contradicts existing literature [158]. For the remaining four participants, P1 (female, 82 years old) reported -2 dioptri in the right eye, where the last cataract surgery was in 2014. P2 (male, 74 years old) was interviewed after recent cataract surgery; during that interview, there was no reported loss of vision. P3 (female, 56 years old) reported < 5% vision remaining in her right eye, where she had cataract surgery last performed in 2008. However, her vision was affected by several other vision impairments. P4 (male, 68 years old), interviewed after recent cataract surgery, reported +2.5 dioptri in both eyes. Unfortunately, none of the participants reported the exact type of cataracts.

Results

During the interviews, we found that all participants reported that the interpolation with gray (see Figure 3.18c) is more representative of the loss of contrast and that the color interpolation (see Figure 3.18e) is perceived to be more representative of the color shift. All participants went through the iterative parameter adjustments until they found that this represented their perception (either in the cataracts-affected eye or while it was affected).

We visualize their reported cataracts progression over time in Figure 3.21. For every parameter respectively, we do not plot those where the parameters were not present. Furthermore, we visualize all parameters' average progression over time using the dashed lines. By doing so, we assume a linear growth of each parameter, as we could not collect data about progression over the span of multiple years[3]. None of our participants report any loss of peripheral vision.

[3]We acknowledge that this might oversimplify the progression of cataracts.



(a) Progression of VA reduction
(b) Progression of contrast reduction
(c) Progression of color shift
(d) Progression of light sensitivity

Figure 3.21: The progression of the different parameters over the years. Each line represents the progression for one interviewee, where the colors are consistently representing the same participant throughout the graphs. Where interviewees did not report a progression of one parameter, this is not represented in the data. The dashed line represents the average linear progression over 10 years.

3.3.7 Minimal Noticeable Differences

While cataracts traditionally present themselves in both eyes, our goal in this study is to visualize minimal progression to (1) investigate if participants notice the changes and (2) investigate if these changes affect eye movements and task performance before the participants become aware of the presence of the simulation. As such, in the following section, we showed the simulation exclusively in the non-dominant eye, as participants are less likely to notice the change [208]. In this and the following study, we chose to simulate cataracts rather than recruiting participants already affected by cataracts for three reasons. (1) Those affected by cataracts are sensitive to light, in VR this would mean exposing them to the LUX output of the headset, which can be uncomfortable for them. Furthermore, (2) simulations allow us to maintain precise control over the progression and severity of visual impairment. Simulations allow us to systematically vary the degree of cataract opacity and closely monitor subtle changes in participants' visual experiences, which would be challenging with individuals naturally affected due to variability in progression and subjective experiences. Finally, (3) simulations enable us to isolate and specifically investigate the early stages of cataract development and their initial impact on visual perception and behavior, offering clearer insights into early detection and management strategies.

Apparatus

We selected a visual search task to assess the impact of the cataract simulation on participants' behavior and performance. Visual search tasks have been employed widely to investigate visual impairments both on patients with visual impairments [52, 198] as well as in simulations of visual impairments [56, 113, 137]. Pollmann et al. [198] showed that participants with Age-Related Macular Degeneration are significantly less deficient in search tasks compared to healthy participants in realistic scenes (such as rooms in a home) due to the possibility of contextual cueing. This shows that displaying realistic scenes is necessary to adequately assess the deficiencies of visual impairment in search tasks. Therefore, we opted to show realistic environments of different rooms in the virtual environment. Similar to David et al. [56], we built eight virtual environments to represent different rooms in a house; see Figure 3.3. We ensured that luminance in all rooms was consistent, see **??**.

For the search items, we used items that could be found in a household, namely, an apple, calculator, rubber duck, first aid kit, flashlight, glass, hairdryer, kettle, medicine bottle, phone, screwdriver, soap bar, or stapler. All rooms and items were chosen to preserve contextual cueing.

We built the virtual environments in Unity3D and deployed it to a desktop with a Ryzan 9 7900X3D, 64GB RAM, an NVIDIA RTX 4080 Super graphics card, and a connected HTC Vive Pro Eye with a built-in Tobii eye tracker. The HMD has two AMOLED screens, with a resolution of 2.880×1.600 pixels in total ($1.440.600$ pixels to each eye, resulting in a pixel density of 615 pixels per inch (PPI)), with a native refresh rate of 90 Hz and a field of view of $110°$. The built-in Tobii eye tracker was accessed using the Vive SRanipal SDK at a sampling frequency of 120 Hz.

We collected the head position and directional 3D vectors from Unity. We recorded eye tracking data using the built-in eye tracker in the HTC Vive Pro Eye (120Hz) through the SRanipal SDK, giving us eye-directional 3D vectors relative to both the world and the head. To facilitate our analysis, we converted the head directional 3D vectors, eye-in-world directional 3D vectors, and directional 3D vectors into 2D Fick angles using the Fick-gimbal method [92].

This transformation applies two sequential rotations (vertical axis followed by a nested horizontal axis) to characterize the position of each vector. These resulting 2D Fick angles for eye and head directions served as the foundation for our subsequent analysis. For the analysis, we focused on the eye tracking data for each trial. Our trials were relatively brief, with a mean duration of 5.3 ± 8.89 seconds, leading us to perform eye tracking analysis at the trial level instead of examining behaviors within individual trials.

We calculated fixations and saccades using pymovements [129], an open-source Python package for analyzing eye-tracking data. We leveraged pymovements' implementation of the ID-T algorithm [217] with specific fixations thresholds set at a minimum fixation duration of 83 ms and a maximum dispersion of $1.8°$, in line with prior work [22, 0, 260]. From this, we derived fixation metrics including: total fixation duration, average fixation duration, fixation count, and latency to the final fixation in each trial. For saccades, we again leverage pymovements, this time for its implementation of the microsaccade algorithm [62]. This algorithm allows us to calculate the saccadic amplitude and saccade frequency. We defined saccadic amplitude as the angular displacement in degrees from onset to offset, and computed saccadic frequency by dividing saccade count per trial by trial duration in minutes. To calculate the average pupil dilation for each trial, we computed the mean of the left and right pupil sizes and subsequently calculated the standard deviation to facilitate our analysis at the trial level. Finally, for computing the Index of Pupillary Activity (IPA), we employed the implementation by Duchowski et al. [59]. Thus, we utilized discrete wavelet transforms to analyze pupil diameter signals, specifically employing a two-level discrete wavelet transform. We normalized the wavelet coefficients to ensure a uniform analysis and identified key extremes in the signal to mark significant changes in pupil diameter. We then applied a universal threshold to filter out noise.

Procedure

The study coordinator first welcomed the participants and, after signing the consent form, asked them to answer a few demographic questions. The coordinator then introduced and explained the study procedure. Namely, they were tasked with searching for household items in 8 rooms. Additionally, the study coordinator told the participant that we would manipulate something throughout the study, and if the participant noticed, they should report it. Afterward, the coordinator and participant determined the dominant eye using the Dolman method [43]. Finally, before entering VR, the study coordinator explained the hardware and controls relevant to the study to the participants.

After putting on the HTC Vive Pro Eye, the study coordinator navigated the participant to the integrated eye-tracking calibration. After which, the experiment was started. For each trial, we present the item the participant was tasked with finding on a gray background. Once the participant confirms, the item spawns at a random location in the room. These spawning locations were randomly assigned throughout the room, and potential collisions with other items were checked. If there was a collision, a different spawning location was selected. This was then repeated for 15 items in each room. We increased the severity of the simulation linearly after each trial, except for when the participant moved from one room to the next for the participant in the *linear increase* condition. For the control condition, there was no change. In the linear increase condition, the participants all started without any cataract simulation and ended with a simulation severity of 1 year as determined in the interviews (see section 3.3.6). As the participants were made aware that we would manipulate something throughout the study, they were asked to report anything they noticed that was off from how they expected it to be. If they did not actively report to the study coordinator before the end of the study, the study coordinator asked them specifically if they noticed something odd.

We applied a drift correction after each room, following Sipatchin et al. [233]. This drift correction was designed to compensate for eye tracking data quality decay in VR due to the participant's movement, which is known to induce drift into the precision of the eye tracker [47]. These drifts can influence the gaze-contingent simulation of the peripheral loss.

Participants

We conducted our study with 33 participants, ages 20 to 71 (mean = 27.6, std = 10), with 14 identifying as male, 19 as female, and none as non-conforming. Six participants wore contacts during the experiment. The remaining participants had normal vision, with 19 being rightly dominant-eyed, and 15 had a left dominant eye, as determined by the Dolman method. 20 participants were part of the *linear increase*, while the remaining participants were part of the control condition. One participant in the *linear increase* condition was excluded as their eye tracking data was incomplete. While we recognize the imbalance between the two conditions, our analysis methods are robust against this imbalance [197]

Results

In this section, we first present the findings of how our simulation affects eye movements, among others, while being influenced by the cataract simulation. We employ the Bayesian Independent Samples T-Test to compare the slopes of fitted linear regressions over the severity.

Preprocessing

We had a total of 3840 trials (120 trials per participant). Following this, we filtered out trials where the participants had excessive head movements as we were interested in gaze. We define excessive head movements as trials where the participant moves their head 15 cm higher or lower than the overall average across all trials. We also filtered out trials where the participants did not look at the object they were supposed to search for. To do this, we took the maximum angular diameter of the object, i.e., if the object was not perfectly square, we took the largest measurement of the height, width, or length. Using the following formula: $\delta = arctan\left(\frac{d}{2D}\right)$, we calculated the maximum angular diameter of each sample δ. Then, we compared the absolute angle of the gaze to the center of the object during the trials and the angle at which the object took us in the virtual environment. If, in the last ten samples of a trial, more than 50% of the absolute gaze angle was higher than 180% of δ, trials were discarded. Lastly, we excluded trials that were more than 20 seconds. After preprocessing, we had a total of 3616 trials remaining.

Bayesian Results

We fit a linear regression for each participant over the different extracted eye-tracking characteristics to compare the two conditions (linear increase of simulated cataract severity and control). As we expect that there will be a learning effect throughout the user study and per room, we compared the slopes of these regressions using Bayesian independent sample t-tests. Thus, we conducted a Bayesian independent sample t-test using JASP. Priors for all tests were set to be Cauchy (location = 0, scale = 0.707), with our alternative hypothesis being that the slopes from linear change and control are not equal. We found $BF_{10} = 0.103$, $error\ \% = 0.166$ when including all the different extracted eye-tracking characteristics. This indicates strong evidence that the slopes of the linear change and control are equal. We further explored the data by looking at individual eye-tracking characteristics. Our results show anecdotal to moderate evidence that the slopes of the linear change and control are equal (H_0). However, when exclusively looking at the fixation slope and fixation rate, we find anecdotal evidence for the sloped being different (H_1). We report the results of all our tests in Table 3.1 and Table 3.2. We checked if normalizing the participants' variance showed differences compared to the slopes reported; however, running the same test with the z-score normalization increases the evidence for the data being from the same population.

Table 3.1: Bayesian Independent Samples T-Test (Time, Pupillometry, and Saccade Measures)

				Saccades	
	Time	IPA	PD	Amplitude	Freq
BF_{10}	0.345	0.350	0.420	0.346	0.338
error %	0.009	0.003	0.010	0.009	0.007

Out of the 19 participants included in the final results with the linear increase, we found that 60% of these participants (12) were unaware of the simulated visual impairment at the end of the user study. This means they were unaware of the progression that normally would happen over a year in 30 minutes. Of those who did notice the simulated cataracts, 6 participants noticed it in the second half of the study. All of those who noticed described the loss of VA as noticeable. At the same time, none of the participants reported having seen any of the other effects, even when explicitly asked for it at the end of the user study.

Table 3.2: Bayesian Independent Samples T-Test (Fixation Measures)

	Count	Disp. Area	Avg. Dur	Std. Dur	Tot. Dur	Var. Dur	Rate
BF_{10}	0.407	0.671	0.400	0.375	0.368	0.552	1.014
error %	0.011	$<.001$	0.012	0.013	0.013	0.017	$<.001$

Interim Discussion

We found that most participants were unaware of the simulation of cataracts. Given how we have developed the virtual environments, this articulated the established finding that contextual cueing influences the ability of those with simulated visual impairment to find objects [198]. This echoes the importance of implementing realistic scenes and items to investigate simulated visual impairments. Furthermore, we found that 60% of our participants were completely unaware of the simulation representing the progression of one year of cataracts in the study of 30 minutes. Since participants were unaware of the progression, we expect that, since the progression happens much slower, individuals who experience cataracts will be unlikely to notice it themselves, even after several years of progression. This is corroborated by our interview findings, in which one of our participants mentioned that they were unaware of their cataracts; however, the doctor diagnosed < 5% of residual vision. This expresses the need for detection methods for visual impairments, as when the individual notices the visual impairment, it might already be in a far progressed state.

When analyzing the eye tracking characteristics, we found no differences between those who experienced a linear increase in severity and those in the control condition. More precisely, we found strong evidence that these are equal, which could be attributed to the learning effect of gradually increasing the simulation. While we manipulated the severity of the cataract simulation over one year, we did not investigate further progression or binocular cataract simulation. As such, in the second user study, we will increase the severity and change the paradigm to binocular.

3.3.8 User Study

In subsection 3.3.7, we saw that the potential learning effect outways the visualization of cataracts when gradually increasing the simulation in the non-dominant eye. In this follow-up, we simulate up to 36 months of progression in 8 conditions using a Latin square balanced experimental paradigm. Varying from no simulation up to 36 months of progression in equal increments following the results of the section 3.3.6. In the remainder of this section, we discuss the user study performed with 32 participants. The apparatus and procedure are the same as subsection 3.3.7.

Participants

We recruited 34 participants for the evaluation. However, we had to exclude two participants (one for technical issues and one who did not follow the instructions). The remaining 32 participants, ages 19 to 68, had a mean age of 27.3 ($SD = 10.2$), 16 identifying as female, and 16 identifying as male. All participants had normal or corrected-to-normal vision through contact lenses. Finally, none of the participants participated in the first user study.

3.3.9 Results

In this section, we first present the results of our preprocessing. After which, we present the findings of how our simulation affects eye movements and task performance while being affected by different severities of our cataract simulation.

Preprocessing

After extracting the relevant eye tracking characteristics following section 3.3.7. After all filtering, we were left with 3149 trials. 402 trials for no visualization, 406 for 5.1 months of progression, 400 for 10.3 months or progression, 410 for 15.4 months of progression, 371 for 20.6 months of progression, 395 trials for 25.7, 401 trials for 30.9, and 364 for 36 months of progression. To investigate if there is an effect in the remaining number of trials after preprocessing, we used a Bayes repeated measure ANOVA in JASP with default prior scales to estimate the likelihood of an effect. This resulted in a $BF_{10} = 0.05$, thereby providing strong evidence for the absence of an effect.

General Linear Mixed Model Results

We use generalized linear mixed-effect models (GLMMs) to analyze the data from our experiment allowing us to account for the complex experiment design by using random effects for individual variability, rooms, objects, spawning locations, or learning effects and is particularly suited for our data as it is non-normally distributed and allows for a repeated measures design [41, 229]. For each dependent variable, we include a brief description of each part, including their observed distribution (i.e., corresponding distribution family). These distribution families are selected based on comparing all possible distributions against the data and selecting the best-fitted one. For each dependent variable, our formula was $DV \sim condition + (1|Room) + (1|Object) + (1|Location) + (1|TrialNr) + (1|Pid)$ where DV is our dependent variable and condition is the 8 different months of cataract progression. We conducted our analysis in R with lme4 [21]. We report conditional R^2 as R^2_c and marginal R^2 as R^2_m.

Time per trial

Our model for time per trial as a function of cataract progression exhibits substantial explanatory power, with a conditional $R^2 = 0.72$ and a marginal $R^2 = 0.50$, using an inverse Gaussian family with an identity link. The intercept of the model, representing no simulated cataract progression, was estimated at 4.41 ($t = 10.54$, $p < .001$). We found that time per trial increased significantly as cataract progression advanced over time. At 5.1 months, the change in time per trial was not significant ($t = 1.46$, $p = 0.143$). However, by 10.3 months, a significant increase was observed ($t = 2.29$, $p = 0.022$), suggesting a growing impact on task performance. A similar pattern emerged at 15.4 months ($t = 1.98$, $p = 0.048$), confirming a gradual decline in efficiency. The most pronounced effects appeared at 20.6 months ($t = 5.08$,

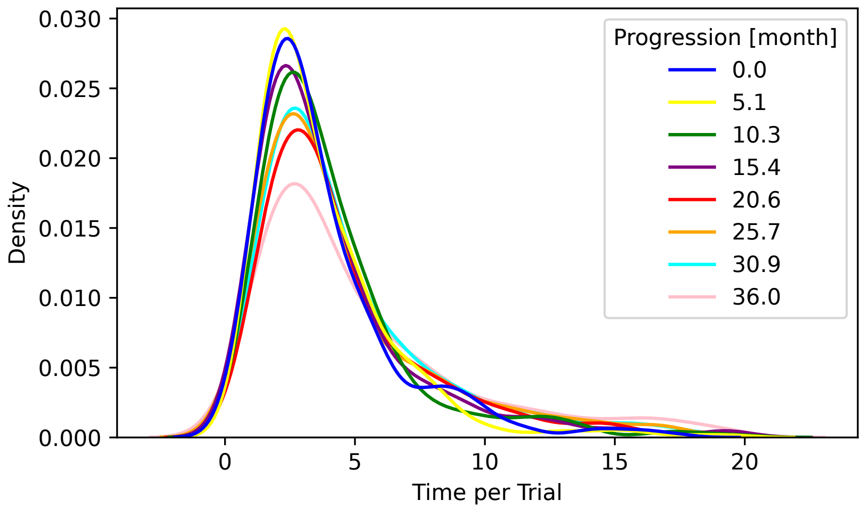


Figure 3.22: Kernel Density Estimate of time per trial for each of the different cataract simulations.

$p < .001$), 25.7 months ($t = 2.62$, $p = 0.009$), and 30.9 months ($t = 4.34$, $p < .001$), demonstrating a significant reduction in performance speed as the simulated cataract worsened. By 36.0 months, the time per trial had increased substantially, with the strongest observed effect ($t = 6.00$, $p < .001$), reinforcing the detrimental impact of extended cataract progression on task completion times, see Figure 3.22.

Mean Fixation Duration

Our model for mean fixation duration as a function of cataract progression exhibits limited explanatory power, with a conditional $R^2 = 0.11$ and a marginal $R^2 = 0.02$, using an inverse Gaussian family with an identity link. The intercept of the model, representing no simulated cataract progression, was estimated at 0.15 ($t = 18.63$, $p < .001$). We found no significant changes in mean fixation duration for most of the cataract progression conditions. At 5.1 months, the change in mean fixation duration was not significant ($t = 1.25$, $p = 0.265$). Similarly, at 10.3 months, there was no significant effect observed ($t = 1.16$, $p = 0.450$). At 15.4 months, the effect remained non-significant ($t = 1.40$, $p = 0.140$), indicating that early cataract progression does not strongly impact fixation duration. Even as cataract severity increased, the changes in mean fixation duration did not reach statistical significance at 20.6 months ($t = 2.11$, $p = 0.135$), 25.7 months ($t = 2.15$, $p = 0.132$), or 30.9 months ($t = 2.20$, $p = 0.128$). By 36.0 months, the effect remained non-significant ($t = 1.28$, $p = 0.201$), reinforcing that mean fixation duration does not appear to be substantially influenced by the simulated progression of cataracts in our study.

Fixation Rate

Our model for fixation rate as a function of cataract progression exhibits moderate explanatory power, with a conditional $R^2 = 0.59$ and a marginal $R^2 = 0.30$, using an inverse Gamma

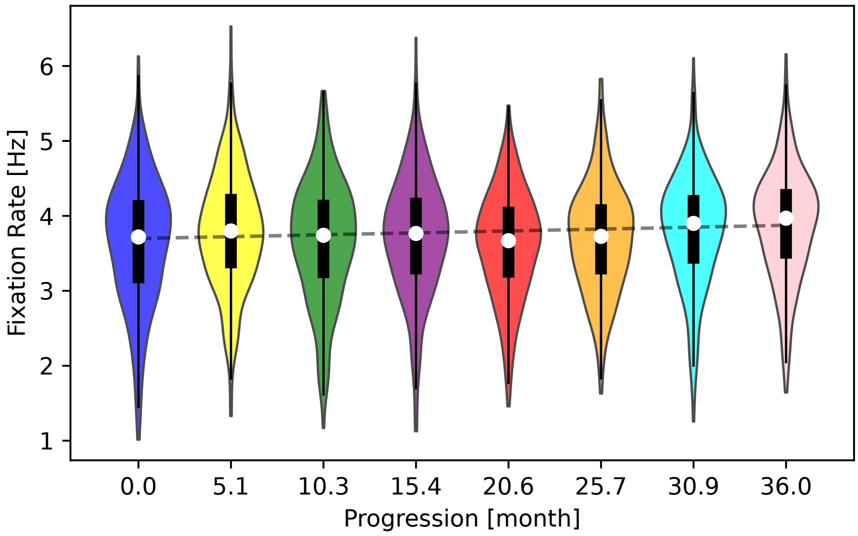


Figure 3.23: Fixation rate for each of the different cataract simulations.

family with an identity link. The intercept of the model, representing no simulated cataract progression, was estimated at 3.65 ($t = 28.08$, $p < .001$). At 5.1 months, the fixation rate showed a significant increase ($t = 3.52$, $p < .001$). However, at 10.3 months, the change was not significant ($t = 1.14$, $p = 0.254$). At 15.4 months, a significant increase was observed ($t = 2.06$, $p = 0.039$). At 20.6 months, the effect was not significant ($t = 0.52$, $p = 0.602$), but a significant increase reappeared at 25.7 months ($t = 2.58$, $p = 0.010$). At 30.9 months, fixation rate increased significantly ($t = 3.86$, $p < .001$), with the strongest observed effect at 36.0 months ($t = 5.05$, $p < .001$). These results suggest that fixation rate is influenced by cataract progression, with significant increases at several time points, particularly in later stages, see Figure 3.23.

Table 3.3: Results of the GLMM analysis for the time per trial, fixation rate and mean fixation duration under different central-field loss conditions. We used the inverse gaussian family and identity link for the time per trial and mean fixation duration. For the fixation rate, we used an inverse Gamma family with an identity link.

	Time Per Trial			Mean Fix. Duration			Fix. Rate		
Progression (month)	**Est.**	**t**	**p**	**Est.**	**t**	**p**	**Est.**	**t**	**p**
0 (intercept)	4.41	10.54	< 0.001	0.15	18.63	< 0.001	3.65	28.08	< 0.001
5.1		1.46	0.143		1.25	0.265		3.52	< 0.001
10.3		2.29	0.022		1.16	0.450		1.14	0.254
15.4		1.98	0.048		1.40	0.140		2.06	0.039
20.6		5.08	< 0.001		2.11	0.135		0.52	0.602
25.7		2.62	0.009		2.15	0.132		2.58	0.010
30.9		4.34	< 0.001		2.20	0.128		3.86	< 0.001
36.0		6.00	< 0.001		1.28	0.201		5.05	< 0.001
	$R^2_c = 0.72$, $R^2_m = 0.50$			$R^2_c = 0.11$, $R^2_m = 0.02$			$R^2_c = 0.59$, $R^2_m = 0.30$		

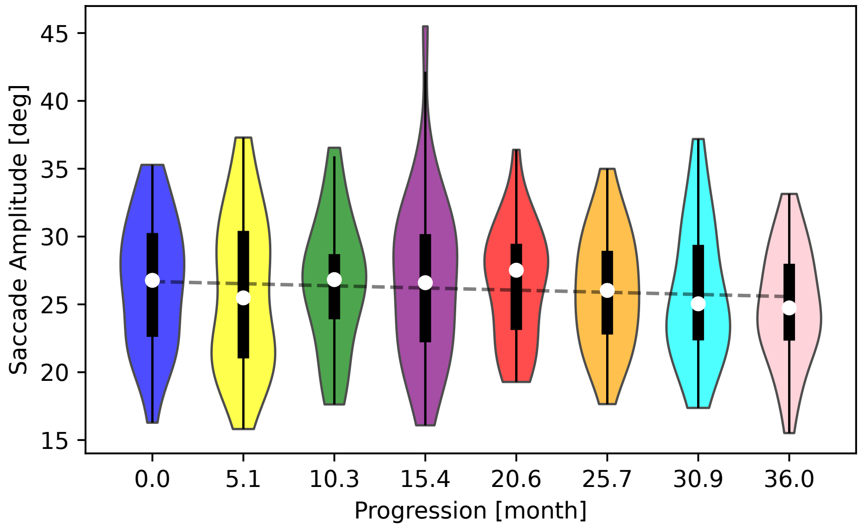


Figure 3.24: Saccade amplitude for each of the different cataract simulations.

Saccade Amplitude

Our model for saccade amplitude as a function of cataract progression exhibits substantial explanatory power, with a conditional $R^2 = 0.99$ and a marginal $R^2 = 0.70$, using an inverse Gamma family with an identity link. The intercept of the model, representing no simulated cataract progression, was estimated at 27.51 ($t = 19.58$, $p < .001$). At 5.1 months, a significant decrease in saccade amplitude was observed ($t = -3.17$, $p = 0.002$). No significant change was found at 10.3 months ($t = -1.53$, $p = 0.125$) or 15.4 months ($t = -1.54$, $p = 0.123$). However, at 20.6 months, a significant reduction in saccade amplitude was observed ($t = -2.70$, $p = 0.007$), followed by another significant decrease at 25.7 months ($t = -2.08$, $p = 0.037$). The most pronounced effects were found at 30.9 months ($t = -3.90$, $p < .001$) and 36.0 months ($t = -5.36$, $p < .001$), indicating a progressive decline in saccade amplitude as cataract severity increased.

Saccade Frequency

Our model for saccade frequency as a function of cataract progression exhibits moderate explanatory power, with a conditional $R^2 = 0.40$ and a marginal $R^2 = 0.13$, using an inverse Gamma family with an identity link. The intercept of the model, representing no simulated cataract progression, was estimated at 2.87 ($t = 28.33$, $p < .001$). At 5.1 months, no significant change in saccade frequency was observed ($t = 0.06$, $p = 0.956$). Similarly, at 10.3 months, the effect remained non-significant ($t = -1.54$, $p = 0.124$). At 15.4 months, a significant decrease was found ($t = -3.08$, $p = 0.002$). No significant changes were found at 20.6 months ($t = -1.33$, $p = 0.183$) or 25.7 months ($t = -1.39$, $p = 0.166$). However, at 30.9 months, saccade frequency significantly decreased ($t = -2.90$, $p = 0.004$), with a further reduction observed at 36.0 months ($t = -2.34$, $p = 0.019$). These results indicate that while early stages of cataract

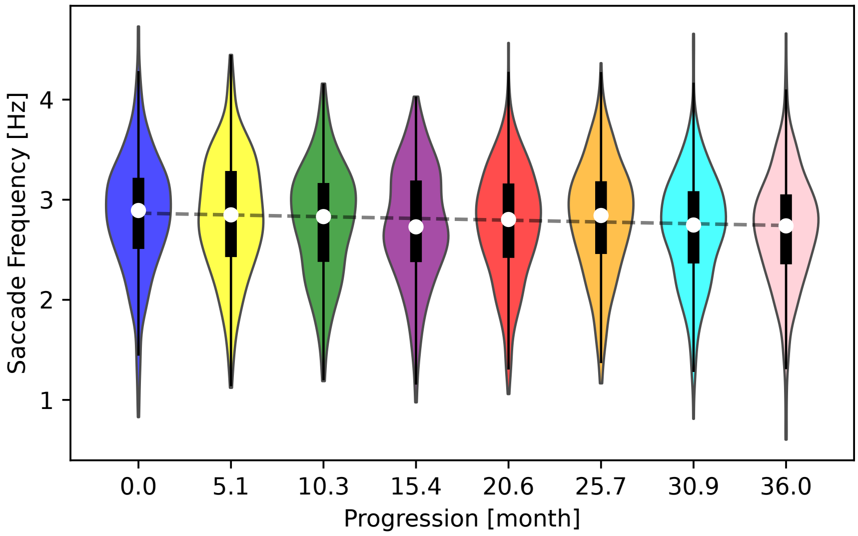


Figure 3.25: Saccade frequency for each of the different cataract simulations.

progression do not significantly impact saccade frequency, later stages are associated with a notable decline.

Table 3.4: GLMM results for saccade amplitude and saccade frequency under different cataract progression conditions.

	Sac. Amplitude			Sac. Frequency		
Progression (months)	**Est.**	**t**	**p**	**Est.**	**t**	**p**
0 (Intercept)	27.51	19.58	< 0.001	2.87	28.33	< 0.001
5.1		-3.17	0.002		0.06	0.956
10.3		-1.53	0.125		-1.54	0.124
15.4		-1.54	0.123		-3.08	0.002
20.6		-2.70	0.007		-1.33	0.183
25.7		-2.08	0.037		-1.39	0.166
30.9		-3.90	< 0.001		-2.90	0.004
36.0		-5.36	< 0.001		-2.34	0.019
	$R_c^2 = 0.99$, $R_m^2 = 0.70$			$R_c^2 = 0.40$, $R_m^2 = 0.13$		

Pupil Size

Our model for pupil size as a function of cataract progression exhibits moderate explanatory power, with a conditional $R^2 = 0.40$ and a marginal $R^2 = 0.13$, using an inverse Gaussian family with an identity link. The intercept of the model, representing no simulated cataract progression, was estimated at 4.66 ($t = 23.50$, $p < .001$). A significant decrease in pupil size was observed at 5.1 months ($t = -12.40$, $p < .001$), indicating an early effect of cataract progression. This decrease continued at 10.3 months ($t = -14.79$, $p < .001$) and became more pronounced at 15.4 months ($t = -23.48$, $p < .001$). At 20.6 months, pupil size further

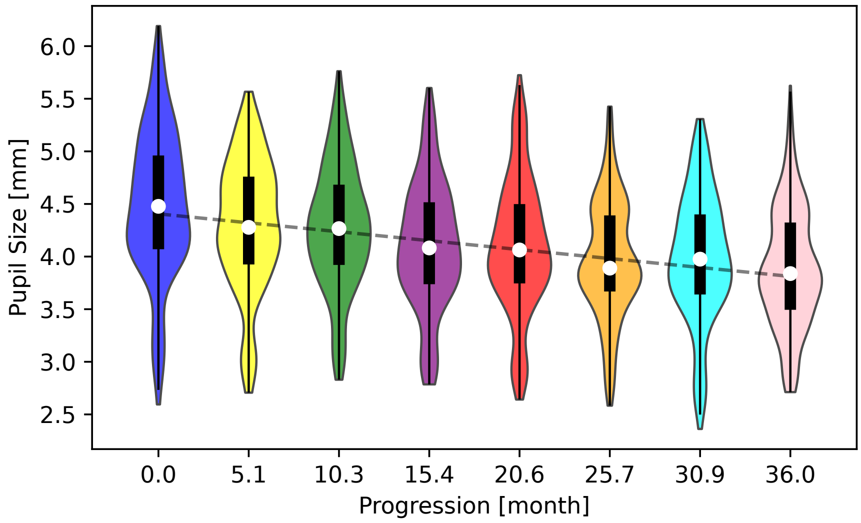


Figure 3.26: Pupil size for each of the different cataract simulations.

decreased significantly ($t = -25.12$, $p < .001$), followed by additional reductions at 25.7 months ($t = -34.14$, $p < .001$), 30.9 months ($t = -37.71$, $p < .001$), and 36.0 months ($t = -42.47$, $p < .001$). These results indicate a consistent and substantial reduction in pupil size as cataract severity increases.

Index of Pupillary Activity (IPA)

Our model for the Index of Pupillary Activity (IPA) as a function of cataract progression exhibits strong explanatory power, with a conditional $R^2 = 0.96$ and a marginal $R^2 = 0.54$,

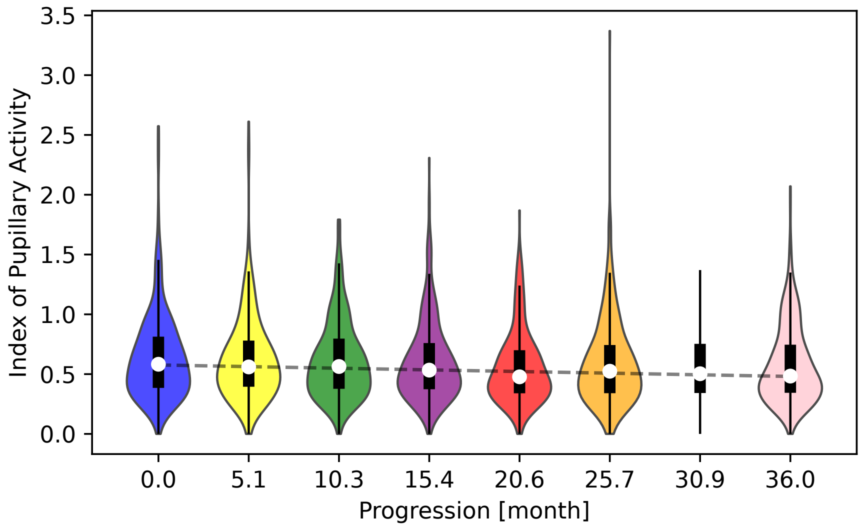


Figure 3.27: Index of Pupillary Activity for each of the different cataract simulations.

using an inverse Gamma family with an identity link. The intercept of the model, representing no simulated cataract progression, was estimated at 0.65 ($t = 21.06$, $p < .001$). No significant change in IPA was observed at 5.1 months ($t = -0.62$, $p = 0.536$) or 10.3 months ($t = -0.92$, $p = 0.358$). At 15.4 months, the decrease in IPA approached significance ($t = -1.67$, $p = 0.094$). A significant reduction was observed at 20.6 months ($t = -1.84$, $p = 0.049$) and at 25.7 months ($t = -2.11$, $p = 0.035$). The most pronounced decreases were found at 30.9 months ($t = -2.92$, $p = 0.004$) and 36.0 months ($t = -3.41$, $p < .001$), indicating a progressive decline in IPA as cataract severity increased. These results suggest that pupillary activity is increasingly affected in later stages of cataract progression.

Table 3.5: GLMM results for pupil size and the index of pupillary activity (IPA) under different cataract progression conditions.

	Pupil Size			**IPA**		
Progression (months)	**Est.**	**t**	**p**	**Est.**	**t**	**p**
0 (Intercept)	4.66	23.50	< 0.001	0.65	21.06	< 0.001
5.1		-12.40	< 0.001		-0.62	0.536
10.3		-14.79	< 0.001		-0.92	0.358
15.4		-23.48	< 0.001		-1.67	0.094
20.6		-25.12	< 0.001		-1.84	0.049
25.7		-34.14	< 0.001		-2.11	0.035
30.9		-37.71	< 0.001		-2.92	0.004
36.0		-42.47	< 0.001		-3.41	< 0.001
	$R^2_c = 0.40$, $R^2_m = 0.13$			$R^2_c = 0.96$, $R^2_m = 0.54$		

3.4 Chapter Summary

In this chapter, we investigated gaze behaviors associated with visual impairments through VR simulations to support preventive healthcare approaches. Timely detection of visual impairments is crucial, as early intervention can improve outcomes. Our analysis particularly focused on cataracts, glaucoma, and AMD, each characterized by distinct visual deficits and progression patterns.

Firstly, we explored cataracts using realistic VR simulations validated through iterative interviews. Participants often remained unaware of mild cataract progression, though objectively measurable changes in gaze patterns emerged at moderate-to-severe levels. We observed notable increases in fixation rate and prolonged visual search times with advanced cataracts. These findings emphasize gaze characteristics as potential early diagnostic indicators, which can enhance preventive healthcare by facilitating earlier recognition and management of cataracts.

Secondly, we examined glaucoma by creating refined simulations based on patient-reported visual experiences, significantly improving upon traditional visualizations. Our user study revealed clear gaze pattern alterations, specifically reduced saccadic amplitudes and increased fixation frequencies, corresponding closely with peripheral vision loss severity. These distinctive eye movement adaptations underscore the potential of gaze analysis for the early detection and monitoring of glaucoma, allowing preventive strategies to intervene before substantial vision loss occurs.

Lastly, our investigation into AMD centered around enhanced simulations of CFL, informed by structured patient interviews. Our refined AMD visualization demonstrated precise correlations between gaze instability, altered saccadic patterns, and varying degrees of central vision impairment. Task performance analysis indicated increased search times and heightened fixation rates in moderate-to-advanced AMD severities. Importantly, our realistic simulations revealed gaze patterns closely aligned with actual patient experiences, significantly advancing preventive healthcare potential through accurate early-stage detection and ongoing monitoring.

We specifically addressed our research question **RQ1**, demonstrating that realistic VR simulations effectively facilitate early-stage detection via gaze pattern analysis. The findings from our studies further identified specific gaze characteristics indicative of different impairment severities and types. Furthermore, we confirmed that gaze indicators become reliably detectable at impairment severities where participants themselves remain largely unaware.

The combined findings from these studies provide compelling evidence that detailed gaze analysis through realistic VR simulations offers critical diagnostic insights into visual impairments at their early stages. By enhancing simulation accuracy and rigorously quantifying gaze patterns, we enable proactive identification and management strategies. This proactive approach aligns closely with preventive healthcare goals, ultimately aiming to reduce the individual and societal impacts of age-related visual impairments and improve overall quality of life.

4 STANDARDIZING

Eye tracking holds considerable promise for preventive healthcare, offering a non-invasive and behaviorally grounded means to detect early signs of cognitive and perceptual decline. With the integration of eye-tracking into real-time applications comes the opportunity to continuously monitor the presence of visual impairments through gaze. However, the effectiveness of such systems hinges on the quality and continuity of the gaze data they rely upon. This is especially important in systems designed to detect subtle gaze cues for early intervention, where even small disruptions in data can compromise diagnostic sensitivity and reliability.

Machine learning models, particularly recurrent architectures like long short-term memory (LSTM) networks, have emerged as a powerful tool for interpreting sequential gaze data in these systems. These models rely on temporal coherence to identify patterns that may signify user intent, cognitive state, or visual impairments. However, their effectiveness is highly sensitive to missing or noisy data, which can degrade performance and introduce biases [82]. When gaze input is incomplete, models may fail to generalize to new users or tasks, reducing both fairness and robustness.

Blinks are an inevitable part of the human visual system—short, involuntary closures of the eyelids that maintain ocular health and protect the eye from environmental stressors [65]. While biologically necessary, these momentary interruptions pose a unique challenge in the context of eye-tracking-based interactive systems. In recent years, eye tracking has become a powerful and widely used tool in human-computer interaction (HCI), enabling systems to interpret user attention, intention, and cognitive state in increasingly nuanced and context-aware ways [83]. As eye tracking is increasingly integrated into real-time applications, from adaptive interfaces to attention-aware systems, the need for continuous and reliable gaze input becomes more critical. However, the robustness and reliability of these systems fundamentally depend on the integrity of the input data stream. When a blink occurs, the gaze signal is interrupted—sometimes briefly, sometimes for extended durations—leading to gaps, artifacts, and discontinuities that can propagate through downstream processes [81].

These disruptions are particularly problematic for intelligent systems that operate on temporal data streams and require coherent sequential input to learn from or respond to user behavior. In particular, recurrent neural networks (RNNs), and more recently long short-term memory (LSTM) architectures, have emerged as promising tools for modeling gaze behavior and predicting user state. However, these models rely heavily on temporal consistency and are highly sensitive to missing or noisy data [82]. Gaps in the signal—whether caused by blinks, signal loss, or hardware limitations—can introduce significant noise into model training and inference. Such missing data may lead to performance degradation, increased error rates, or overfitting, especially when systems attempt to generalize to unseen tasks or users [82].

Despite this, the field lacks a unified and empirically grounded strategy for handling blink-induced data loss. Existing approaches vary widely: some researchers remove all affected segments, while others attempt to interpolate, impute, or otherwise reconstruct missing samples. These decisions are often made ad hoc, depending on task-specific constraints or device limitations. As a result, current practices exhibit a high degree of methodological inconsistency, which undermines both reproducibility and comparability across studies [83].

Beyond the challenge of missing samples, blinks also introduce subtle physiological artifacts that extend beyond the actual duration of eyelid closure. Eyelid motion before and after the blink can distort gaze signals due to partial occlusion and muscular interference, affecting as much as 70 milliseconds before and 118 milliseconds after a blink [83]. These blink-adjacent artifacts are rarely accounted for in existing systems, despite their potential to skew results and introduce noise even in apparently valid samples. Consequently, systems that fail to handle both the absence and the distortion of gaze data risk producing unreliable outputs, especially in high-stakes applications such as driver monitoring or medical diagnostics. A growing body of work highlights the importance of addressing not only the missing data itself but also the surrounding signal, in order to build robust and fair gaze-based systems.

In this Chapter, we investigate the consequences of blink-related data interruptions for the design of interactive systems and gaze-based prediction models. We begin by reviewing how prior research has handled blink-induced data loss and examine the diversity and limitations of existing approaches. We then explore the magnitude of the problem across multiple open datasets and evaluate the impact of different interpolation and artifact removal techniques. In doing so, we aim to quantify the influence of pre-processing strategies on the performance of gaze-based models. Through these investigations, we propose a standard pipeline for blink-aware preprocessing of eye-tracking data and offer design recommendations for building more robust, transparent, and replicable gaze-driven systems[1].

We address the following research questions with our investigation:

RQ 2: *How should we deal with gaps in eye-tracking data for machine learning models?*

[1]The following sections are copied with minor changes from the publications: subsection 4.1.1 until subsection 4.1.2, subsection 4.2.1 until subsection 4.2.5, and subsection 4.3.1 until subsection 4.3.2

This chapter is based on the following publications.

Grootjen, Jesse W. et al. 'Highlighting the Challenges of Blinks in Eye Tracking for Interactive Systems.' In: *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications*. ETRA '23. Tubingen, Germany: Association for Computing Machinery, 2023. DOI: 10.1145/3588015.3589202

Grootjen, Jesse W. et al. 'Uncovering and Addressing Blink-Related Challenges in Using Eye Tracking for Interactive Systems.' In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. CHI '24. New York, NY, USA: Association for Computing Machinery, January 1, 2024. DOI: 10.1145/3613904.3642086

Grootjen, Jesse W. et al. 'Investigating the Effects of Eye-Tracking Interpolation Methods on Model Performance of LSTM.' In: *Proceedings of the 2024 Symposium on Eye Tracking Research and Applications*. ETRA '24. Glasgow, United Kingdom: Association for Computing Machinery, January 1, 2024. DOI: 10.1145/3649902.3656353

4.1 Identifying the Challenge of Blinks

Blinks introduce missing data and artifacts in eye-tracking streams, posing significant challenges for interactive systems. In this section, we highlight how prior work has addressed blinks and demonstrate that existing strategies—such as data removal or interpolation—are often inconsistent, limiting reproducibility and potentially impairing system performance [81].

4.1.1 Initial Literature Search

We identified several ways of blink detection methods applied in related work, which we found through our search. While for the SMI and EyeLink trackers, there are exciting parsers available, most of the identified literature simply reports on the missing data being the detector for an indication of a blink. For example, Li et al. [149] reported on missing data using the EyeLink and others [27, 119, 128] report on missing data from SMI eye tracker. As such, we identified that the blink detection methods do not leverage the available parsers, and improving this has the potential for a positive impact on research quality.

Although various solutions for detecting blinks and handling missing data have been proposed, there are no readily available solutions integrated into the current eye-tracking pipeline. Since eye-tracking data is a time-series, traditional imputation methods, such as replacing missing data with the mean or median, are not appropriate for handling missing data caused by blinks. Therefore, the majority of previous studies identified with our search have resorted to removing the data containing blinks. Nevertheless, alternative methods for handling missing data caused by blinks have been explored in the literature, such as interpolation and imputation. Despite being less commonly used, these approaches have the advantage of preserving the applicability of eye-tracking data in interactive systems.

Remove

Most of the identified literature on blink detection and handling involves removing the data that contains blinks. However, some studies have proposed criteria for retaining the blinks in the data. For instance, Ishii et al. [108] and Nakano and Ishii [174] remove data when blinks last longer than 200 milliseconds, while they merge data before and after the blink if it is shorter than 200 milliseconds. Similarly, Bixler and D'Mello [28] allow for up to 20% of the data points to be missing during a specific time period, whereas Gwizdka [90] excludes trials where gaps in data are longer than 1 second.

Interpolate

Interpolation is a frequently employed technique for replacing the gaps in the data introduced by blinks, as reported in the identified literature. The work utilized various forms of interpolation, such as linear, polynomial, spline, and cubic spline. For instance, Kinnunen et al. [121] applied linear interpolation by assuming the continuity of the data for the parts with missing data and independently applying interpolation to all coordinate time series. Another example is Wang et al. [269], where the authors set missing data when the pupil size was outside a range and then applied spline interpolation to correct the missing values.

Linear Interpolation

Linear interpolation is, of the identified literature, the most commonly used method to deal with missing data in eye-tracking studies after removing the data. This technique assumes that the underlying signal is a linear function and fills in the missing data by connecting the available data points with straight lines. While linear interpolation is a simple and efficient method, it can produce artifacts in the interpolated data and can introduce errors in the analysis of the eye-tracking data.

An example of the use of the linear interpolation method is in Bafna et al. [15]. In this work, the authors removed 200 ms before and after a blink using the Tobii 4C Eye Tracker[2] during a typing task. Another example of this is Saluja et al. [216], here the authors also used linear interpolation. However, this was during a reading task and without removing additional data before and after the blink.

Spline Interpolation

Spline interpolation is another method we have seen in the identified literature more than once. Unlike linear interpolation, which assumes a constant rate of change between data points, spline interpolation uses piecewise polynomials to interpolate data, resulting in a smoother fit. Cubic spline interpolation is a specific type of spline interpolation that uses cubic polynomials to interpolate between data points. Spline interpolation has been used in eye-tracking studies to fill in missing data due to blinks, head movements, or other sources of noise.

For example, Wang et al. [269] used spline interpolation to correct for missing data when the pupil samples were outside of a pre-determined range during a driving simulator experiment. For cubic spline interpolation, we identified Morales et al. [170]. Here the authors used the EyeLink and its online parser software to identify blinks, after which they applied cubic splines to remove blink data.

Polynomial Interpolation

Polynomial interpolation is another method that can be used for infilling missing values in eye-tracking data. In this method, a polynomial function is used to fit the known data points, and the polynomial is then used to estimate the values of the missing data points. An advantage of polynomial interpolation is that it is easy to implement, and it can be used to estimate missing data points with a high degree of accuracy. However, polynomial interpolation is also sensitive to outliers in the data, and it can produce unrealistic estimates if the data

[2] `https://www.tobii.com/`

contains extreme values or abrupt changes in eye movement patterns. An example of the use of polynomial interpolation we identified in our search is the work from Keshava et al. [119], here the authors employ the interpolation method to the eye movements from a VR task.

Other

Finally, we identified a range of alternative methods to deal with missing data caused by blinks, including imputation, WEKA, aggregating, and winsoring. For instance, Li et al. [149] employed the unsupervised Expectation-Maximization algorithm to impute missing values in their eye-tracking data. Likewise, Cole et al. [48] filled in the missing data using observed session transition probabilities. Despite the various approaches identified, there is no consensus in the literature on which method is the most appropriate for dealing with blinks. While most reviewed studies remove the affected data, this often leads to a significant loss of data.

4.1.2 Results

In this work, we identified several ways past work has employed to identify blinks and algorithms used to infill the missing data and found that these are not consistent throughout the existing literature. In the following, we will showcase the challenges of the different interpolation methods and the impact having incorrect cut-off times can have on the interpolation.

(a) (b) (c) (d)

Figure 4.1: An example of **linear interpolation** over a blink during a reading experiment. Here, the black dots represent fixations, the black line eye movements between the fixations and the red lines represent the interpolation. In (a), we visualized the interpolation without removing any of the information pre or post-blink. In (b), we visualize the interpolation with the removal of information pre and post-blink, however, not enough of the artifacts have been removed. In (c), we visualize the correct removal of artifacts pre and post-blink. Finally, in (c), we showcase the removal of too much of the data pre and post-blink.

Linear interpolation Linear interpolation is a simple and commonly used infilling method for missing data in the identified literature. The method involves drawing a straight line between pairs of data points that flank a gap and then interpolating data points along the lines to fill in the missing values. From the identified infilling methods used, linear interpolation was the most common one, which we attribute to its simplicity and low computational cost. However, as we visualize in Figure 4.1, the method assumes that the data points before and after the gap are linearly related, which is not always the case. Additionally, the method can produce abrupt changes in velocity or acceleration, as most predominantly visualized in Figure 4.1a and Figure 4.1b. As such, when using linear interpolation, a best practice would be to visually inspect each individual blink to see where artifacts of blink behavior are introduced into the data. Removing these artifacts will allow for the best possible interpolation between data gaps, as visualized in Figure 4.1c. However, overdoing the cut-off before and after the artifacts will introduce new noise into the data, as visualized in Figure 4.1d.

(a) (b) (c) (d)

Figure 4.2: An example of **spline interpolation** over a blink during a reading experiment. Here, the black dots represent fixations, the black line eye movements between the fixations and the red lines represent the interpolation. In (a), we visualized the interpolation without removing any of the information pre or post-blink. In (b), we visualize the interpolation with the removal of information pre and post-blink, however, not enough of the artifacts have been removed. In (c), we visualize the correct removal of artifacts pre and post-blink. Finally, in (c), we showcase the removal of too much of the data pre and post-blink.

Spline interpolation Spline interpolation is a more sophisticated way of interpolating compared to linear interpolation. This method fits a piecewise polynomial function to the data points where the resulting function is smooth and passes through all data points, including those with missing data. It allows for considering the velocity and direction of the data points pre and post-missing data to ensure the curve will not abruptly change velocity or acceleration. While this method is not as popular in the identified literature as linear interpolation, it could provide a more smooth and more accurate interpolation compared to linear interpolation. However, spline interpolation requires more computational resources than linear interpolation. Still, this interpolation method struggles with the artifacts introduced pre and post-blink as visualized in Figure 4.2a and Figure 4.2b. When visualizing the exact cut-off where the artifacts from blinks start in Figure 4.2c, the infill method results in a smooth curve taking the velocity and acceleration into account before and after the gap. However, when cutting off too much of the data as visualized in Figure 4.2d, we can see that the infilling method misses quite a few of the words that would otherwise be in the path of the eye path.

(a) (b) (c) (d)

Figure 4.3: An example of **polynomial interpolation** over a blink during a reading experiment. Here, the black dots represent fixations, the black line eye movements between the fixations and the red lines represent the interpolation. In (a), we visualized the interpolation without removing any of the information pre or post-blink. In (b), we visualize the interpolation with the removal of information pre and post-blink, however, not enough of the artifacts have been removed. In (c), we visualize the correct removal of artifacts pre and post-blink. Finally, in (c), we showcase the removal of too much of the data pre and post-blink.

Polynomial interpolation Polynomial interpolation is the last method we identified in multiple papers throughout the identified literature. It fits a single polynomial function to all the data points, which is different from spline interpolation, which divides all data points into segments. The resulting function is smoother than linear interpolation but not as smooth as spline interpolation. Similarly to linear interpolation, it is a simple and computationally efficient method for dealing with missing data. However, it has limited accuracy in case the underlying data is highly nonlinear, which results in spurious oscillations near the endpoints of the interpolated data, as visualized in Figure 4.3a and Figure 4.3b. Despite this, when the cut-off periods are chosen adequately, it can neatly estimate the missing data, as visualized in Figure 4.3c. On the other hand, removing too much of the data before and after a blink can result in an error compared to the original data, see Figure 4.3d.

4.2 Uncovering and Addressing Blink-Related Challenges

While previous work has acknowledged the challenges posed by blinks in eye-tracking data, comprehensive evaluations of their impact remain scarce. In this section, we conducted a systematic literature review and large-scale analysis of public datasets. Our work revealed significant inconsistencies in how missing data and blink artifacts are handled across studies, as well as the lack of standardized preprocessing pipelines.

We conducted a structured literature review to identify the wide range of approaches used in dealing with eye tracking data in interactive systems. In detail, we aim to review the blink detection methods and algorithms used. For this, we follow the four-phase procedure of the PRISMA [187] guidelines on reporting systematic reviews. Figure 4.4 visualizes the PRISMA flowchart.

4.2.1 Method

We follow the PRISMA guidelines to review prior work, which is in line with other papers [13, 26, 96] in the HCI domain. The review focuses on three key aspects: the *Method* used for handling blinks, the use of blink *Detectors*, and the *Task* that users carry out.

Identification

We defined the eligibility criteria, namely, exclusion criteria as shown in Figure 4.4. We defined the inclusion criteria as papers involving eye tracking data and their handling of missing data. From the databases, we selected the ACM digital library and IEEE as they are representative of high-quality and mature research published in the field of interactive systems and eye tracking. At the same time, we acknowledge that this excludes other venues interested in eye tracking data (e.g., ARVO, Journal of Vision and Thieme Medical Publishers, and Journal of Academic Ophthalmology). However, the ACM Digital Library and IEEE Xplore are major libraries for interactive systems and, thus, best fit the aim of this review. We conducted our search given our inclusion criteria, selecting papers with terms relevant to eye tracking and terms that indicate the presence of missing data. Specifically, we used the following search string:

```
("eye tracking" OR "gaze tracking" OR "eye movements" OR
"gaze movements")
AND ("missing data" OR "missing value" OR "data gaps")
AND E-Publication Date: (* TO 12/31/2022)
```

We manually saved all resulting records in a CSV file for screening.

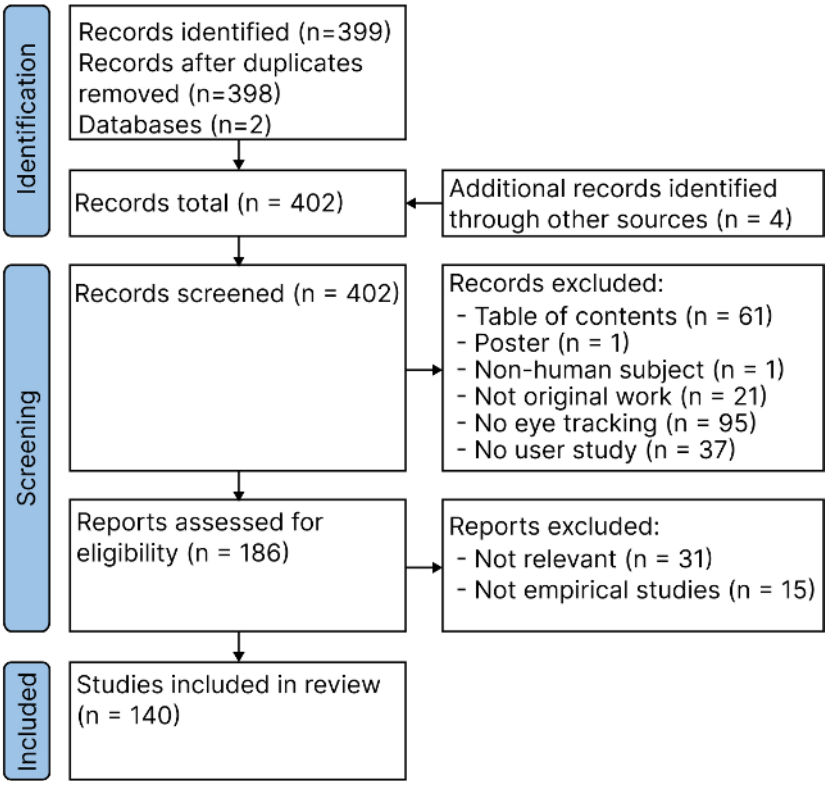


Figure 4.4: Literature search and inclusion phases and rates using the PRISMA flowchart.

Screening

We excluded all tables of content and posters, papers that do not include a study, and papers that do a study with non-human subjects as shown in Figure 4.4. For this, the first author then screened each paper while not using automated tools. The goal of the initial screening is to keep all papers that could help us understand how current papers deal with missing data. Thus, we excluded papers based on the exclusion criteria in Figure 4.4. We excluded 1) *table of contents* ($n = 61$); 2) *posters* ($n = 1$); 3) papers with *non-human subjects* ($n = 1$) as they do not add to our investigation; 4) *not original work* as they only review or comment on others' work but do not process eye tracking data ($n = 21$); 5) even though eye tracking was in the keywords, we had to exclude works, which is *not eye tracking* ($n = 95$); and 6) we excluded results that do not entail a *user study* as no validation was done ($n = 37$). In this step, we reduced the number of included papers from 402 to 186.

For the second step, the first two authors then read the remaining 186 papers after the initial screening. Here, we excluded papers based on two criteria: *not relevant* ($n = 31$) to the inclusion criteria and *not an empirical study* ($n = 15$), see Figure 4.4. The two people who screened the papers again did this independently to minimize potential bias. The inter-rater reliability was 97% on the exclusion criteria and discrepancies were resolved through discussions.

Included Papers

We included the remaining 140 publications in the review[3]. The first two authors of the paper this section is published on independently coded each of these papers without the use of automated tools.

Without specifying a codebook beforehand, the two authors each coded the *Method* used for handling blinks, the possible use of blink *Detectors*, and the *Task* the users carried out. As we did not establish a codebook beforehand, we did not expect high inter-rater reliability on the open-ended text. Despite this, we had an inter-rater reliability of 79.5% for coding the *Detectors*, 53.4% for the *Method*, and 13.7% for the *Task*. As before, we resolved all discrepancies in discussions, resulting in the final codes reported in Table 4.1, Table 4.2 and Appendix A.

Selected Papers

Table 4.1 and Table 4.2 give an overview of how the 140 papers dealt with missing data. The earliest paper in our selected papers is from 1993. We note that 60 papers reported insufficient information on how they handled missing data, see Table 4.1; thus, we cannot include them in further analyses. However, they provide evidence for the need for better reporting guidelines. We showcase a subset of 81 papers in Appendix A to highlight the work that deals with missing data.

Papers with Insufficient Information An overwhelming majority of the papers (44/140) did not report how they dealt with missing data from their experiments or they acknowledged the existence of missing data but did not elaborate on how they handled the missing data, see Table 4.1. Some of these papers mention the removal of participants who are missing over a certain amount of data (e.g., [69, 138, 155, 181]) or trials where missing data reached a threshold (e.g., [58, 126]). However, they did not elaborate on what they did with trials and participants who had missing data but were not excluded in the analysis, i.e., how they dealt with remaining missing data was not described. Moreover, we identified 15 papers that did not handle missing data while acknowledging the issue around missing data. Finally, one paper did remove blinks but gave no information on the criteria for removal or how they were removed.

Papers with Helpful Information Sixty papers had insufficient information to be fully included in this review. In the following, we will categorize the remaining 81 papers on how they dealt with missing data.

Although all reviewed work comes from either the ACM digital library or IEEE Xplore, the work reviewed was published in several venues. From the work reported in Appendix A and reports on dealing with blinks, the most predominant venue was the ACM Symposium on Eye Tracking Research and Applications (ETRA), with 12 papers. Beyond that there were venues

[3]These 140 papers are marked with a • in the references of this thesis.

Table 4.1: Different methods for handling missing data, these papers have insufficient information to be included in the detailed review.

Method	Count	Publications
Nothing reported	44	[5, 11, 17, 27, 40, 44, 46, 53, 69, 77, 99, 111, 112, 124, 125, 135, 140, 154, 155, 169, 176, 181, 188, 191, 192, 193, 195, 199, 203, 220, 221, 227, 235, 249, 251, 256, 257, 266, 267, 271, 276, 281, 293]
No handling of missing data	15	[50, 60, 61, 66, 94, 104, 109, 138, 142, 153, 189, 218, 228, 285, 286]
Removing blinks, not defining them	1	[126]

like ICMI (6), IEEE EMBC (4), CHI (3), IEEE Access (3), and several others. Regarding eye tracking systems, the most popular brand is the Tobii brand, where 36 papers in the reviewed work used one of the Tobii eye trackers for their research; out of these, the Tobii 1750 (4) is the most popular. After Tobii, the SMI brand is also well-represented with 15 papers. Here the SMI RED250 is the most common (4). Studies that have used a mobile eye tracker also seem to favor the EyeTribe (6) and the Tobii Pro Glasses 2 (5).

The tasks we identified in the reviewed work are even more diverse than the selection of eye trackers. The most common task is visual search (15) and driving simulators (15), closely followed by free viewing (14), video watching (9), reading (8), and input method (4). Here, using eye tracking as an input method is interesting as it is the only task (n > 1) where all reviewed work used an interpolation method to deal with the missing data from the eye tracker.

Table 4.2: Different methods for handling missing data, these papers have sufficient information to be included in the detailed review.

Method	Count	Publications
Removed missing data	16	[1, 9, 10, 25, 31, 35, 58, 88, 114, 120, 130, 173, 186, 212, 265, 291]
Interpolation – linear	19	[2, 3, 14, 15, 24, 31, 102, 106, 110, 121, 127, 128, 179, 180, 212, 216, 246, 263, 288]
Interpolation – polynomial	2	[119, 252]
Interpolation – bilinear	1	[150]
Interpolation – spline	3	[31, 269, 290]
Interpolation – cubic spline	4	[95, 170, 201, 231]
Interpolation – Other	14	[9, 55, 73, 78, 97, 151, 159, 200, 230, 254, 282, 287, 289, 292]
Imputation	7	[48, 149, 162, 164, 165, 247, 283]
WEKA	2	[103, 204]
Aggregating	4	[108, 152, 174, 211]
Winsoring	1	[28]
Reconstructing	1	[182]
Averaging	2	[90, 115]
Extrapolate	2	[29, 242]
Other	6	[49, 64, 196, 238, 272, 284]

4.2.2 Findings on Detecting Blinks

Most (73/81) of the reviewed literature reports that they classified missing data as a blink. For the Tobii eye trackers without a blink detector, authors often classified blinks as points where the pupil size is outside a pre-determined range or when the tracker loses the pupil temporarily, e.g., [61, 88, 269]. Other papers (3/81) mention excluding data before and after the missing data. For example, Bafna et al. [15] specifies blinks as missing data 75-500 ms long and, additionally, they remove a further 200 ms before and after the missing data to combat the artifacts before and after the blink. Appel et al. [9] removed data up to 100 ms before and after a blink to counter artifacts, and Appel et al. [8] removed parts from the pupil signal that "had an unreasonably large slope right before and after a sequence of missing data" [8].

EyeLink provides users with a built-in parser[1] and SMI provides the BeGaze parser[2]. Thus, both have their own integrated parsers for blinks; however, not all studies we found during our literature review using those eye trackers report on using the respective parser. More specifically, out of the surveyed papers, over half of the work using an EyeLink (e.g., [149, 192]) and using an SMI (e.g., [8, 9, 27, 40, 119, 128, 164]) did not report on using the built-in parsers or any other parser. However, they report the missing data, which they handled with use-case specificity. As such, we identified that *the methods used for blink detection are inconsistent in the current literature and improving this has a potential positive impact on research replicability and quality*.

4.2.3 Findings on Dealing With Blinks

As discussed in subsection 4.2.2, there are integrated solutions for detecting blinks (e.g., EyeLink parser[1] and BeGaze parser[2]). However, there are no out-of-the-box solutions integrated into the current eye trackers that handle blinks once they are identified. Traditionally, researchers dealing with missing data have so far proposed a set of methods, such as replacing by mean/median [215] and last observed carried forward [226]. However, this does not work for time-series data due to the underlying speed and possible acceleration of the eye movement. As such, the most common method identified in our literature review applies linear interpolation (19/81) closely followed by removing the data containing missing data (16/81), see Appendix A. However, our review revealed numerous use-case-specific methods for handling missing data. Here, the most common method was interpolation (43) of various forms and imputation (7), followed by a wide range of adapted approaches. While these methods are less prevalent and more widespread in the reviewed literature, they have the advantage of retaining the data to be used in interactive systems.

Interpolation Interpolation is the most popular option in our review as 43/81 papers employed a form of interpolation to replace the gaps in the data. In the reviewed work, we identified different kinds of interpolation, including linear (19), polynomial (2), bilinear (1), spline (3), cubic spline (4), and others (14). For example, Kinnunen et al. [121] assumed the continuity of the data and, thus, applied linear 1-D interpolation independently for both axes. In Wang et al. [269], the authors removed data when the pupil size was outside a pre-defined range, after which they applied spline interpolation to infill the missing values. As there is no consensus on how to interpolate eye tracking data, the effects of such methods also need to be more adequately understood. *This can lead to inaccurate or even wrong interactions in interactive systems with low reproducibility chances.*

Remove Blinks Removing the data that contain blinks is the second most popular among the reviewed work (16/81). However, some papers reported elaborate criteria to be met to retain parts or all of the data affected by a blink. For example, Ishii et al. [108] and Nakano and Ishii [174] did not analyze samples with blinks longer than 200 ms; however, if the blink was shorter, then they cut out the missing values. Others allowed for 20% missing data during a trial [28, 203] or they retained trials where the missing data was less than 1 s [90]. *While it is clear that removing all missing data helps the overall performance, this approach is not useful for interactive systems.*

Uncommon Methods Lastly, we identified a series of other methods using a variety of tools and methods, e.g., imputation (7), WEKA (2), aggregating (4), and winsoring (1). To impute missing values, Li et al. [149] used an unsupervised Expectation-Maximization algorithm on their data, and Cole et al. [48] used the observed session transition probabilities to fill in the missing data. We could not identify which specific method is most suitable for dealing with blinks based on the reviewed literature. A large portion of the reviewed work (16) removes missing data where this exists, additionally to the work the removes whole participants or trials. *This results in significant data loss; therefore, other ways of dealing with the missing data could be more appropriate.*

4.2.4 Evaluating the Fitness of Approaches

Although our literature review uncovered trends that researchers use more often, there is no overall consensus for the “best” approach. Additionally, we did not find any evaluations comparing the many approaches to establish guidelines for future interactive systems. Thus, in the following, we compare the different identified approaches. First, we collect a wide range of open-source eye tracking datasets, see Table 4.3. Second, we showcase the implications of the most common approaches to detecting and dealing with blinks. After that we use the acquired data to run the previously identified infilling methods (see Table 4.2) and evaluate against each other.

Datasets

To motivate the importance of our work and evaluate our identified infilling methods, we acquired 11 different open-source eye tracking datasets. All datasets retrieved are part of

Table 4.3: Overview of the used datasets, listed from the oldest to the newest (and alphabetically for authors from the same year)

			Eye Tracker			Screen	
	Author(s)	**Venue**	**Company**	**Device**	**Freq. [Hz]**	**Inch**	**Aspect Ratio**
S01	Foster et al. [72]	Psychol Sci.	EyeLink	1000 Plus	1000	17	16 : 9
S02	Marzecová et al. [161]	Biolo. Psychol	EyeLink	1000	500	19	4 : 3
S03	Krstić et al. [134]	EJPE	SMI	RED-m	60	15.6	16 : 9
S04	Marzecová et al. [160]	Scientific Rep.	EyeLink	1000	500	19	4 : 3
S05	Schuetz et al. [223]	ACM CHI	EyeLink	1000 Plus	1000	113	8 : 5
S06	Annerer-Walcher et al. [7]	Cogn. Sci.	SMI	RED250	250	24	16 : 9
S07	Felßberg and Dombrowe [68]	Vision Res.	EyeLink	1000 Plus	1000	27	16 : 9
S08	Hollenstein et al. [100]	LREC	EyeLink	1000 Plus	1000	27	16 : 9
VR01	Stein et al. [242]	IEEE VR	Tobii	Pro	90	3.5	9 : 10
VR02	Steil et al. [239]	ACM ETRA	Pupil Labs	Add-on	30	5.7	8 : 9
M01	Steil et al. [240]	ACM ETRA	Pupil Labs	Pro	30	-	–

published and peer-reviewed work from various venues. An overview of the acquired datasets is provided in Table 4.3. We acquired all 11 datasets independently of the literature review. They are all available online via `https://osf.io/` and `https://github.com/`. For direct links, see `https://eyetrackingguidelines.github.io/` where we provide links to the datasets on our Eye Tracking Guidelines page, allowing us to extend the list with future published datasets.

Pre-Processing

We first processed the datasets so that they all had the same format, which enabled us to work with the data more easily. If needed, we used the parsers for the EyeLink[1] and SMI[2] data to create the initial tabular files. To format all data equally, we turned all x and y screen gaze coordinates into degrees of visual angle, allowing us to compare them independently from the specific apparatus used (e.g., distance to the screen and screen size). We included only data from the left eye whenever binocular eye tracking data was available. We sampled all time in milliseconds (ms); if there were gaps in timestamps that were bigger than one second (e.g., those created through pausing an experiment), then we split the data into different parts to ensure we do not associate pre- and post-gap data. Where there was missing data (i.e., zero or *Not a Number*), we consider the data as part of a blink. However, we did not consider the missing data of only one sample (e.g., 1 ms for 1000 Hz and 33.3 ms for 30 Hz) as part of a blink as prior work reported blinks to last about one-third of a second [65]. Thus, we did not analyze such short occurrences[4]. With our pre-processing, we aim to reduce external factors not caused by a human eye blink (e.g., breaks in the experiment, looking away from the tracker). Hence, we argue that blinks primarily cause the remaining gaps.

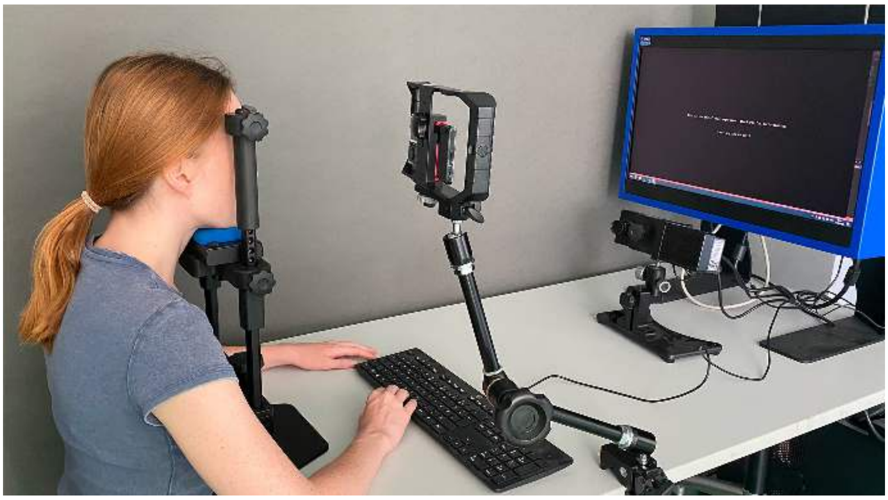


Figure 4.5: Experimental setup used for the verification of data gaps.

Blink Verification

By comparing a high-speed RGB video stream of the eye to recorded eye tracking data, we can ensure that blinks indeed cause the gaps. To maintain high ecological validity, we first searched for publicly available datasets of paired data containing blinks. However, none of the datasets of eyes contain blinks. Thus, the data required for such a comparison are not publicly available. Thus, we decided to conduct an experiment to capture the real eye movement and the eye tracker data simultaneously.

In this experiment, we used an EyeLink 1000 plus from SR Research to capture the eye tracking data at 1000 Hz and a Motorola 30 Ultra to capture the RGB video stream from the blinks at 240 fps as shown in Figure 4.5. After a nine-point calibration and validation, we presented a dot on the center of a ViewPixx (22.5 inch, 1920×1200, 120 Hz) monitor and asked the participant to blink one time once this dot turned green. We had one participant perform 10 trials. The data from this experiment was used to verify that the missing data and the preceding and following artifacts follow a similar trajectory as real blinks[5].

4.2.5 Results

In this section, we first report on the statistics from the surveyed datasets, see Table 4.3. Namely, we investigate the blink frequency, duration of missing data, and inter-blink interval[6]. Second, we report on the effective data loss that occurs through the common rolling window

[4]We acknowledge that this could include data missing for other reasons; however, due to the vast amount of data considered in the analysis this would only amplify our findings.

[5]We acknowledge that these are all voluntary blinks and that there are differences between voluntary and involuntary blinks, e.g., duration [101].

[6]In the following, we will assume all gaps in the data, i.e., missing data, as to be part of a blink.

approach and the fact that windows containing missing data cannot be processed (i.e., zero or *Not a Number*). Third, we investigate the behavior before and after a gap to understand the impact of the eye closing and opening. Finally, we investigate the impact of the five most common infilling methods found in the literature review, see Table 4.2.

Analysis

In this work, we adopt a Bayesian approach for data analysis, specifically employing Bayesian linear mixed models (BLMM). This approach has gained recent prominence [116, 145, 219] and offers several advantages over classical statistics. One of the advantages, as discussed by Kay et al. [117], is the incorporation of prior knowledge from eye tracking data. Additionally, Bayesian statistics facilitate effect estimation in small sample sizes and allow readers to evaluate effect sizes, including those close to zero, rather than solely determining the presence or absence of effects. Consequently, we utilize Bayesian parameter estimation to estimate effect sizes and quantify uncertainty surrounding these estimates by leveraging the information in our data and the applied priors. For all our models, we use the package *brms* to compute 10 Hamilton-Monte-Carlo chains with 20,000 iterations each and 10% warm-up samples. All Rubin-Gelman [74] statistics were well below 1.1 for effective sample size.

Table 4.4: Bayesian statistics results of the priors for blink frequency (contrasted against the whole dataset)

Type	**Freq.**	p_b	Med.	$HDI_{95\%}$
Mobile	30	$<.001$	1.549	$[0.457, 2.980]$
Stationary	60	$<.001$	1.334	$[0.356, 2.689]$
	250	$<.001$	1.282	$[0.354, 2.521]$
	500	$<.001$	1.139	$[0.294, 2.342]$
	1000	$<.001$	2.509	$[1.491, 3.344]$
VR	30	$<.001$	1.472	$[0.427, 2.848]$
	90	$<.001$	1.358	$[0.354, 2.726]$

We explored the effect of different weakly informative priors on the data. None affected statistical inference. As a result, priors were chosen to resemble only weakly informative priors when standardizing with a prior on the Gamma distribution of the data of ($\alpha = 5, \beta = 3$) without allowing negative numbers ($\gamma > 0$). Additionally, we accounted for potential variability across datasets by incorporating a random factor on the shape parameter. This approach acknowledges that different datasets may exhibit varying characteristics and allows for more nuanced modeling. By explicitly modeling the dataset-specific effects, we capture the heterogeneity and better account for the underlying structure of the data.

Effects were considered meaningful when there was a particularly low probability ($p_b <= 2.5\%$) of the effect being zero or the opposite. We calculated p_b through the relative proportion of posterior samples being zero or opposite to the median. This metric has similar properties to the classical p-value and is an accepted substitution cf. [122, 147, 219]. Still, it quantifies the proportion of probability that the effect is zero or the opposite, given the data observed. Note that this is the reverse of the classical approach to inferential statistics, where one measures data probability given the test statistic's null hypothesis. In addition to the median of the parameter, we calculated the High-Density Interval (HDI) at 95% of the posterior distribution for all parameters, which indicates the possible range of effects given the data alongside the median of the respective parameter. Simple mean comparisons were made on standardized outcome variables. Therefore, all $\tilde{b}$ represent an effect size in standard deviations from the mean (corresponding to Cohen's d for simple effects of categorical predictors with two levels).

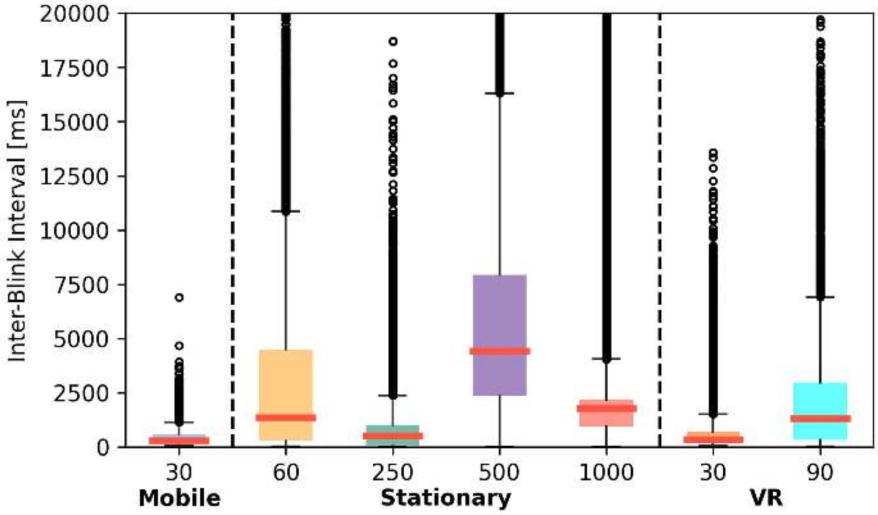


Figure 4.6: We visualized the inter-blink interval for the different types of eye trackers gathered in the dataset, i.e., mobile, VR, and stationary, and their respective frequencies. In the datasets we analyzed, we see that participants have no inter-blink interval while participants from the 30 Hz mobile datasets have the inter-blink interval. Error bars indicate standard errors.

Blink Frequency & Inter-Blink Interval

In Figure 4.7, we present the results of the blink frequency in blinks per minute across the different eye tracker frequencies and eye tracking modalities, i.e., stationary, mobile, or VR. Eye trackers with a low frequency had a generally higher blink frequency; here, the 30 Hz eye tracker had a mean of 49.3 blinks per minute ($SD = 30.5$) while the mean of the others

Table 4.5: Bayesian statistics results of the priors for inter-blink interval (contrasted against the whole dataset)

Type	**Freq.**	p_b	Med.	$HDI_{95\%}$
Mobile	30	$<.001$	1.476	$[0.341, 2.843]$
Stationary	60	$<.001$	2.046	$[0.569, 4.015]$
	250	$<.001$	2.816	$[0.830, 4.821]$
	500	$<.001$	1.392	$[0.316, 2.803]$
	1000	$<.001$	3.597	$[1.078, 5.658]$
VR	30	$<.001$	1.385	$[0.402, 2.828]$
	90	$<.001$	1.497	$[0.469, 2.973]$

Table 4.6: Bayesian statistics results of the priors for closed-eye time (contrasted against the whole dataset)

Type	**Freq.**	p_b	Med.	$HDI_{95\%}$
Mobile	30	$<.001$	1.388	$[0.390, 2.785]$
Stationary	60	$<.001$	1.714	$[0.503, 3.029]$
	250	$<.001$	3.693	$[1.873, 5.323]$
	500	$<.001$	1.433	$[0.408, 3.032]$
	1000	.018	1.734	$[0.164, 3.216]$
VR	30	$<.001$	1.397	$[0.373, 2.730]$
	90	$<.001$	1.520	$[0.442, 2.822]$

is 25.9 blinks per minute ($SD = 37.8$). These results verify the findings that a lot of factors influence blink frequency, as is well established in the literature cf. subsection 2.4.1. We visualize the *inter-blink interval* in milliseconds for the different frequencies and types of eye trackers in Figure 4.6. Next, we investigate how the different types and frequency as fixed effect affects blink frequency in a mixed effects model with the previously described intercepts and priors in section 4.2.5. We found that all combinations of eye tracker TYPE and eye tracker FREQUENCY had a distinguishable effect on the *blink frequency*, see Table 4.4.

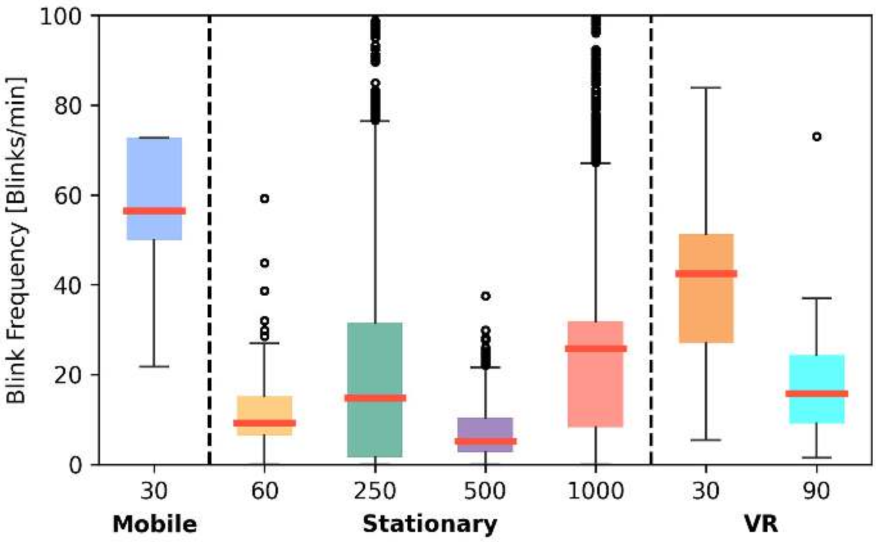


Figure 4.7: We visualized the blink frequencies for the different types of eye trackers gathered in the dataset, i.e., mobile, VR, and stationary, and their respective frequencies. In the datasets we analyzed, we see that participants from the 30 Hz mobile dataset have a high blink frequency and participants from the 500 Hz stationary datasets have the lowest blink frequency. Error bars indicate standard errors.

Impact of Eyelid on Eye Tracking Quality

As illustrated in Figure 4.8, the eyelid's movement can influence the eye tracking quality even before the tracker reaches the state that the eye tracker cannot recognize the pupil anymore.

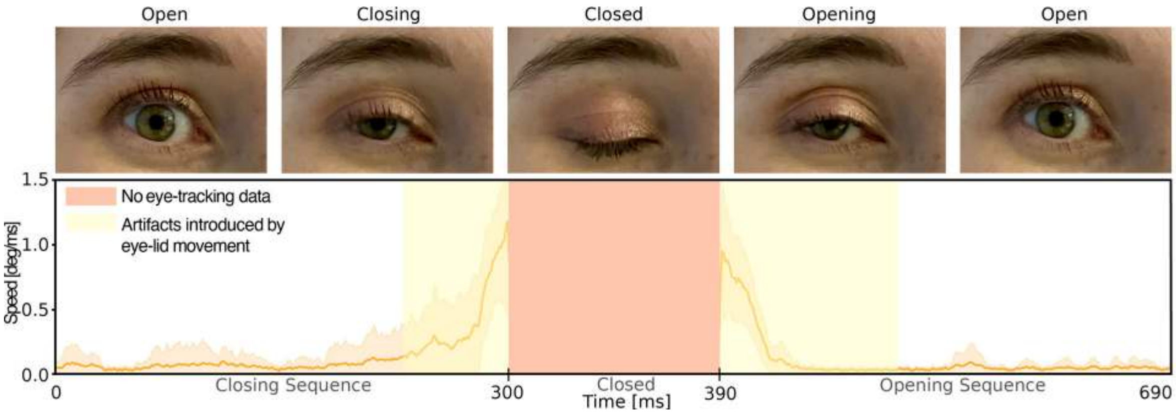


Figure 4.8: Raw eye movement speed before and after a blink with a gap of 90 ms where the eye tracker cannot obtain data. Highlighted in yellow are the areas before and after a blink that contain artifacts introduced by eyelid movements.

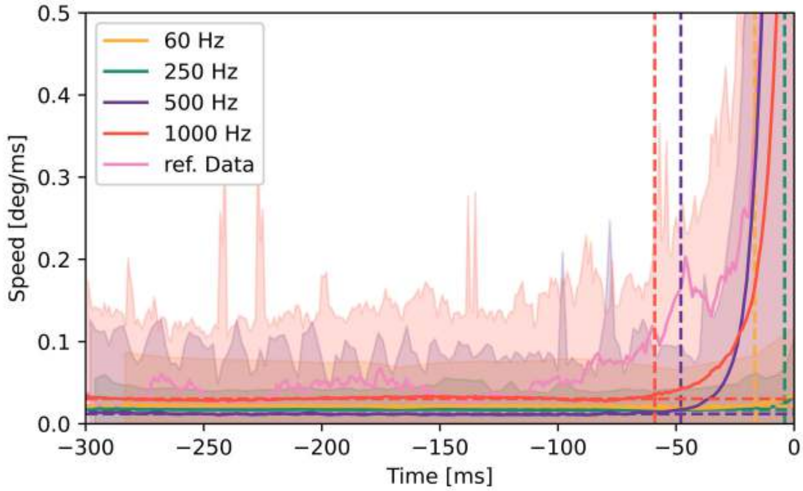
In this process, the eyelid might cover the pupil partly, but the eye tracker still assumes a circular pupil shape in its tracking algorithm. Thus, we investigate the potential impacts of the eyelid before and after a blink on the tracking quality. Uncovering such an influence will allow us to derive cut-off timings before and after the blink to facilitate better overall tracking accuracy.

For this, we investigate the eye movement speed and the speed variation before and after a blink, see Figure 4.9, 4.10, and 4.11. First, we looked at the change in speed before and after, see Figure 4.9a and 4.10a. We found that the speed increases surrounding the blink, which diverges from the time before and after. To study the behavior, we fitted horizontal lines to the speed trajectory (visualized as dashed lines) to the data -300 to -150 ms before and 150 to 300 ms after the blink. This illustrates the difference between with and without blinks. Next, we determined the points of divergence and visualized them as vertical dashed lines. Thus, when the average speed exceeds the fitted average linked plus epsilon for the first time, we determine this to be the cut-off, where the eyelid impacts the tracking resulting in inaccurate tracking, see Table 4.7.

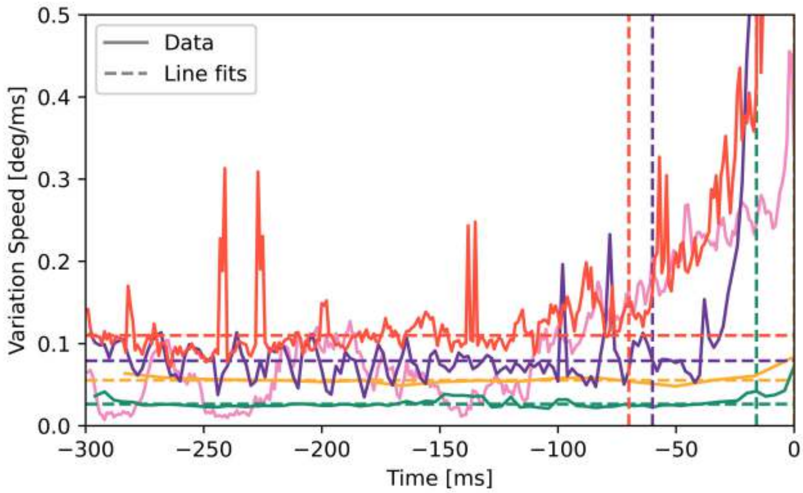
Table 4.7: Identified time [ms] where data preceding and following a blink diverges from normal movement

	Closing Sequence		Opening Sequence	
Freq. [Hz]	**Speed**	**Var. Speed**	**Speed**	**Var. Speed**
60	-16.6	-0.0	66.6	50.0
250	-4.0	-16.0	60.0	36.0
500	-48.0	-44.0	74.0	74.0
1000	-59.0	-70.0	59.0	118.0

As the variation, expressed by the standard deviation in Figure 4.9a and 4.10a, shows a similar trend, we next analyze the speed variation using the same methods. First, we fitted a horizontal line and then a vertical line to determine the point of divergence. The results show that the variation of the eye movement speed following a blink is high directly after a blink and decreases over time until it stabilizes around 0.1. This point depends on the frequency and eye tracker and, thus, must be individually determined.

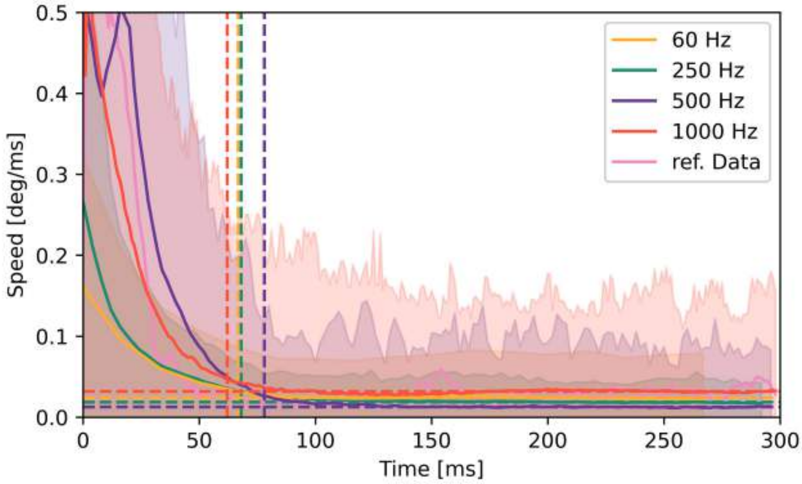


(a)

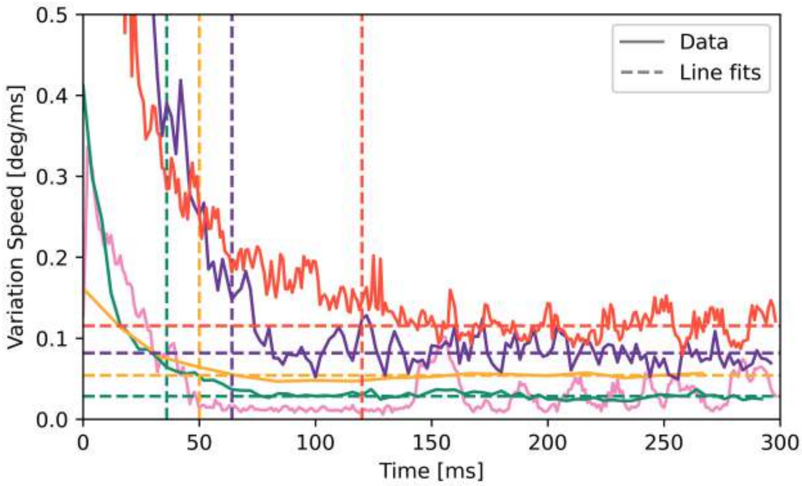


(b)

Figure 4.9: (a) Mean of the eye movement speed preceding a blink for the different frequencies of the stationary eye trackers in the data, including our reference data. We observe a steep increase in speed before missing data appears, around -60 ms. **(b)** Variation of the eye movement speed preceding a blink for the different frequencies of the stationary eye trackers in the data, including our reference data. We observe a steep increase in variation before missing data appears, around -70 ms. The vertical dashed lines represent the points identified as the first time the line crossed the fitted function plus epsilon for both **(a)** and **(b)**.



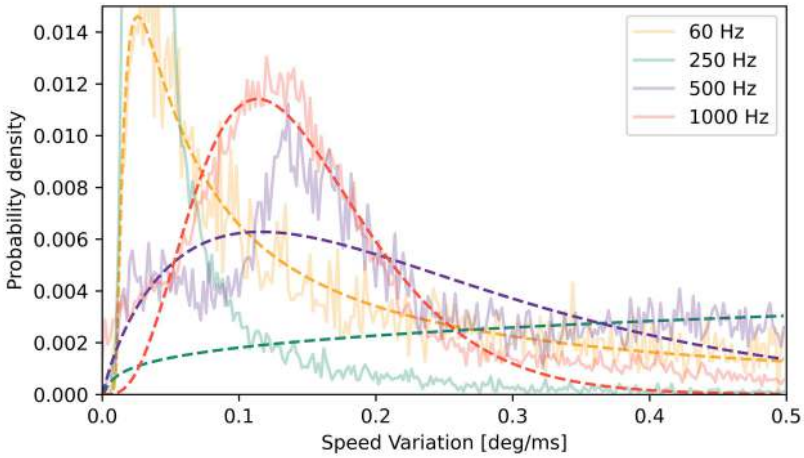
(a)



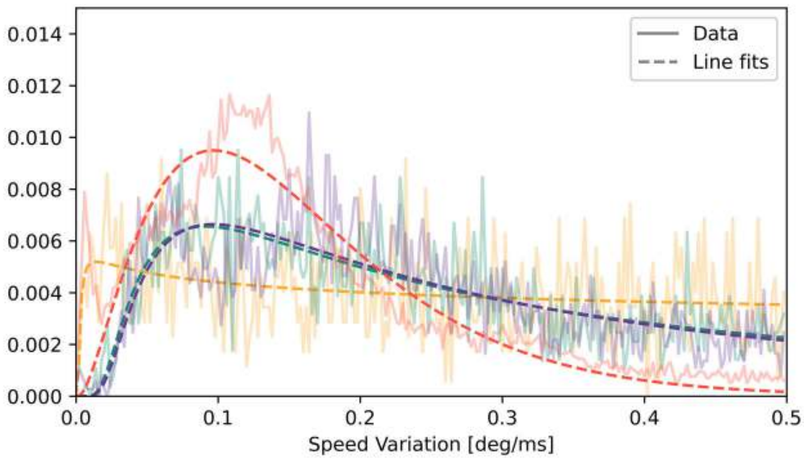
(b)

Figure 4.10: (a) Mean of the eye movement speed following a blink for the different frequencies of the stationary eye trackers in the data, including our reference data. We observe a steep decrease in speed after missing data appear, with values normalizing around 60 ms. **(b)** Variation of the eye movement speed following a blink for the different frequencies of the stationary eye trackers in the data, including our reference data. We observe a steep decrease in variation after missing data appear, normalizing around 120 ms. The vertical dashed lines represent the points identified as the first time the line crossed the fitted function plus epsilon for both **(a)** and **(b)**.

Closed-Eye Time & Length of Missing Data



(a)



(b)

Figure 4.11: (a) Distribution of eye movement speed [deg/ms] from 50 ms preceding missing data for the different frequencies of the stationary eye trackers in the dataset. **(b)** Distribution of the variation of speed [deg/ms] from 50 ms following missing data for the different frequencies of the stationary eye trackers in the dataset. All dashed lines represent an inverse Gaussian distribution fitted to the data with an $R^2 > 0.98$.

Figure 4.12 presents the results of the blink length across the different eye tracker frequencies and eye tracking modalities. Our results show that between frequencies and modalities, the lengths are distinguishable. In Figure 4.13, we show the normalized length distribution for the stationary eye tracker frequencies. We use a Generalized Inverse Gaussian distribution [194] to model the distributions of the lengths, see Figure 4.13. Our regression models yielded an R^2 value of 0.93 for 1000 Hz, 0.98 for 500 Hz, 0.99 for 250 Hz, and 0.67 for 60 Hz.

We investigated how the different types and frequencies as fixed effects affect blink frequency in a mixed effects model with the previously described intercepts and priors in

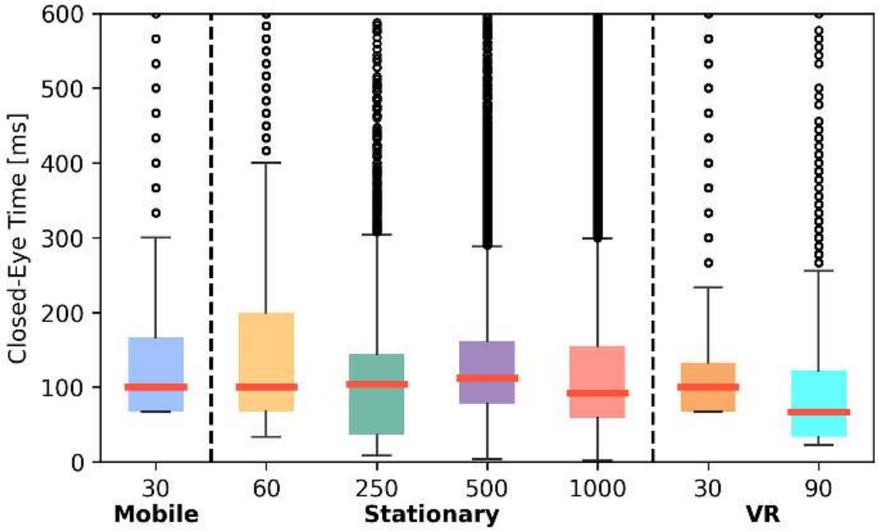


Figure 4.12: We visualized the closed-eye time for the different types of eye trackers gathered in the dataset, i.e., mobile, VR, and stationary, and their respective frequencies. Error bars represent the standard error.

subsection 4.2.5. We found that all combinations of frequency and eye tracker type (stationary, mobile, and VR) had a distinguishable effect on the blink frequency. Our findings are reported in Table 4.4.

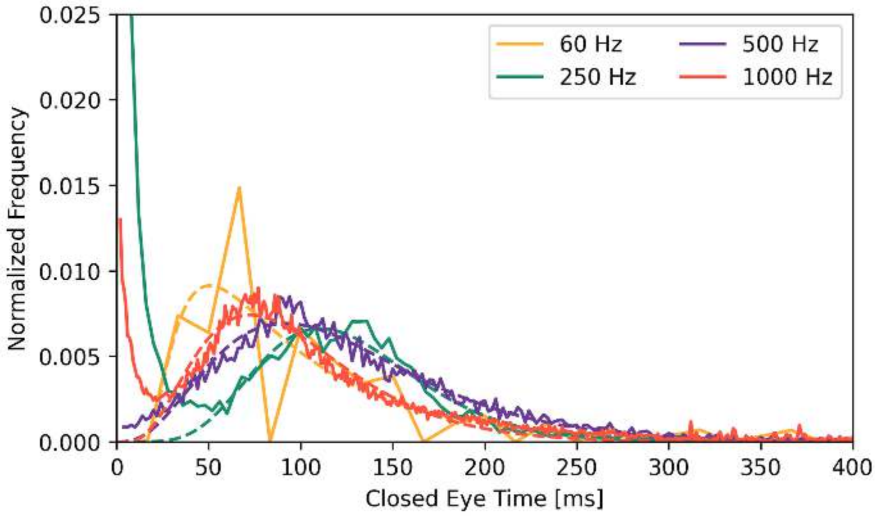


Figure 4.13: We visualized the normalized frequency of closed-eye time for the different frequencies for the stationary eye trackers. The dashed line represents an inverse Gaussian probability density function fitted to the data. The closed-eye time is very similar independent of the frequency or type of eye tracker used, which suggests it is independent of task.

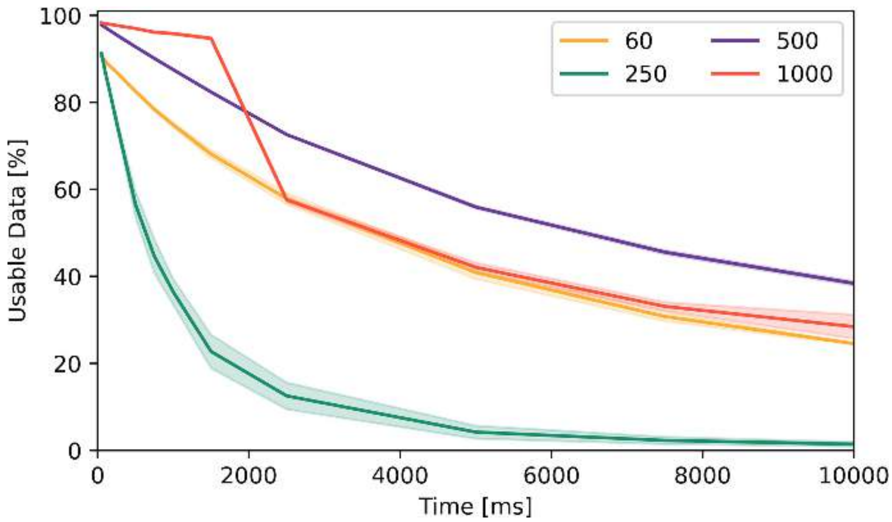


Figure 4.14: We visualize the usable data for the different frequencies available in the stationary eye trackers. We observe a steep decrease in usable windows, where more than half of the data becomes not usable when creating windows of 10 seconds, independent of frequency. The filled area represents the standard deviation.

Baseline Analysis of Removing Sample with Missing Data

In Figure 4.14, we show the relation between the usable data (i.e., data without gaps) and window length. We show that as the window length increases the chance that a window contains one or more gaps increases. We note that the most predominant method for dealing with missing data in a window is to ignore it. We show that applying this method results in a decrease in usable data for further analysis. From an interactive systems point of view, this would hinder user interaction. The data from the different stationary frequencies follow a similar trajectory of a decrease in usable data as the window size increases. We also evaluate the impact of different step sizes[7] on the usable data. These follow all the same trajectory, which suggests that there is no impact of step size on the usable data[8].

Evaluating Infilling Methods

To evaluate the different infilling methods, we used the previously generated distributions from Figure 4.13 to create artificial blinks in our dataset where there were no natural blinks present. We used the most extreme value as an additional cut-off from Table 4.7 to simulate the additional data we should remove when dealing with blinks as this data would be data that contain artifacts from the blink. We applied this to our data from the stationary eye trackers and for each frequency individually to generate roughly 40.000 blinks evenly distributed throughout the data where there are no blinks naturally present, i.e., 750 ms before or after the artificial blink. The generated blinks are about half the actual blinks in the dataset.

[7]Evaluate the window every X samples.

[8]We acknowledge that there are adaptive window size algorithms to deal with gaps; however, adaptive sizes are not traditionally compatible with RNN and LSTM neural network models.

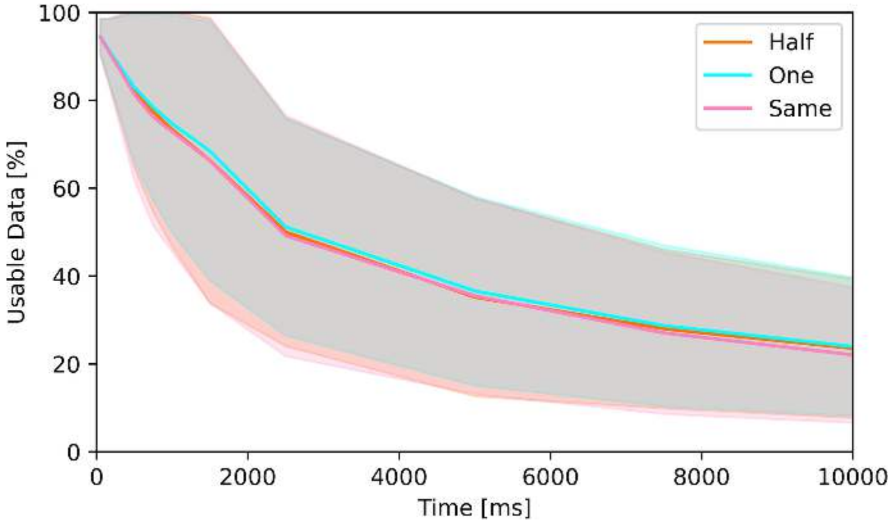


Figure 4.15: We visualize the usable data for different step sizes, where "half" represents half the size of the window, "same" represents the same step size as the window size, and "one" represents one step. The filled area represents the standard deviation. We observe that step size has a similar result in how much data is available across conditions.

We calculated the error as the mean distance between the points generated by the interpolation methods and the actual data in degrees (ground truth). Following this, we applied linear, polynomial (3rd- and 4th-order), cubic spline, and spline interpolation methods over each generated blink and calculated the error. Our results show that the linear interpolation achieves the lowest mean error for 60 and 500 Hz data (0.43 and 0.09 mean degrees error, respectively) and that cubic spline interpolation achieves the lowest mean error for 250 and 1000 Hz data (0.18 and 0.54 mean degrees error, respectively). However, the results between linear and cubic spline are close. Polynomial interpolation on the 4th-order performs worst over all frequencies. We have visualized these findings in Figure 4.16. For more details of all the different errors, see Table 4.8 and Table 4.9.

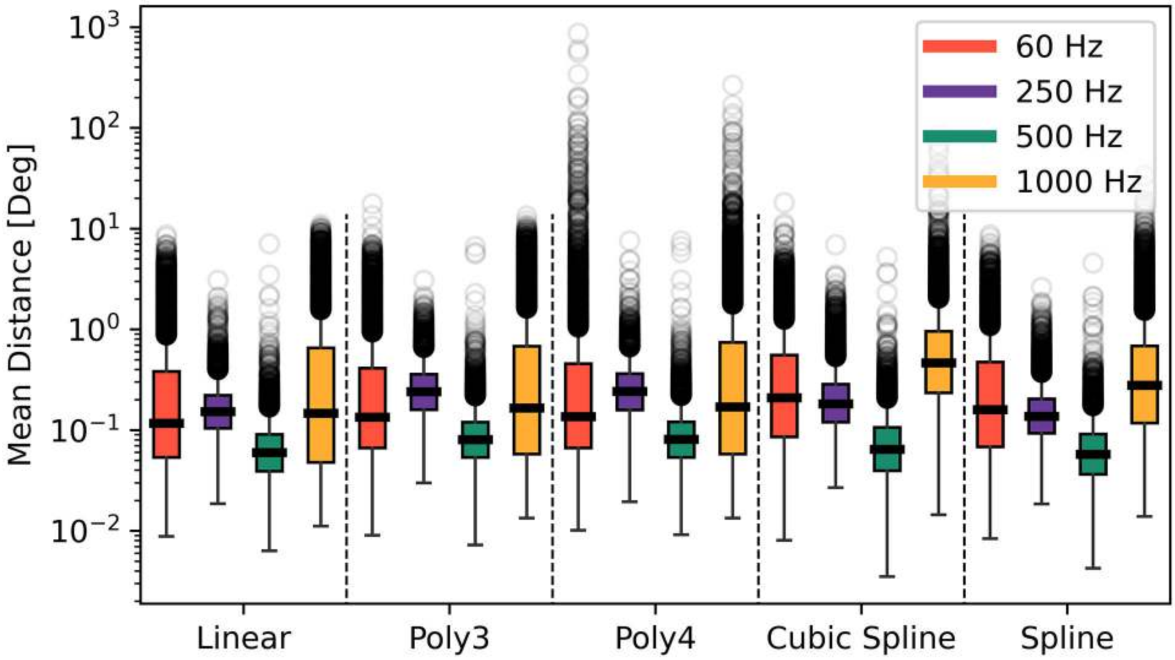


Figure 4.16: Error in mean distance from the ground truth in artificially introduced blinks into the data over different infill methods and frequencies for all stationary eye trackers. We see that the infilling using linear or cubic spline interpolation overall results in the least amount of mean distance; additionally, 4th-order polynomial infilling results in the worse.

Table 4.8: Mean error in degrees for Linear, Poly 3 (3rd order Polynomial), and Poly 4 (4th order Polynomial) infilling methods across different stationary eye tracker frequencies. M = Mean, SD = Standard Deviation, LL = Lower Limit, UP = Upper Limit for 95% CI.

Freq [Hz]	Linear				Poly 3				Poly 4			
	M	SD	LL	UP	M	SD	LL	UP	M	SD	LL	UP
60	0.43	0.79	0.41	0.45	0.48	1.02	0.46	0.51	1.68	15.83	1.31	2.04
250	0.19	0.21	0.19	0.2	0.3	0.3	0.29	0.3	0.31	0.74	0.29	0.33
500	0.09	0.08	0.09	0.09	0.11	0.1	0.11	0.11	0.11	0.21	0.11	0.12
1000	0.56	1.17	0.53	0.58	0.6	1.34	0.58	0.63	1.2	6.74	1.06	1.33
Avg.	0.32	0.56	0.3	0.33	0.37	0.69	0.36	0.39	0.82	5.88	0.69	0.96

Table 4.9: Mean error in degrees for Cubic Spline and Spline infilling methods across different stationary eye tracker frequencies. M = Mean, SD = Standard Deviation, LL = Lower Limit, UP = Upper Limit for 95% CI.

Freq [Hz]	Cubic Spline				Spline			
	M	SD	LL	UP	M	SD	LL	UP
60	0.49	0.86	0.47	0.51	0.61	1.24	0.58	0.64
250	0.18	0.18	0.17	0.18	0.25	0.27	0.25	0.26
500	0.09	0.09	0.09	0.09	0.11	0.12	0.11	0.12
1000	0.54	0.93	0.52	0.56	0.73	1.08	0.70	0.75
Avg.	0.33	0.52	0.32	0.34	0.42	0.68	0.41	0.44

4.3 Eye-Tracking Interpolation Effects on LSTM Model Performance

While previous work has explored methods to handle missing eye-tracking data, the effects of these choices on machine learning performance remain underexplored. In this section, we evaluate how different interpolation methods impact LSTM-based classification models. Our findings show that preprocessing decisions can significantly affect model accuracy and generalizability, underscoring the need for standardized approaches.

4.3.1 Method

We base our analyses on the data and scripts from Annerer-Walcher et al. [7] to evaluate the different infilling methods. We selected this dataset as, from their work, they provide both their data and model structure open-source on `https://osf.io/scmry/`. Their dataset consisted of binocular eye tracking data, including x and y-screen-based coordinates and pupil dilation. It contains information on conditions internal and external focus vs. verbal, numerical, and visuospatial tasks (two conditions $\times$ three tasks). They investigated how consistently different eye parameters respond to internal versus external attentional focus across the three task modalities (verbal, numerical, and visuo-spatial). They report that classifying the focus of attention worked well across participants but that generalizing it across the different tasks had proven challenging.

We use their data and model structure to evaluate the impact of the different infilling methods from peer-reviewed work. As such, we ran their scripts in 3 fold, once without changes, once where we linearly interpolated the gaps and once where the interpolation was done using a cubic spline method, without altering their scripts to preserve the validity of our work. These are the recommended infilling methods as by Grootjen et al. [83]. Following those guidelines, we removed 70 ms preceding and 118 ms following missing data, as the blink can affect these. We used scripts containing the interpolation methods and removed the data preceding and following a blink from Grootjen et al. [83], as these are openly available on `https://eyetrackingguidelines.github.io/`.

Pre-Processing the Data

To preprocess the data, we leveraged the existing scripts from Annerer-Walcher et al. [7]. These existing scripts allowed us to read the different files that are part of the main task. They also provided training before the main task on their open-science framework repository. Even though we could not find in their code that this explicitly was excluded, we assumed it was and thus only used the files from the main task. The SMI RED250mobile system (SensoMotoric Instruments, Germany) with a temporal resolution of 250 Hz, a spatial resolution of $0.03°$, and gaze position accuracy of $0.4°$ that they used for their experiment, writes 0 in the file whenever the eye tracker cannot find the pupil[2]. To leverage the existing scripts for interpolation from Grootjen et al. [83], we replaced all of the “zeros” with “Not a Number’s” (NaN).

Missing Data

Figure 4.17 presents the results of the consecutive NaN's logged. In the work of Grootjen et al. [83], they showed that the closed eye time does not go beyond 1 second. As such, we have split the dataset into different parts once this happens, as the assumption here would be that there is missing values for other reasons than a blink. When looking at the remaining consecutive NaN's in the data we can see that this follows a similar pattern as the one visualized in Grootjen et al. [83] and Holmqvist et al. [101]. We use a Generalized Inverse Gaussian distribution [194] to model the distributions of the lengths, see Figure 4.18. Our regression models yielded an R^2 value of 0.51.

Gaps in the Recording due to Eye-Tracker

When visualizing the data, we found jumps in time that do not follow the 8 ms gap of the 125 Hz recording. Jumps in time between samples were on average 11.9 ms long with a standard deviation of 234 ms. In Figure 4.19, we visualize the number of seconds between two lines being logged into the different files in seconds that are over 8 ms (+ 10%). In total, the dataset contains 78.755 gaps in recording that are over this limit (m = 1324.9 ms, std = 4076.2 ms).

While sensors, such as eye trackers, deliver samples at a given frequency, the time is typically not precise. It can even happen that one or multiple consecutive samples are arbitrarily dropped. For this reason, we standardize the sampling frequency to perfect 125 Hz to counteract these gaps in the eye tracking data. Thus, we resampled the data to be exactly $1/frequency = 8msec$ apart. This allows us to investigate potential missing samples and gaps in the data and, thus, test infilling methods. During this process, we did not infill data; we merely ensured a perfect sample frequency of 125 Hz. If a sample was missing in the original dataset, the data was added at the correct time and marked as NaN for later processing using

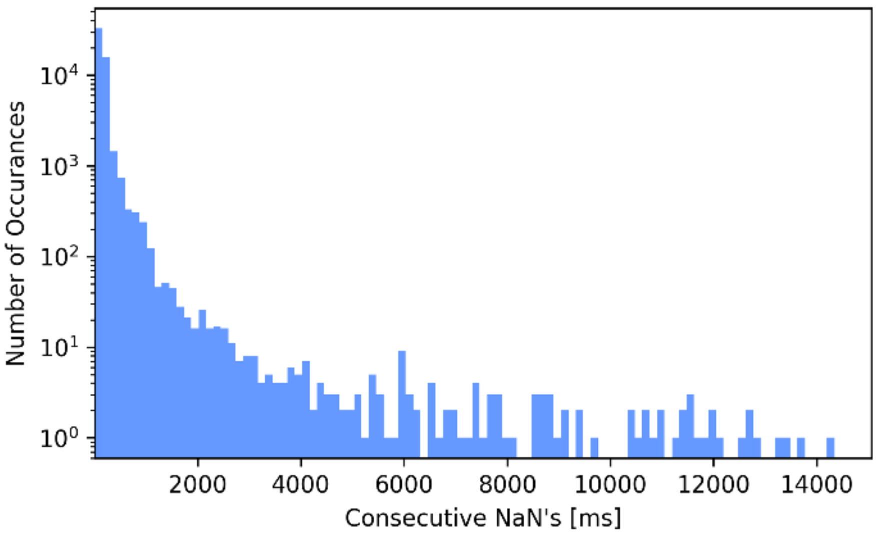


Figure 4.17: The log-scaled distribution of consecutive NaN's appearing in the data.

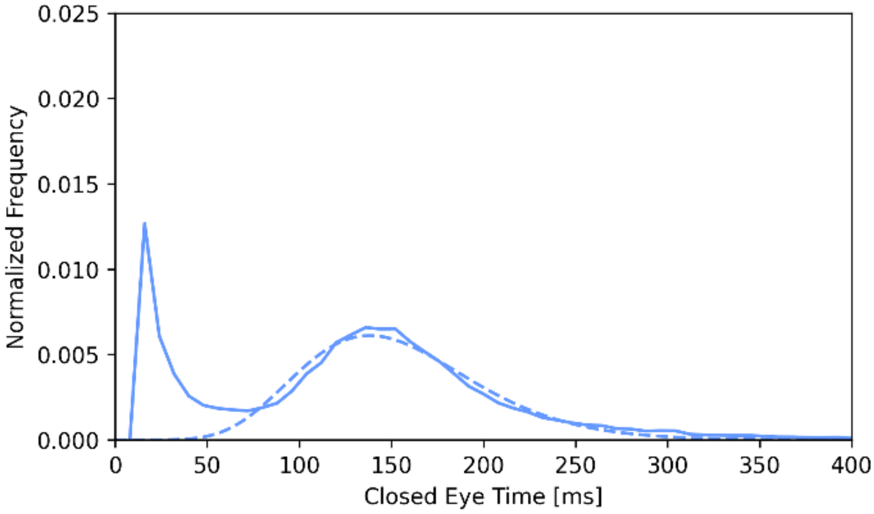


Figure 4.18: We visualize the normalized frequency of closed-eye time for the dataset. The dashed line represents an inverse Gaussian probability density function fitted to the data ($R^2 = 0.51$).

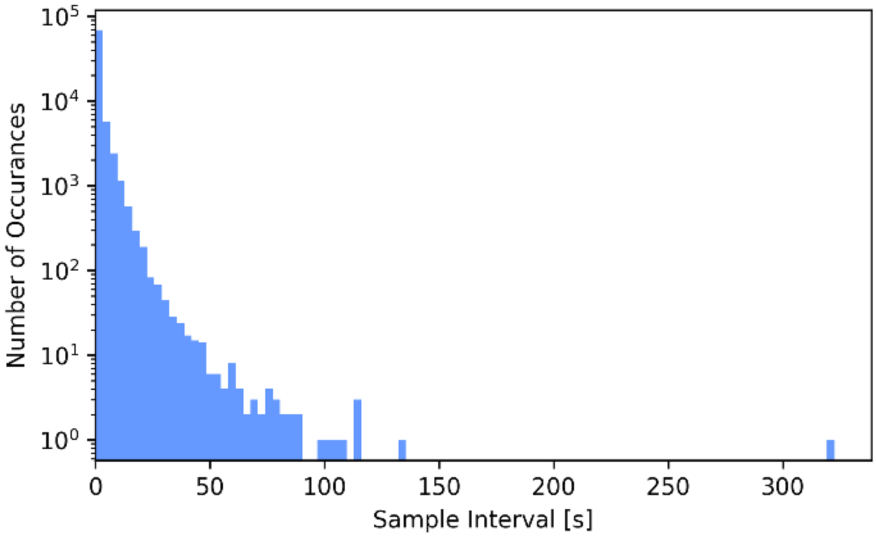


Figure 4.19: Illustration highlighting the interval (time between) of the logged samples over 8.8 ms (which should be the normal interval of a 125 Hz recording + allowance of 10% in variation).

the different infilling methods.

4.3.2 Results

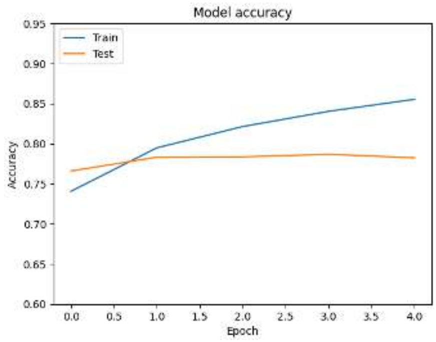
In total, the dataset has 78.755 gaps in the recording where the eye tracker did not log any data and 52.219 sections of data where there were consecutive 0's (by us converted to NaN's) logged because the eye tracker was not able to recognize the pupil. We then resampled the data frame to allow for consistent steps in time according to the 125 Hz logging in the remaining of data. We then split the dataset in parts where there were more than 1500 consecutive NaN's in the dataset as the long short-term memory (LSTM) neural network (NN) provided by Annerer-Walcher et al. [7] takes in a window of 1500 samples. In the initial dataset, having 1500 or more consecutive 0's happening only existed nine times; after our pre-processing, this happened 1198 times.

We infilled the dataset linearly and used cubic spline to fill in the remaining gaps in the data. For the cubic spline infilling, we take three samples at the start and end into consideration to allow the cubic spline interpolation to consider the velocity. If there was a gap of a single sample during the cubic spline interpolation, we linearly interpolated this, as cubic spline interpolation is not warranted in these scenarios. If there are NaN's in the 3 samples before or after the gap, we recusively go back or forward, respectively, until we find the allotted samples without NaN's. After interpolating, we had 3 datasets, one as provided and processed exactly by Annerer-Walcher et al. [7], one linearly interpolated, and one cubic spline interpolated. These interpolation methods were applied over x and y screen-based coordinates for both eyes and the pupil dilation values for both eyes. Both of the interpolated datasets were without any missing samples or NaN's remaining.

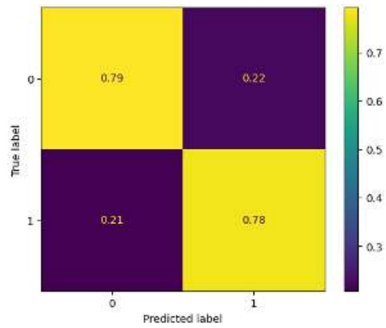
In the work of Annerer-Walcher et al. [7], they have used 157 participants from the whole dataset and excluded 9. Of this data, 135 sessions were used as training data to adjust the model parameters, and the remainder were used as testing data, pooling data from all tasks. Unfortunately, we could not identify in their publication or from the OSF page which participants were excluded and which sessions were used for which purposes. As such, we split the whole dataset, using a fixed seed, based on all participants (166), using 60% of the participants (99) for training, 20% for testing (33), and 20% for validation (33). We used the existing processing scripts after the infilling to generate the input for the LSTM model. We left the model's hyperparameters unchanged in the code.

As such, our model starts with the input layer of 1500×16, followed by an LSTM layer of 64 units, after which there was a dense layer present of 64 units with a ReLu activation function with a dropout layer of 0.45 following the dense layer. After the dense layer, the model contains a convolutional layer followed by a max pooling layer as 1D. The final two layers contain another LSTM layer of 32 units and a final dense layer of 2 for the output. The model uses a nadam optimizer with a fixed learning rate of 0.001. The models were trained over 40 epochs in batches of 30 with an early stopping rule and a patience of 3 on the validation loss.

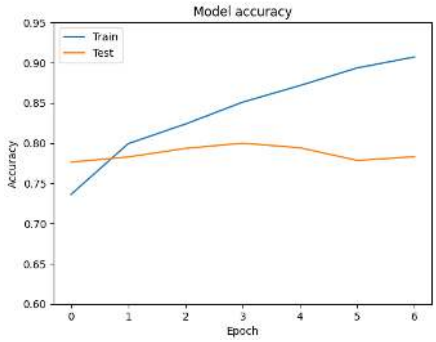
Our reference model yielded a training accuracy after five epochs of 0.8713, a test accuracy of 0.7699, and a validation accuracy of 0.7863. We achieved this with 6813 sets of windows for training, 2490 sets for testing, and 2176 sets for validation for 11,479 sets of windows. The model after linear interpolation yields a training accuracy after six epochs of 0.8834, a test accuracy of 0.7732, and a validation accuracy of 0.7946. We achieved this with 8765 sets of windows for training, 3157 sets for testing, and 2795 sets for validation for 14,717 sets of windows (28.2% more data over default). Our final model using spline interpolation yields a training accuracy after seven epochs of 0.9023, a test accuracy of 0.7555, and a validation accuracy of 0.7599. For this, we had 8732 sets of windows for training, 3141 for testing, and 2790 sets for validating (27.7% more data over default).



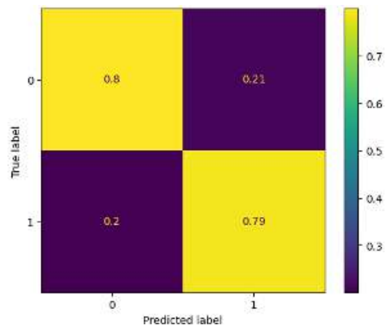
(a) Default



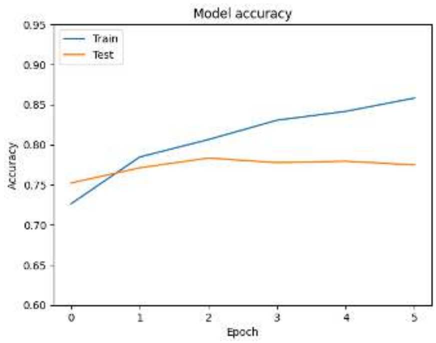
(b) Default



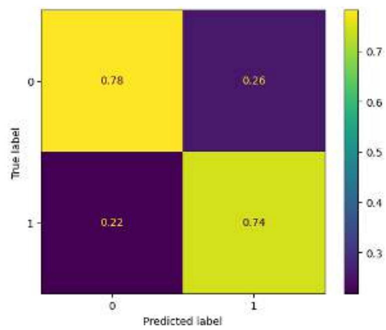
(c) Linear



(d) Linear



(e) Cubic Spline



(f) Cubic Spline

Figure 4.20:

Figure 4.21: Left column: (**a**, **c** and **e**): Train and test accuracy of the LSTM model for three models. Right column: (**b**, **d** and **f**): Corresponding confusion matrices from predictions on the validation set.

4.4 Chapter Summary

In this Chapter, we investigated the impact of blinks on eye-tracking data quality and the downstream consequences for interactive systems and predictive modeling. Blinks are short, involuntary closures of the eyelids, necessary for ocular health, yet disruptive in the context of gaze-based interaction. As eye-tracking becomes more embedded in real-time applications—such as attention-aware systems or adaptive user interfaces—the presence of missing or distorted gaze data caused by blinks poses a critical challenge. We observed that these disruptions are rarely handled in a consistent manner, threatening the reproducibility and reliability of gaze-based research and applications.

To better understand current practices, we conducted a systematic literature review following PRISMA guidelines. The goal was to assess how previous work detects and handles blink-related missing data. We focused on three aspects: the method used to deal with missing data, the detection mechanism for identifying blinks, and the task context in which gaze data was collected. From 140 reviewed papers, we found that nearly half did not report sufficient detail on their blink-handling strategies. Of the remaining 81 papers, we identified a wide range of approaches, including removal, linear interpolation, spline and polynomial interpolation, as well as other forms of imputation and signal reconstruction. The review uncovered a high degree of methodological inconsistency across papers and tasks, providing options for answering **RQ 2:** *How should we deal with gaps in eye-tracking data for machine learning models?* .

To further address **RQ 2**, we explored the use and implementation of blink detectors across studies and eye tracker types. We found that while commercial systems like EyeLink and SMI offer integrated blink detection software, these are often not utilized or reported. Instead, researchers frequently treat missing data points as implicit indicators of blinks. However, the lack of precise blink parsing, especially for detecting artifacts before and after the blink, leads to suboptimal data treatment and possible retention of distorted signal segments.

After this, we evaluated the performance and implications of different interpolation techniques using real and synthetic datasets. Specifically, we visualized and compared linear, spline, cubic spline, and polynomial interpolation under various blink scenarios. Our results showed that linear and cubic spline interpolation produced the lowest mean error in reconstructing gaze trajectories. Polynomial interpolation, particularly of higher order, was more prone to artifacts. We also quantified the influence of removing too little or too much data around the blink, finding that poor cut-off choices significantly impacted interpolation quality.

We complemented these analyses with an empirical study using 11 open eye-tracking datasets, from both stationary and VR devices, where we investigated blink frequency, duration, and inter-blink intervals. We also conducted a high-speed video validation study to confirm the relationship between signal gaps and eyelid movements. From this, we derived frequency-specific thresholds for artifact trimming before and after blinks, extending previous work and providing empirical grounding for preprocessing pipelines.

To assess how preprocessing decisions influence machine learning outcomes, we applied the identified methods in a practical modeling context. We retrained an LSTM-based classification model on versions of a benchmark dataset that had undergone different preprocessing steps. Models trained on data interpolated with linear or cubic spline methods not only improved validation accuracy but also benefited from increased usable data—over 27% more samples—demonstrating the practical value of retaining blink-afflicted data through interpolation rather than removal.

Our work shows that while blinks are a natural part of human vision, they are often mishandled in gaze interaction systems. This chapter presents the first comprehensive overview and empirical assessment of blink-related challenges in eye tracking. By analyzing the prevalence of different methods in the literature, evaluating their effects on both data integrity and model performance, and proposing empirically validated preprocessing recommendations, we provide a standardized pipeline for blink-aware gaze analysis.

We conclude that standardizing blink detection, artifact trimming, and interpolation can substantially improve the robustness, fairness, and reproducibility of gaze-based systems. The recommendations derived in this chapter serve as a foundation for more reliable research practices and higher-quality interactive systems—laying the groundwork for future work in gaze-based modeling, attention-aware interaction, and physiological computing.

5
APPLICATIONS

Central Field Loss (CFL) is a hallmark symptom of progressive retinal diseases such as macular degeneration, often emerging subtly and remaining unnoticed in early stages [71]. Detecting CFL early and monitoring its progression continuously is critical for timely intervention and rehabilitation. Traditionally, such detection relies on clinical tests that are periodic and sparse, limiting their capacity to capture the temporal dynamics of functional vision loss. To make testing more timely and accessible, we investigate how eye-tracking, collected during a visually guided task, can support the continuous detecting of CFL. For this, we emphasize low-latency and cost-effective classification methods.

To achieve reproducibility and isolate the effects of CFL, we use an available dataset from Barraza-Bernal et al. [20], which employs a simulated gaze-contingent scotoma. This approach is widely accepted in vision science (e.g., [232]) and allows for controlled experimentation with a fixed scotoma size of 6°, a threshold known to sometimes go unnoticed [71]. Using this data, we compare traditional machine learning models (Support Vector Machines and Random Forests) with deep learning approaches (Long Short-Term Memory networks) in terms of classification accuracy and computational cost.

The primary goal of this chapter is to answer the following research question:

RQ 3: *Is real-time detection of central field loss possible using eye-tracking data?*

To answer it, we preprocess the gaze data using established algorithms, extract relevant features such as fixation duration and saccade amplitude, and train the models using leave-one-participant-out cross-validation. We then analyze feature importance (e.g., pupil size; see Figure 5.2), model performance (Table 5.1), and cost-benefit trade-offs (Figure 5.4b). The findings provide insights into how CFL alters gaze behavior and whether these changes can be reliably and efficiently detected using unobtrusive eye tracking technology[1].

This chapter is based on the following publication.

Grootjen, Jesse W. et al. ‘Assessing Eye Tracking for Continuous Central Field Loss Monitoring.’ In: *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia*. MUM ’23. Vienna, Austria: Association for Computing Machinery, 2023, pp. 54–64. DOI: 10.1145/3626705.3627776

[1]The following sections are copied with minor changes from the publications: subsection 5.1.1 until subsection 5.1.2

5.1 Eye Tracking for Continuous Central Field Loss Monitoring

This section outlines our approach for using eye tracking to detect central field loss (CFL). We describe the dataset, preprocessing steps, and models used—ranging from traditional machine learning to deep learning—alongside an evaluation framework that quantifies the trade-off between classification performance and computational cost across varying temporal resolutions.

5.1.1 Methods

This section details the steps involved in the proposed approach for determining the cost-benefit trade-off and estimation accuracy of central field loss. First, we describe the data acquisition process. Next, we describe the pre-processing. This is followed by describing the model used to train and validate both the machine learning and deep learning approaches. Finally, we explain how we evaluated the cost associated with the different models and how this relates to variable window sizes for the deep learning approach.

Dataset and Experimental Description

Our experiment used eye-tracking data from Barraza-Bernal et al. [20]. We decided to use a dataset where participants had a simulated CFL to ensure the reproducibility of our results across participants. This ensures that the CFL does not vary as would with participants affected by CFL (i.e., shape and size of CFL may vary among affected participants). Using a dataset from a study where participants use a simulated visual impairment is commonly used to study phenomena in vision research e.g., Sipatchin et al. [232]. Five volunteers ($\bar{x}$ = 28.8 years) with normal or corrected-to-normal vision participated in the study. Every participant was trained to acquire a peripheral retinal locus of fixation after four training sessions using a simulated gaze-contingent CFL of six degrees, totaling between two to three hours of training per participant [19]. Eye movements were recorded using an SR Research EyeLink 1000 Plus eye tracker with a spatial resolution of $0.01°$ and a sampling rate of 1000 Hz [20]. The participants were asked to perform a visual discrimination task where the participant was asked to discriminate between the presence of having more red or blue dots on the screen by pressing the arrow up or down on the keyboard (see Figure 5.1). We used a subset of the data consisting of two CFL simulations: one simulation without CFL (i.e., the baseline) and a simulation with a simulated gaze-contingent CFL of $6°$, representing CFL that can go unnoticed [71].

Machine Learning

Before training our machine learning models, we prepared the eye-tracking data by extracting meaningful features from raw gaze signals. The following section details the preprocessing steps applied to identify and segment fixations, saccades, and blinks, forming the input for our classifiers.

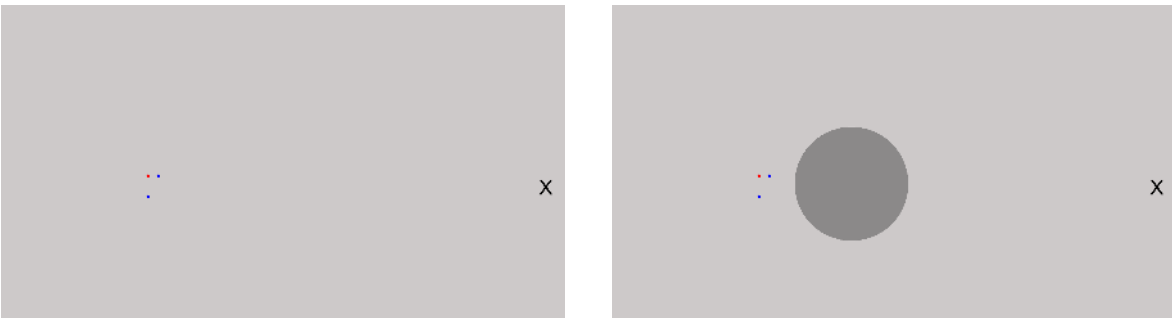
Dataset Preprocessing

We identified fixations, saccades, and eye blinks using the EyeLink parsing algorithm and adopted the following cutoffs based on previous related work [262]: saccadic velocity threshold of $30°/s$, a saccadic acceleration threshold of $8000°/s^2$, and motion threshold of $0.1°$. Eye movement data below the velocity threshold and acceleration threshold criteria were classified as a fixation. Otherwise, it was labeled a saccade eye movement. For the saccade, we collected the following features: duration, average position of the gaze in X and Y coordinates, and average pupil dilation. For the fixations, we extracted the duration of the saccade, the start and end points in x and y coordinates, the amplitude in degrees, and the peak velocity in degrees per second. For the blinks, we simply extracted the duration of the blink. All cells containing no values (e.g., NaN) were set to an arbitrary number that does not occur (-1) to enable these rows of data to be fed into the models described below.

Model Description

We decided to use a Support Vector Machine (SVM) and Random Forests (RFs) to evaluate the CFL prediction accuracy. CFL leads to a linear gaze deviation. SVMs and RFs are linear statistical models and, hence, suitable to predict CFL. We analyzed the eye movement features mentioned above using hyperparameter optimization for RF and SVM models. SVMs were our first choice as a model as it has been extensively used in classification problems, especially in eye-tracking research. Thus, using SVM as a baseline method allows comparisons with previous and future work. SVM maps its input vectors into a high dimensional feature space through a chosen non-linear mapping and then finds an optimal hyperplane to separate the classes with a maximal margin, which reduces the generalization error [171]. This work uses the SVM implementation in the Python library Scikit-learn [36].

RF is a classifier comprising an ensemble of randomized decision trees, which make a joint decision on the class [32]. Like SVM, RF has been used with good results in various tasks. First, we applied RF to compare SVM with at least one other popular method. In addition to its demonstrated practical usefulness, RF models tend to be robust, which is important



(a) (b)

Figure 5.1: (a): Baseline without simulated CFL. **(b):** Simulated gaze contingent CFL of $6°$ of the experiment.

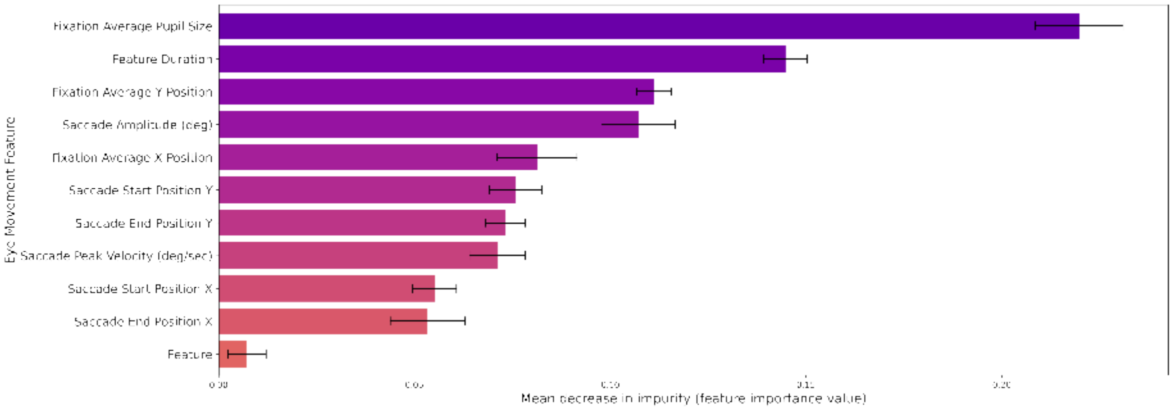


Figure 5.2: Feature importances measured by the mean decrease of Gini-impurity for the leave-one-participant-out nested cross-validation. The pupil size achieves the highest importance among all features.

as this work could generalize to health diagnostics, generalization capability, and intrinsic feature selection opportunities embedded in decision tree-based methods [54]. Here, we applied the RF and SVM implementation from *scikit-learn*[2] [36].

An appropriate configuration of the hyperparameters is necessary to produce a model with the best performance for the problem. For this, we used grid search for exploring the hyperparameter space, using *scikit-learn*. We opted to do an exhaustive search of the most suitable kernel type, C (i.e., the regularization parameters), and γ (i.e., the radial basis width) for the SVM.

Similarly, we did an exhaustive search to find the optimal number of trees for the RF (i.e., denoted as $n\_estimators$), the criterion, maximal depth of the trees in the forest, the minimal sample split, and the $max\_features$ parameter. Having a larger number of trees in the forest increases the classification efficiency at the cost of increasing the computational time required to train a model. When splitting a node in the decision tree, the feature used for the split is selected from a random subset of features. The number of features chosen for this subset is determined by the $max\_feature$ parameter. The other RF parameters remained in their default settings.

Deep Learning

To train our deep learning models on raw gaze sequences, we first segmented the continuous eye-tracking data into overlapping temporal windows. The following section outlines the preprocessing steps used to construct these windows and normalize the input data for effective temporal modeling.

[2] `www.scikit-learn.org/stable`

Table 5.1: Accuracy, precision, recall, weighted F_1 scores and AUC PR of the leave-one-participant-out cross-validation for the best models from the hyper-parameters optimization for the feature-based models.

Model	Parameter	Accuracy	Precision	Recall	Weighted F_1-Score	AUC PR
SVM	C: 1000 Kernel: rbf	.730	.736	.722	.726	.821
RF	Criterion: Gini Max Depth: None Max Features: Auto Min Sample Split: 6 Number of Estimators: 900	**.769**	**.758**	.750	**.766**	**.863**
RF	Criterion: Gini Max Depth: None Max Features: Auto Min Sample Split: 9 Number of Estimators: 300	.766	.756	**.752**	.765	.860

Dataset Preprocessing

For the deep learning models, one second of data was removed from the eye movement data before and after the calibration and validation sequence. The following windows represent the data interval that the model needs to estimate the likelihood of CFL. The window sizes vary from 50 msec to 10,000 msec of data, increasing in steps of 50 msec. The overlap between windows is always exactly half of the window size, e.g., if the window size is 50 msec, the window is moved 25 msec at a time. The window is removed if the window contains invalid values (e.g., because of blinks). Following this, we normalized all windows by moving the series of data points at the start of each window to the coordinate 0,0. Afterward, we turned the window such that the data point following exactly points up. Each window contains the x and y screen gaze coordinates (i.e., the pixel coordinates) and the pupil size value in the sum of the number of pixels inside the detected pupil contour.

Model Description

We use a Long Short-Term Memory (LSTM) network for learning temporal long-term dependencies, especially when predicting time-series which is usually the case when classifying gaze data. The input layer of our model varied in size to accommodate the window size, e.g., when there was a window size of 50 msec, the input layer was 50×3 (x, y, pupil dilation). After the input layer, we used an average pool layer with a pool size of 3. The third layer is our first LSTM layer with 32 nodes. The fourth layer was another average pool layer with a pool size of 3. The fifth layer was a dropout layer with the dropout set to 0.5. The sixth layer is our second LSTM layer with 20 nodes. After which, we have a second dropout layer, again set to 0.5. Followed by the output layer with 2 nodes representing the no-CFL or CFL condition.

We trained the model using an exponential decay learning rate scheduler, with an initial learning date of 0.0001, decay steps of 20,000, and a decay rate of 0.98. We enabled early stopping while monitoring the validation precision-recall with a patience 10. We set the maximum number of epochs to 500 and the batch size of 128. All models were trained with a loss function for binary cross-entropy and precision-recall area under the curve for the metrics, using one GPU per run (Tesla V100), in TensorFlow using the Keras API [45].

Evaluation

All models were evaluated using a leave-one-participant-out nested cross-validation. All run times of final models are done on a MacBook Pro 13" (i7 2.8 GHz, 16 GB, Intel 655). For the cost function, we take the run time into account that a single entry requires to be evaluated by our models. We denote the cost as the following:

$$Cost = \begin{cases} t, & \text{if } t < w \\ 2t, & \text{if } t \geq w \end{cases} \tag{5.1}$$

t is the time in msec needed to evaluate a single entry, and w is the duration of the feature or window size in msec, depending on whether we evaluate our approach via machine or deep learning. The cost function is a linear function of the time required to evaluate an entry. However, if the time required for evaluating the entry is larger than the duration or window size, there will be a penalty of a factor of two for the time.

5.1.2 Results

We compared the classification accuracy between traditional machine learning (i.e., SVM, RF) and deep learning (i.e., LSTM), after which we compared these models using the cost-benefit function. We evaluate a general classification model by investigating the classification performance through leave-one-participant-out cross-validation. This means that we use all participant data for training except for one participant, which we use for evaluation. Semantically, this approach learns a model without knowing anything about the person in advance and predicts the presence of CFL independent of individual context and differences. We describe the results in the following.

Machine Learning

We applied the methods of section 5.1.1, which resulted in 11 features and 33,583 entries for these features combined. We then trained and validated the model with 100 cycles for precision and recall and using a leave-one-participant-out cross-validation for the two classification techniques and their combinations of parameters. Table 5.1 shows an overview of the best results. The column *model* indicates the machine learning method used to produce the model. The column "parameter" column indicates the parameters used to initialize the model. The precision and recall columns contain the average score with their standard deviation.

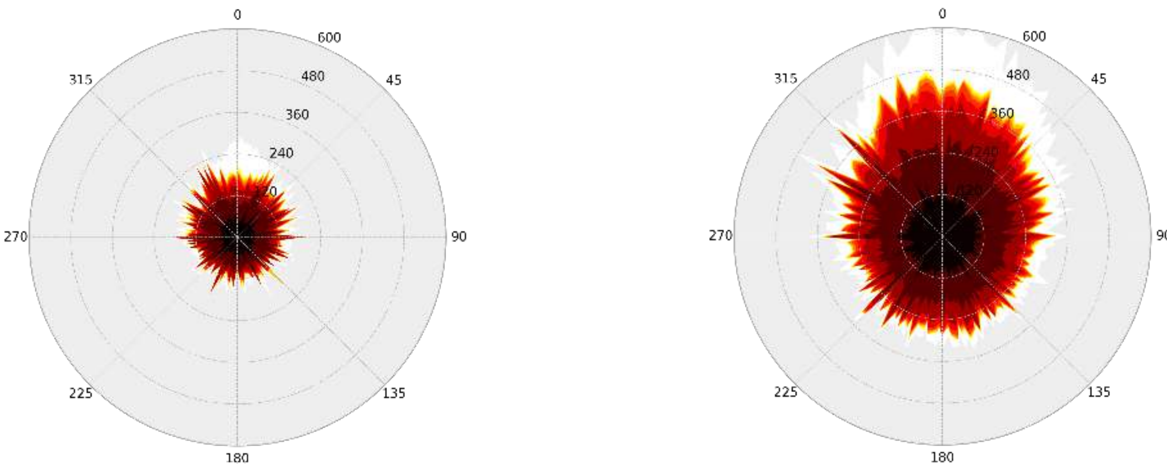
Support Vector Machine

The SVM showed the best performance by using $C = 1000$ and with a radial basis function kernel, resulting in an accuracy of $.73$ (precision: $.72$, recall: $.73$, weighted F_1 score: $.73$). We found a single set of parameters that performed best for precision and recall.

Random Forests

The RF classifier outperformed the SVM classifier for optimizing for precision and recall with two sets of parameters as shown in Table 5.1. We found the difference in the two models on two features, namely, the minimum sample split and the number of estimators. While the best-performing model for precision uses a minimum sample split of 6 and 900 estimators, the best model for recall uses values of 9 and 300, respectively.

Analyzing the most relevant features shows us that the average pupil size accounts for .238 of the Mean Decrease in Impurity (MDI), which can be seen in Figure 5.2. Followed by the saccadic amplitude, the duration of the feature and the average y position during a fixation. Together, these four features make up over half of the MDI.



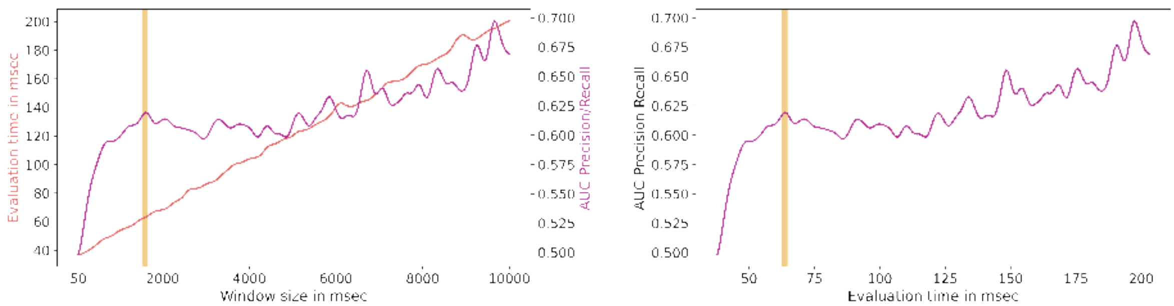
(a) (b)

Figure 5.3: Rose plot of eye movements after pre-processing for deep learning for participant 5 for a window size of 1250ms. It shows where eye movements are going towards and how many points are overlapping. Darker is more. **(a)** Without simulated CFL and **(b)** with simulated CFL.

Deep Learning

We applied the methods described in section 5.1.1 to train our LSTM models. We trained and validated our models using the leave-one-participant-out cross-validation, resulting in a total of 1000 trained models. Each is trained on a different combination of window sizes and participants. As visualized in Figure 5.4a, we can see a general increase in mean Area Under the Curve (AUC) for precision and recall as we increase the window size for our LSTM models. This is to be expected.

More interestingly, the steepest AUC benefit is achieved at ca. 1,600 msec. Accuracy increases are minimal and variable after 5 seconds of window size, although increasing the window size to 9950 msec generated accuracy benefits of up to .725. However, increasing the window size also tends to capture more blinks. This resulted in missing values, resulting in sparser data. We visualize the eye-tracking data for windows sizes of 1250 msec in Figure 5.3. Here we see all the individual data points visualized after processing for participant five for both conditions. The data points scatter more during the CFL simulation conditions than those without CFL.



(a) (b)

Figure 5.4: (a): Mean evaluation time plotted on the left-y-axis, mean AUC Precision-Recall plotted on the right-y-axis,the accompanying window size on the x-axis. The yellow line denotes the trade-off between the required window size and classification accuracy. A reasonable trade-off is achieved at ca. 1600 msec, since longer window sizes degrade the classification performance (e.g., due to loss of data through eye blinks). **(b):** Mean AUC Precision/Recall plotted on the y-axis and the accompanying evaluation time on the x-axis.

Evaluation

We applied the methods from section 5.1.1. We ran each model 100 times, evaluating a single data entry. In Figure 5.4a, we show a linear relation between mean evaluation time and the window size, denoted as $Evaluation\ Time = 0.00415 \times window\ size + 37$. We also see that the penalty function described in section 5.1.1 never applies, as the evaluation time stays well below the window size.

When visualizing the cost described by the linear equation above and the benefit from the increase in window size as visualized in Figure 5.4b we can see that there is a lot of benefit at the start while the evaluation time only increases slightly. After this, there is little gain in the AUC PR while the evaluation time keeps linearly increasing. Given that the information becomes sparser with the increase in window size, we recommend using a window size of 1600 msec for the best cost-benefit trade-off (see Figure 5.4b).

5.2 Chapter Summary

In this chapter, we investigated whether eye tracking can be used to continuously and efficiently detect the presence of central field loss (CFL), with the goal of enabling unobtrusive and real-time monitoring. We addressed this by comparing traditional feature-based machine learning models with deep learning approaches trained on raw temporal sequences of gaze and pupil data. Our approach utilized a dataset with simulated CFL from Barraza-Bernal et al. [20], which allowed for reproducible evaluation across controlled conditions.

RQ 3: *Is real-time detection of central field loss possible using eye-tracking data?*

This question was addressed by developing and evaluating multiple classification models. We found that both machine learning and deep learning approaches are capable of distinguishing between normal and CFL-affected gaze behavior with reasonable accuracy, with Random Forests reaching an accuracy of 0.769 and LSTMs achieving comparable precision with window sizes around 1600 msec.

In the first part of the chapter (section 5.1.1), we extracted a set of 11 eye movement features (e.g., fixation duration, saccade amplitude, and pupil size). These were used to train SVM and RF classifiers using leave-one-participant-out cross-validation. Random Forests outperformed SVMs in terms of accuracy and area under the precision-recall curve (AUC PR), with pupil size emerging as the most important predictive feature (Figure 5.2).

In the second part of the chapter (section 5.1.1), we trained LSTM models on raw gaze coordinates and pupil size. We varied the input window size from 50 msec to 10,000 msec and evaluated performance across all settings. We found that performance improves with larger window sizes, peaking around 1600 msec, after which gains plateau or degrade due to increased data sparsity (e.g., from blinks; see Figure 5.4a).

In the final part (section 5.1.1), we introduced a cost model based on evaluation time relative to window size. By plotting AUC PR against runtime (Figure 5.4b), we identified a cost-efficient sweet spot at a window size of 1600 msec. Beyond this, the marginal benefit in accuracy diminishes while evaluation cost continues to rise.

Overall, this chapter demonstrates the feasibility of using eye tracking to support continuous, low-latency monitoring of central field loss and provides design recommendations for window-based gaze analysis in real-time applications.

6 DISCUSSION

This dissertation contributes to the dynamic environment of preventive healthcare by demonstrating that gaze behavior can serve as a non-invasive, cost-efficient, and scalable biomarker for early detection of visual impairments. Across three interrelated studies, we show that changes in oculomotor dynamics (such as fixation instability, altered saccadic patterns, and pupil responses) can reveal subtle signs of functional vision loss before individuals become consciously aware of symptoms. This aligns with global healthcare priorities that emphasize early intervention to reduce the burden of chronic, progressive diseases [168, 278].

Unlike conventional ophthalmological screening, which is episodic, equipment-intensive, and typically initiated only after symptomatic onset, gaze-based methods offer a passive and ubiquitous alternative. Our approaches demonstrates that simulated impairments, validated preprocessing pipelines, and machine learning can jointly support unobtrusive monitoring suitable for home or community-based settings. This is especially important in marginalized populations or aging societies where early clinical access may be limited.

By addressing both technical feasibility and behavioral validity, the findings in this dissertation advance the integration of ocular health into broader preventive care frameworks. The models and methodologies introduced herein not only highlight the diagnostic potential of eye tracking but also set the stage for future systems that can identify visual impairments silently progressing in the background—thereby enabling timely referrals, targeted interventions, and improved long-term outcomes.

Chapter 3 addressed **RQ 1:** *Can gaze patterns in VR simulations reveal signs of visual impairments?* By analyzing gaze behavior during simulated visual impairments—specifically cataracts, glaucoma, and age-related macular degeneration (AMD)—we found that gaze metrics are sensitive to impairment severity, particularly in later stages, and can provide early indicators before participants become consciously aware of perceptual degradation.

In the glaucoma and cataract simulations, we observed distinct and condition-specific gaze adaptations. For glaucoma, increasing peripheral vision loss was associated with longer task durations, reduced saccadic amplitudes, and lower pupillary activity. These effects mirror real-world deficits observed in glaucoma patients and align with prior simulation work [234, 244]. In cataract simulations, we found similar signs of disruption: task completion time increased, fixation rate rose, and saccade frequency dropped with greater simulated severity. Strikingly, over 60% of participants remained unaware of the simulated monocular cataract progression, even as these measurable gaze changes emerged. This dissociation between subjective awareness and objective behavior highlights the potential of gaze-based systems to detect early impairment stages—critical for preventive healthcare [278].

The AMD simulation, however, revealed more limited gaze-based effects in early-stage central field loss. While prior studies have documented compensatory strategies such as PRLs and fixation instability in real patients [39, 51], our findings suggest that subtle CFL simulations may not sufficiently disrupt behavior in short visual search tasks to produce robust gaze signatures. This may reflect both the strength of contextual cueing in our realistic VR scenes and the need for longer or more cognitively demanding tasks to elicit AMD-specific adaptations.

A key contribution of this chapter was the validation and refinement of simulation fidelity. Through structured interviews with individuals who had experienced cataracts, we learned that common simulation methods—such as uniform color filters or blur overlays—were often described as unrealistic. In response, we implemented alternatives, including luminance compression for contrast loss and perceptually grounded color shifts. Participants reported that these techniques more accurately reflected their lived visual experiences, increasing the ecological validity of our simulation pipeline [133]. These validated parameters were then used in user studies to probe how different severities of simulated impairments affect gaze.

Overall, our results show that when VR simulations are perceptually validated and combined with gaze analysis, they can reveal subtle signs of visual impairments. The strongest effects emerge in moderate to severe impairment conditions, but early markers—such as increased fixation rate or saccadic changes—can appear before participants themselves notice changes. These findings offer a foundation for low-burden, gaze-based screening tools that align with preventive health goals. By identifying impairments through naturalistic tasks and without requiring explicit input, such systems could support earlier interventions and reduce the societal burden of undiagnosed vision loss.

Chapter 4 addressed **RQ 2:** *How should we deal with gaps in eye-tracking data for machine learning models?* through a combination of a literature review, empirical validation, and machine learning analysis, we provide a comprehensive answer that not only quantifies the challenges posed by blink-induced gaps but also identifies effective strategies for mitigating their impact on gaze-based prediction systems.

Taken together, **Chapter 4** answers **RQ 2** by providing concrete, evidence-based guidelines for detecting, trimming, and interpolating blink-related gaps in eye-tracking data. These contributions advance the state of the art in both gaze-based machine learning and real-time interactive systems and pave the way for more robust and reproducible models of visual attention and impairment detection.

Our findings demonstrate that blink-related interruptions are both prevalent and methodologically under-addressed in the literature. As established in the related work [83, 101], blinks are physiologically complex events that not only result in missing gaze data but also distort the signal before and after the closure. This aligns with prior research on blink dynamics [65, 259], which shows that cognitive and environmental factors modulate blink characteristics, complicating their treatment in gaze data pipelines.

The structured literature review conducted in this **Chapter 4** confirms that the field currently lacks consensus on how to detect and handle blink-related gaps. Many studies rely on implicit blink indicators—such as pupil disappearance—without using built-in parser functionalities, as also observed by Holmqvist et al. [101] and Appel et al. [9]. While tools provided by the eye trackers from EyeLink and BeGaze include native blink detection features, these are inconsistently reported or leveraged, which undermines replicability.

In addressing these gaps, this **Chapter 4** contributes an empirically grounded pipeline for blink-aware preprocessing. Through high-speed video verification and extensive analysis of eleven open datasets, we identified precise artifact windows surrounding blinks that must be trimmed to ensure signal fidelity. The resulting time thresholds—ranging from 59 to 118 ms post-blink depending on sampling frequency—offer a novel contribution to preprocessing recommendations for eye-tracking studies. These findings extend the methodological discussions in Grootjen et al. [81] and operationalize their call for more rigorous artifact management.

The comparison of interpolation strategies further clarifies best practices. Linear and cubic spline interpolation emerged as optimal across most conditions, aligning with the practical recommendations by Grootjen et al. [83]. Conversely, polynomial interpolation—especially of higher orders—often introduced distortions, corroborating concerns from prior work [170, 269]. Importantly, our results also showed that over-removal of pre- and post-blink samples can degrade interpolation accuracy, suggesting the need for careful calibration rather than aggressive trimming.

Finally, we demonstrated how these preprocessing decisions propagate into downstream machine learning outcomes. Applying the different infilling methods to a benchmark LSTM model revealed that models trained on interpolated datasets not only improved validation accuracy but also gained over 27% more usable training samples. This directly responds to concerns raised in related work [33, 242] about data sparsity and overfitting in gaze-based sequence models. Moreover, it supports the argument that standardizing preprocessing improves not only fairness and reproducibility but also model robustness—a critical goal for real-time applications in HCI and preventive healthcare [57, 280].

In **Chapter 5** we aim to answer **RQ 3:** *Is real-time detection of central field loss possible using eye-tracking data?* by exploring whether gaze-based machine learning and deep learning models can accurately and efficiently detect central field loss (CFL) in continuous monitoring contexts. Our results provide compelling evidence that both model classes are capable of distinguishing CFL-affected gaze behavior from unimpaired viewing with substantial accuracy. Random Forests (RF) achieved the highest performance among feature-based classifiers, while Long Short-Term Memory (LSTM) models reached comparable results using only raw gaze trajectories and pupil size—highlighting the potential for real-time and low-burden CFL detection using relatively short input windows.

This work builds directly upon findings in prior research that demonstrate how CFL alters gaze characteristics, particularly through increased fixation instability and reliance on eccentric fixation strategies [51, 222]. Our analysis further supports these observations by showing that pupil size and saccadic amplitude were among the most informative features in distinguishing CFL from non-CFL trials (see Figure 5.2). These results align with previous work indicating that pupil size and saccadic dynamics reflect compensatory mechanisms in individuals with simulated or real macular degeneration [39, 210]. Importantly, our findings show that even subtle changes—introduced by a 6° gaze-contingent scotoma—are sufficient to produce detectable signal patterns when analyzed with appropriate models.

The use of simulated impairments, such as those from Barraza-Bernal et al. [20], has been validated in earlier studies [232] and allows for reproducible testing while maintaining ecological validity. These simulations closely mimic the perceptual experience of central field loss, enabling experimental control that is often unattainable in clinical studies. Our results echo the conclusions of Pollmann et al. [198], who emphasized the alignment between simulation-based findings and patient behavior, thereby reinforcing the value of simulated datasets for machine learning development.

A notable contribution of **Chapter 5** lies in modeling the trade-off between classification performance and evaluation cost. While deep learning models often require substantial computational resources, our window size analysis revealed a cost-benefit sweet spot around 1600 ms—where precision-recall performance is high without incurring excessive latency. This contributes to emerging research that emphasizes runtime efficiency and privacy preservation in gaze-based systems [241, 280]. Moreover, the finding that robust classification is possible with relatively short windows challenges the assumption that long observation periods are needed for detecting early-stage visual impairments [33, 237].

The methodological contrast between traditional feature engineering (SVM/RF) and end-to-end learning (LSTM) also offers insight into how different modeling paradigms leverage gaze data. While feature-based models benefit from interpretability—e.g., ranking feature importance—the LSTM approach allows for fully automated analysis of raw gaze and pupil time series, avoiding potential biases introduced during manual feature selection [93, 242].

To summarize, **Chapter 5** advances the case for gaze-based detection of central field loss in real-world scenarios. It demonstrates that unobtrusive eye tracking, combined with appropriate machine learning pipelines, can serve as a viable tool for continuous vision monitoring. These findings reinforce the growing relevance of eye tracking in preventive healthcare, aligning with global efforts to detect age-related diseases like AMD before they impair quality of life [205, 278].

6.1 Limitations

While **Chapter 3** shows that gaze patterns in VR simulations can reveal early signs of visual impairments, several limitations must be considered. Although the simulations were grounded in validated parameters and informed by structured interviews with individuals affected by cataracts and AMD, they cannot fully replicate the lived experience of progressive vision loss. Participants with real impairments often develop long-term compensatory behaviors, such as using preferred retinal loci (PRLs) in AMD or adopting scanning strategies in glaucoma, which are unlikely to emerge during short, simulated exposures. Additionally, individuals adapt their daily environments to mitigate vision-related challenges—for example, placing frequently used items in consistent locations. Such strategies, which reduce the visual search load in everyday life, are not observable in controlled experimental tasks and may limit the ecological validity of findings.

The simulations were evaluated with young, normally sighted participants to ensure controlled comparisons, which may not generalize to older adults who are more likely to be affected by these impairments. The weaker effects observed in the AMD simulation suggest that either perceptual fidelity, task sensitivity, or adaptation time was insufficient to capture realistic behavior. Despite these constraints, the use of immersive VR combined with eye tracking provides a powerful method to study early perceptual changes under tightly controlled yet visually rich conditions.

In **Chapter 4** we offer a detailed analysis of blink-related artifacts in eye-tracking data and their impact on machine learning models, however, several limitations should be noted. First, although the literature review adhered to PRISMA guidelines and focused on high-impact venues in HCI (ACM and IEEE), it excluded vision science and clinical journals such as ARVO or Journal of Vision. Consequently, some medically oriented approaches to blink detection and correction may have been overlooked. Additionally, the review revealed inconsistencies in reporting blink handling methods, which may have limited the reliability of the extracted metadata despite our structured coding and inter-rater validation.

Second, our empirical evaluations primarily used screen-based datasets and focused on stationary setups with standard lighting and head positioning. As a result, the generalizability of the proposed artifact cut-offs and interpolation strategies to mobile or wearable contexts (e.g., AR or real-world tracking) may be constrained. While the blink verification experiment using high-speed video adds ecological validity, it was limited to a single participant and voluntary blinks, which are known to differ in duration and dynamics from spontaneous blinks [101]. Furthermore, our machine learning experiments focused on one dataset and a specific LSTM architecture. Although the results support the effectiveness of blink-aware preprocessing, future studies should explore broader model classes, task types, and real-time applications to establish more generalizable guidelines.

While **Chapter 5** demonstrates the feasibility of CFL from gaze data using both traditional and deep learning methods, several limitations constrain the generalizability of these findings. Most notably, the study relies on simulated impairments rather than clinical populations. While gaze-contingent scotoma simulations have been validated in prior work [198, 232], they cannot fully reproduce the variability, perceptual adaptation, and long-term compensatory strategies seen in real-world cases. Individuals with CFL often learn to manage their environment by establishing routines or spatial anchors—such as placing keys in a fixed location—that are unlikely to manifest in short-term experimental tasks [131]. As such, the behavioral signals captured in this study may underestimate the diversity of gaze patterns in individuals with progressive vision loss.

Additionally, the small sample size and task-specific nature of the dataset from Barraza-Bernal et al. [20] may limit the external validity of the results. The visual discrimination task provides a controlled testing ground but lacks the ecological complexity of daily visual activities such as reading, navigation, or face recognition [98, 214]. Moreover, although leave-one-participant-out cross-validation reduces overfitting, inter-individual variability remains a confounding factor. Lastly, computational cost was benchmarked on high-performance hardware, and further investigation is needed to understand how model latency and energy consumption translate to consumer-grade or wearable devices intended for continuous health monitoring.

6.2 Future Work

While the findings across this dissertation demonstrate the feasibility and promise of gaze-based techniques for detecting visual impairments in virtual and real-world settings, several limitations must be acknowledged that span simulation design, data quality, methodological generalizability, and ecological validity. The following sections outline these limitations in greater detail across each major contribution, highlighting the need for more individualized, longitudinal, and multimodal approaches. Together, they emphasize the importance of validating gaze-based systems not only in controlled experimental conditions but also in clinically diverse populations and real-life environments. Addressing these constraints will be critical for ensuring that gaze-driven models are both scientifically robust and practically deployable in preventive healthcare contexts.

While **Chapter 3** demonstrates the feasibility of using gaze patterns in VR simulations to reveal signs of visual impairments, future work should extend this line of research toward more personalized and longitudinal approaches. The current simulations modeled generalized impairment progressions based on interview data, but they are based on single pointed interviews without longitudinal objective data, as such individual differences in impairment type, progression speed, and gaze adaptation strategies remain unaccounted for. Future studies could integrate clinical imaging data or patient-specific vision profiles to create individualized simulations that more closely mirror real-world impairments. In parallel, longitudinal studies with individuals diagnosed with cataracts, glaucoma, or AMD could reveal how gaze characteristics evolve over time, offering stronger validation for early-stage detection models.

Further, research should broaden the scope of tasks and contexts under investigation. The current studies focused on structured visual search in familiar environments, but daily life often includes more dynamic, open-ended, and unpredictable scenarios. Incorporating tasks such as indoor navigation, object manipulation, or mobile interactions could uncover additional gaze markers of impairment. Moreover, future systems may benefit from multimodal data streams—such as blink dynamics, head movements, or luminance-driven pupil responses—to build more robust, real-time detection models. These expansions would not only improve the accuracy and applicability of gaze-based diagnostics but also move the field closer to practical deployment in preventive healthcare settings.

Building on the findings presented in **Chapter 4**, future work should explore the generalizability of blink-aware preprocessing across a wider range of eye-tracking modalities and application domains. While we in this **Chapter 4** focused on stationary, screen-based setups, mobile and head-mounted systems introduce additional noise sources such as motion blur, occlusion, and environmental variability. Investigating how blink-related artifact trimming and interpolation behave under such conditions is essential for extending these methods to real-world settings, including AR/VR, in-the-wild cognition studies, or gaze-based accessibility tools. Moreover, the blink cut-off thresholds established here could be refined through larger-scale video validation studies involving both spontaneous and voluntary blinks across diverse participant samples.

In addition, future research should evaluate how different interpolation strategies affect not just model performance but also interpretability and trust in gaze-based systems. While linear and cubic spline methods yielded the lowest reconstruction error and highest LSTM accuracy, other models—such as transformers or convolutional networks—may be differentially sensitive to infilling artifacts. Evaluating how preprocessing choices influence explainability, robustness to adversarial noise, or fairness across user subgroups could offer deeper insight into model behavior. Finally, integrating blink-aware preprocessing into real-time pipelines and benchmarking latency-performance trade-offs would advance the applicability of these methods in adaptive interfaces, diagnostic tools, and continuous monitoring systems.

Future work should prioritize the validation of gaze-based CFL detection models with individuals diagnosed with macular degeneration or other clinical sources of central vision loss to continue from the contributions mentioned in **Chapter 5**. While simulated scotomata offer a standardized testbed, studies with real patients would enable investigation of long-term behavioral adaptations, the stability of preferred retinal loci (PRLs), and the impact of individualized compensation strategies [210, 222]. Such work could also explore how visual training interventions alter gaze behavior and whether these adaptations improve or hinder model-based detection. Including clinical populations would further clarify the relationship between impairment severity and classifier sensitivity, enabling models that support individualized risk assessment or therapy planning.

Beyond participant diversity, expanding the experimental scope to more complex and naturalistic visual tasks represents another key direction. Future studies could explore gaze behavior in activities such as reading, visual search in cluttered scenes, or interaction with mobile devices, potentially uncovering richer features or temporal patterns. Incorporating multimodal data—including head movement, scene context, or even audio cues—could enhance the robustness of gaze-based classification. Finally, integrating these models into real-time, privacy-aware applications—such as in-home screening tools or adaptive user interfaces—would advance the goal of gaze-based preventive diagnostics. Doing so requires continued focus on cost-efficiency, runtime constraints, and user transparency, ensuring that such systems remain usable, ethical, and scalable in real-world healthcare settings [241, 280].

7 CONCLUSION

This dissertation set out to investigate how gaze behavior can be used to detect early signs of visual impairment, and to explore how such detection could support the goals of preventive healthcare. While vision loss remains one of the most underdiagnosed non-communicable conditions globally, gaze-based approaches offer a promising route toward early, unobtrusive, and cost-effective screening. By combining high-fidelity simulations, machine learning, and artifact-aware preprocessing, this work contributes a multidisciplinary framework for developing gaze-based diagnostic systems that are both technically robust and behaviorally grounded.

The chapters of this thesis followed a three-part structure aligned with the overarching research questions. First, we examined how different visual impairments influence gaze behavior in virtual environments using perceptually validated simulations. Second, we addressed methodological challenges related to data reliability—particularly the handling of blink-induced gaps in eye-tracking sequences. Finally, we explored real-time gaze-based classification of CFL using both traditional machine learning and deep learning models, with special attention to cost-efficiency and scalability. Each of these parts contributes to a growing body of research that repositions eye tracking from a niche research tool to a practical component of future healthcare systems.

7.1 Insights

This thesis produced several key insights. From the behavioral perspective, we found that gaze metrics such as fixation rate, saccadic amplitude, and pupillary response are sensitive to the presence and severity of visual impairments—even before participants become consciously aware of the simulated degradation. In particular, simulations of cataracts and glaucoma yielded distinct, measurable gaze adaptations, supporting the hypothesis that early-stage impairments can be inferred from naturalistic gaze behavior. These findings reinforce prior research on perceptual compensation [51, 210], while also offering new empirical support for gaze-based screening in preventive care.

In the second part of this work, we contributed a systematic analysis of blink artifacts—one of the most common sources of noise in eye-tracking data. We reviewed current practices, defined empirical trimming thresholds across multiple datasets, and evaluated how different infilling methods affect machine learning performance. This work addressed an important gap in the literature [83, 101] and demonstrated that careful preprocessing is essential for building reliable gaze-based models, particularly when deployed in real-time systems.

The final part of the thesis evaluated the feasibility of continuous CFL detection using eye-tracking data. Both feature-based classifiers (SVM, RF) and sequence models (LSTM) were able to differentiate between baseline and CFL conditions. Notably, we introduced a cost-benefit framework that quantified the trade-off between classification accuracy and computational latency. This analysis revealed that robust classification is possible with relatively short input windows (ca. 1600 ms), making real-time deployment feasible even on resource-constrained devices. These contributions align with broader efforts in applied AI to balance predictive performance with deployment constraints and reinforce the potential of gaze-based screening for scalable, at-home use.

7.2 Closing Remarks

This thesis presents a vertical evaluation of gaze-based vision loss detection, examining the behavioral foundations, methodological barriers, and computational strategies required to enable real-world applications. While prior research has often focused on technical feasibility or clinical validation in isolation, we argue that effective diagnostic systems must integrate simulation realism, preprocessing transparency, and runtime pragmatism. By bringing these dimensions together, this dissertation lays the groundwork for gaze-based screening tools that are clinically meaningful, grounded, and scalable in practice.

I envision a future where subtle behavioral signals like gaze patterns while texting, scrolling through social media, quietly inform systems that support our health. Eye tracking, once limited to labs, will be embedded into everyday devices like smartphones, enabling continuous, low-burden monitoring. A brief hesitation, altered fixation, or change in viewing behavior could trigger an early alert or silent referral, prompting care before symptoms are even noticed. These systems will not replace clinicians but will extend their reach—bringing preventive diagnostics into daily routines. By integrating vision health into the technologies we already use, gaze-based systems could shift age-related vision loss from inevitable decline to manageable change. This thesis is an early step toward that future—one where early detection is seamless, privacy-respecting, and deeply embedded into how we live.

Beyond the scope of vision health, the methodologies and insights developed in this work extend to other areas of preventive care, behavioral diagnostics and beyond. Gaze data is inherently passive, privacy-preserving, and rich in cognitive and perceptual signals. As healthcare systems continue to shift toward proactive, data-driven models, gaze-based technologies hold potential not only for detecting visual decline but also for supporting broader cognitive, neurological, and interaction-based monitoring.

In closing, this work contributes to the design of future health technologies that are embedded, adaptive, and human-centered. By demonstrating that subtle changes in gaze behavior can reveal meaningful health information, we offer a new perspective on how everyday interactions can be transformed into diagnostic opportunities. This is a step toward making preventive healthcare more accessible, more intelligent, and ultimately, more personalized.

Table 7.1: Clarification of contributions for all publications that contributed to this thesis.

	My Contribution	Contribution of Co-authors
[85]	I was the first author of the publication and led the writing. I collaborated with **Sven Mayer** to develop the study design and doing the analysis.	**Sairam Yadla** implemented the VR visualizations and conducted the study as part of his Bachelor's thesis.
[86]	I was the first author of the publication and led the writing. I collaborated with **Sven Mayer** to develop the study design and doing the analysis.	**Sairam Yadla** implemented the VR visualizations, conducted the study and did the initial analysis as part of his Bachelor's thesis.
[84]	I was the first author of the publication and led the writing. I collaborated with **Sven Mayer** to develop the study design and doing the analysis.	**Tobias Daniel** implemented the VR visualizations, conducted the study and did the initial analysis as part of his Master's thesis. **Fabian Kahmann** implemented additional the VR rooms, implemented the spawning of items, and conducted the study as parto of his Bachelor's thesis. **Yannick Weiss** co-supervised both **Tobias Daniel** and **Fabian Kahmann** and contributed to the writing.
[81]	I led the project and was the first author of the resulting publication. I led the writing and collaborated on the initial idea.	**Henrike Weingärtner** contributed to the writing and visualizations. **Sven Mayer** contributed to the initial ideation and writing.
[83]	I led the project and was the first author of the resulting publication. I led the writing, literature review, collection of datasets, analysis and collaborated on the initial idea.	**Henrike Weingärtner** contributed to the literature review, writing and visualizations. **Sven Mayer** contributed to the initial ideation, analysis and writing.
[82]	I led the project and was the first author of the resulting publication. I led the writing, collection of dataset, implementing and running analysis and collaborated on the initial idea.	**Henrike Weingärtner** contributed to the writing and visualizations. **Sven Mayer** contributed to the initial ideation, analysis and writing.
[79]	I was the project lead and first author of the resulting publication. I conducted the final data analysis, led the study design, and led the writing.	**Alexandra Sipatchin** contributed to the writing and provided the dataset. **Siegfried Wahl** provided the dataset. **Tonja Machulla** was providing the research environment and advising. **Lewis Chuang** contributed to the writing, ideation, and was providing the research environment. **Thomas Kosch** contributed to the ideation, the analysis and to the writing.

References

[1] Abdrabou, Yasmeen, Abdelrahman, Yomna, Khamis, Mohamed, and Alt, Florian. 'Think Harder! Investigating the Effect of Password Strength on Cognitive Load during Password Creation.' In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI EA '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–7. DOI: 10.1145/3411763.3451636.

[2] Aftab, Abdul Rafey, von der Beeck, Michael, and Feld, Michael. 'You Have a Point There: Object Selection Inside an Automobile Using Gaze, Head Pose and Finger Pointing.' In: *Proceedings of the 2020 International Conference on Multimodal Interaction*. ICMI '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 595–603. DOI: 10.1145/3382507.3418836.

[3] Aftab, Abdul Rafey, Von Der Beeck, Michael, Rohrhirsch, Steven, Diotte, Benoit, and Feld, Michael. 'Multimodal Fusion Using Deep Learning Applied to Driver's Referencing of Outside-Vehicle Objects.' In: *2021 IEEE Intelligent Vehicles Symposium (IV)*. New York: IEEE, 2021, pp. 1108–1115. DOI: 10.1109/IV48863.2021.9575815.

[4] Al Copolillo, Ann Christopher and Lyons, Amanda. 'Effects of Simulated Low Vision on Postural Adjustment to Changes in Center of Mass in Older Adults.' In: *Occupational Therapy In Health Care* 31.2 (2017), pp. 115–125. DOI: 10.1080/07380577.2016.1278295. eprint: https://doi.org/10.1080/07380577.2016.1278295.

[5] Ali, Qasim, Dæhlen, Are, Heldal, Ilona, and Helgesen, Carsten Gunnar. 'Development of Vision Screening Tool in Virtual Reality: A Usability Assessment Study.' In: *2022 E-Health and Bioengineering Conference (EHB)*. New York, NY, USA: IEEE, 2022, pp. 1–5. DOI: 10.1109/EHB55594.2022.9991705.

[6] Almutleb, Essam S. and Hassan, Shirin E. 'The Effect of Simulated Central Field Loss on Street-crossing Decision-Making in Young Adult Pedestrians.' In: *Optometry and Vision Science* 97.4 (2020).

[7] Annerer-Walcher, Sonja, Ceh, Simon M, Putze, Felix, Kampen, Marvin, Körner, Christof, and Benedek, Mathias. 'How reliably do eye parameters indicate internal versus external attentional focus?' In: *Cognitive Science* 45.4 (2021), e12977. DOI: 10.1111/cogs.12977.

[8] Appel, Tobias, Scharinger, Christian, Gerjets, Peter, and Kasneci, Enkelejda. 'Cross-Subject Workload Classification Using Pupil-Related Measures.' In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. ETRA '18. Warsaw, Poland: Association for Computing Machinery, 2018. DOI: 10.1145/3204493.3204531.

[9] Appel, Tobias, Sevcenko, Natalia, Wortha, Franz, Tsarava, Katerina, Moeller, Korbinian, Ninaus, Manuel, Kasneci, Enkelejda, and Gerjets, Peter. 'Predicting Cognitive

Load in an Emergency Simulation Based on Behavioral and Physiological Measures.' In: *2019 International Conference on Multimodal Interaction*. ICMI '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 154–163. DOI: 10.1145/3340555.3353735.

[10] Arefin, Mohammed Safayet, Swan II, J. Edward, Cohen Hoffing, Russell A., and Thurman, Steven M. 'Estimating Perceptual Depth Changes with Eye Vergence and Interpupillary Distance Using an Eye Tracker in Virtual Reality.' In: *2022 Symposium on Eye Tracking Research and Applications*. ETRA '22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 1–7. DOI: 10.1145/3517031.3529632.

[11] Arianezhad, Majid, Camp, L. Jean, Kelley, Timothy, and Stebila, Douglas. 'Comparative Eye Tracking of Experts and Novices in Web Single Sign-On.' In: *Proceedings of the Third ACM Conference on Data and Application Security and Privacy*. CODASPY '13. San Antonio, Texas, USA: Association for Computing Machinery, 2013, pp. 105–116. DOI: 10.1145/2435349.2435362.

[12] Asish, Sarker Monojit, Kulshreshth, Arun K., and Borst, Christoph W. 'User Identification Utilizing Minimal Eye-Gaze Features in Virtual Reality Applications.' In: vol. 1. 1. 2022, pp. 42–61. DOI: 10.3390/virtualworlds1010004.

[13] Babaei, Ebrahim, Tag, Benjamin, Dingler, Tilman, and Velloso, Eduardo. 'A Critique of Electrodermal Activity Practices at CHI.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan: Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445370.

[14] Babiker, Areej, Faye, Ibrahima, and Malik, Aamir. 'Pupillary Behavior in Positive and Negative Emotions.' In: *2013 IEEE International Conference on Signal and Image Processing Applications*. New York, NY, USA: IEEE, 2013, pp. 379–383. DOI: 10.1109/ICSIPA.2013.6708037.

[15] Bafna, Tanya, Hansen, John Paulin Paulin, and Baekgaard, Per. 'Cognitive Load during Eye-typing.' In: *ACM Symposium on Eye Tracking Research and Applications*. ETRA '20 Full Papers. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–8. DOI: 10.1145/3379155.3391333.

[16] Bank, World. The Evidence Base for Preventive Health. 2020.

[17] Bär, Tobias, Reuter, Jan Felix, and Zöllner, J. Marius. 'Driver head pose and gaze estimation based on multi-template ICP 3-D point cloud alignment.' In: *2012 15th International IEEE Conference on Intelligent Transportation Systems*. New York, NY, USA: IEEE, 2012, pp. 1797–1802. DOI: 10.1109/ITSC.2012.6338678.

[18] Bar-Haim, Yair, Ziv, Talee, Lamy, Dominique, and Hodes, Richard M. 'Nature and nurture in own-race face processing.' In: *Psychological science* 17.2 (2006), pp. 159–163. DOI: 10.1111/j.1467-9280.2006.01679.x.

[19] Barraza-Bernal, Maria J, Ivanov, Iliya V, Nill, Svenja, Rifai, Katharina, Trauzettel-Klosinski, Susanne, and Wahl, Siegfried. 'Can positions in the visual field with high attentional capabilities be good candidates for a new preferred retinal locus?' In: *Vision research* 140 (2017), pp. 1–12. DOI: 10.1016/j.visres.2017.07.009.

[20] Barraza-Bernal, Maria J., Rifai, Katharina, and Wahl, Siegfried. 'The retinal locus of fixation in simulations of progressing central scotomas.' In: *Journal of Vision* 18.1 (2018), p. 7. DOI: 10.1167/18.1.7.

[21] Bates, Douglas, Mächler, Martin, Bolker, Ben, and Walker, Steve. 'Fitting Linear Mixed-Effects Models Using lme4.' In: *Journal of Statistical Software* 67.1 (2015), pp. 1–48. DOI: 10.18637/jss.v067.i01.

[22] Beintema, Jaap A, Loon, Editha M van, and Berg, Albert V van den. 'Manipulating saccadic decision-rate distributions in visual search.' In: *Journal of vision* 5.3 (2005), pp. 1–1. DOI: 10.1167/5.3.1.

[23] Bek, Judith, Poliakoff, Ellen, and Lander, Karen. 'Measuring emotion recognition by people with Parkinson's disease using eye-tracking with dynamic facial expressions.' In: *Journal of neuroscience methods* 331 (2020), p. 108524. DOI: 10.1016/j.jneumeth.2019.108524.

[24] Bekele, E., Wade, J., Bian, D., Fan, J., Swanson, Amy, Warren, Z., and Sarkar, N. 'Multimodal Adaptive Social Interaction in Virtual Environment (MASI-VR) for Children with Autism Spectrum Disorders (ASD).' In: *2016 IEEE Virtual Reality (VR)*. New York, NY, USA: IEEE, 2016, pp. 121–130. DOI: 10.1109/VR.2016.7504695.

[25] Bekele, Esubalew, Zheng, Zhi, Swanson, Amy, Crittendon, Julie, Warren, Zachary, and Sarkar, Nilanjan. 'Understanding How Adolescents with Autism Respond to Facial Expressions in Virtual Reality Environments.' In: *IEEE Transactions on Visualization and Computer Graphics* 19.4 (2013), pp. 711–720. DOI: 10.1109/TVCG.2013.42.

[26] Bergström, Joanna, Dalsgaard, Tor-Salve, Alexander, Jason, and Hornbæk, Kasper. 'How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. , Yokohama, Japan, Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445193.

[27] Bhattacharya, Nilavra, Rakshit, Somnath, and Gwizdka, Jacek. 'Towards Real-Time Webpage Relevance Prediction UsingConvex Hull Based Eye-Tracking Features.' In: *ACM Symposium on Eye Tracking Research and Applications*. ETRA '20 Adjunct. Stuttgart, Germany: Association for Computing Machinery, 2020. DOI: 10.1145/3379157.3391302.

[28] Bixler, Robert E. and D'Mello, Sidney K. 'Crossed Eyes: Domain Adaptation for Gaze-Based Mind Wandering Models.' In: *ACM Symposium on Eye Tracking Research and*

Applications. ETRA '21 Full Papers. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–12. DOI: 10.1145/3448017.3457386.

[29] Bodala, Indu P., Abbasi, Nida I., Sun, Yu, Bezerianos, Anastasios, Al-Nashash, Hasan, and Thakor, Nitish V. 'Measuring Vigilance Decrement Using Computer Vision Assisted Eye Tracking in Dynamic Naturalistic Environments.' In: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. New York, NY, USA: IEEE, 2017, pp. 2478–2481. DOI: 10.1109/EMBC.2017.8037359.

[30] Borji, Ali and Itti, Laurent. 'Defending Yarbus: Eye movements reveal observers' task.' In: *Journal of vision* 14.3 (2014), pp. 29–29. DOI: 10.1111/j.1467-9280.2006.01679.x.

[31] Brambilla, Riccardo, Onorati, Francesco, Russo, Vincenzo, Mauri, Maurizio, Magrassi, Lorenzo, Mainardi, Luca T., and Barbieri, Riccardo. 'A Stimulus-Response Processing Framework for Pupil Dynamics Assessment during Iso-Luminant Stimuli.' In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. New York, NY, USA: IEEE, 2018, pp. 400–403. DOI: 10.1109/EMBC.2018.8512325.

[32] Breiman, Leo. 'Random forests.' In: *Machine learning* 45 (2001), pp. 5–32. DOI: 10.1023/A:1010933404324.

[33] Bremer, Gianni, Stein, Niklas, and Lappe, Markus. 'Machine Learning Prediction of Locomotion Intention from Walking and Gaze Data.' In: *International Journal of Semantic Computing* 17.01 (2023), pp. 119–142. DOI: 10.1142/S1793351X22490010.

[34] Brookings, Jeffrey B, Wilson, Glenn F, and Swain, Carolyne R. 'Psychophysiological responses to changes in workload during simulated air traffic control.' In: *Biological psychology* 42.3 (1996), pp. 361–377. DOI: 10.1016/0301-0511(95)05167-8.

[35] Broz, Frank, Lehmann, Hagen, Nehaniv, Chrystopher L., and Dautenhahn, Kerstin. 'Mutual Gaze, Personality, and Familiarity: Dual Eye-Tracking during Conversation.' In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. New York, NY, USA: IEEE, 2012, pp. 858–864. DOI: 10.1109/ROMAN.2012.6343859.

[36] Buitinck, Lars, Louppe, Gilles, Blondel, Mathieu, Pedregosa, Fabian, Mueller, Andreas, Grisel, Olivier, Niculae, Vlad, Prettenhofer, Peter, Gramfort, Alexandre, Grobler, Jaques, Layton, Robert, Vanderplas, Jake, Joly, Arnaud, Holt, Brian, and Varoquaux, Gaël. API design for machine learning software: experiences from the scikit-learn project. 2013. DOI: 10.48550/ARXIV.1309.0238.

[37] Bullimore, Mark A and Bailey, Ian L. 'Reading and eye movements in age-related maculopathy.' In: *Optometry and Vision Science* 72.2 (1995), pp. 125–138.

[38] Burton, Robyn, Smith, Nicholas D, and Crabb, David P. 'Eye movements and reading in glaucoma: observations on patients with advanced visual field loss.' In: *Graefe's Archive for Clinical and Experimental Ophthalmology* 252 (2014), pp. 1621–1630. DOI: 10.1007/s00417-014-2752-x.

[39] Calabrèse, Aurélie, Bernard, Jean-Baptiste, Faure, Géraldine, Hoffart, Louis, and Castet, Eric. 'Eye movements and reading speed in macular disease: the shrinking perceptual span hypothesis requires and is supported by a mediation analysis.' In: *Investigative ophthalmology & visual science* 55.6 (2014), pp. 3638–3645. DOI: 10.1167/iovs.13-13408.

[40] Castner, Nora, Kasneci, Enkelejda, Kübler, Thomas, Scheiter, Katharina, Richter, Juliane, Eder, Thérése, Hüttig, Fabian, and Keutel, Constanze. 'Scanpath Comparison in Medical Image Reading Skills of Dental Students: Distinguishing Stages of Expertise Development.' In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. ETRA '18. Warsaw, Poland: Association for Computing Machinery, 2018. DOI: 10.1145/3204493.3204550.

[41] Catrysse, Leen, Gijbels, David, Donche, Vincent, De Maeyer, Sven, Lesterhuis, Marije, and Van den Bossche, Piet. 'How are learning strategies reflected in the eyes? Combining results from self-reports and eye-tracking.' In: *British Journal of Educational Psychology* 88.1 (2018), pp. 118–137. DOI: 10.1111/bjep.12181.

[42] Chen, Zenghai, Fu, Hong, Lo, Wai-Lun, and Chi, Zheru. 'Strabismus Recognition Using Eye-Tracking Data and Convolutional Neural Networks.' In: *Journal of Healthcare Engineering* 2018 (2018). Ed. by Weide Chang. Publisher: Hindawi, p. 7692198. DOI: 10.1155/2018/7692198.

[43] Cheng, Ching-Yu, Yen, May-Yung, Lin, Hsin-Yi, Hsia, Wei-Wei, and Hsu, Wen-Ming. 'Association of Ocular Dominance and Anisometropic Myopia.' In: *Investigative Ophthalmology & Visual Science* 45.8 (2004), pp. 2856–2860. DOI: 10.1167/iovs.03-0878.

[0] Chiossi, Francesco, Gruenefeld, Uwe, Hou, Baosheng James, Newn, Joshua, Ou, Changkun, Liao, Rulu, Welsch, Robin, and Mayer, Sven. 'Understanding the Impact of the Reality-Virtuality Continuum on Visual Search Using Fixation-Related Potentials and Eye Tracking Features.' In: *Proc. ACM Hum.-Comput. Interact.* 8.MHCI (2024). DOI: 10.1145/3676528.

[44] Chitalkina, Natalia. 'When You Don't See What You Expect: Incongruence in Music and Source Code Reading.' In: *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. ETRA '19. Denver, Colorado: Association for Computing Machinery, 2019. DOI: 10.1145/3314111.3322866.

[45] Chollet, François et al. Keras. https://keras.io. 2015.

[46] Cierro, Alessandro, Philippette, Thibault, Francois, Thomas, Nahon, Sebastien, and Watrin, Patrick. 'Eye-Tracking for Sense of Immersion and Linguistic Complexity in the Skyrim Game: Issues and Perspectives.' In: *ACM Symposium on Eye Tracking Research and Applications*. ETRA '20 Short Papers. Stuttgart, Germany: Association for Computing Machinery, 2020. DOI: 10.1145/3379156.3391836.

[47] Clay, Viviane, König, Peter, and König, Sabine U. 'Eye tracking in virtual reality.' In: *Journal of Eye Movement Research* 12.1 (2019). DOI: 10.16910/jemr.12.1.3.

[48] Cole, Michael J., Hendahewa, Chathra, Belkin, Nicholas J., and Shah, Chirag. 'User Activity Patterns During Information Search.' In: *ACM Trans. Inf. Syst.* 33.1 (2015). DOI: 10.1145/2699656.

[49] Couceiro, Ricardo, Duarte, Gonçalo, Durães, João, Castelhano, João, Duarte, Catarina, Teixeira, Cesar, Castelo Branco, Miguel, Carvalho, Paulo, and Madeira, Henrique. 'Pupillography as Indicator of Programmers' Mental Effort and Cognitive Overload.' In: *2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. New York, NY, USA: IEEE, 2019, pp. 638–644. DOI: 10.1109/DSN.2019.00069.

[50] Coutinho, Flávio Luiz and Morimoto, Carlos H. 'Augmenting the Robustness of Cross-Ratio Gaze Tracking Methods to Head Movement.' In: *Proceedings of the Symposium on Eye Tracking Research and Applications*. ETRA '12. Santa Barbara, California: Association for Computing Machinery, 2012, pp. 59–66. DOI: 10.1145/2168556.2168565.

[51] Crossland, Michael D, Culham, Louise E, and Rubin, Gary S. 'Fixation stability and reading speed in patients with newly developed macular disease.' In: *Ophthalmic and Physiological Optics* 24.4 (2004), pp. 327–333. DOI: 10.1111/j.1475-1313.2004.00213.x.

[52] Culham, Louise E., Chabra, Anthony, and Rubin, Gary S. 'Clinical performance of electronic, head-mounted, low-vision devices.' In: *Ophthalmic and Physiological Optics* 24.4 (2004), pp. 281–290. DOI: https://doi.org/10.1111/j.1475-1313.2004.00193.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-1313.2004.00193.x.

[53] Currie, Jonathan, Bond, Raymond R., McCullagh, Paul, Black, Pauline, Finlay, Dewar D., and Peace, Aaron. 'Eye Tracking the Visual Attention of Nurses Interpreting Simulated Vital Signs Scenarios: Mining Metrics to Discriminate Between Performance Level.' In: *IEEE Transactions on Human-Machine Systems* 48.2 (2018), pp. 113–124. DOI: 10.1109/THMS.2017.2754880.

[54] Cutler, Adele, Cutler, D. Richard, and Stevens, John R. 'Random Forests.' In: *Ensemble Machine Learning: Methods and Applications*. Boston, MA: Springer US, 2012, pp. 157–175. DOI: 10.1007/978-1-4419-9326-7_5.

[55] Dan, Emanuela L., Dînşoreanu, Mihaela, and Mureşan, Raul C. 'Accuracy of Six Interpolation Methods Applied on Pupil Diameter Data.' In: *2020 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*. New York, NY, USA: IEEE, 2020, pp. 1–5. DOI: 10.1109/AQTR49680.2020.9129915.

[56] David, Erwan, Beitner, Julia, and Võ, Melissa Le-Hoa. 'Effects of Transient Loss of Vision on Head and Eye Movements during Visual Search in a Virtual Environment.' In: *Brain Sciences* 10.11 (2020), p. 26. DOI: 10.3390/brainsci10110841.

[57] Dechant, Martin Johannes, Frommel, Julian, and Mandryk, Regan Lee. 'The Development of Explicit and Implicit Game-Based Digital Behavioral Markers for the Assessment of Social Anxiety.' In: *Frontiers in Psychology* 12 (2021). DOI: 10.3389/fpsyg.2021.760850.

[58] Dechterenko, Filip and Lukavsky, Jiri. 'Models of Eye Movements in Multiple Object Tracking with Many Objects.' In: *2014 5th European Workshop on Visual Information Processing (EUVIP)*. New York, NY, USA: IEEE, 2014, pp. 1–6. DOI: 10.1109/EUVIP.2014.7018375.

[59] Duchowski, Andrew T, Krejtz, Krzysztof, Krejtz, Izabela, Biele, Cezary, Niedzielska, Anna, Kiefer, Peter, Raubal, Martin, and Giannopoulos, Ioannis. 'The index of pupillary activity: Measuring cognitive load vis-à-vis task difficulty with pupil oscillation.' In: *Proceedings of the 2018 CHI conference on human factors in computing systems*. 2018, pp. 1–13.

[60] Duchowski, Andrew T., House, Donald H., Gestring, Jordan, Congdon, Robert, undefinedwirski, Lech, Dodgson, Neil A., Krejtz, Krzysztof, and Krejtz, Izabela. 'Comparing Estimated Gaze Depth in Virtual and Physical Environments.' In: *Proceedings of the Symposium on Eye Tracking Research and Applications*. ETRA '14. Safety Harbor, Florida: Association for Computing Machinery, 2014, pp. 103–110. DOI: 10.1145/2578153.2578168.

[61] Ekman, Inger, Poikola, Antti, Mäkäräinen, Meeri, Takala, Tapio, and Hämäläinen, Perttu. 'Voluntary Pupil Size Change as Control in Eyes Only Interaction.' In: *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*. ETRA '08. Savannah, Georgia: Association for Computing Machinery, 2008, pp. 115–118. DOI: 10.1145/1344471.1344501.

[62] Engbert, Ralf and Kliegl, Reinhold. 'Microsaccades uncover the orientation of covert attention.' In: *Vision research* 43.9 (2003), pp. 1035–1045. DOI: 10.1016/S0042-6989(03)00084-1.

[63] Esteva, Andre, Robicquet, Alexandre, Ramsundar, Bharath, Kuleshov, Volodymyr, DePristo, Mark, Chou, Katherine, et al. 'A guide to deep learning in healthcare.' In: *Nature Medicine* 25.1 (2019), pp. 24–29. DOI: 10.1038/s41591-018-0316-z.

[64] Fajnzylber, Victor, González, Larry, Maldonado, Pedro, Del Villar, Rafael, Yáñez, Rodrigo, Madariaga, Samuel, Magdics, Milán, and Sbert, Mateu. 'Augmented Film Narrative by Use of Non-Photorealistic Rendering.' In: *2017 International Conference on 3D Immersion (IC3D)*. New York, NY, USA: IEEE, 2017, pp. 1–8. DOI: 10.1109/IC3D.2017.8251912.

[65] Fatt, Irving and Weissman, Barry A. Physiology of the Eye: An Introduction to the Vegetative Functions. Oxford, United Kingdom: Butterworth-Heinemann, 2013.

[66] Feder, Sascha, Bendixen, Alexandra, and Einhäuser, Wolfgang. 'A hybrid control strategy for capturing cognitive processes in virtual reality (VR) in a natural and efficient way.' In: *2022 IEEE 9th International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*. New York, NY, USA: IEEE, 2022, pp. 1–6. DOI: 10.1109/CIVEMSA53371.2022.9853646.

[67] Feger, Sebastian S., Windl, Maximiliane, Grootjen, Jesse, and Schmidt, Albrecht. 'ConnectivityControl: Providing Smart Home Users with Real Privacy Configuration Options.' In: *End-User Development: 9th International Symposium, IS-EUD 2023, Cagliari, Italy, June 6–8, 2023, Proceedings*. Cagliari, Italy: Springer-Verlag, 2023, pp. 180–188. DOI: 10.1007/978-3-031-34433-6_11.

[68] Felßberg, Anna-Maria and Dombrowe, Isabel. 'The effect of different brightness conditions on visually and memory guided saccades.' In: *Vision Research* 142 (2018), pp. 20–26. DOI: 10.1016/j.visres.2017.10.004.

[69] Feng, Mingwang, Fan, Rong, Liu, Chao, Gao, Tianyi, Wang, Zheng, and Zheng, Rencheng. 'Influence of Vehicle Speeds in Curve Driving on Pupil Diameters of Drivers.' In: *2020 4th CAA International Conference on Vehicular Control and Intelligence (CVCI)*. New York, NY, USA: IEEE, 2020, pp. 777–780. DOI: 10.1109/CVCI51460.2020.9338528.

[70] Fleming, Phillip, Whitlock, Evelyn P., Beil, Timothy L., and Lederle, Frank A. 'Screening for and Treatment of Age-Related Macular Degeneration.' In: *Evidence Report/Technology Assessment* 191 (2010).

[71] Fletcher, Donald C, Schuchard, Ronald A, and Renninger, Laura. 'Patient awareness of binocular central visual field defects in age related macular degeneration (AMD).' In: *Investigative Ophthalmology & Visual Science* 52.14 (2011), pp. 4236–4236.

[72] Foster, Joshua J, Sutterer, David W, Serences, John T, Vogel, Edward K, and Awh, Edward. 'Alpha-band oscillations enable spatially and temporally resolved tracking of covert spatial attention.' In: *Psychological science* 28.7 (2017), pp. 929–941. DOI: 10.1177/0956797617699167.

[73] Gavas, Rahul, Chatterjee, Debatri, and Sinha, Aniruddha. 'Estimation of Cognitive Load Based on the Pupil Size Dilation.' In: *2017 IEEE International Conference on Systems,*

Man, and Cybernetics (SMC). New York, NY, USA: IEEE, 2017, pp. 1499–1504. DOI: 10.1109/SMC.2017.8122826.

[74] Gelman, Andrew and Rubin, Donald B. 'Inference from iterative simulation using multiple sequences.' In: *Statistical science* 7.4 (1992), pp. 457–472. DOI: 10.1214/ss/1177011136.

[75] Gheorghe, Andreea, Mahdi, Labib, and Musat, Ovidiu. 'Age-related macular degeneration.' In: *Maedica - a Journal of Clinical Medicine* 10.3 (2015), pp. 256–261.

[76] Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2021 (GBD 2021) Results. Accessed: 2025-05-20. 2024.

[77] Gonçalves, Martinho, Rocha, Tânia, Magalhães, Luís, Peres, Emanuel, Bessa, Maximino, and Chalmers, Alan. 'Identifying Different Visual Patterns in Web Users Behaviour.' In: *Proceedings of the 27th Spring Conference on Computer Graphics*. SCCG '11. Viničné, Slovak Republic: Association for Computing Machinery, 2011, pp. 65–70. DOI: 10.1145/2461217.2461231.

[78] Greiter, Lukas, Strauch, Christoph, and Huckauf, Anke. 'Pupil Responses Signal Less Inhibition for Own Relative to Other Names.' In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. ETRA '18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 1–5. DOI: 10.1145/3204493.3204576.

[79] Grootjen, Jesse W., Sipatchin, Alexandra, Wahl, Siegfried, Machulla, Tonja-Katrin, Chuang, Lewis, and Kosch, Thomas. 'Assessing Eye Tracking for Continuous Central Field Loss Monitoring.' In: *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia*. MUM '23. Vienna, Austria: Association for Computing Machinery, 2023, pp. 54–64. DOI: 10.1145/3626705.3627776.

[80] Grootjen, Jesse W., Thalhammer, Philipp, and Kosch, Thomas. 'Your Eyes on Speed: Using Pupil Dilation to Adaptively Select Speed-Reading Parameters in Virtual Reality.' In: *Proc. ACM Hum.-Comput. Interact.* 8.MHCI (2024). DOI: 10.1145/3676531.

[81] Grootjen, Jesse W., Weingärtner, Henrike, and Mayer, Sven. 'Highlighting the Challenges of Blinks in Eye Tracking for Interactive Systems.' In: *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications*. ETRA '23. Tubingen, Germany: Association for Computing Machinery, 2023. DOI: 10.1145/3588015.3589202.

[82] Grootjen, Jesse W., Weingärtner, Henrike, and Mayer, Sven. 'Investigating the Effects of Eye-Tracking Interpolation Methods on Model Performance of LSTM.' In: *Proceedings of the 2024 Symposium on Eye Tracking Research and Applications*. ETRA '24. Glasgow, United Kingdom: Association for Computing Machinery, 1, 2024. DOI: 10.1145/3649902.3656353.

[83] Grootjen, Jesse W., Weingärtner, Henrike, and Mayer, Sven. 'Uncovering and Addressing Blink-Related Challenges in Using Eye Tracking for Interactive Systems.' In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. CHI '24. New York, NY, USA: Association for Computing Machinery, 1, 2024. DOI: 10.1145/3613904.3642086.

[84] Grootjen, Jesse W., Weiss, Yannick, Daniel, Tobias, Kahman, Fabian, and Mayer, Sven. 'Effects of Cataracts Severities on Eye Movements and Task Performance During a Visual Search Task Through Virtual Reality.' In: *Mobile and Ubiquitous Multimedia 2025*. New York, NY, USA location = Enna, Italy: Association for Computing Machinery, 1, 2025. DOI: 10.1145/3771882.3771884.

[85] Grootjen, Jesse W., Yadla, Sairam Narsimha Reddy, and Mayer, Sven. 'The Effects of AMD Severities on Eye Movements in Virtual Environments.' In: *Mensch und Computer 2025*. New York, NY, USA location = Chemnitz, Germany: Association for Computing Machinery, 1, 2025. DOI: 10.1145/3743049.3748561.

[86] Grootjen, Jesse W., Yadla, Sairam Narsimha Reddy, and Mayer, Sven. 'Understanding the Impact of Glaucoma Severities Through Virtual Reality Simulations.' In: *Proceedings of Mensch und Computer 2025*. MuC2́5. Chemnitz, Germany: Association for Computing Machinery, 31, 2025. DOI: 10.1145/10.1145/3743049.3743082.

[87] Guadron, Leslie, Titchener, Samuel A, Abbott, Carla J, Ayton, Lauren N, Opstal, John van, Petoe, Matthew A, and Goossens, Jeroen. 'The saccade main sequence in patients with retinitis pigmentosa and advanced age-related macular degeneration.' In: *Investigative Ophthalmology & Visual Science* 64.3 (2023), pp. 1–1. DOI: 10.1167/iovs.64.3.1.

[88] Gunawardena, Nishan, Matscheko, Michael, Anzengruber, Bernhard, Ferscha, Alois, Schobesberger, Martin, Shamiyeh, Andreas, Klugsberger, Bettina, and Solleder, Peter. 'Assessing Surgeons' Skill Level in Laparoscopic Cholecystectomy Using Eye Metrics.' In: *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. ETRA '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–8. DOI: 10.1145/3314111.3319832.

[89] Gupta, Anshul, Mesik, Juraj, Engel, Stephen A., Smith, Rebecca, Schatza, Mark, Calabrèse, Aurélie, van Kuijk, Frederik J., Erdman, Arthur G., and Legge, Gordon E. 'Beneficial Effects of Spatial Remapping for Reading with Simulated Central Field Loss.' In: *Investigative Ophthalmology & Visual Science* 59.2 (2018), pp. 1105–1112. DOI: 10.1167/iovs.16-21404.

[90] Gwizdka, Jacek. 'Characterizing Relevance with Eye-Tracking Measures.' In: *Proceedings of the 5th Information Interaction in Context Symposium*. IIiX '14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 58–67. DOI: 10.1145/2637002.2637011.

[91] Häkkilä, J., Colley, A., Väyrynen, J., and Yliharju, A.-J. 'Introducing Virtual Reality Technologies to Design Education.' In: *Seminar. net* 14 (2018), pp. 1–12. DOI: 10.7577/seminar.2584.

[92] Haslwanter, Thomas. 'Mathematics of three-dimensional eye rotations.' In: *Vision Research* 35.12 (1995), pp. 1727–1739. DOI: https://doi.org/10.1016/0042-6989(94)00257-M.

[93] Hassan, Ahmad, Fan, Wei, Hu, Xiaoyu, Wang, Wenhao, and Li, Hanxi. 'LSTM-based eye-movement trajectory analysis for reading behavior classification.' In: *International Conference on Image, Signal Processing, and Pattern Recognition (ISPP 2022)*. Ed. by Michael Opoku Agyeman and Seppo Sirkemaa. Vol. 12247. International Society for Optics and Photonics. SPIE, 2022, p. 1224715. DOI: 10.1117/12.2636952.

[94] He, Kanghang, Yang, Cheng, Stankovic, Vladimir, and Stankovic, Lina. 'Graph-based clustering for identifying region of interest in eye tracker data analysis.' In: *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*. New York, NY, USA: IEEE, 2017, pp. 1–6. DOI: 10.1109/MMSP.2017.8122264.

[95] Hettiarachchi, Imali T., Hanoun, Samer, Veerabhadrappa, Rakesh, Jia, Dawei, Hosking, Simon, and Bhatti, Asim. 'Performance Quantification and Heart Rate Analysis in A Repeated-trial Simulation-based Training Task.' In: *2021 IEEE International Systems Conference (SysCon)*. New York, NY, USA: IEEE, 2021, pp. 1–7. DOI: 10.1109/SysCon48628.2021.9447147.

[96] Hirsch, Linda, Li, Jingyi, Mayer, Sven, and Butz, Andreas. 'A Survey of Natural Design for Interaction.' In: *Proceedings of Mensch Und Computer 2022*. MuC '22. Darmstadt, Germany: Association for Computing Machinery, 2022, pp. 240–254. DOI: 10.1145/3543758.3543773.

[97] Hirzle, Teresa, Fischbach, Fabian, Karlbauer, Julian, Jansen, Pascal, Gugenheimer, Jan, Rukzio, Enrico, and Bulling, Andreas. 'Understanding, Addressing, and Analysing Digital Eye Strain in Virtual Reality Head-Mounted Displays.' In: *ACM Trans. Comput.-Hum. Interact.* 29.4 (2022). DOI: 10.1145/3492802.

[98] Ho, Elton, Boffa, Jack, and Palanker, Daniel. 'Performance of Complex Visual Tasks Using Simulated Prosthetic Vision via Augmented-Reality Glasses.' In: *Journal of Vision* 19.13 (2019), pp. 22–22. DOI: 10.1167/19.13.22.

[99] Hofmann, Kajta, Mitra, Bhaskar, Radlinski, Filip, and Shokouhi, Milad. 'An Eye-Tracking Study of User Interactions with Query Auto Completion.' In: *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. CIKM '14. Shanghai, China: Association for Computing Machinery, 2014, pp. 549–558. DOI: 10.1145/2661829.2661922.

[100] Hollenstein, Nora, Barrett, Maria, and Björnsdóttir, Marina. 'The Copenhagen Corpus of Eye Tracking Recordings from Natural Reading of Danish Texts.' In: *Proceedings of the Thirteenth Language Resources and Evaluation Conference*. Marseille, France: European Language Resources Association, 2022, pp. 1712–1720. DOI: 10.48550/ARXIV.2204.13311.

[101] Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., and Weijer, J. van de. Eye Tracking: A comprehensive guide to methods and measures. Oxford, United Kingdom: OUP Oxford, 2011.

[102] Huang, Michael Xuelin and Bulling, Andreas. 'SacCalib: Reducing Calibration Distortion for Stationary Eye Trackers Using Saccadic Eye Movements.' In: *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. ETRA '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–10. DOI: 10.1145/3317956.3321553.

[103] Hutt, Stephen, Mills, Caitlin, Bosch, Nigel, Krasich, Kristina, Brockmole, James, and D'Mello, Sidney. '"Out of the Fr-Eye-ing Pan": Towards Gaze-Based Models of Attention during Learning with Technology in the Classroom.' In: *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*. UMAP '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 94–103. DOI: 10.1145/3079628.3079669.

[104] Hutt, Stephen, Stewart, Angela E.B., Gregg, Julie, Mattingly, Stephen, and D'Mello, Sidney K. 'Feasibility of Longitudinal Eye-Gaze Tracking in the Workplace.' In: *Proc. ACM Hum.-Comput. Interact.* 6.ETRA (2022). DOI: 10.1145/3530889.

[105] Hwang, Alex D. and Peli, Eli. 'An Augmented-Reality Edge Enhancement Application for Google Glass.' In: *Optometry and Vision Science* 91.8 (2014), pp. 1021–1030. DOI: 10.1097/opx.0000000000000326.

[106] Ioannou, Constantina, Bækgaard, Per, Kindler, Ekkart, and Weber, Barbara. 'Towards a Tool for Visualizing Pupil Dilation Linked with Source Code Artifacts.' In: *2020 Working Conference on Software Visualization (VISSOFT)*. New York, NY, USA: IEEE, 2020, pp. 105–109. DOI: 10.1109/VISSOFT51673.2020.00016.

[107] Irons, Jessica, McKone, Elinor, Dumbleton, Rachael, Barnes, Nick, He, Xuming, Provis, Jan, Ivanovici, Callin, and Kwa, Alisa. 'A New Theoretical Approach to Improving Face Recognition in Disorders of Central Vision: Face Caricaturing.' In: *Journal of Vision* 14.2 (2014), pp. 12–12. DOI: 10.1167/14.2.12. eprint: https://arvojournals.org/arvo/content_public/journal/jov/932814/i1534-7362-14-2-12.pdf.

[108] Ishii, Ryo, Nakano, Yukiko I., and Nishida, Toyoaki. 'Gaze Awareness in Conversational Agents: Estimating a User's Conversational Engagement from Eye Gaze.' In: *ACM Trans. Interact. Intell. Syst.* 3.2 (2013). DOI: 10.1145/2499474.2499480.

[109] Istance, Howell and Hyrskykari, Aulikki I. 'Supporting Making Fixations and the Effect on Gaze Gesture Performance.' In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, pp. 3022–3033. DOI: 10.1145/3025453.3025920.

[110] Jerčič, Petar, Sennersten, Charlotte, and Lindley, Craig. 'The Effect of Cognitive Load on Physiological Arousal in a Decision-Making Serious Game.' In: *2017 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. New York, NY, USA: IEEE, 2017, pp. 153–156. DOI: 10.1109/VS-GAMES.2017.8056587.

[111] Jermann, Patrick and Nüssli, Marc-Antoine. 'Effects of Sharing Text Selections on Gaze Cross-Recurrence and Interaction Quality in a Pair Programming Task.' In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. CSCW '12. Seattle, Washington, USA: Association for Computing Machinery, 2012, pp. 1125–1134. DOI: 10.1145/2145204.2145371.

[112] Jia, Shaohua, Koh, Do Hyong, Seccia, Amanda, Antonenko, Pasha, Lamb, Richard, Keil, Andreas, Schneps, Matthew, and Pomplun, Marc. 'Biometric Recognition Through Eye Movements Using a Recurrent Neural Network.' In: *2018 IEEE International Conference on Big Knowledge (ICBK)*. New York, NY, USA: IEEE, 2018, pp. 57–64. DOI: 10.1109/ICBK.2018.00016.

[113] Jones, Pete R., Somoskeöy, Tamás, Chow-Wing-Bom, Hugo, and Crabb, David P. 'Seeing other perspectives: evaluating the use of virtual and augmented reality to simulate visual impairments (OpenVisSim).' In: *npj Digital Medicine* 3.1 (2020), p. 32. DOI: 10.1038/s41746-020-0242-6.

[114] Jun, Jiang, Zhao, Boxin, Zhang, Peng, Chen, Zhenkun, and Peng, Fang. 'Research on UAV Control Method Based on Eye Tracking.' In: *2021 33rd Chinese Control and Decision Conference (CCDC)*. New York, NY, USA: IEEE, 2021, pp. 3281–3286. DOI: 10.1109/CCDC52312.2021.9602071.

[115] Karthik, G, Amudha, J, and Jyotsna, C. 'A Custom Implementation of the Velocity Threshold Algorithm for Fixation Identification.' In: *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*. New York, NY, USA: IEEE, 2019, pp. 488–492. DOI: 10.1109/ICSSIT46314.2019.8987791.

[116] Kay, Matthew, Haroz, Steve, Guha, Shion, and Dragicevic, Pierre. 'Special Interest Group on Transparent Statistics in HCI.' In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '16. San Jose, Califor-

nia, USA: Association for Computing Machinery, 2016, pp. 1081–1084. DOI: 10.1145/2851581.2886442.

[117] Kay, Matthew, Nelson, Gregory L., and Hekler, Eric B. 'Researcher-Centered Design of Statistics: Why Bayesian Statistics Better Fit the Culture and Incentives of HCI.' In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. San Jose, California, USA: Association for Computing Machinery, 2016, pp. 4521–4532. DOI: 10.1145/2858036.2858465.

[118] Keel, S, Lee, A, Scheetz, J, Li, Z, Kotowicz, MA, MacIsaac, RJ, Taylor, HR, and Xie, J. 'Implementation of artificial intelligence-based screening for diabetic retinopathy in primary care: protocol for a pragmatic cluster randomized trial.' In: *BMJ Open* 8.5 (2018), e017195.

[119] Keshava, Ashima, Aumeistere, Anete, Izdebski, Krzysztof, and Konig, Peter. 'Decoding Task From Oculomotor Behavior In Virtual Reality.' In: *ACM Symposium on Eye Tracking Research and Applications*. ETRA '20 Short Papers. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–5. DOI: 10.1145/3379156.3391338.

[120] Khan, Ahsan Raza, Bokhari, Syed Mohsin, Khosravi, Sara, Hussain, Sajjad, Ghannam, Rami, Imran, Muhammad Ali, and Zoha, Ahmed. 'Feature Selection Mechanism for Attention Classification Using Gaze Tracking Data.' In: *2022 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*. New York, NY, USA: IEEE, 2022, pp. 1–4. DOI: 10.1109/ICECS202256217.2022.9970936.

[121] Kinnunen, Tomi, Sedlak, Filip, and Bednarik, Roman. 'Towards Task-Independent Person Authentication Using Eye Movement Signals.' In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*. ETRA '10. New York, NY, USA: Association for Computing Machinery, 2010, pp. 187–190. DOI: 10.1145/1743666.1743712.

[122] Kohm, Kristopher, Porter, John, and Robb, Andrew. 'Sensitivity to Hand Offsets and Related Behavior in Virtual Environments over Time.' In: *ACM Trans. Appl. Percept.* 19.4 (2022). DOI: 10.1145/3561055.

[123] Komogortsev, Oleg V., Ryu, Young Sam, Koh, Do Hyong, and Gowda, Sandeep M. 'Instantaneous saccade driven eye gaze interaction.' In: *Proceedings of the International Conference on Advances in Computer Enterntainment Technology - ACE '09*. Athens, Greece: ACM Press, 2009, p. 140. DOI: 10.1145/1690388.1690412.

[124] Kontogiorgos, Dimosthenis, Sibirtseva, Elena, Pereira, Andre, Skantze, Gabriel, and Gustafson, Joakim. 'Multimodal Reference Resolution In Collaborative Assembly Tasks.' In: *Proceedings of the 4th International Workshop on Multimodal Analyses Enabling Artificial Agents in Human-Machine Interaction*. MA3HMI'18. Boulder, CO, USA: Association for Computing Machinery, 2018, pp. 38–42. DOI: 10.1145/3279972.3279976.

[125] Koochaki, Fatemeh and Najafizadeh, Laleh. 'A Data-Driven Framework for Intention Prediction via Eye Movement With Applications to Assistive Systems.' In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29 (2021), pp. 974–984. DOI: 10.1109/TNSRE.2021.3083815.

[126] Kornrumpf, Benthe, Niefind, Florian, Sommer, Werner, and Dimigen, Olaf. 'Neural Correlates of Word Recognition: A Systematic Comparison of Natural Reading and Rapid Serial Visual Presentation.' In: *Journal of Cognitive Neuroscience* 28.9 (2016), pp. 1374–1391. DOI: 10.1162/jocn_a_00977. eprint: https://direct.mit.edu/jocn/article-pdf/28/9/1374/1951600/jocn_a_00977.pdf.

[127] Korotin, Alexander, Khromov, Nikita, Stepanov, Anton, Lange, Andrey, Burnaev, Evgeny, and Somov, Andrey. 'Towards Understanding of eSports Athletes' Potentialities: The Sensing System for Data Collection and Analysis.' In: *2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*. New York, NY, USA: IEEE, 2019, pp. 1804–1810. DOI: 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00319.

[128] Koskinen, Jani and Bednarik, Roman. 'Gaze - Grabber Distance in Expert and Novice Forest Machine Operators: The Effects of Automatic Boom Control.' In: *ACM Symposium on Eye Tracking Research and Applications*. ETRA '20 Adjunct. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–7. DOI: 10.1145/3379157.3391414.

[129] Krakowczyk, Daniel G., Reich, David R., Chwastek, Jakob, Jakobi, Deborah N., Prasse, Paul, Süss, Assunta, Turuta, Oleksii, Kasprowski, Paweł, and Jäger, Lena A. 'Pymovements: A Python Package for Eye Movement Data Processing.' In: ETRA '23. Tubingen, Germany: ACM, 2023. DOI: 10.1145/3588015.3590134.

[130] Krieger, Laura, Heidemann, Gunther, and Schöning, Julius. 'Object of Interest Segmentation in Video Sequences with Gaze Data.' In: *2018 IEEE International Conference on Image Processing, Applications and Systems (IPAS)*. New York, NY, USA: IEEE, 2018, pp. 104–109. DOI: 10.1109/IPAS.2018.8708873.

[131] Krishnan, Arun Kumar and Bedell, Harold E. 'Functional Changes at the Preferred Retinal Locus in Subjects with Bilateral Central Vision Loss.' In: *Graefe's Archive for Clinical and Experimental Ophthalmology* 256.1 (1, 2018), pp. 29–37. DOI: 10.1007/s00417-017-3818-3.

[132] Krösl, Katharina, Elvezio, Carmine, Hürbe, Matthias, Karst, Sonja, Wimmer, Michael, and Feiner, Steven. 'ICthroughVR: Illuminating Cataracts through Virtual Reality.' In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. New York, NY, USA: IEEE, 2019, pp. 655–663. DOI: 10.1109/VR.2019.8798239.

[133] Krösl, Katharina, Elvezio, Carmine, Luidolt, Laura R., Hürbe, Matthias, Karst, Sonja, Feiner, Steven, and Wimmer, Michael. 'CatARact: Simulating Cataracts in Augmented Reality.' In: *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. New York, NY, USA: IEEE, 2020, pp. 682–693. DOI: 10.1109/ISMAR50242.2020.00098.

[134] Krstić, Ksenija, Šoškić, Anđela, Ković, Vanja, and Holmqvist, Kenneth. 'All good readers are the same, but every low-skilled reader is different: an eye-tracking study using PISA data.' In: *European Journal of Psychology of Education* 33.3 (2018), pp. 521–541. DOI: 10.1007/s10212-018-0382-0.

[135] Krukar, Jakub, Mavros, Panagiotis, and Hoelscher, Christoph. 'Towards Capturing Focal/Ambient Attention during Dynamic Wayfinding.' In: *ACM Symposium on Eye Tracking Research and Applications*. ETRA '20 Adjunct. Stuttgart, Germany: Association for Computing Machinery, 2020. DOI: 10.1145/3379157.3391417.

[136] Kumar, Girish and Chung, Susana T. L. 'Characteristics of Fixational Eye Movements in People with Macular Disease.' In: *Investigative Ophthalmology & Visual Science* 55.8 (2014), pp. 5125–5133. DOI: 10.1167/iovs.14-14608. eprint: https://arvojournals.org/arvo/content_public/journal/iovs/932996/i1552-5783-55-8-5125.pdf.

[137] Kwon, Miyoung, Ramachandra, Chaithanya, Satgunam, Premnandhini, Mel, Bartlett W, Peli, Eli, and Tjan, Bosco S. 'Contour enhancement benefits older adults with simulated central field loss.' In: *Optom Vis Sci* 89.9 (2012), pp. 1374–1384. DOI: 10.1097/OPX.0b013e3182678e52.

[138] Ladouce, Simon, Mustile, Magda, Ietswaart, Magdalena, and Dehais, Frédéric. 'Capturing Cognitive Events Embedded in the Real World Using Mobile Electroencephalography and Eye-Tracking.' In: *Journal of Cognitive Neuroscience* 34.12 (2022), pp. 2237–2255. DOI: 10.1162/jocn_a_01903.

[139] Laeng, Bruno and Falkenberg, Liv. 'Women's pupillary responses to sexually significant others during the hormonal cycle.' In: *Hormones and behavior* 52.4 (2007), pp. 520–530. DOI: 10.1016/j.yhbeh.2007.07.013.

[140] Lai, Kecheng, Liu, Yuqi, He, Qikun, Yi, Ming, and Fujinami, Tsutomu. 'Saliva Secretion as Indicator of Appetite.' In: *Proceedings of the 5th International Conference on Medical and Health Informatics*. ICMHI '21. Kyoto, Japan: Association for Computing Machinery, 2021, pp. 249–253. DOI: 10.1145/3472813.3473198.

[141] Lam, Alexander K N, To, Elaine, Weinreb, Robert N, Yu, Marco, Mak, Heather, Lai, Gilda, Chiu, Vivian, Wu, Ken, Zhang, Xiujuan, Cheng, Timothy P H, Guo, Philip Yawen, and Leung, Christopher K S. 'Use of virtual reality simulation to identify vision-related disability in patients with glaucoma.' In: *JAMA Ophthalmol.* 138.5 (2020), pp. 490–498.

[142] Lander, Christian and Krüger, Antonio. 'EyeSense: Towards Information Extraction on Corneal Images.' In: *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. UbiComp '18. Singapore, Singapore: Association for Computing Machinery, 2018, pp. 980–987. DOI: 10.1145/3267305.3274121.

[143] Lang, Florian, Grootjen, Jesse W., Chuang, Lewis L., and Machulla, Tonja. 'IDeA: A Demonstration of a Mixed Reality System to Support Living with Central Field Loss.' In: *Proceedings of Mensch Und Computer 2022*. MuC '22. Darmstadt, Germany: Association for Computing Machinery, 2022, pp. 611–614. DOI: 10.1145/3543758.3547521.

[144] Lawson, ML, Crewther, DP, Duke, CC, Henry, L, Kiely, PM, West, SJ, and Crewther, SG. 'Attentional blink in global versus local attentional modes.' In: *Australian and New Zealand Journal of Ophthalmology* 26 (1998), S88–S90. DOI: 10.1111/j.1442-9071.1998.tb01384.x.

[145] Lee, Jisoo, Walker, Erin, Burleson, Winslow, Kay, Matthew, Buman, Matthew, and Hekler, Eric B. 'Self-Experimentation for Behavior Change: Design and Formative Evaluation of Two Approaches.' In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, pp. 6837–6849. DOI: 10.1145/3025453.3026038.

[146] Lee, Samantha Sze-Yee, Black, Alex A., and Wood, Joanne M. 'Scanning Behavior and Daytime Driving Performance of Older Adults With Glaucoma.' In: *Journal of Glaucoma* 27.6 (2018), pp. 558–565. DOI: 10.1097/ijg.0000000000000962.

[147] Leusmann, Jan, Oechsner, Carl, Prinz, Johanna, Welsch, Robin, and Mayer, Sven. 'A Database for Kitchen Objects: Investigating Danger Perception in the Context of Human-Robot Interaction.' In: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI EA '23. Hamburg, Germany: Association for Computing Machinery, 2023. DOI: 10.1145/3544549.3585884.

[148] Lewis, James, Brown, David, Cranton, Wayne, and Mason, Robert. 'Simulating visual impairments using the Unreal Engine 3 game engine.' In: *2011 IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH)*. 2011, pp. 1–8. DOI: 10.1109/SeGAH.2011.6165430.

[149] Li, Beibin, Barney, Erin, Hudac, Caitlin, Nuechterlein, Nicholas, Ventola, Pamela, Shapiro, Linda, and Shic, Frederick. 'Selection of Eye-Tracking Stimuli for Prediction by Sparsely Grouped Input Variables for Neural Networks: Towards Biomarker Refinement for Autism.' In: *ACM Symposium on Eye Tracking Research and Applications*. ETRA '20 Full Papers. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–8. DOI: 10.1145/3379155.3391334.

[150] Li, Beibin, Nuechterlein, Nicholas, Barney, Erin, Foster, Claire, Kim, Minah, Mahony, Monique, Atyabi, Adham, Feng, Li, Wang, Quan, Ventola, Pamela, Shapiro, Linda, and Shic, Frederick. 'Learning Oculomotor Behaviors from Scanpath.' In: *Proceedings of the 2021 International Conference on Multimodal Interaction*. ICMI '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 407–415. DOI: 10.1145/3462244.3479923.

[151] Li, Fan, Chen, Chun-Hsien, Xu, Gangyan, and Khoo, Li-Pheng. 'Hierarchical Eye-Tracking Data Analytics for Human Fatigue Detection at a Traffic Control Center.' In: *IEEE Transactions on Human-Machine Systems* 50.5 (2020), pp. 465–474. DOI: 10.1109/THMS.2020.3016088.

[152] Li, Weifeng, Nüssli, Marc-Antoine, and Jermann, Patrick. 'Gaze Quality Assisted Automatic Recognition of Social Contexts in Collaborative Tetris.' In: *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*. ICMI-MLMI '10. New York, NY, USA: Association for Computing Machinery, 2010, pp. 1–8. DOI: 10.1145/1891903.1891914.

[153] Liu, Zaijia, Zheng, Xianjun Sam, Wu, Mingyang, Dong, Rui, and Peng, Kaiping. 'Culture Influence on Aesthetic Perception of Chinese and Western Paintings: Evidence from Eye Movement Patterns.' In: *Proceedings of the 6th International Symposium on Visual Information Communication and Interaction*. VINCI '13. Tianjin, China: Association for Computing Machinery, 2013, pp. 72–78. DOI: 10.1145/2493102.2493111.

[154] Loschky, Lester C. and Wolverton, Gary S. 'How Late Can You Update Gaze-Contingent Multiresolutional Displays without Detection?' In: *ACM Trans. Multimedia Comput. Commun. Appl.* 3.4 (2007). DOI: 10.1145/1314303.1314310.

[155] Lutz, Otto Hans-Martin and Krüger, Jörg. 'Assessing visual attention in virtual reality: Automatic one-point calibration for eye-tracking.' In: *2017 International Conference on Virtual Rehabilitation (ICVR)*. New York, NY, USA: IEEE, 2017, pp. 1–2. DOI: 10.1109/ICVR.2017.8007455.

[156] Lystbæk, Mathias N, Rosenberg, Peter, Pfeuffer, Ken, Grønbæk, Jens Emil, and Gellersen, Hans. 'Gaze-hand alignment: Combining eye gaze and mid-air pointing for interacting with menus in augmented reality.' In: *Proceedings of the ACM on Human-Computer Interaction* 6.ETRA (2022), pp. 1–18. DOI: 10.1145/3530886.

[157] Maciosek, Michael V., Coffield, Ashley B., Edwards, Nicole M., Flottemesch, Thomas J., Goodman, Michael J., and Solberg, Leif I. 'Greater Use of Preventive Services in U.S. Health Care Could Save Lives at Little or No Cost.' In: *Health Affairs* 29.9 (2010), pp. 1656–1660.

[158] Magno, Benjamin V., Datiles, Manuel B., and Lasa, Maria Susan M. 'Progression of lens opacities in cataract patients after one year.' In: *Acta Ophthalmologica Scandinavica* 73.1 (1995), pp. 45–49. DOI: 10.1111/j.1600-0420.1995.tb00012.x.

[159] Malladi, Sai Phani Kumar, Mukherjee, Jayanta, Larabi, Mohamed-Chaker, and Chaudhury, Santanu. 'EG-SNIK: A Free Viewing Egocentric Gaze Dataset and Its Applications.' In: *IEEE Access* 10 (2022), pp. 129626–129641. DOI: 10.1109/ACCESS.2022.3228484.

[160] Marzecová, Anna, Schettino, Antonio, Widmann, Andreas, SanMiguel, Iria, Kotz, Sonja A, and Schröger, Erich. 'Attentional gain is modulated by probabilistic feature expectations in a spatial cueing task: ERP evidence.' In: *Scientific Reports* 8.1 (2018), pp. 1–14. DOI: 10.1038/s41598-017-18347-1.

[161] Marzecová, Anna, Widmann, Andreas, SanMiguel, Iria, Kotz, Sonja A, and Schröger, Erich. 'Interrelation of attention and prediction in visual processing: Effects of task-relevance and stimulus probability.' In: *Biological Psychology* 125 (2017), pp. 76–90. DOI: 10.1016/j.biopsycho.2017.02.009.

[162] McIntire, Lindsey K., McIntire, John P., McKinley, R. Andy, and Goodyear, Chuck. 'Detection of Vigilance Performance with Pupillometry.' In: *Proceedings of the Symposium on Eye Tracking Research and Applications*. ETRA '14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 167–174. DOI: 10.1145/2578153.2578177.

[163] McKone, Elinor, Robbins, Rachel A., He, Xuming, and Barnes, Nick. 'Caricaturing faces to improve identity recognition in low vision simulations: How effective is current-generation automatic assignment of landmark points?' In: *PLOS ONE* 13.10 (2018), pp. 1–18. DOI: 10.1371/journal.pone.0204361.

[164] Merenda, Coleman, Kim, Hyungil, Gabbard, Joseph L., Leong, Samantha, Large, David R., and Burnett, Gary. 'Did You See Me? Assessing Perceptual vs. Real Driving Gains Across Multi-Modal Pedestrian Alert Systems.' In: *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 40–49. DOI: 10.1145/3122986.3123013.

[165] Merenda, Coleman, Kim, Hyungil, Tanous, Kyle, Gabbard, Joseph L., Feichtl, Blake, Misu, Teruhisa, and Suga, Chihiro. 'Augmented Reality Interface Design Approaches for Goal-directed and Stimulus-driven Driving Tasks.' In: *IEEE Transactions on Visualization and Computer Graphics* 24.11 (2018), pp. 2875–2885. DOI: 10.1109/TVCG.2018.2868531.

[166] Michael, R. and Bron, A. J. 'The ageing lens and cataract: a model of normal and pathological ageing.' In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 366.1568 (2011), pp. 1278–1292. DOI: 10.1098/rstb.2010.0300.

[167] Michael, Ralph, Van Rijn, Laurentius J., Van Den Berg, Thomas J. T. P., Barraquer, Rafael I., Grabner, Günther, Wilhelm, Helmut, Coeckelbergh, Tanja, Emesz, Martin, Marvan, Patrik, and Nischler, Christian. 'Association of lens opacities, intraocular straylight, contrast sensitivity and visual acuity in European drivers.' In: *Acta Ophthalmologica* 87.6 (2009), pp. 666–671. DOI: 10.1111/j.1755-3768.2008.01326.x.

[168] Milstein, Arnold and Homer, Jack. 'Will Prevention Save Money?' In: *New England Journal of Medicine* 358.7 (2008), pp. 661–663.

[169] Moacdieh, Nadine Marie and Sarter, Nadine. 'The Effects of Data Density, Display Organization, and Stress on Search Performance: An Eye Tracking Study of Clutter.' In: *IEEE Transactions on Human-Machine Systems* 47.6 (2017), pp. 886–895. DOI: 10.1109/THMS.2017.2717899.

[170] Morales, Aythami, Costela, Francisco M., Tolosana, Ruben, and Woods, Russell L. 'Saccade Landing Point Prediction: A Novel Approach Based on Recurrent Neural Networks.' In: *Proceedings of the 2018 International Conference on Machine Learning Technologies*. ICMLT '18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 1–5. DOI: 10.1145/3231884.3231890.

[171] Moulin, LS, Da Silva, AP Alves, El-Sharkawi, MA, and Marks, Robert J. 'Support vector machines for transient stability analysis of large-scale power systems.' In: *IEEE Transactions on Power Systems* 19.2 (2004), pp. 818–825.

[172] Munoz, DP, Broughton, JR, Goldring, JE, and Armstrong, IT. 'Age-related performance of human subjects on saccadic eye movement tasks.' In: *Experimental brain research* 121.4 (1998), pp. 391–400. DOI: 10.1007/s002210050473.

[173] Muñoz, Jorge, Yannakakis, Georgios N., Mulvey, Fiona, Hansen, Dan Witzner, Gutierrez, German, and Sanchis, Araceli. 'Towards Gaze-Controlled Platform Games.' In: *2011 IEEE Conference on Computational Intelligence and Games (CIG'11)*. New York, NY, USA: IEEE, 2011, pp. 47–54. DOI: 10.1109/CIG.2011.6031988.

[174] Nakano, Yukiko I. and Ishii, Ryo. 'Estimating User's Engagement from Eye-Gaze Behaviors in Human-Agent Conversations.' In: *Proceedings of the 15th International Conference on Intelligent User Interfaces*. IUI '10. New York, NY, USA: Association for Computing Machinery, 2010, pp. 139–148. DOI: 10.1145/1719970.1719990.

[175] Nations, United, Economic, Department of, and Affairs, Social. World population prospects 2019. 2019.

[176] Navvab, Mojtaba, Bisegna, Fabio, and Gugliermetti, Franco. 'Dynamic roadways and in-vehicle lighting conditions for determining mesopic adaptation luminance.' In: *2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC)*. New York, NY, USA: IEEE, 2015, pp. 2005–2010. DOI: 10.1109/EEEIC.2015.7165482.

[177] Nishimura, Junko and Hashimoto, Hideki. Empirical Study on the Utilization and Effects of Health Checkups in Japan. Tech. rep. 17-E-082. Research Institute of Economy, Trade and Industry (RIETI), 2017.

[178] Ogata, Nara G., Boer, Erwin R., Daga, Fábio B., Jammal, Alessandro A., Stringham, James M., and Medeiros, Felipe A. 'Visual Crowding in Glaucoma.' In: *Investigative Opthalmology & Visual Science* 60.2 (2019), p. 538. DOI: 10.1167/iovs.18-25150.

[179] Ojha, Amitash, Ervas, Francesca, and Gola, Elisabetta. 'Emotions as Intrinsic Cognitive Load: An Eye Movement Analysis of High and Low Intelligent Individuals.' In: *2017 3rd IEEE International Conference on Cybernetics (CYBCONF)*. New York, NY, USA: IEEE, 2017, pp. 1–6. DOI: 10.1109/CYBConf.2017.7985776.

[180] Oliveira, Flavio T.P., Aula, Anne, and Russell, Daniel M. 'Discriminating the Relevance of Web Search Results with Measures of Pupil Size.' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. New York, NY, USA: Association for Computing Machinery, 2009, pp. 2209–2212. DOI: 10.1145/1518701.1519038.

[181] Olson, Wes, Kaliardos, Bill, Zuschlag, Michael, and Kendra, Andrew. 'Impact of Traffic Symbol Directional Cues on Pilot Performance during TCAS Events.' In: *2009 IEEE/AIAA 28th Digital Avionics Systems Conference*. New York, NY, USA: IEEE, 2009, pp. 5.D.2–1–5.D.2–10. DOI: 10.1109/DASC.2009.5347460.

[182] Onorati, Francesco, Barbieri, Riccardo, Mauri, Maurizio, Russo, Vincenzo, and Mainardi, Luca. 'Reconstruction and Analysis of the Pupil Dilation Signal: Application to a Psychophysiological Affective Protocol.' In: *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. New York, NY, USA: IEEE, 2013, pp. 5–8. DOI: 10.1109/EMBC.2013.6609423.

[183] Organisation for Economic Co-operation and Development. 'The Heavy Burden of Obesity: The Economics of Prevention.' In: OECD/European Union report highlighting that preventive interventions improve early diagnosis, reduce complications, and yield long-term cost savings. Paris: OECD Publishing, 2019.

[184] Organization, World Health. Global Action Plan for the Prevention and Control of NCDs 2013-2020. 2013.

[185] Ousler, GW, Emory, TB, Welch, D, and Abelson, MB. 'Factors that influence the inter-blink interval (IBI) as measured by the ocular protection index (OPI).' In: *Investigative Ophthalmology & Visual Science* 43.13 (2002), pp. 56–56.

[186] Owens, Justin W., Chaparro, Barbara S., and Palmer, Evan M. 'Text Advertising Blindness: The New Banner Blindness?' In: *J. Usability Studies* 6.3 (2011), pp. 172–197.

[187] Page, Matthew J., McKenzie, Joanne E., Bossuyt, Patrick M., Boutron, Isabelle, Hoffmann, Tammy C., Mulrow, Cynthia D., Shamseer, Larissa, Tetzlaff, Jennifer M., Akl,

Elie A., Brennan, Sue E., Chou, Roger, Glanville, Julie, Grimshaw, Jeremy M., Hróbjartsson, Asbjørn, Lalu, Manoj M., Li, Tianjing, Loder, Elizabeth W., Mayo-Wilson, Evan, McDonald, Steve, McGuinness, Luke A., Stewart, Lesley A., Thomas, James, Tricco, Andrea C., Welch, Vivian A., Whiting, Penny, and Moher, David. 'The PRISMA 2020 statement: an updated guideline for reporting systematic reviews.' In: *Systematic Reviews* 10.1 (2021), p. 89. DOI: 10.1186/s13643-021-01626-4.

[188] Park, Hyun Soo, Jain, Eakta, and Sheikh, Yaser. 'Predicting Primary Gaze Behavior Using Social Saliency Fields.' In: *2013 IEEE International Conference on Computer Vision*. New York, NY, USA: IEEE, 2013, pp. 3503–3510. DOI: 10.1109/ICCV.2013.435.

[189] Parr, Nez and Zeng, Biao. 'British Sign Language (BSL) User's Gaze Patterns Between Hands and Face During Online Communication.' In: *Companion Publication of the 13th ACM Web Science Conference 2021*. WebSci '21 Companion. Virtual Event, United Kingdom: Association for Computing Machinery, 2021, pp. 10–14. DOI: 10.1145/3462741.3466645.

[190] Patel, Sudi, Henderson, Ross, Bradley, L, Galloway, B, and Hunter, L. 'Effect of visual display unit use on blink rate and tear stability.' In: *Optom Vis Sci* 68.11 (1991), pp. 888–892. DOI: 10.1097/00006324-199111000-00010.

[191] Payne, Alexander R., Plimmer, Beryl, and Davies, T. Claire. 'Repeatability of Eye-Hand Movement Onset Asynchrony Measurements and Cerebral Palsy: A Case Study.' In: *Proceedings of the 15th New Zealand Conference on Human-Computer Interaction*. CHINZ 2015. Hamilton, New Zealand: Association for Computing Machinery, 2015, pp. 31–38. DOI: 10.1145/2808047.2808058.

[192] Peng, Xiaoqing. 'An Eye-Movement Study of Pronoun Resolutions by Chinese-Speaking Learners.' In: *Proceedings of the 2022 3rd International Conference on Control, Robotics and Intelligent System*. CCRIS '22. Virtual Event, China: Association for Computing Machinery, 2022, pp. 80–84. DOI: 10.1145/3562007.3562023.

[193] Peng, Xiaoqing. 'The Resolution of English Pronoun by Chinese-speaking Learners: Evidence from Eye Movement.' In: *2022 International Conference on Frontiers of Artificial Intelligence and Machine Learning (FAIML)*. New York, NY, USA: IEEE, 2022, pp. 79–82. DOI: 10.1109/FAIML57028.2022.00025.

[194] Perreault, L., Bobée, Bernard, and Rasmussen, Peter. 'Halphen Distribution System. I: Mathematical and Statistical Properties.' In: *Journal of Hydrologic Engineering - J HYDROL ENG* 4 (1999). DOI: 10.1061/(ASCE)1084-0699(1999)4:3(189).

[195] Pfleging, Bastian, Fekety, Drea K., Schmidt, Albrecht, and Kun, Andrew L. 'A Model Relating Pupil Diameter to Mental Workload and Lighting Conditions.' In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. New York,

NY, USA: Association for Computing Machinery, 2016, pp. 5776–5788. DOI: 10.1145/2858036.2858117.

[196] Pillai, Prarthana, Balasingam, Balakumar, Jaekel, Arunita, and Biondi, Francesco. 'Kalman Filtering to Track Changes in Pupil Size for Automated Driving Systems.' In: *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*. New York, NY, USA: IEEE, 2021, pp. 1–6. DOI: 10.1109/VTC2021-Fall52928.2021.9625045.

[197] Pinheiro, José C. 'Linear Mixed Effects Models for Longitudinal Data.' In: *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd, 2014. DOI: https://doi.org/10.1002/9781118445112.stat05514. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118445112.stat05514.

[198] Pollmann, Stefan, Rosenblum, Lisa, Linnhoff, Stefanie, Porracin, Eleonora, Geringswald, Franziska, Herbik, Anne, Renner, Katja, and Hoffmann, Michael B. 'Preserved Contextual Cueing in Realistic Scenes in Patients with Age-Related Macular Degeneration.' In: *Brain Sciences* 10.12 (2020). DOI: 10.3390/brainsci10120941.

[199] Proudfoot, Jeffrey G., Jenkins, Jeffrey L., Burgoon, Judee K., and Nunamaker, Jay F. 'Deception is in the eye of the communicator: Investigating pupil diameter variations in automated deception detection interviews.' In: *2015 IEEE International Conference on Intelligence and Security Informatics (ISI)*. New York, NY, USA: IEEE, 2015, pp. 97–102. DOI: 10.1109/ISI.2015.7165946.

[200] Qian, Ming, Aguilar, Mario, Zachery, Karen N., Privitera, Claudio, Klein, Stanley, Carney, Thom, and Nolte, Loren W. 'Decision-Level Fusion of EEG and Pupil Features for Single-Trial Visual Detection Analysis.' In: *IEEE Transactions on Biomedical Engineering* 56.7 (2009), pp. 1929–1937. DOI: 10.1109/TBME.2009.2016670.

[201] Qin, Yinuo, Zhang, Weijia, Lee, Richard, Sun, Xiaoxiao, and Sajda, Paul. 'Predictive Power of Pupil Dynamics in a Team Based Virtual Reality Task.' In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. New York, NY, USA: IEEE, 2022, pp. 592–593. DOI: 10.1109/VRW55335.2022.00147.

[202] Quigley, Harry A and Broman, Aimee T. 'The number of people with glaucoma worldwide in 2010 and 2020.' In: *British journal of ophthalmology* 90.3 (2006), pp. 262–267. DOI: 10.1136/bjo.2005.081224.

[203] Qvarfordt, Pernilla and Lee, Matthew. 'Gaze Patterns during Remote Presentations While Listening and Speaking.' In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. ETRA '18. Warsaw, Poland: Association for Computing Machinery, 2018. DOI: 10.1145/3204493.3204540.

[204] Raisi, Seema F. Al and Edirisinghe, Eran. 'A Machine Learning Based Approach to Human Observer Behaviour Analysis in CCTV Video Analytics & Forensics.' In: *Proceedings of the 1st International Conference on Internet of Things and Machine Learning*.

IML ’17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 1–10. DOI: 10.1145/3109761.3158376.

[205] Rein, David B., Wittenborn, John S, Burke-Conte, Zeb, Gulia, Rohit, Robalik, Toshana, Ehrlich, Joshua R., Lundeen, Elizabeth A., and Flaxman, Abraham D. ‘Prevalence of Age-Related Macular Degeneration in the US in 2019.’ In: *JAMA Ophthalmology* 140.12 (2022), pp. 1202–1208. DOI: 10.1001/jamaophthalmol.2022.4401. eprint: https://jamanetwork.com/journals/jamaophthalmology/articlepdf/2797921/jamaophthalmology_rein_2022_oi_220068_1670963675.41862.pdf.

[206] Rein, David B., Zhang, Ping, Wirth, Kathleen E., Honeycutt, Amanda, Hoerger, Thomas J., Phillips, Brandon, Saaddine, Jinan, and Lundeen, Elizabeth A. ‘The Economic Burden of Major Adult Visual Disorders in the United States.’ In: *Archives of Ophthalmology* 124.12 (2009), pp. 1754–1760.

[207] Reinhard, Jens, Messias, André, Dietz, Klaus, MacKeben, Manfred, Lakmann, Raimund, Scholl, Hendrik PN, Apfelstedt-Sylla, Eckart, Weber, Bernhard HF, Seeliger, Mathias W, Zrenner, Eberhart, et al. ‘Quantifying fixation in patients with Stargardt disease.’ In: *Vision research* 47.15 (2007), pp. 2076–2085. DOI: 10.1016/j.visres.2007.04.012.

[208] Reiss, Michael and Reiss, Gilfe. ‘Ocular Dominance: Some Family Data.’ In: *Laterality* 2.1 (1997), pp. 7–16. DOI: 10.1080/713754254.

[209] Rello, Luz and Ballesteros, Miguel. ‘Detecting Readers with Dyslexia Using Machine Learning with Eye Tracking Measures.’ In: *Proceedings of the 12th International Web for All Conference*. W4A ’15. Florence, Italy: Association for Computing Machinery, 2015. DOI: 10.1145/2745555.2746644.

[210] Renninger, Laura and Ma-Wyatt, Anna. ‘Recalibration of eye and hand reference frames in age-related macular degeneration.’ In: *Journal of Vision* 11.11 (2011), pp. 954–954. DOI: 10.1167/11.11.954.

[211] Romberg, Alexa, Zhang, Yayun, Newman, Benjamin, Triesch, Jochen, and Yu, Chen. ‘Global and Local Statistical Regularities Control Visual Attention to Object Sequences.’ In: *2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*. New York, NY, USA: IEEE, 2016, pp. 262–267. DOI: 10.1109/DEVLRN.2016.7846829.

[212] Rosa, Pedro J., Esteves, Francisco, and Arriaga, Patrícia. ‘Beyond Traditional Clinical Measurements for Screening Fears and Phobias.’ In: *IEEE Transactions on Instrumentation and Measurement* 64.12 (2015), pp. 3396–3404. DOI: 10.1109/TIM.2015.2450292.

[213] Rousek, J.B. and Hallbeck, M.S. 'The use of simulated visual impairment to identify hospital design elements that contribute to wayfinding difficulties.' In: *International Journal of Industrial Ergonomics* 41.5 (2011), pp. 447–458. DOI: https://doi.org/10.1016/j.ergon.2011.05.002.

[214] Rubin, Gary S. and Feely, Mary. 'The Role of Eye Movements During Reading in Patients with Age-Related Macular Degeneration (AMD).' In: *Neuro-Ophthalmology* 33.3 (2009), pp. 120–126. DOI: 10.1080/01658100902998732.

[215] Rustum, Rabee and Adeloye, Adebayo J. 'Replacing Outliers and Missing Values from Activated Sludge Data Using Kohonen Self-Organizing Map.' In: *Journal of Environmental Engineering* 133.9 (2007), pp. 909–916. DOI: 10.1061/(ASCE)0733-9372(2007)133:9(909).

[216] Saluja, Kamalpreet Singh, Dv, JeevithaShree, Arjun, Somnath, Biswas, Pradipta, and Paul, Teena. 'Analyzing Eye Gaze of Users with Learning Disability.' In: *Proceedings of the 3rd International Conference on Graphics and Signal Processing*. ICGSP '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 95–99. DOI: 10.1145/3338472.3338481.

[217] Salvucci, Dario D and Goldberg, Joseph H. 'Identifying fixations and saccades in eye-tracking protocols.' In: *Proceedings of the 2000 symposium on Eye tracking research & applications*. 2000, pp. 71–78. DOI: 10.1145/355017.355028.

[218] Sangin, Mirweis, Molinari, Gaëlle, Nüssli, Marc-Antoine, and Dillenbourg, Pierre. 'How Learners Use Awareness Cues about Their Peer's Knowledge? Insights from Synchronized Eye-Tracking Data.' In: *Proceedings of the 8th International Conference on International Conference for the Learning Sciences - Volume 2*. ICLS'08. Utrecht, The Netherlands: International Society of the Learning Sciences, 2008, pp. 287–294.

[219] Sarma, Abhraneel and Kay, Matthew. 'Prior Setting in Practice: Strategies and Rationales Used in Choosing Prior Distributions for Bayesian Analysis.' In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI, USA: Association for Computing Machinery, 2020, pp. 1–12. DOI: 10.1145/3313831.3376377.

[220] Schlösser, Christian, Harrer, Andreas, and Kienle, Andrea. 'Supporting Dyadic Chat Communication with Eye Tracking Based Reading Awareness.' In: *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*. New York, NY, USA: IEEE, 2018, pp. 149–151. DOI: 10.1109/ICALT.2018.00042.

[221] Schubert, Rebekka S., Hartwig, Josephine, Müller, Mathias, Groh, Rainer, and Pannasch, Sebastian. 'Are Age Differences Missing in Relative and Absolute Distance Perception of Stereoscopically Presented Virtual Objects?' In: *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*. VRST '16. Munich, Germany:

Association for Computing Machinery, 2016, pp. 307–308. DOI: 10.1145/2993369.2996334.

[222] Schuchard, Ronald A. 'Preferred retinal loci and macular scotoma characteristics in patients with age-related macular degeneration.' In: *Canadian Journal of Ophthalmology* 40.3 (2005), pp. 303–312. DOI: 10.1016/S0008-4182(05)80073-0.

[223] Schuetz, Immo, Murdison, T. Scott, MacKenzie, Kevin J., and Zannoli, Marina. 'An Explanation of Fitts' Law-like Performance in Gaze-Based Selection Tasks Using a Psychophysics Approach.' In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow Scotland Uk: ACM, 2019, pp. 1–13. DOI: 10.1145/3290605.3300765.

[224] Seiple, William, Rosen, Richard B, and Garcia, Patricia MT. 'Abnormal fixation in individuals with age-related macular degeneration when viewing an image of a face.' In: *Optometry and Vision Science* 90.1 (2013), pp. 45–56. DOI: 10.1097/OPX.0b013e3182794775.

[225] Seychell, Martin. Towards better prevention and management of chronic diseases. 70% to 80% of all healthcare costs in the EU are currently spent on chronic diseases. 2020.

[226] Shao, Jun and Zhong, Bob. 'Last observation carry-forward and last observation analysis.' In: *Statistics in Medicine* 22.15 (2003), pp. 2429–2441. DOI: https://doi.org/10.1002/sim.1519. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.1519.

[227] Sharafi, Zohreh, Huang, Yu, Leach, Kevin, and Weimer, Westley. 'Toward an Objective Measure of Developers' Cognitive Activities.' In: *ACM Trans. Softw. Eng. Methodol.* 30.3 (2021). DOI: 10.1145/3434643.

[228] Sidenmark, Ludwig, Parent, Mark, Wu, Chi-Hao, Chan, Joannes, Glueck, Michael, Wigdor, Daniel, Grossman, Tovi, and Giordano, Marcello. 'Weighted Pointer: Error-aware Gaze-based Interaction through Fallback Modalities.' In: *IEEE Transactions on Visualization and Computer Graphics* 28.11 (2022), pp. 3585–3595. DOI: 10.1109/TVCG.2022.3203096.

[229] Silva, Breno B., Orrego-Carmona, David, and Szarkowska, Agnieszka. 'Using linear mixed models to analyze data from eye-tracking research on subtitling.' In: *Translation Spaces* 11.1 (2022), pp. 60–88. DOI: https://doi.org/10.1075/ts.21013.sil.

[230] Simione, Luca, Vabba, Alisha, Raffone, Antonino, and Mirolli, Marco. 'Pupil Dilation and Self-Reported Emotional Response to IAPS Pictures: The Role of Emotional Regulation and Trait Mindfulness.' In: *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*. New

York, NY, USA: IEEE, 2022, pp. 471–476. DOI: 10.1109/MetroXRAINE54828.2022.9967538.

[231] Sinha, Aniruddha, Roy, Dibyendu, Chaki, Rikayan, De, Bikram Kumar, and Saha, Sanjoy Kumar. 'Readability Analysis Based on Cognitive Assessment Using Physiological Sensing.' In: *IEEE Sensors Journal* 19.18 (2019), pp. 8127–8135. DOI: 10.1109/JSEN.2019.2917834.

[232] Sipatchin, Alexandra, García, Miguel García, and Wahl, Siegfried. 'Assistance for macular degeneration (MD): Different strategies for different augmentations.' In: *Investigative Ophthalmology & Visual Science* 63.7 (2022), 714–F0442.

[233] Sipatchin, Alexandra, García García, Miguel, and Wahl, Siegfried. 'Target Maintenance in Gaming via Saliency Augmentation: An Early-Stage Scotoma Simulation Study Using Virtual Reality (VR).' In: *Applied Sciences* 11.15 (2021), p. 7164. DOI: 10.3390/app11157164.

[234] Sippel, Katrin, Kasneci, Enkelejda, Aehling, Kathrin, Heister, Martin, Rosenstiel, Wolfgang, Schiefer, Ulrich, and Papageorgiou, Elena. 'Binocular Glaucomatous Visual Field Loss and Its Impact on Visual Exploration - A Supermarket Study.' In: *PLoS ONE* 9.8 (2014). Ed. by Ranji Cui, e106089. DOI: 10.1371/journal.pone.0106089.

[235] Sliman, Layth. 'Ocular Guided Robotized Wheelchair for Quadriplegics Users.' In: *2018 TRON Symposium*. TRONSHOW. New York, NY, USA: IEEE, 2018, pp. 1–7.

[236] Smith, Nicholas D, Crabb, David P, and Garway-Heath, David F. 'An exploratory study of visual search performance in glaucoma.' In: *Ophthalmic and Physiological Optics* 31.3 (2011), pp. 225–232. DOI: 10.1111/j.1475-1313.2011.00836.x.

[237] Smith, Nicholas D., Glen, Fiona C., Mönter, Vera M., and Crabb, David P. 'Using Eye Tracking to Assess Reading Performance in Patients with Glaucoma: A Within-Person Study.' In: *Journal of Ophthalmology* 2014 (2014). Ed. by Stefanie I. Becker. Publisher: Hindawi Publishing Corporation, p. 120528. DOI: 10.1155/2014/120528.

[238] Souchet, Alexis D., Diallo, Mamadou Lamarana, and Lourdeaux, Domitile. 'Cognitive Load Classification with a Stroop Task in Virtual Reality Based on Physiological Data.' In: *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. New York, NY, USA: IEEE, 2022, pp. 656–666. DOI: 10.1109/ISMAR55827.2022.00083.

[239] Steil, Julian, Hagestedt, Inken, Huang, Michael Xuelin, and Bulling, Andreas. 'Privacy-Aware Eye Tracking Using Differential Privacy.' In: *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. arXiv:1812.08000 [cs]. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–9. DOI: 10.1145/3314111.3319915.

[240] Steil, Julian, Huang, Michael Xuelin, and Bulling, Andreas. 'Fixation detection for head-mounted eye tracking based on visual similarity of gaze targets.' In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. Warsaw Poland: ACM, 2018, pp. 1–9. DOI: 10.1145/3204493.3204538.

[241] Steil, Julian, Koelle, Marion, Heuten, Wilko, Boll, Susanne, and Bulling, Andreas. 'PrivacEye: Privacy-Preserving Head-Mounted Eye Tracking Using Egocentric Scene Image and Eye Movement Features.' In: *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. ETRA '19. Denver, Colorado: Association for Computing Machinery, 2019. DOI: 10.1145/3314111.3319913.

[242] Stein, Niklas, Bremer, Gianni, and Lappe, Markus. 'Eye Tracking-based LSTM for Locomotion Prediction in VR.' In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. New York, NY, USA: IEEE, 2022, pp. 493–503. DOI: 10.1109/VR51125.2022.00069.

[243] Stern, John A, Boyer, Donna, and Schroeder, David. 'Blink rate: a possible measure of fatigue.' In: *Human factors* 36.2 (1994), pp. 285–297.

[244] Stock, Simon Claus, Erler, Christina, Stork, Wilhelm, Labuz, Grzegorz, Son, Hyeck Soo, Khoramnia, Ramin, and Auffarth, Gerd U. 'Suitability of Virtual Reality for Vision Simulation – A Case Study Using Glaucomatous Visual Fields.' In: *Investigative Ophthalmology & Visual Science* 60.9 (22, 2019), pp. 2441–2441.

[245] Strasburger, Hans, Rentschler, Ingo, and Jüttner, Martin. 'Peripheral Vision and Pattern Recognition: A Review.' In: *Journal of Vision* 11.5 (2011), pp. 13–13. DOI: 10.1167/11.5.13.

[246] Stuart, Samuel, Galna, Brook, Lord, Sue, Rochester, Lynn, and Godfrey, Alan. 'Quantifying Saccades While Walking: Validity of a Novel Velocity-Based Algorithm for Mobile Eye Tracking.' In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. New York, NY, USA: IEEE, 2014, pp. 5739–5742. DOI: 10.1109/EMBC.2014.6944931.

[247] Subburaj, Shree Krishna, Stewart, Angela E.B., Ramesh Rao, Arjun, and D'Mello, Sidney K. 'Multimodal, Multiparty Modeling of Collaborative Problem Solving Performance.' In: *Proceedings of the 2020 International Conference on Multimodal Interaction*. ICMI '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 423–432. DOI: 10.1145/3382507.3418877.

[248] Sun, Michelle J., Rubin, Gary S., Akpek, Esen K., and Ramulu, Pradeep Y. 'Impact of Glaucoma and Dry Eye on Text-Based Searching.' In: *Translational Vision Science & Technology* 6.3 (2017), p. 24. DOI: 10.1167/tvst.6.3.24.

[249] Takabi, Hassan, Hashem, Yessir, and Dantu, Ram. 'Prediction of human error using eye movements patterns for unintentional insider threat detection.' In: *2018 IEEE 4th*

International Conference on Identity, Security, and Behavior Analysis. ISBA. New York, NY, USA: IEEE, 2018, pp. 1–8. DOI: 10.1109/ISBA.2018.8311479.

[250] Takeuchi, Masato, Shinozaki, Tomohiro, and Kawakami, Koji. 'Effectiveness of Specific Health Check-ups in Japan for the primary prevention of obesity-related diseases: a protocol for a target trial emulation.' In: *BMJ open* 13.7 (2023), e070417. DOI: 10.1136/bmjopen-2022-070417.

[251] Tangmanee, Chatpong and Ruchikachorn, Puripant. 'Online Forms: An Eye Tracking Exploration into Fixation across Three Label Alignments.' In: *2018 22nd International Computer Science and Engineering Conference (ICSEC)*. New York, NY, USA: IEEE, 2018, pp. 1–6. DOI: 10.1109/ICSEC.2018.8712777.

[252] Taşkın, Kemal and Gökçay, Didem. 'Investigation of Risk Taking Behavior and Outcomes in Decision Making with Modified BART (m-BART).' In: *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*. New York, NY, USA: IEEE, 2015, pp. 302–307. DOI: 10.1109/ACII.2015.7344587.

[253] Thepass, G., Pel, J. J. M., Vermeer, K. A., Creten, O., Bryan, S. R., Lemij, H. G., and Van Der Steen, J. 'The Effect of Cataract on Eye Movement Perimetry.' In: *Journal of Ophthalmology* 2015 (2015), pp. 1–9. DOI: 10.1155/2015/425067.

[254] Tien, Geoffrey, Atkins, M. Stella, Jiang, Xianta, Zheng, Bin, and Bednarik, Roman. 'Verbal Gaze Instruction Matches Visual Gaze Guidance in Laparoscopic Skills Training.' In: *Proceedings of the Symposium on Eye Tracking Research and Applications*. ETRA '14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 331–334. DOI: 10.1145/2578153.2578217.

[255] Ting, Daniel SW, Cheung, Cyrus Y, Lim, Gilbert, Tan, Gabriel SW, Quang, Nghi Q, Gan, Andrew, Hamzah, Hazlin, Garcia-Franco, Rodrigo, Yeo, Ian Y, Lee, Sui, et al. 'Artificial intelligence and deep learning in ophthalmology.' In: *British Journal of Ophthalmology* 103.2 (2019), pp. 167–175.

[256] Toker, Dereck, Conati, Cristina, Steichen, Ben, and Carenini, Giuseppe. 'Individual User Characteristics and Information Visualization: Connecting the Dots through Eye Tracking.' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Paris, France: Association for Computing Machinery, 2013, pp. 295–304. DOI: 10.1145/2470654.2470696.

[257] Tonomura, Ryuta, Tadamitsu, Tadamitsu, Manji, Atsushi, Kubota, Naoyuki, and Obo, Takenori. 'Rehabilitation Support System for Attentional Deficits Based on Trail-Making Test.' In: *2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS)*. New York, NY, USA: IEEE, 2018, pp. 1127–1132. DOI: 10.1109/SCIS-ISIS.2018.00177.

[258] Topol, Eric J. 'High-performance medicine: the convergence of human and artificial intelligence.' In: *Nature Medicine* 25.1 (2019), pp. 44–56. DOI: 10.1038/s41591-018-0300-7.

[259] Tsai, Yi-Fang, Viirre, Erik, Strychacz, Christopher, Chase, Bradley, and Jung, Tzyy-Ping. 'Task performance and eye activity: predicting behavior relating to cognitive workload.' In: *Aviation, space, and environmental medicine* 78.5 (2007), B176–B185.

[260] Van den Berg, AV and Van Loon, EM. 'An invariant for timing of saccades during visual search.' In: *Vision Research* 45.12 (2005), pp. 1543–1555. DOI: 10.1016/j.visres.2004.12.018.

[261] Van Den Berg, T. J. T. P. 'Importance of pathological intraocular light scatter for visual disability.' In: *Documenta Ophthalmologica* 61.3-4 (1986), pp. 327–333. DOI: 10.1007/BF00142360.

[262] Van der Stigchel, Stefan, Bethlehem, Richard AI, Klein, Barrie P, Berendschot, Tos TJM, Nijboer, Tanja, and Dumoulin, Serge O. 'Macular degeneration affects eye movement behavior during visual search.' In: *Frontiers in psychology* 4 (2013), p. 579. DOI: 10.3389/fpsyg.2013.00579.

[263] Veneri, Giacomo, Pretegiani, Elena, Federighi, Pamela, Rosini, Francesca, Federico, Antonio, and Rufa, Alessandra. 'Evaluating Human Visual Search Performance by Monte Carlo Methods and Heuristic Model.' In: *Proceedings of the 10th IEEE International Conference on Information Technology and Applications in Biomedicine*. New York, NY, USA: IEEE, 2010, pp. 1–4. DOI: 10.1109/ITAB.2010.5687697.

[264] Vora, Sourabh, Rangesh, Akshay, and Trivedi, Mohan Manubhai. 'Driver gaze zone estimation using convolutional neural networks: A general framework and ablative analysis.' In: *IEEE Transactions on Intelligent Vehicles* 3.3 (2018), pp. 254–265. DOI: 10.1109/TIV.2018.2843120.

[265] Vrzakova, Hana, Amon, Mary Jean, Rees, McKenzie, Faber, Myrthe, and D'Mello, Sidney. 'Looking for a Deal? Visual Social Attention during Negotiations via Mixed Media Videoconferencing.' In: *Proc. ACM Hum.-Comput. Interact.* 4.CSCW3 (2021). DOI: 10.1145/3434169.

[266] Vrzakova, Hana, Amon, Mary Jean, Stewart, Angela E. B., and D'Mello, Sidney K. 'Dynamics of Visual Attention in Multiparty Collaborative Problem Solving Using Multidimensional Recurrence Quantification Analysis.' In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–14. DOI: 10.1145/3290605.3300572.

[267] Vrzakova, Hana, Bednarik, Roman, Nakano, Yukiko I., and Nihei, Fumio. 'Speakers' Head and Gaze Dynamics Weakly Correlate in Group Conversation.' In: *Proceedings of*

the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications. ETRA '16. Charleston, South Carolina: Association for Computing Machinery, 2016, pp. 77–84. DOI: 10.1145/2857491.2857522.

[268] Wan, Yu, Yang, Jiarui, Ren, Xiaotong, Yu, Zitong, Zhang, Rong, and Li, Xuemin. 'Evaluation of Eye Movements and Visual Performance in Patients with Cataract.' In: *Scientific Reports* 10.1 (18, 2020), p. 9875. DOI: 10.1038/s41598-020-66817-w.

[269] Wang, Manhua, Lee, Seul Chan, Kamalesh Sanghavi, Harsh, Eskew, Megan, Zhou, Bo, and Jeon, Myounghoon. 'In-Vehicle Intelligent Agents in Fully Autonomous Driving: The Effects of Speech Style and Embodiment Together and Separately.' In: *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 247–254. DOI: 10.1145/3409118.3475142.

[270] Wang, Yingli, Kung, LeeAnn, and Byrd, Terry Anthony. 'Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations.' In: *Technological Forecasting and Social Change* 126 (2018), pp. 3–13. DOI: 10.1016/j.techfore.2015.12.019.

[271] Wang, Yixin, Qiu, Shuang, Li, Dan, Du, Changde, Lu, Bao-Liang, and He, Huiguang. 'Multi-Modal Domain Adaptation Variational Autoencoder for EEG-Based Emotion Recognition.' In: *IEEE/CAA Journal of Automatica Sinica* 9.9 (2022), pp. 1612–1626. DOI: 10.1109/JAS.2022.105515.

[272] Wang, Youmian, Hu, Bin, Chen, Zhenxiang, Jiang, Xiaoqing, Liu, Wenjuan, and Wang, Peicheng. 'INSIDER: A Framework for Assessing Confidence in Psychological Scales Based on Multi-Modal Physiological Signal Fusion.' In: *2022 IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta)*. New York, NY, USA: IEEE, 2022, pp. 1298–1303. DOI: 10.1109/SmartWorld-UIC-ATC-ScalCom-DigitalTwin-PriComp-Metaverse56740.2022.00199.

[273] Weiss, Yannick, Villa, Steeven, Grootjen, Jesse W, Hoppe, Matthias, Kale, Yasin, and Müller, Florian. 'Exploring Redirection and Shifting Techniques to Mask Hand Movements from Shoulder-Surfing Attacks during PIN Authentication in Virtual Reality.' In: *Proc. ACM Hum.-Comput. Interact.* 8.MHCI (2024). DOI: 10.1145/3676502.

[274] Wensveen, Janice M., Bedell, Harold E., and Loshin, David S. 'Reading Rates With Artificial Central Scotomata With and Without Spatial Remapping of Print.' In: *Optometry and Vision Science* 72.2 (1995), pp. 100–114. DOI: 10.1097/00006324-199502000-00009.

[275] Whittaker, Stephen G, Budd, James, and Cummings, RW. 'Eccentric fixation with macular scotoma.' In: *Investigative ophthalmology & visual science* 29.2 (1988), pp. 268–278.

[276] Wolf, Julian, Lohmeyer, Quentin, Holz, Christian, and Meboldt, Mirko. 'Gaze Comes in Handy: Predicting and Preventing Erroneous Hand Actions in AR-Supported Manual Tasks.' In: *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. New York, NY, USA: IEEE, 2021, pp. 166–175. DOI: 10.1109/ISMAR52148.2021.00031.

[277] Wood, Joanne, Chaparro, Alex, Carberry, Trent, and Chu, Byoung Sun. 'Effect of simulated visual impairment on nighttime driving performance.' In: *Optom Vis Sci* 87.6 (2010), pp. 379–386. DOI: 10.1097/OPX.0b013e3181d95b0d.

[278] World Health Organization. World Report on Vision. 2021.

[279] World Health Organization. World report on vision. Licence: CC BY-NC-SA 3.0 IGO. 2019.

[280] Xu, Yanyu, Dong, Yanbing, Wu, Junru, Sun, Zhengzhong, Shi, Zhiru, Yu, Jingyi, and Gao, Shenghua. 'Gaze Prediction in Dynamic 360° Immersive Videos.' In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA: IEEE, 2018, pp. 5333–5342. DOI: 10.1109/CVPR.2018.00559.

[281] Yaneva, Victoria, Ha, Le An, Eraslan, Sukru, Yesilada, Yeliz, and Mitkov, Ruslan. 'Detecting Autism Based on Eye-Tracking Data from Web Searching Tasks.' In: *Proceedings of the 15th International Web for All Conference*. W4A '18. Lyon, France: Association for Computing Machinery, 2018. DOI: 10.1145/3192714.3192819.

[282] Yekhshatyan, Lora and Lee, John D. 'Changes in the Correlation Between Eye and Steering Movements Indicate Driver Distraction.' In: *IEEE Transactions on Intelligent Transportation Systems* 14.1 (2013), pp. 136–145. DOI: 10.1109/TITS.2012.2208223.

[283] Zahabi, Maryam, Wang, Yinsong, and Shahrampour, Shahin. 'Classification of Officers' Driving Situations Based on Eye-Tracking and Driver Performance Measures.' In: *IEEE Transactions on Human-Machine Systems* 51.4 (2021), pp. 394–402. DOI: 10.1109/THMS.2021.3090787.

[284] Zaid, Ala'a Abu, Sobuh, Mohammad, and Yaman, Musa Al. 'Development of an Algorithm for Automating Gaze Data Analysis Gathered While Using Upper Limb Prostheses.' In: *2017 7th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO)*. New York, NY, USA: IEEE, 2017, pp. 1–5. DOI: 10.1109/ICMSAO.2017.7934914.

[285] Zhang, Kun, Liu, Chenxin, Chen, Jingying, Liu, Xiaodi, Wang, Guangshuai, and Zhang, Rujing. 'Research on the Influence of Target Features on the Visual Tracking of Chil-

dren with Autism.' In: *2021 International Symposium on Educational Technology (ISET)*. New York, NY, USA: IEEE, 2021, pp. 202–206. DOI: 10.1109/ISET52350.2021.00049.

[286] Zheng, Rencheng, Nakano, Kimihiko, Ishiko, Hiromitsu, Hagita, Kenji, Kihira, Makoto, and Yokozeki, Toshiya. 'Eye-Gaze Tracking Analysis of Driver Behavior While Interacting With Navigation Systems in an Urban Area.' In: *IEEE Transactions on Human-Machine Systems* 46.4 (2016), pp. 546–556. DOI: 10.1109/THMS.2015.2504083.

[287] Zheng, Zhaobo, Akash, Kumar, Misu, Teruhisa, Krishnamoorthy, Vidya, Dong, Miaomiao, Lee, Yuni, and Huang, Gaojian. 'Identification of Adaptive Driving Style Preference through Implicit Inputs in SAE L2 Vehicles.' In: *Proceedings of the 2022 International Conference on Multimodal Interaction*. ICMI '22. Bengaluru, India: Association for Computing Machinery, 2022, pp. 468–475. DOI: 10.1145/3536221.3556637.

[288] Zheng, Zhaobo K., Akash, Kumar, and Misu, Teruhisa. 'Detection of Perceived Discomfort in SAE L2 Automated Vehicles through Driver Takeovers and Physiological Spikes.' In: *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. New York, NY, USA: IEEE, 2022, pp. 1717–1722. DOI: 10.1109/ITSC55140.2022.9922161.

[289] Zhou, Zhiwen, Ma, Jianxiao, Lu, Tao, Li, Gen, Fang, Song, and Tan, Ting. 'An Evaluation Method for Visual Search Stability in Urban Tunnel Entrance and Exit Sections Based on Markov Chain.' In: *IEEE Access* 8 (2020), pp. 68559–68569. DOI: 10.1109/ACCESS.2020.2986272.

[290] Zhu, Huadi, Jin, Wenqiang, Xiao, Mingyan, Murali, Srinivasan, and Li, Ming. 'BlinKey: A Two-Factor User Authentication Method for Virtual Reality Devices.' In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4.4 (2020), 164:1–164:29. DOI: 10.1145/3432217.

[291] Zhu, Jing, Wang, Ying, La, Rong, Zhan, Jiawei, Niu, Junhong, Zeng, Shuai, and Hu, Xiping. 'Multimodal Mild Depression Recognition Based on EEG-EM Synchronization Acquisition Network.' In: *IEEE Access*. Vol. 7. New York, NY, USA: IEEE, 2019, pp. 28196–28210. DOI: 10.1109/ACCESS.2019.2901950.

[292] Zhu, Wenchao, Kucyi, Aaron, Kramer, Arthur F., and Lin, Yingzi. 'Multimodal Physiological Assessment of the Task-related Attentional States in a VR Driving Environment.' In: *2022 28th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*. New York, NY, USA: IEEE, 2022, pp. 1–5. DOI: 10.1109/M2VIP55626.2022.10041103.

[293] Zyrianov, Vlas, Guarnera, Drew T., Peterson, Cole S., Sharif, Bonita, and Maletic, Jonathan I. 'Automated Recording and Semantics-Aware Replaying of High-Speed Eye Tracking and Interaction Data to Support Cognitive Studies of Software Engineering

Tasks.' In: *2020 IEEE International Conference on Software Maintenance and Evolution (ICSME)*. New York, NY, USA: IEEE, 2020, pp. 464–475. DOI: 10.1109/ICSME46990.2020.00051.

List of Figures

List of Tables

A APPENDIX

Author(s)	Task	Method	Detector
Qian et al. [200]	Visual Search Sequence	Interpolate	Missing Data
Oliveira et al. [180]	Visual Search	Interpolation	Missing Data
Kinnunen et al. [121]	Video Watching	Interpolation	Tobii Studio
Li et al. [152]	Game	Aggregate	Missing Data
Nakano and Ishii [174]	Wizard of Oz	Aggregate or Split	Missing Data
Veneri et al. [263]	Visual Search	Interpolate	Missing Data
Muñoz et al. [173]	Game	Remove	Missing Data
Owens et al. [186]	Semantic Search	Remove	Missing Data
Broz et al. [35]	Holding a Conversation	Remove	Missing Data
Babiker et al. [14]	Audio Stimuli	Interpolate	Missing Data
Bekele et al. [25]	Virtual Reality	Remove	Missing Data
Ishii et al. [108]	Holding a Conversation	Aggregate or split	Missing Data
Onorati et al. [182]	Holding a Conversation	Reconstruct	Missing Data
Yekhshatyan and Lee [282]	Driving Simulator	Interpolation	Missing Data
Dechterenko and Lukavsky [58]	Visual Search	Remove	Pupil Size
Gwizdka [90]	Visual Search	Average	Missing Data
McIntire et al. [162]	Screen Watching	Imputation	Missing Data
Stuart et al. [246]	Free Viewing	Interpolate	0, 0 coordintes
Tien et al. [254]	Visual Attention	Interpolation	Missing Data
Cole et al. [48]	Visual Search	Imputation	Missing Data
Rosa et al. [212]	Video Watching	Remove / Interpolate	Missing Data
Taşkın and Gökçay [252]	Game	Interpolate	Missing Data
Bekele et al. [24]	Social Task	Interpolate	Missing Data
Romberg et al. [211]	Free Viewing	Aggregate	Missing Data
Bodala et al. [29]	Driving Simulator	Extrapolate	Missing Data
Fajnzylber et al. [64]	Video Watching	"filtered"	Missing Data
Gavas et al. [73]	Memory Task	Interpolate	Missing Data
Hutt et al. [103]	Learning	WEKA	Missing Data
Jerčič et al. [110]	Game	Interpolate	Missing Data
Merenda et al. [164]	Driving Simulator	Imputation	Missing Data
Ojha et al. [179]	Reading	Interpolate	Missing Data
Raisi and Edirisinghe [204]	Video Watching	WEKA	Missing Data
Zaid et al. [284]	Manual Task	Ignore	Missing Data

Brambilla et al. [31]	Free Viewing	Interpolation / Remove	Missing Data
Greiter et al. [78]	Go/NoGo	Interpolation	Missing Data
Jia et al. [112]	Free Viewing	Only Clean Data	Noise / Tracking Loss
Krieger et al. [130]	Watching Video	Remove	Tobii Eye Tracker
Merenda et al. [165]	Driving Simulator	Imputation	Missing Data
Morales et al. [170]	Video Watching	Interpolation	EyeLink Parser
Appel et al. [9]	Game	Interpolate / Remove	Missing Data
Couceiro et al. [49]	Programming / Coding	Resampling	Missing Data
Gunawardena et al. [88]	Surgical Intervention	Remove	Missing Data
Huang and Bulling [102]	Input Method	Interpolation	Missing Data
Karthik et al. [115]	Visual Search	Replace	Missing Data
Korotin et al. [127]	Game	Interpolation	Missing Data
Saluja et al. [216]	Reading	Interpolation	Missing Data
Sinha et al. [231]	Reading	Interpolation	Missing Data
Zhu et al. [291]	Free Viewing	Remove / Replace	Missing Data
Aftab et al. [2]	Input Method	Interpolation	Missing Data
Bafna et al. [15]	Typing Task	Interpolation	Missing Data
Dan et al. [55]	Visual Search	Interpolation	Missing Data
Ioannou et al. [106]	Programming / Coding	Interpolation	Missing Data
Keshava et al. [119]	Align Objects in VR	Interpolation	Missing Data
Koskinen and Bednarik [128]	Operate Joystick	Interpolation	BeGaze Parser
Li et al. [149]	Free Viewing	Imputation	EyeLink Parser
Li et al. [151]	Target Tracking	Interpolation	Missing Data
Subburaj et al. [247]	Game	Imputation	Missing Data
Zhou et al. [289]	Driving Simulator	Interpolation	Missing Data
Zhu et al. [290]	Input Method	Interpolation	Missing Data
Abdrabou et al. [1]	Typing Task	Remove	Missing Data
Aftab et al. [3]	Driving Simulator	Interpolation	Missing Data
Bixler and D'Mello [28]	Free Viewing	Winsorization	Missing Data
Hettiarachchi et al. [95]	Game	Interpolation	Missing Data
Jun et al. [114]	Drone Flying	Remove	Missing Data
Li et al. [150]	Learning	Interpolation	Missing Data
Pillai et al. [196]	Driving Simulator	Moving Window Infilling	Missing Data
Vrzakova et al. [265]	Video Watching	Remove	Missing Data
Wang et al. [269]	Driving Simulator	Interpolation	Missing Data
Zahabi et al. [283]	Driving Simulator	Approximated	Missing Data
Arefin et al. [10]	Visual Discriminaton	Remove	Missing Data
Hirzle et al. [97]	Virtual Reality	Interpolation	Missing Data
Khan et al. [120]	Video Watching	Remove	Missing Data
Malladi et al. [159]	Free Viewing	Interpolation	Missing Data
Qin et al. [201]	VR Task	Interpolation	Missing Data
Simione et al. [230]	Free Viewing	Remove / Interpolate	Missing Data

Souchet et al. [238]	Stroop task	Tobii Pro I-VT software	Missing Data
Stein et al. [242]	VR Task	Extrapolate	Missing Data
Wang et al. [272]	Questionaire	Infill with 0	Missing Data
Zheng et al. [288]	Driving Simulator	Interpolation	Missing Data
Zheng et al. [287]	Input Method	Interpolation	Missing Data
Zhu et al. [292]	Driving Simulator	Interpolation	Missing Data

Table A.1: Overview of the 81 papers that reported on how they dealt with missing data, listed from the oldest to the newest (and alphabetically for authors from the same year)

A 4

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbständig und nur mit den angegebenen Hilfsmitteln verfasst habe. Alle Passagen, die ich aus der Literatur oder aus anderen Quellen übernommen habe, habe ich deutlich als Zitat mit Angabe der Quelle kenntlich gemacht.

München, 8th July 2025

Jesse W. Grootjen