Eye tracking measures of bicyclists' behavior and perception: a systematic review

Shiyu Ma[a], Wenwen Zhang[a], Robert B. Noland[a], Clinton J. Andrews[a]

[a] *Edward J. Bloustein School of Planning and Public Policy, Rutgers University, New Brunswick, NJ, the United States*

Shiyu Ma: shiyu.ma@rutgers.edu ORCID: 0009-0001-8999-8203

Wenwen Zhang (corresponding author): wenwen.zhang@ejb.rutgers.edu ORCID: 0000-0002-2517-8601

Robert B. Noland: rnoland@rutgers.edu ORCID: 0000-0003-0775-0624

Clinton J. Andrews: CJA1@rutgers.edu ORCID: 0000-0002-2989-8091

Eye tracking measures of bicyclists' behavior and perception: a systematic review

Abstract

With improved portability and affordability, eye tracking devices have facilitated an expanding range of cycling experiments, offering valuable insights into cycling gaze behaviors and states of mind. Given the complexity of cyclists' visual behavior and gaze measurements, the field would benefit from a comprehensive review. We aim to bridge this gap with three key focuses: 1) the adoption and interpretations of various gaze metrics derived from cycling experiments, 2) a summary of the findings of those experiments, and 3) identifying areas for future research. A systematic review of three databases yielded thirty-five articles that met our inclusion criteria. The review results show eye tracking technology aided cycling experiments can provide cyclist-center perspectives to understand the impact of factors, including built environment, human factors, mode comparison, and methodology assessment, on navigation behavior and mental workload and/or stress levels. The results suggest the selection of eye-tracking devices, cycling experiment design, and gaze metrics adoption/interpretation vary by research objectives. A variety of general gaze metrics and gaze measurements related to Areas of Interest (AOI) are applied to infer cyclists' mental workload/stress levels and attention allocation respectively. The diversity in gaze metrics design and interpretation, however, highlights the need for standardization to facilitate cross-study comparisons. Areas for future research, especially potential integration with latest computer vision and digital twin technologies, are also discussed.

Keywords: eye-tracking, cycling experiments, gaze metric, safety, stress

1. Introduction

Mobile eye tracking devices are a powerful tool to capture cyclists' vision in naturalistic cycling experiments. Cyclists are subject to many external stimuli while cycling, such as motor-vehicles, pedestrians, potholes, and other things that may be a safety hazard. But they are also sensitive to positive features of the road environment that make cycling more pleasant and safer. Mobile eye-tracking devices are one technique for capturing what features of their environment cyclists are watching as they ride. Our objective is to review the literature on eye-tracking device instrumented cycling experiments to determine how these devices have been used, what metrics are typically analyzed, and how this information can be used to enhance understandings in cycling safety and comfort.

Visual cues trigger emotions and how people look at objects in their environment has been used as a way to decipher emotional responses (Strange & Dolan, 2006; Lu & Pesarakli, 2023). Eye tracking devices provide information on what cyclists look at, which can be linked to various biomarkers, for example galvanic skin response and heart rate, to measure feelings and stress. Cyclists' perceptions of safety and comfort (PSC) is a major determinant of travel satisfaction and a key component in evaluating low-stress bicycling facilities (Mekuria et al., 2012). Perceived safety and comfort are mainly measured by stated preference (SP) and revealed preference (RP) surveys, which are subject to response biases and challenges in data resolution (Bigazzi et al., 2022). To address this limitation, collecting data with eye tracking devices, have been proposed as objective, in situ, and high-resolution alternatives for cyclists' stress.

This lit review, with its focus on cyclists, is embedded in the broader literature of vision study and eye tracking device applications. Understanding types of oculomotor event and their functions is important for the use and interpretation of eye tracking. We use the term "oculomotor event" to include distinct eye movements, eyelid movements (blink), and changes in pupil size. According to (Duchowski, 2017), there are five distinct types of eye movements that involves repositioning fovea (moving the eyes to see clearly). Three are gaze-orienting movements: saccades are rapid eye movements repositioning the fovea to visual targets; smooth pursuit is involved when visually tracking a moving target; vergence movement involves depth detection when focusing on distant targets. The other two types of eye movements function as gaze-stabilizing: vestibule-ocular (VOR) stabilizes gaze during head rotation; opto-kinetic nystagmus (OKN) stabilizes gaze in a moving scene. Other relevant oculomotor events include fixation, blink, and pupil dilation. Fixations are periods when eyes are relatively stationary, with the retina stabilized over objects of interest. Blink is the closing and reopening of eye lids, and involuntary reflexive blinks is a form of protection from external stimuli. Pupil dilation is the change of pupil size in response low-light and emotional stimuli.

For the interest of cycling study, we will focus on oculomotor events that could infer cognition, attention, and internal state. Literature suggests that fixation, saccade, and smooth pursuit are considered as "the only three types of movements need be modeled to gain insight into the overt localization of visual attention" (Duchowski, 2017 p.45). Fixation-based metrics, such as fixation count and fixation duration, could infer cognitive processing and attention engagement (Duchowski, 2017). Metrics on fixation variability correlate to workload, stress level, and emotions (Shiferaw et al., 2019). Saccade metrics, such as saccade velocity and scan path, reveal changing focuses of attention and the amount of processed information (Berto et al., 2008). Changes in pupil sizes are used to indicate arousal levles to visual stimuli and intensity of attention (Pedrotti et al., 2014). Blink duration and eye openness indicators reveal cognitive load and visual attention, especially helpful in detecting driver's drowsiness (Siegle et al., 2008).

In the field of transportation, researchers first used eye trackers to study drivers dating back to the 1970s (Mourant & Rockwell, 1972). The application of eye tracking in driver studies covers topics such as understanding hazard detection, detecting distraction and fatigue, enhancing human-machine interface, evaluating the impacts of infrastructure, and assessing automation monitoring (Acerra, Lantieri, et al., 2023; Ahlstrom et al., 2013; Benedetto et al., 2011; Brome et al., 2021; Hergeth et al., 2016). With the rising awareness on promoting active travel and protecting vulnerable road users, studies on pedestrians using eye tracking devices started to emerge in early 2000s. Research topics on pedestrian study focused on the effects of urban design and infrastructure, safety and risk assessment, and distractions and cognition load (Gruden et al., 2021; Jiang et al., 2018; Simpson et al., 2019). Comparing with drivers and pedestrians, cyclists have distinct visual characteristics. Cyclists typically move faster than pedestrians and slower than auto vehicles. The speed affects cyclists' field of views and renders them more vulnerable to injuries in cases of distractions. Often sharing the road or riding adjacent to vehicles, the road composition and infrastructure put higher attentional requirements on bicyclists for traffic monitoring. Moreover, maintaining balance while travelling is a cognition load unique to cyclists. Due to cyclists' distinct attentional requirements and visual behaviors, a targeted review on the application of eye trackers in cyclist study is deemed necessary.

Using mobile eye tracking devices in naturalistic cycling experiments is a relatively new method. The first study of this kind dated to 2013 in Belgium (Vansteenkiste et al.). Over the past decade, the field has witnessed significant growth in both quantity and diversity. The selection and interpretations of oculomotor metrics are essential parts of data analysis that affect the understanding of cyclist's behavior and perception. Owing to the complexity of oculomotor events and eye tracking data, a proliferation of metrics have been adopted in the cycling studies and there is need for aggregated knowledge. (Kapitaniak et al., 2015) reviewed the application of eye tracking in drivers, emphasizing on drivers' visual strategies and the conspicuity phenomenon. (Mahanama et al., 2022) reviewed various measures of eye movement and pupillary activities and their applications in neuroscience, human-computer interaction, and psychology. Our review uniquely contributes to the literature with a special focus naturalistic cycling experiments, with detailed reviews on experiment characteristics, specific AOIs annotations related to cycling task, gaze metrics and their interpretations for cycling behavior. Our aim is to answer the following three research questions: (1) What gaze metrics are used to interpret bicyclists' behavior and perception? (2) What cycling treatments have been studied with eye tracking and what are the findings? (3) Where are the needs for further research?

The remainder of our paper is organized as follows: In the "Search method" section we describe the eligibility criteria, search procedures, and search results to refine our literature review. In the "Experiment characteristics" section we summarize key features of the experiment designs, including the eye tracking devices, routes, and participants based on our review of the literature. In the "Gaze metrics" section we present a systematic and critical review of gaze metrics, examining how they are defined and interpreted and used. We then synthesize experimental findings grouped by four major research topics. Finally, we discuss the limitations and identify areas that merit future research.

2. Search method

We specify four criteria for inclusion in our review. Firstly, they need to be peer-reviewed journal papers published in English. Secondly, the cycling experiment must involve participants actively riding a bike such that participants' sight and motion are synchronized to emulate a natural cycling experience. Studies that only investigate cyclists' gaze behavior from watching videos on screens are excluded from the review. Thirdly, mobile eye-tracking devices, either head-mounted or integrated with glasses, must be utilized for data collection. Finally, quantitative gaze metrics must be employed to describe eye movements. Studies that discuss gaze location heatmaps without incorporating quantitative analysis are excluded.

We combined phrases incorporating two key search terms to search for pertinent studies. The first key term is "cycling", "bicycling", "cyclist", or "bicyclist" to state the topic of cycling-related studies. The second key term is "eye tracking" or "gaze" to filter the search to studies that used eye tracking devices.

The search was conducted in January 2024 using three databases. The first database is the Transportation Research International Documentation (TRID), which specializes in transportation research literature covering all modes of transportation[1]. The initial search using our specific phrases yielded 60 articles, of which 12 met the inclusion criteria. The second database is ScienceDirect, which provides scientific, technical, and medical research literature. We narrowed down the search to transportation-related journals and reviewed 432 articles. After screening titles and abstracts, 14 additional articles were included. The third database is Google Scholar, which covers scholarly literature across various disciplines. The search generated 64,420 results, and we screened the first 400 records sorted by relevance. Eight more articles were added for review. Finally, the reference lists of the included full-text studies were screened,

[1] https://trid.trb.org

and one additional article was identified to be included. In total, 35 articles were selected for this review.

The earliest study we identified was conducted in 2013 in Belgium. As eye-tracking devices became more available and more affordable in subsequent years, the field witnessed consistent growth from 2013 to 2020 and rapid growth after the COVID-19 pandemic. Eleven studies were published in 2023, indicating increased research interest. Most early studies were conducted in Western and Northern Europe, especially Belgium and Sweden. Studies in Eastern Asia and the United States are more recent. As the geographic diversity of studies increases, researchers will be able to assess how different visual cues associated with different road characteristics, built environments, and cultural perceptions interact and affect bicycling behavior.

3. Experiment characteristics

Studies in our review have a variety of different designs and characteristics associated with the eye-tracking devices used, the selected routes, and the sample of participants. We summarize these differences in this section.

3.1 Eye tracking devices

A total of six brands (SMI, ASL, Tobii, Pupil Labs, HoloLens, and FOVE) and 12 models of eye tracking devices are employed in the reviewed studies, among which eight models are used for naturalistic cycling experiments, three models for Virtual Reality experiments, and one model for an Augmented Reality experiment. These are listed in Table 1 along with specifications for each instrument. Six models are currently available on the market, two have been updated to newer versions, and four have been discontinued. Older eye tracking devices are mounted on headwear or glasses frame, whereas more recent devices increasingly resemble regular glasses. All the devices track eyes with three key components: infrared illuminators, eye cameras, and a scene camera. The infrared illuminators emit infrared light which generates corneal reflections. Pupils are rendered dark due to the reflections and pupil locations are recorded by eye cameras. The scene camera is positioned to face forward and records the road scene. These data are processed in proprietary software with varied eye movement detection algorithms that add fixation points onto the scene video. Figure 1 is an example of an eye tracking device (a) and a frame of processed video showing the fixation in the red point and saccade path in the red line (b).

(a) Example of an eye tracking device

(b) A frame of processed video
(red dot is fixation, red lines are saccade paths)

Figure 1 Example of an eye tracking device and a processed video frame

Table 1. Summary of eye tracking devices, their specifications, and the studies that used them

ET devices	**Use**	**Status**	**Accuracy (degree)**	**No. eye camera**	**Eye camera (Hz)**	**Scene camera (fps)**	**Calibration (point)**	**Weight (g)**	**Software**	**Studies applied**
Pupil Invisible	Naturalistic	O	4	2	200	60	0	46.9	Pupil Player & Pupil Capture	Aasvik & Fyhri, 2022; Gadsby et al., 2022; Kircher & Ahlström, 2023
Pupil Core	Naturalistic	O	0.6	2	200	30/60/120	5	22.75	Pupil Player & Pupil Capture	Acerra et al., 2023
Pupil Labs VR Add-ons + HTC Vive	VR	O	1	2	120	90	8	470	Unity & Pupil Labs software	Zeuwts et al., 2021, 2023
Tobii Pro add ons + HTC Vive	VR	O	0.5-1.1	4	120	90	1	550	Unity & Tobii XR SDK	Bishop et al., 2023; Guo et al., 2023; Ramirez Juarez et al., 2023
HoloLens 2 HMD	AR	O	1.5	2	30		several	566	Python and Unity	Zhao et al., 2023
FOVE VR Headset	VR	O	1.15	2	120	70	1	520	FOVE Unity SDK	van Paridon et al., 2021
Pupil Pro	Naturalistic	U	0.6	1	30	30	9	22.75	Pupil Player & Pupil Capture	Stelling-Konczak et al., 2018
Tobii Pro Glasses 2	Naturalistic	U	NA	4	100	25	1	45	Tobii Pro Lab	Gay et al., 2023; Jang & Kim, 2019; Jiang et al., 2021; Pashkevich et al., 2022; Pfeifer et al., 2023 Ryerson et. al, 2021;

ASL Mobile Eye-XG	Naturalistic	D	0.5-1	1	30	30	10/15	76	EyeVision	Abadi et al., 2022; Jashami et al., 2023; Mantuano et al., 2017; Rupi & Krizek, 2019; Scott-Deeter et al., 2023
SMI iView ETG v1	Naturalistic	D	0.5	2	60	30	3	47	BeGaze	van Paridon et al., 2019&2021; von Stülpnagel, 2020;
SMI iView ETG v2	Naturalistic	D	0.5	2	50/60	30	5	47	BeGaze	Ahlstrom et al., 2016; Kircher & Ahlström, 2020; Nygårdhs et al., 2018; Vansteenkiste et al., 2017; Zeuwts et al., 2016;
SMI iViewX HED	Naturalistic	D	1	1	50	25	5	79	BeGaze	Vansteenkiste et al., 2013, 2014a, 2014b, 2015a, 2015b

O: On Market; U: Updated; D: Discontinued.

Critical specifications that affect device reliability include accuracy rate, number of illuminators and eye cameras, and camera sampling rate. Most of the manufacturer-reported accuracy rates have a deviation of less than 1.5 degrees from the real fixation points. These reported accuracy rates are tested during indoor sedentary tasks, the precise accuracy rates remain untested for outdoor and motion-based tasks (Onkhar & Dodou, 2023). Having at least two sets of illuminators and eye cameras and recording the movements of both eyes reduces the likelihood of data loss. Regarding the eye camera, a higher sampling rate enables the capture of shorter-duration fixations and reduces errors in the detected fixation time. As for the scene camera, the sampling rate determines the length of frame duration and the number of frames used for subsequent frame-by-frame fixation analysis. Currently available or upgraded devices offer eye camera sampling rates above 100 Hz and scene camera sampling rates above 50 Hz. The discontinued devices have eye cameras below 60 Hz and scene cameras below 30 Hz. Prior literature suggests that the eye trackers' sampling frequency should be twice the speed of the particular eye movement, ideally reaching 120 Hz for studying fixation and approximately 600 Hz for micro-saccade (Andersson et al., 2010). When a high-frequency eye tracker is unavailable, quadrupling the collected data is equivalent to doubling the device sampling frequency (Andersson et al., 2010).

Other factors that influence experiment implementation are weather limitations, calibration method, battery life, and weight of the head unit. As sunlight contains large amount of infrared radiation and eye tracking devices use infrared light to illuminate the eyes, direct sunlight causes interference and makes it difficult for eye camera to properly record the pupils. The devices are not water-resilient, and the outdoor experiment cannot be carried out in rainy days. Overcast weather is most ideal for outdoor experiments, and the use of shaded glasses and hat is recommended to shield from glares. Due to the natural variations in the shape of each person's eye and various other properties, eye tracking devices require calibration to optimize gaze estimation for each user. During the calibration procedure, the participant views a card with a varying number of points, and the eye tracking device collects data on participant's gaze at those points. The number of calibration points varies across devices, ranging from 10 to 15 points to zero points. Most devices are calibrated at the start of the experiment, while studies utilizing SMI iView ETG reported calibrating the device at the beginning, during the middle of the ride, and after the completion. Simplifying the calibration process can expedite the experiment and minimize sample exclusion due to calibration failure. Battery life limits the length of routes and duration of the experiments. The ASL device offers a recording time of 1 hour, while SMI, Pupil, and Tobii devices provide 2 hours and more recording time, enhancing flexibility for more extended experiments. The device's comfort is crucial for participants to cycle naturally during the experiments. Most eye tracking devices are designed to be lightweight, with the headset weighing approximately 50g. Pupil Core features a no-lens design, which weighs only 22.75g but also looks different from conventional glasses. When combined with VR or AR headsets, the overall weight of the device increases to about 500g, posing challenges for prolonged experiment duration.

Based on the specifications mentioned above, devices with higher sampling frequency, more eye cameras and illuminators, a streamlined calibration procedure, longer battery life, and lighter weight tend to reduce data loss and capture shorter fixations. This likely improves experimental implementation and the ecological validity of results. While the data collection device is critical, successful research also requires a good experimental design, sufficient recruitment of participants, and the valid interpretation of gaze metrics. We discuss these issues next.

3.2 Experimental settings and route selection

The reviewed studies showcased five types of experiment settings: indoor, naturalistic without routes, naturalistic with routes, immersive virtual environment (IVE), Virtual Reality, and Augmented Reality. Three early experiments were conducted inside a gymnasium. Among the 19 experiments that had participants cycle in a naturalistic outdoor environment, 16 assigned specific cycling routes and three without routes. Six studies used the IVE setting, where participants rode a stationary bike simulator facing a large screen on which the synchronized virtual cycling scenes are projected. The field of view (FOV) available from the screens depended on the screen size and distance from the simulator. To achieve a wider FOV, researchers designed concave screen and combined multiple screens laterally (Acerra, Shoman, et al., 2023; Gay et al., 2023). Virtual Reality experiments have people wearing VR headsets while riding a bike simulator. Although VR headsets like the HTC Vive (110 degrees) and FOVE (100 degrees) have a narrower FOV than the peripheral vision (210 degrees horizontally), they provide the flexibility to expand horizontal search through head and eye rotation, and “over the shoulder" checks. Augmented Reality experiments have participants wear AR headsets while cycling in the real world. Six studies used VR and one used AR. The IVE, VR and AR settings are all affected by motion sickness, which exclude individuals with severe symptoms from participating, also limiting each cycling session to less than 10 minutes.

The routes of early experiments conducted inside gymnasiums and the latest Augmented Reality experiment were under 60 meters in length. In the naturalistic experiments, the routes typically ranged from 2.5 to 5 kilometers or 15 to 30 minutes in duration. Routes for young cyclists were generally shorter, around 1.5 to 2 kilometers, to accommodate their energy levels and cycling capabilities. These route lengths and durations are comparable to typical real-world trips, which average 16 minutes or 1 mile (National Household Travel Survey, 2017) and provide sufficient time for participants to become familiar with the experimental settings and apparatus. Those studies that examined built environment features required route lengths long enough to encompass a diverse range of features, such as a variety of different intersections, pavement conditions, and bike infrastructure. Those focusing on the impact of phone use and detection of hazards required longer durations. In these latter studies participants would ride the same route multiple times and were instructed to carry out multiple distraction or hazard detection tasks.

In addition to the use of eye tracking devices, the reviewed cycling experiments also applied other instruments, including cycling behavior detection sensors, such as speed, braking, and head movement, as well as physiological signal sensors, such as heart rate and Galvanic Skin Response (GSR). It is also common to supplement objective sensor data with stated preference surveys to understand cyclists' subjective perceptions of safety and comfort.

3.3 Participant recruitment and sample sizes

Convenience sampling is the most common recruiting method used in the research reported in the reviewed studies. Eligibility criteria generally include some level of cycling competence to ensure the participant can safely ride a bicycle during an on-road experiment. If participants are put in more challenging situations a higher level of self-reported cycling experience are required (von Stülpnagel, 2020). Secondly, participants must have normal or corrected-to-normal vision (e.g., wearing contact lenses) to ensure compatibility with the eye tracking devices. The convenience sampling method and the eligibility criteria introduce sampling biases. University students, university affiliated personnel, and experienced cyclists are likely overrepresented. Framed lenses are incompatible with the eye tracking glasses and people who wear them are excluded from the sample. The challenges of calibrating the device with seniors' eyes, combined with the difficulty of recruiting senior cyclists, result in a lower representation of this group.

The number of participants recruited is dependent on experimental design. Notably, experiments with IVE and VR that mitigate real-world riding risks while maintaining the authenticity of naturalistic cycling experiences, recruited the most participants. Generally, recent studies have recruited more than 20 participants for naturalistic experiments and 40 participants for experiments with IVE and VR settings. Larger and more diverse groups of participants can help reduce bias, increase generalizability, and enhance statistical power of any analyses (Stelling-Konczak et al., 2018; von Stülpnagel, 2020). However, many studies need to omit some participants due to failure of calibration, low eye tracking rates, and withdrawal due to motion sickness, a particular problem with VR, AR, and IVE settings.

The studies we reviewed recruited participants of different genders and a range of ages, though most studies focused on adults, a few recruited young children. Twenty-eight studies on adult cyclists reported mean ages mostly under 30 years, with an age range to up to 75 years (Scott-Deeter et al., 2023). Seven studies of child cyclists report participant mean ages around 10 years old, with an age range between 6 to 18 years. Twenty-seven studies reported participants' gender distribution, half of which have male-to-female ratios between 0.75 and 1.25.

4. Gaze metrics

In the Introduction section, we reviewed the taxonomy and roles of various oculomotor events. Among these, fixation, saccades, smooth pursuit, blink, and pupil dilation have been instrumental in revealing cognition processes and attention. In our review of the 35 articles, measures of oculomotor events are predominately limited to fixation and saccades. This limited scope may stem from device capabilities and experimental constraints. All reviewed eye-tracking devices can detect fixation and saccades but identifying the other three types requires specific algorithms to analyze eye imagery.

Smooth pursuit, which reveals how eyes follow on moving objects, is particularly pertinent for cyclist to monitor traffic cues. However, without dedicated algorithm, this activity is usually classified as fixations interspersed with short saccades (Mital et al., 2011). Accurate detection of smooth pursuit typically necessitates clinical level high frequency eye trackers, such as the 1250 HZ devices used by (Larsson et al., 2015), significantly higher than the mobile eye tracking devices reviewed(up to 200 Hz by Pupil Invisible and Pupil Core).

Blinks can be both voluntary and involuntary, and those particularly reflexive and spontaneous are robustly affected by mental workload and level of attention (Cori et al., 2019). Without algorithms specialized at identifying blink, the event is recorded as missing data or noise. Tobii developed the eye openness (EO) signal based on the sphere between upper and lower eyelids but is only applicable to their screen-based products not mobile wearables (Miseviciute, n.d.). Pupil Labs's Blink Detector plugin, which uses onset and offset thresholds associated with 2D pupil confidence to detect blink, is applicable to their Core and Invisible products ("Blink Detector", n.d.). One article mentioned about the Pupil Blink Detector feature (Zeuwts et al., 2021), and another reported the percentage of blink time during an predefined event for one participant as an example (van Paridon et al., 2021). No other blink parameters have been assessed.

Changes in pupil size is another metric that can indicate cognition processes and emotion, but is also affected by ambient light conditions. Where user manuals are available, devices from Pupil Labs and Tobii are capable of reporting pupil sizes. However, outdoor naturalistic cycling experiments present changing light conditions along the route which affect pupil sizes. Though the environmental lighting could potentially be controlled for with an illuminance sensor, none of the reviewed articles have used pupil measures.

As we focus solely on fixation and saccades measures in the subsequent session, we will refer to these collectively as “gaze metrics”. Table 2 summarizes the category, type, definition, and interpretations of gaze metrics.

Table 2. Summary of gaze metrics

Category	Type	Measure	Definition	Interpretation	Studies applied
General metrics	Fixation Count	Total number of fixations	The total number of fixations during a ride	Compared between participants to examine different visual search strategy.	(Mantuano et al., 2017; Pashkevich et al., 2022)
		Number of fixations per minute/second	The number of fixations per minute or per second during a ride	Compared between different modes of travellers or different experiment settings to examine visual search strategy.	(Gay et al., 2023; Pashkevich et al., 2022)
	Fixation Duration	Total fixation time	The sum of fixation time, measured in seconds	Compared with total saccade time to examine visual search strategy.	(Mantuano et al., 2017)
		Mean fixation duration	The average duration of all fixations, measured in seconds	Shorter fixation durations infer increased visual tasks and higher hazard estimations.	(Guo et al., 2023; Mantuano et al., 2017; Stelling-Konczak et al., 2018; von Stülpnagel, 2020)
		% Fixation time	The percentage of fixation time to the total trip time	Higher total fixation percentage infers increased cognitive processes and visual workload.	(Mantuano et al., 2017; Vansteenkiste et al., 2013, 2015a)
	Fixation Dispersion	Horizontal and vertical variability	The standard deviation of the X and Y coordinates of gaze locations	Varied explanations associated with mental workload and stress.	(Guo et al., 2023; Ryerson et al., 2021; Vansteenkiste et al., 2013, 2014b; Zeuwts et al., 2016)
		Stationary gaze entropy	Defined on uncertainties of choices and calculated with Shannon's entropy equation	Larger entropy indicates greater randomness in the transition behavior and higher task complexity in visual information acquisition.	(Guo et al., 2023; van Paridon et al., 2019)
		Gaze transition entropy	A conditional entropy considering the temporal dependency between different fixations	Increased values from the optimal indicate stress and anxiety.	(Guo et al., 2023)

		Gaze angular velocity	The angular degree of fixations between frames, measured in degrees per second	Rapid eye movements reflect high cognitive workload which could lead to stress and error.	(Ryerson et al., 2021; Zhao et al., 2023)
	Fixation Distance	Sight vector length	The distance between cyclists’ current body location and their gaze location, measured in meters	Longer fixation distances indicate cyclists’ capability of hazards anticipation. Shorter fixation distances indicate cyclists' focus on immediate surroundings.	(von Stülpnagel, 2020)
	Fixation Angle	Gaze angle from travel	The angular degree between fixation direction and travel direction/face-forward.	Larger gaze angles indicate higher needs of hazard detection from various directions.	(von Stülpnagel, 2020; Zhao et al., 2023)
	Saccade Count	Number of saccades per second	The number of saccades per second during a task	Scanning frequency reflects cyclists' visual search strategies and is influenced by different distraction treatments.	(Jiang et al., 2021)
	Saccade Duration	Total saccade time	The sum of saccade time, measured in seconds	Compared with the total fixation time to reflect cyclists' visual search strategy.	(Mantuano et al., 2017)
AOI-related metrics	Fixation/dwell Count	Number of fixations per s/min per AOI	The number of fixations per second or minute on a specific AOI	AOIs receiving more fixation counts capture more attention.	(Pashkevich et al., 2022; Vansteenkiste et al., 2017)
		% fixation counts per AOI	The percentage of fixation on an AOI to the total number of fixations	AOIs with higher percentage of fixation counts capture more attention	(Ahlstrom et al., 2016; Jiang et al., 2021; Kircher & Ahlström, 2020; Mantuano et al., 2017; Nygårdhs et al., 2018; Pashkevich et al., 2022; Rupi & Krizek, 2019; Van Paridon et al., 2019; Vansteenkiste et al., 2014b; Zeuwts et al., 2021)
		Fixation rate	The percentage of AOIs being fixated on to the total number of AOIs.	Participants with higher fixation rates show higher safety awareness and better hazard detection skills.	(Bishop et al., 2023; Gadsby et al., 2022; Zeuwts et al., 2023)

Fixation/dwell Duration	Total fixation time per AOI	The sum of fixation time on a specific AOI, measured in seconds	AOIs with longer summed fixation time capture more attention	(Abadi et al., 2022; Acerra et al., 2023; Ahlstrom et al., 2016; Gay et al., 2023; Jashami et al., 2023; Rupi & Krizek, 2019; Scott-Deeter et al., 2023; Zeuwts et al., 2021)
	% of fixation time per AOI	The percentage of fixation time on an AOI to the total fixation time	AOIs with higher percentage of fixation time capture more attention	(Aasvik & Fyhri, 2022; Acerra et al., 2023; Ahlstrom et al., 2016; Gadsby et al., 2022; Guo et al., 2023; Jang & Kim, 2019; Mantuano et al., 2017; Pfeifer et al., 2023; Ramirez Juarez et al., 2023; Rupi & Krizek, 2019; Vansteenkiste et al., 2013, 2014a, 2015a, 2015b, 2017; Zeuwts et al., 2021; Zhao et al., 2023)
	Mean/ median fixation duration per AOI	The average duration of fixations on an AOI	AOIs with longer mean/median fixation time capture more attention	(Ahlstrom et al., 2016; Gadsby et al., 2022; Gay et al., 2023; Jashami et al., 2023; Nygårdhs et al., 2018; Pashkevich et al., 2022; van Paridon et al., 2021; Vansteenkiste et al., 2017; Zeuwts et al., 2016, 2023)
	Maximum fixation duration per AOI	The maximum duration of fixations on an AOI	Longer maximum fixation durations (on the phone) indicate cyclists' possibility to prepare and plan ahead.	(Ahlstrom et al., 2016)
Fixation Distance	Fixation distance	The distance between the cyclist and the first fixation of an AOI, measured in meters	AOIs with longer fixation distances indicate higher demand on safety concerns.	(Pashkevich et al., 2022; Rupi & Krizek, 2019)
	Time to arrival	The duration between the first fixation of an AOI and arrival, measured in seconds	AOIs with longer time to arrival indicate its visual salience or cyclists' needs for longer reaction time.	(Gadsby et al., 2022; Zeuwts et al., 2021, 2023)

	Saccade Count	Total number of saccades per AOI	The total number of saccades on an AOI during a ride	AOIs receiving more saccade counts capture more attention.	(Ramirez Juarez et al., 2023)

4.1 General metrics

After identifying fixations and saccades, descriptive statistics on general gaze metrics are generated directly by software provided with the eye-tracking glasses. General metrics, also known as global measures (van Paridon et al., 2019), are gaze indicators that offer an overall description of visual behaviors without specifying the objects being observed. Thirteen studies utilized general metrics before specifying the targets of fixations, and three studies used general metrics only.

Fixation duration and fixation dispersion are the two most commonly used metrics to infer workload and stress, which are essential proxies for cyclists' travel satisfaction and comfort. Measurements of fixation duration and fixation dispersion are applied to a wide range of research topics, such as the impact of different bike infrastructure, pavement quality, intersection layout, lane width, experimental settings, and the influence of listening to music while cycling. Regarding measurements of fixation duration, four studies used mean fixation duration (i.e., the average duration of all fixations) (Guo et al., 2023; Mantuano et al., 2017; Stelling-Konczak et al., 2018; von Stülpnagel, 2020) and three used percentage of fixation time (i.e., the percentage of summed fixation time to the total trip duration) (Mantuano et al., 2017; Vansteenkiste et al., 2013, 2015a). Interpretations of fixation duration measurements are consistent: shorter mean fixation durations and a larger percent of total fixation time are associated with higher cognitive workload and higher levels of stress. Fixation dispersion refers to the variability and randomness of gaze locations. Five studies measured fixation dispersion using horizontal or vertical fixation variability (i.e., the standard deviation of gaze locations on the X or Y axis) (Guo et al., 2023; Ryerson et al., 2021; Vansteenkiste et al., 2013, 2014b; Zeuwts et al., 2016), two studies applied entropy measurements (i.e., Stationary Gaze Entropy and Gaze Transition Entropy) (Guo et al., 2023; van Paridon et al., 2019), and another two studies used gaze angular velocity (i.e., the angular degree of fixations between frames) (Ryerson et al., 2021; Zhao et al., 2023). There are different interpretations of fixation dispersion measurements. For example, less horizontal variability is explained as increased workload on low-quality pavement and narrow lanes (Vansteenkiste et al., 2015a, 2017), but also explained as decreased workload when cycling on more protected bike facilities (Guo et al., 2023). We elaborate on these contradictions in our discussion section.

Other types of general metrics used less frequently are fixation counts, saccade counts, fixation angle, and fixation distance. Fixation and saccade counts are used to describe visual search patterns, such as scanning a smartphone and switching to look at a cycling lane. Researchers also normalize the total counts by time to calculate fixation or saccade frequency. These measurements are often used to compare between transportation modes or different experimental settings to demonstrate different visual behaviors (Gay et al., 2023; Mantuano et al., 2017; Pashkevich et al., 2022). In particular, one study uses saccade frequency to examine how using a phone influences cyclists' scan of the environment (Jiang et al., 2021). Fixation distance and fixation angle are used to infer the difficulty of detecting hazards on the road. Unlike the previous measurements generated from software programs, these entail manual estimations of the distance and angle between cyclists' current body location and their fixation location, as illustrated in Figure 2. The hypothesis is that longer fixation distances and larger fixation angles are associated with the increasing need to detect safety `hazards from further away and in varied directions (von Stülpnagel, 2020; Zhao et al., 2023). Fixation distance and angle are less commonly applied, as these can introduce subjective variation and require manual estimation.

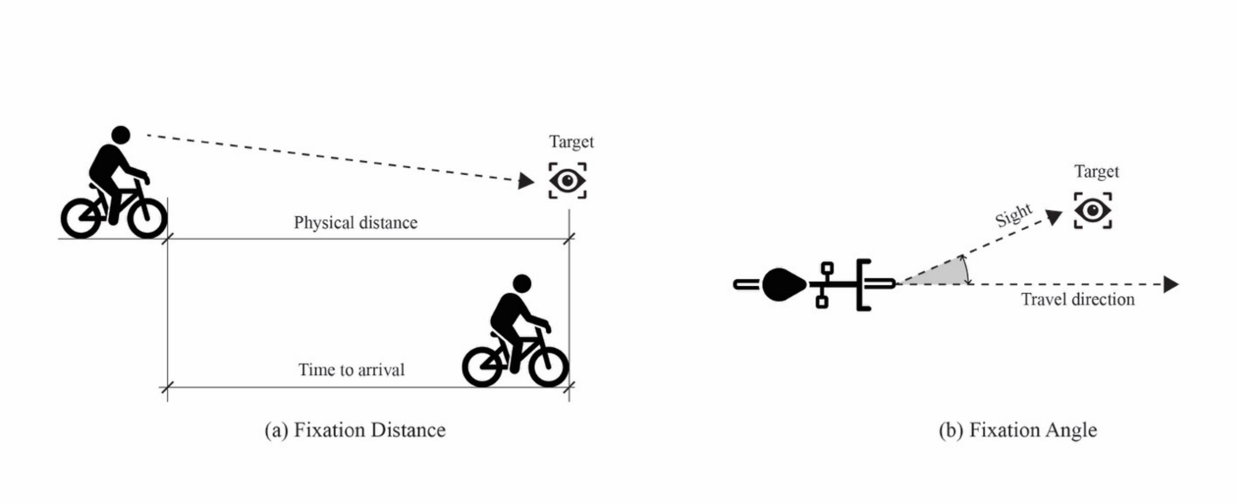
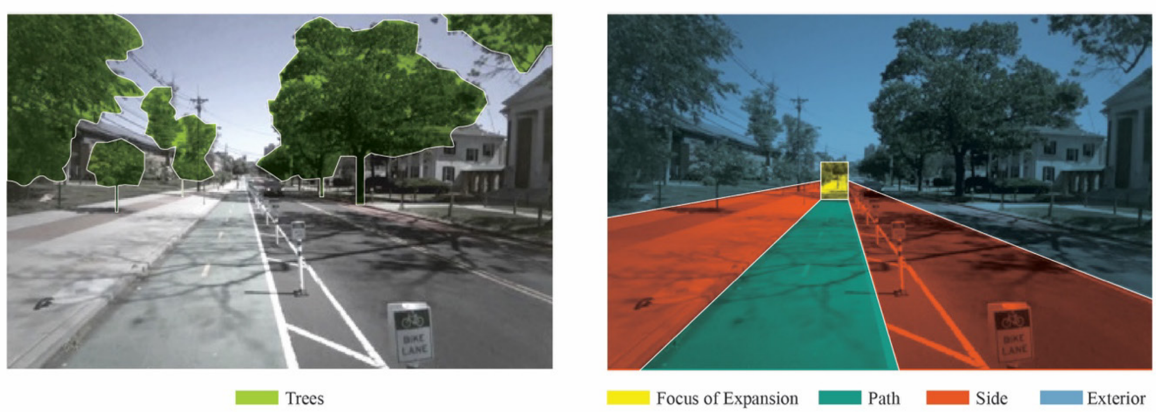


Figure 2 Diagrams of fixation distance and fixation angle

4.2 Area-Of-Interest (AOI) related metrics

Compared with general gaze measurements, AOI-related gaze measurements require the specification of fixation targets, which is crucial for understanding where attention is allocated. AOI-related metrics are also more commonly used than general metrics, thirty-two of the thirty-five reviewed papers employed gaze metrics related to specific AOIs. This section will first review AOI annotations and the metrics applied.

AOIs are identified based on research objectives and could be specific objects or zones of sight. Studies with object-based AOIs usually extract only the objects of interest, such as traffic signal lights (Rupi & Krizek, 2019), other road users (Zeuwts et al., 2021), and pavement issues (Gadsby et al., 2022). Identifying objects that are small in size and short in fixation duration require higher precision of AOI labelling. Zone-based AOIs are demarcated under the hypothesis that each zone provides different information needed for cycling. The most frequently applied AOI zones include "path" as the cycling track for lane keeping and pavement monitoring; "goal" or “focus of expansion” as the intersection of cycling trajectory and the horizon for navigation and wayfinding; "sides" as areas next to the cycling track with potential hazards from other road users; "external" as areas outside the cycling path with little cycling-related information; "behind left" and "behind right" as indications of cyclists checking over shoulders; and “phones” when smartphone tasks were assigned. The number of zones varies from three (Nygårdhs et al., 2018) to seven (Kircher & Ahlström, 2020). The more zones and smaller zone sizes, the more a zone resembles an object. It is also possible to combine zones of AOI and objects of AOI. For instance, van Paridon et al. (2019) annotated both “path” to indicate general pavement areas and “pothole” to identify the specific pavement problem. In addition to objects and zones, Kircher & Ahlström (2023) argued that AOI annotations should be more purposefully related to the required attention needed for the cycling task, and categorized the fixated areas into “necessary”, “useful”, and “not required”.



(a) Object based AOIs (b) Zone based AOIs

Figure 3 Examples of object based AOI and Zone based AOI

In most studies AOIs are delineated and annotated manually. Six studies reported recruiting two raters for AOI annotation and validated their reliability on a selected portion of the data. Cohen’s kappa (Kircher & Ahlström, 2020; Stelling-Konczak et al., 2018) and Pearson correlation (van Paridon et al., 2019; Vansteenkiste et al., 2013, 2014b, 2015a) are calculated to indicate agreements between raters. Only one study applied computer vision techniques to automate the process (Aasvik & Fyhri, 2022). The reliability concerns of the AOI annotation method are discussed further in Section 6.

AOI-based analysis introduces new terms such as dwell, visit, and glance, which describe gaze behaviors associated with fixations and saccades within an Area of Interest (AOI). These terms often include varying definitions concerning the inclusion of the initial saccade into the AOI, blinks, and invalid data. Such inconsistencies arise mostly from different eye tracking devices and software but also from researchers adopting varying definitions across publications. Reporting proprietary metric rather than a standard metric and not offering sufficient technical details of processing methods can lead to confusion.

For example, “dwell” is defined by SMI BeGaze as the sum of all fixations and saccades that hit the AOI. However, its usage varies: some follow SMI BeGaze’s definition (Vansteenkiste et al., 2014); some consider only fixations within the AOI (Rupi & Krizek, 2019; van Paridon et al., 2019; Zeuwts et al., 2023); others include fixations, saccades and also blinks (Mantuano et al., 2017; van Paridon et al., 2021). Moreover, (Vansteenkiste et al., 2015b) proposed a fixation-by-fixation method to calculate dwell, which is argued to be more time efficient than the classic frame-by-frame analysis and highly correlated with the latter. Although the term “dwell” is still used, the fixation-by-fixation method essentially only counts fixations, excluding saccades and blinks. “Visit” is defined by Tobii Pro Lab as “the data from the start of the first fixation inside and AOI until the last fixation in the AOI, including saccades, blinks or invalid gaze data”. The same is adopted by (Gay et al., 2023). “Glance” is defined by SMI BeGaze as dwell plus the first saccade leading to the AOI. Tobii Pro Lab has a similar definition but incorporates blinks and invalid data. Researchers using SMI products likely adopt SMI BeGaze's definition on glance (Ahlstrom et al., 2016; Kircher & Ahlström, 2020; Nygårdhs et al., 2018). However, some researchers set thresholds of minimum glance durations based on fixation studies, causing confusion about whether their glance consist only of fixations (Kircher & Ahlström, 2023; Stelling-Konczak et al., 2018). Despite differences in whether the first saccade, subsequent saccades, or blinks are included, the count metric for glance/dwell/visit remains consistent: all eye movements from entering to leaving the AOI are counted as a single event. This count differs from the of number of fixations, where two successive fixations within an AOI are counted as two events, making it inappropriate for quantitative comparison across studies that use different counting methods. The duration metric, on the other hand, varies with whether saccades and blinks time are counted in. The impact of these variations on glance/dwell duration depends on the characteristics of the specific AOI, such as its size and location. The magnitude of this duration difference is examined to be insignificant in the study by (Vansteenkiste et al., 2015b) where the sight was divided into 5 AOI zones. Clearer reporting of the components included in these terms could improve discussions and understanding within the field.

Count and duration over fixation, dwell, and glance are the most used AOI-related metrics for indicating attention. Their interpretations remain consistent: more counts and longer durations on a specific AOI are associated with more attention paid to the AOI. Measurements of counts include the percentage of event counts per AOI and the number of events per minute per AOI. Measurements of duration include the percentage of the event time per AOI, mean or median event duration per AOI, and total event time per AOI. Especially, the percentage of fixation/dwell/glance time per AOI is the mostly used metrics of all. Nineteen out of the thirty-five studies used this metric, indicating its acceptance and reliability. When the AOIs are related to traffic hazards that need to be observed to ensure safety, such as crossing pedestrians and door-openings of parked cars, researchers apply the measurement fixation rate (i.e., the percentage of AOIs being fixated on to the total number of safety-related AOIs) to describe hazard detection capabilities (Bishop et al., 2023; Gadsby et al., 2022; Zeuwts et al., 2023). In addition, one study used maximum fixation duration on smartphones to indicate how well-prepared cyclists are to perform secondary tasks at preferred locations (Ahlstrom et al., 2016).

Fixation distance is another type of AOI-related metric less commonly applied. Fixation distance is measured both in terms of the physical distance between the cyclist and the first fixation of an AOI and the time difference between the first fixation of an AOI and the cyclist arriving at the AOI. Selected AOIs to apply fixation distance include traffic lights, pavement issues, and traffic hazards. Interpretations of fixation distance are less consistent. In some studies it was stated that AOIs fixated from a longer fixation distance lead to more caution and require longer reaction times (Rupi & Krizek, 2019; Zeuwts et al., 2021, 2023), others ascribe it to how prominent the AOI is to the cyclist instead of how urgent it is (Gadsby et al., 2022).

5 Findings by research topic

5.1 Built environment features

The built environment is the physical surroundings that cyclists encounter when riding, such as bike facilities, pavement quality, lane characteristics, and intersection layouts. Its composition, design, and quality can directly impact cyclists. Eighteen studies investigated the impact of built environment features.

Studies on bike facilities compared cycling in mixed traffic with various features of bike lanes, such as painted bike lanes, bike lanes protected from vehicles by bollards or flowerpots, and raised bike lanes alongside sidewalks. Cycling on painted and separated bike lanes, compared with mixed traffic, is associated with less dispersion in horizontal gazes, less dispersed fixations, longer mean fixation duration, and higher percentage of fixation count on the road center, implying an increased focus on the area directly ahead rather than lateral eye movements to the side (Guo et al., 2023). Cyclists on bike lanes protected by bollards are also found to have a higher percentage of fixation time on distractions such as street furniture and buildings, likely due to reduced task difficulty when using more protected bike lanes (Jang & Kim, 2019). When cycling in mixed traffic, debris and potholes are noticed less, with medium fixation duration, and shorter fixation time to arrival, indicating the chance of missing safety cues in a less protected cycling environment (Gadsby et al., 2022).

Pavement conditions were investigated for their general surface quality and specific pavement problems. Compared to high-quality surfaces, cycling on low-quality surfaces is associated with significantly higher fixation frequency, a larger percentage of fixation time on areas surrounding the cycle lane, lower mean fixation duration on the distant environment, and less dispersed horizontal gaze distributions. These gaze metrics reflect cyclists' adaptation to the increased task demands when riding on low-quality pavement, but an increased attention to the road is at the expense of fewer visual searches for safety hazards in the surrounding area (Vansteenkiste et al., 2014b, 2017). Uneven pavements tend to have higher fixation rate than potholes and debris on the road. This implies that uneven pavements attract greater attention from cyclists. The researchers surveyed their participants and found that uneven pavement is rated less harmful to cycling safety and suggested that the significance of attention to unevenness revealed in gaze metrics is likely due to it being very visible and noticeable to cyclists (Gadsby et al., 2022).

Studies on lane characteristics examined cycling lane width, cycling lane curvature, truck loading zone marking type, and loading zone width. Cyclists riding on wider lanes have a larger percentage of fixation time on the end of lanes and external regions and less on the near pathway, suggesting that decreased task demand on steering allows for looking more on non-task-relevant areas (Vansteenkiste et al., 2013, 2015a). When riding on a curvy path, cyclists adjust where they look to have better steering control. The inside and the center of the curvy lane received a higher percentage of fixation time when entering the curve and-a lower percentage of fixation time when leaving the curve (Vansteenkiste et al., 2014a). Examining the effects of different pavement markings of the truck loading zone (i.e., a designated marked area next to the cycling lane where trucks park), cyclists have a longest total fixation duration on dashed green markings compared with white lane markings and solid green markings, showing that dashed green markings are more successful in arousing attention (Abadi et al., 2022). Compared with a wide truck loading zone, cyclists passing by trucks in a no-loading zone and a minimal loading zone have a longer total fixation duration on the truck, indicating more alertness (Jashami et al., 2023).

Features associated with intersection layouts include bike lane treatments at intersections and intersection openness. Cyclists entering intersections without continuous bike lanes fixate on traffic lights further ahead due to the increased need to anticipate risks (Rupi & Krizek, 2019). Comparing the effect of bike signals and bike box (i.e., a designated area at the head of a traffic lane at a signalized intersection that allows cyclists to get ahead of queuing traffic during the red light), both treatments shorten cyclist's total fixation time on turning vehicles that pose a potential conflict for the cyclists. This could be explained by the increased ratings of perceived safety, but cyclists' lowered attention may increase crash risk for errant drivers not yielding

(Scott-Deeter et al., 2023). Cyclists riding at spatially complex intersections with larger visibility exhibit shorter fixation durations, longer fixation distances, and larger angular differences between gaze and motion direction, interpreted as an increase in perceived risk (von Stülpnagel, 2020).

5.2 Human factors

Human factors refer to cyclists' characteristics, capabilities, and interactions with secondary tasks. Seventeen studies involving human factors examined the impact of age, gender, cycling experience, route familiarity, cycling speed, mental fatigue, and smartphone distractions.

Cyclist age is studied by comparing children and adults and children of different ages. When tested over short indoor routes, children show similar gaze patterns as adults on medium and wide lanes when asked to cycle at their personal preferable and high cycling speeds, demonstrating that children are able to adopt a similar visual-motor strategy as adults for simple precision steering tasks (Vansteenkiste et al., 2015a). In outdoor naturalistic settings, children are found to have more fixations per second and a larger percentage of dwell time on areas not related to the cycling task such as objects along the side of the road and the surrounding area. These gaze patterns reveal children's lower capability to prioritize safety cues and process information from peripheral sight (Vansteenkiste et al., 2017). Comparing child cyclists aged between 6 to 12 years old and 13 to 19 years old, the fixation distribution across AOIs does not vary between age groups, and both age groups manage to monitor more than 80% of the safety targets (Kircher & Ahlström, 2023).

Two studies examined the impact of gender on cyclists' visual behavior. One found no gender difference when riding on different bike facilities (Guo et al., 2023). Another found that men exhibit longer total fixation duration on the pavement markings of truck loading zones, but both genders show a similar amount of attention on trucks (Abadi et al., 2022). Cyclists more familiar with the route and with higher levels of skill have longer mean fixation duration, longer fixation distance, and more gazes to all sides when riding in challenging locations, such as the end of a cycling track and a complex intersection (von Stülpnagel, 2020). Experienced cyclists also exhibit longer total fixation duration on traffic lights when crossing intersections (Rupi & Krizek, 2019). Cyclists with mental fatigue fixate on hazards 1.5 seconds later, indicating attention deterioration and increased danger (Zeuwts et al., 2021). Cyclists rating the test location as more dangerous present shorter mean fixation duration, shorter fixation distances, and larger fixation angles (von Stülpnagel, 2020).

Cycling speed also affects where cyclists fixate on. When asked to cycle at a speed lower than their personal preference, they had a higher percent of dwell time on the near path and road canter (Vansteenkiste et al., 2013, 2014b). When cycling at higher speeds, a higher percent of fixation time is placed on the distant cycling trajectory when the lane is straight (Vansteenkiste et al., 2013, 2015a) and the inside of the road when riding on a curvy path (Vansteenkiste et al., 2014a).

Studies of secondary tasks provided an analysis of the influence of music, phone calls, texting, searching websites, checking bike computers, and user interface displays. One study found that listening to music slightly reduces fixations in the front road area and left area of sight (Jiang et al., 2021), while three others found no significant influence of listening to music on cyclists' visual behaviors (Ahlstrom et al., 2016; Nygårdhs et al., 2018; Stelling-Konczak et al., 2018). When engaged in phone calls, texting, and web-searching tasks, cyclists allocate fixations to the phone mainly at the expense of reduced fixations on the regions less relevant to safety (Ahlstrom et al., 2016) and decreased saccades frequency (Jiang et al., 2021). Tasks initiated by cyclists themselves result in longer total fixation duration on the phone and maximum fixation duration on the phone compared to receiving tasks, as cyclists have more time to plan and choose preferable locations to interact with their phones (Ahlstrom et al., 2016). Higher complexity in texting tasks is associated with fewer fixations on the road ahead (Jiang et al., 2021). Using a bike computer to monitor power output and cadence does not significantly reduce the percentage of dwell time on traffic, suggesting little influence on traffic hazard detection (Pfeifer et al., 2023). Compared with reading messages from a smartphone mounted on the handlebar, using AR interfaces that display the message in a fixed location of the sight or snapped onto moving objects reduces gaze angular velocity and angular differences between fixation and cycling directions, indicating calmer gaze patterns and higher chances to detect safety hazards in the front (Zhao et al., 2023).

5.3 Mode comparison

Attention demand and gaze patterns differ across transportation modes. Two studies compared the gaze behaviors of cyclists to that of drivers, pedestrians, and E-scooter riders.

Comparing how cyclists and drivers attend to safety-related visual cues at urban intersections, cyclists have a significantly lower percentage of fixation counts on the road ahead and a higher percentage to the sides, suggesting that cyclists have place more attention on monitoring traffic than drivers. (Kircher & Ahlström, 2020). Comparing the fixation distribution of pedestrians, cyclists, and E-scooter riders on a shared road, pedestrians frequently look at the sides (40.3% of fixation counts, compared to 14.6% for cyclists and 15.3% for e-scooter riders), and cyclists observe the road ahead more (42.5% of fixation counts, compared to 25.2% for pedestrians and 38.6% for e-scooter riders). This indicates a comparable amount of effort put into visual attention for cycling and e-scooter riding (Pashkevich et al., 2022).

5.4 Methodology assessment

Studies on methodology assessed the robustness and validity of using eye-tracking for assessing cyclist behavior. One study compared two AOI annotation methods, and four studies investigated the ecological validity of experiment settings.

AOI annotation is a critical step in gaze data analysis requiring labour-intensive manual labelling. Comparing annotating AOIs frame-by-frame and fixation-by-fixation, it is found that the latter can reduce the analysis time by a factor of nine while maintaining high consistency with the classic frame-by-frame method. However, AOIs that are small in size, not frequently focused on, or inconsistently specified exhibited larger discrepancies between the two methods (Vansteenkiste et al., 2015b). For instance, when identifying fixations on the curb, which is narrow and can be specified as part of the sidewalk or part of the bike lane by different people, the traditional frame by frame method is suggested as more appropriate.

Ecological validity examines whether the simulated environments resemble naturalistic cycling. Three experiment settings are examined: watching a cycling video (i.e., riding a stationary bike and facing a big screen, motion, and sight unsynchronized), immersive virtual environment (IVE) (i.e., riding a stationary bike and facing a big screen, syncing motion and sight), and Virtual Reality (i.e., riding a stationary bike with VR glasses, syncing motion and sight). Compared to naturalistic cycling, participants watching a cycling video showed less vertical fixation dispersion and less dwell time on other pedestrians and cyclists, reflecting decreased visual search when watching a video. Increasing the cycling task complexity, such as watching a video of or cycling on low-quality pavement paths, reduces the discrepancy between the two experiment settings (Zeuwts et al., 2016). Compared with the real world, cyclists in an immersive virtual environment (IVE) presented longer mean fixation duration and less vertical fixation dispersion. These less attentive gaze behaviors are likely due to the perceived safety of cycling in simulated environments (Acerra et al., 2023; Gay et al., 2023). When cycling past a parked bus, cyclists in VR settings fixated on the bus for 20% longer than naturalistic cycling, albeit the traffic volumes were not controlled for in the two scenarios (van Paridon et al., 2021). Although there is no previous study that directly compares these three experiment settings with each other, all use stationary bikes to eliminate the risks of cycling in real traffic, which increase perceived safety of the cycling task. While the effect of less attentive visual search can be reduced by increasing cycling task complexity (e.g., cycling in more complex traffic), researchers should also be careful about other aspects of restrictions, including the limited use of peripheral vision when using the screen, the fidelity of virtual environments and the impact of motion sickness.

6. Discussion

The increasing use of eye tracking devices in cycling experiments has demonstrated their usefulness and effectiveness at deciphering cyclist attention and perception. However, we also uncovered inconsistencies and limitations which we focus on now.

The first issue is the confusion stemming from unclear reporting of terms and definitions. As noted in section 4.2, there are varied definitions concerning whether first saccade, subsequent saccades, and blinks are included in terms such as glance and dwell, which especially affects the duration metric. Furthermore, even when focusing solely on fixations, parameters set to define a fixation are not always disclosed, and the influence of setting different fixations parameters has not been adequately explored. Four studies define a fixation as a gaze on an identical location for at least three consecutive frames. However, due to the sampling rates of different devices, two studies set the minimum fixation duration at 120ms (Vansteenkiste et al., 2013, 2015a) and the other two at 100ms (Mantuano et al., 2017; Rupi & Krizek, 2019). Three studies cited the same literature on the relationship between fixation duration and attentive processing (Velichkovsky et al., 2002, 2003) but adopted different values. One study set the minimum fixation duration at 200ms to infer extensive processing (Stelling-Konczak et al., 2018), another used 150ms to indicate hazard perception (Gadsby et al., 2022), and the third adopted 90ms to include shorter fixations in the pre-attentive processing phases (Zeuwts et al., 2021). These discrepancies in fixation parameters pose challenges when comparing findings quantitively between studies. As standardizing reporting of measures is crucial in enhancing the replicability, applicability, and comparability of the findings, the definitions and parameters need to be described in sufficient details that the measure can be replicated (Green, 2012). Future studies could further examine how different composition of glance/dwell and fixation parameters affect analysis results.

Secondly, the assumption that fixation-related eye movement events indicate attention deserves careful investigation. Attention is a cognitive process that can be conceptualized from multiple perspectives. Overt attention refers to directing fixations toward the AOI, whereas covert attention indicates processing information from peripheral vision without making accompanying eye movements (Posner, 1980). Evidence from driver studies suggests that peripheral vision without direct fixation is sufficient for lane-keeping and velocity estimation (Lamble et al., 1999). Studies of smartphone treatments also discuss the role of peripheral vision in navigation when cyclists fixate on phones (Ahlstrom et al., 2016; Nygårdhs et al., 2018). The excessive focus on fixations, neglecting insights from peripheral visual information, overlooks chances to explore covert attention. Additionally, attention can be classified as bottom-up or top-down. Bottom-up attention corresponds to the object's salience (i.e., the quality of being visible or noticeable), and top-down attention refers to conscious deliberations affected by subjective experience and evaluations (Aasvik & Fyhri, 2022; Connor et al., 2004). This distinction affects how gaze metrics are interpreted. For instance, pavement unevenness is fixated on earliest and for the longest time compared to potholes and cracks, which is explained by the size and visibility of unevenness instead of higher perceived danger (Gadsby et al., 2022). One method to assess fixation and top-down attention allocation is adopting the thinking-aloud verbal protocol (Ericsson & Simon, 1980), asking the participants to articulate their thoughts during cycling. The downside of the thinking-aloud method is that it lengthens cyclists' fixation duration and deviates from the natural riding experience (Hertzum et al., 2009). Future research may explore more on the effect of visual salience, the role of peripheral vision, and the influence of saccade patterns and gaze sequences.

The third issue is the ambiguity in concepts such as workload and stress. This can be noted from the inconsistencies of gaze metrics interpretations. Less horizontal variability is explained as both increased workload when, cycling on low-quality pavement and narrow lanes (Vansteenkiste et al., 2015a, 2017); and decreased workload, when cycling on more protected bike facilities (Guo et al., 2023). While both findings could be valid, the conflicted interpretations of horizontal fixation variability result from lack of distinctions on workload types. Cycling related tasks can be categorized into navigation, guidance, and control; the complexity and demand of the tasks are affected by varying factors (i.e., route familiarity affects navigation demand, traffic complexity affects guidance demand, and pavement smoothness affects the demand for control) (Bigazzi et al., 2022). As different cycling tasks have different attentional requirements, increased workload in control and decreased workload in guidance could lead to the same gaze pattern (i.e., increased concentration on the road). Future studies would benefit from precise distinctions on workload types when interpreting gaze metrics.

Furthermore, workload is sometimes used interchangeably with stress in the reviewed studies (Guo et al., 2023; Ryerson et al., 2021). Workload and stress are two related concepts derived from different theoretical frameworks. Stress refers to a state of a disorganized physiological system that hinders well-being, but high workload and increased arousal could also be associated with enhanced task performance (Gaillard, 1993). According to the Task Capacity Interface (TCI) model (Fuller, 2005; Fuller et al., 2008), cyclist behavior is related to a balance between task demand and cyclist's capabilities. Cyclists adapt their mental workload to cope with task complexity change. When the task demand is low (e.g., cycling on protected facilities in low traffic environment), cyclists' capability exceeds the task demand, and the spared capacity allows for conducting secondary tasks (e.g., using smartphones or enjoying the surrounding scenery). When the task complexity rises to a point where cycling task demands exceed the cyclist's capacity, stress is triggered as an emotional response. One possible solution to better infer stress is to combine multiple gaze metrics. For example, complex traffic situations put a higher demand on traffic monitoring, which could be quantified as both shorter mean fixation durations and higher horizontal fixation dispersions. In addition, future studies could consider triangulating eye-tracking data with additional physiological sensors, such as Galvanic Skin Response (GSR), Skin Temperature (ST), Heart Rate (HR), and Heart Rate Variability (HRV) (Bigazzi et al., 2022; Caviedes & Figliozzi, 2018; Cobb et al., 2021; Lim et al., 2022). By triangulating data from eye-tracking, physiological sensors, and self-reported surveys, researchers can better understand cyclists' emotions and perceptions.

Fourthly, optimization of AOI analysis is needed to address the time-consuming and reliability challenges regarding the current manual annotation method. Due to the large amount of data collected by eye-tracking devices, assigning fixation to AOIs remains a labor-intensive and time-consuming task despite efforts to streamline the process. The stake of reliability is especially high when labelling AOIs that are small in size, short in fixation durations, and intertwined with proximate AOIs (Vansteenkiste et al., 2013). The reliability concern is also evident in distance-related metrics, as raters subjectively estimated the distance between the participants' location and the fixated objects based on scene videos and reference maps (von Stülpnagel, 2020). Optimization in analysis techniques is needed to address these concerns. Regarding the zone-based AOIs, annotate the zones based on required safety attention, especially at critical locations such as intersections, would be beneficial to streamline the analysis. For instance, looking left or right before crossing a street remains crucial regardless of the presence of other road users. Tracking these areas can ensure safety-relevant behaviors are captured even when specific objects are absent. Regarding the object-based AOIs, computer vision algorithms in image segmentation, such as PSPNET and Segment Anything (Kirillov et al., 2023; Zhao et al., 2017), prove to be of great potential to automate AOI annotation. The automation process not only saves time but also increases reliability. When applying one algorithm to specify fixation targets, the AOIs are specified based on the same method without subjective viewpoints. In addition to annotating the targets of fixations, these algorithms could also be applied to the whole scene video to recognize all objects appearing in sight, which opens new opportunities to study peripheral vision.

Lastly, future research could explore innovative ways to leverage the data collected from the eye-tracking devices. One promising approach involves transforming the existing 2D gaze metrics analysis into 3D analysis by combining gaze data with other sensor inputs, such as GPS and gyro sensor data with ground-based LiDAR data. Moreover, as digital twins become increasingly detailed and realistic, there is a pressing need to incorporate human agents for better simulations (Fotheringham, 2023). Researchers can utilize the comprehensive cyclists' gaze and behavioral data to address this gap and create more lifelike cycling agents. Furthermore, the data could be utilized to develop generative artificial intelligence models capable of generating new, realistic cycling behavior data based on input information. These innovations can potentially enhance cycling behavior modelling and improve cycling satisfaction.

This review has its limitations. Despite the efforts to conduct a thorough search, some articles may be missed, given the extensive and multidisciplinary nature of the literature. This review only included peer-reviewed journal articles published in English, which may have excluded relevant articles published in other venues and languages. Although caution is exercised during the extraction and summary of descriptive information, reviewer bias and errors may still be present. Our review delineates how various eye tracking metrics are used, addressing the "what" and "how". However, determining which metric performs better or is more appropriate under specific conditions is complex and falls outside the scope of this review. This is a crucial question for researchers using eye trackers, and addressing it requires careful experimental design and direct comparison. Furthermore, although our review highlights the quantitative capabilities of various gaze metrics, the findings predominantly reflect qualitative outcomes (i.e., showing directions of correlations rather than magnitudes of intensity). The considerable diversity in experimental treatments across studies and the absence of standardized metrics reporting make quantitative conclusions mostly incommensurable. Additionally, our review does not encompass critical oculomotor events such as pupil dilation, blinking, and smooth pursuit. These events could offer valuable insights into cyclists' visual behavior from angles other than fixations. Unfortunately, they were not featured in the articles available for this review.

7. Conclusions

Using eye tracking devices in cycling experiments to study cyclists' behavior and perceptions is a relatively novel approach. The experiment design and selection of gaze metrics are vital for generating insights that can enhance safety measures and inform bicycle infrastructure planning. We reviewed 35 peer-reviewed studies and summarized experiment characteristics and gaze metrics. A total of six brands of devices appeared in our review. Apart from VR and AR headsets, only two brands (Pupil Labs and Tobii) are currently offering mobile wearables on the market. In addition to outdoor cycling, experiment settings such as immersive virtual environments (IVE) and virtual reality (VR) circumvent weather restriction and mitigate the outdoor cycling risk, at the expense of motion sickness over prolonged sessions and restricted field of view. We introduce a two-category framework to organize the variety of gaze metrics: general gaze metrics, which reveals cyclists' cognition workload, and AOI-related gaze metrics, which help to understand cyclists' attention allocation to visual cues. Under both categories, fixation count, duration, and distance comprise major measurements. We investigate the use and interpretation of these metrics around four areas: built environment features, human factors, mode comparisons, and methodology assessment. Five research gaps are identified that merit future endeavors: standardizing the reporting of terms and parameters, cautious interpretations of fixation and attention, clarifying concepts on workload and stress, optimizing the AOI annotation method, and applying cyclists' eye tracking data in innovative fields.

With raising awareness on inclusion and technological advancement, some eye-tracking devices have been developed to be compatible with prescriptive lenses for people in need of vision correction, and to be customized for vulnerable demographic groups including children and seniors. Some articles we reviewed have explored cycling safety factors and methods especially targeting children and young cyclists (Kircher & Ahlström, 2023; Stelling-Konczak et al., 2018; van Paridon et al., 2019; Vansteenkiste, Cardon, & Lenoir, 2015; Vansteenkiste et al., 2017; Zeuwts et al., 2021, 2023). We look forward to learning more from studies that embrace these advancements to improve the inclusiveness of cycling experiments for more equitable findings.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Acknowledgements

This work is partially supported by Rutgers Research Council.

References

Aasvik, O., & Fyhri, A. (2022). Distracting or informative? Examining signage for cyclists using eye-tracking. *Traffic Safety Research*, *2*, 000013–000013. https://doi.org/10.55329/wxcy5694

Abadi, M. G., Maloney, P., & Hurwitz, D. (2022). Exploring Bicyclists' Visual Attention during Conflicts with Truck Traffic. *Transportation Research Record*, *2676*(11), 137–144. https://doi.org/10.1177/03611981221091709

Acerra, E. M., Shoman, M., Imine, H., Brasile, C., Lantieri, C., & Vignali, V. (2023). The Visual Behavior of the Cyclist: Comparison between Simulated and Real Scenarios. *Infrastructures*, *8*(5), Article 5. https://doi.org/10.3390/infrastructures8050092

Ahlstrom, C., Kircher, K., Thorslund, B., & Adell, E. (2016). Bicyclists' visual strategies when conducting self-paced vs. System-paced smartphone tasks in traffic. *Transportation Research Part F: Traffic Psychology and Behavior*, *41*, 204–216. https://doi.org/10.1016/j.trf.2015.01.010

Andersson, R., Nyström, M., & Holmqvist, K. (2010). Sampling frequency and eye-tracking measures: How speed affects durations, latencies, and more. *Journal of Eye Movement Research*, *3*(3), Article 3. https://doi.org/10.16910/jemr.3.3.6

Berto, R., Massaccesi, S., & Pasini, M. (2008). Do eye movements measured across high and low fascination photographs differ? Addressing Kaplan's fascination hypothesis. *Journal of Environmental Psychology*, *28*(2), 185–191. https://doi.org/10.1016/j.jenvp.2007.11.004

Bigazzi, A., Ausri, F., Peddie, L., Fitch, D., & Puterman, E. (2022). Physiological markers of traffic-related stress during active travel. *Transportation Research Part F: Traffic Psychology and Behavior*, *84*, 223–238. https://doi.org/10.1016/j.trf.2021.12.003

Bishop, D. T., Daylamani-Zad, D., Dkaidek, T. S., Fukaya, K., & Broadbent, D. P. (2023). A brief gamified immersive intervention to improve 11–14-year-olds' cycling-related looking behavior and situation awareness: A school-based pilot study. *Transportation Research Part F: Traffic Psychology and Behavior*, *97*, 17–30. https://doi.org/10.1016/j.trf.2023.06.019

Caviedes, A., & Figliozzi, M. (2018). Modeling the impact of traffic conditions and bicycle facilities on cyclists' on-road stress levels. *Transportation Research Part F: Traffic Psychology and Behavior*, *58*, 488–499. https://doi.org/10.1016/j.trf.2018.06.032

Cobb, D. P., Jashami, H., & Hurwitz, D. S. (2021). Bicyclists' behavioral and physiological responses to varying roadway conditions and bicycle infrastructure. *Transportation Research Part F: Traffic Psychology and Behavior*, *80*, 172–188. https://doi.org/10.1016/j.trf.2021.04.004

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual Attention: Bottom-Up Versus Top-Down. *Current Biology*, *14*(19), R850–R852. https://doi.org/10.1016/j.cub.2004.09.041

Duchowski, A. T. (2017). *Eye Tracking Methodology*. Springer International Publishing. https://doi.org/10.1007/978-3-319-57883-5

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215–251. https://doi.org/10.1037/0033-295X.87.3.215

Fotheringham, A. S. (2023). Digital twins: The current "Krays" of urban analytics? *Environment and Planning B: Urban Analytics and City Science*, *50*(4), 1020–1022. https://doi.org/10.1177/23998083231169159

Fuller, R. (2005). Towards a general theory of driver behavior. *Accident Analysis & Prevention*, *37*(3), 461–472. https://doi.org/10.1016/j.aap.2004.11.003

Fuller, R., McHugh, C., & Pender, S. (2008). Task difficulty and risk in the determination of driver behavior. *European Review of Applied Psychology*, *58*(1), 13–21. https://doi.org/10.1016/j.erap.2005.07.004

Gadsby, A., Tsai, J., & Watkins, K. (2022). Understanding the Influence of Pavement Conditions on Cyclists' Perception of Safety and Comfort Using Surveys and Eye Tracking. *Transportation Research Record*, 03611981221090936. https://doi.org/10.1177/03611981221090936

Gay, N., Trefzger, M., Sawilla, S., & Schlegel, T. (2023). How realistic is the gaze behavior in a cycling simulator? A comparative study between lab and field. *Transportation Research Procedia*, *72*, 1061–1068. https://doi.org/10.1016/j.trpro.2023.11.536

Guo, X., Tavakoli, A., Angulo, A., Robartes, E., Chen, T. D., & Heydarian, A. (2023). Psycho-physiological measures on a bicycle simulator in immersive virtual environments: How protected/curbside bike lanes may improve perceived safety. *Transportation Research Part F: Traffic Psychology and Behavior*, *92*, 317–336. https://doi.org/10.1016/j.trf.2022.11.015

Healey, P. (1992). Planning through debate: The communicative turn in planning theory. *Town Planning Review*, *63*(2), 143. https://doi.org/10.3828/tpr.63.2.422x602303814821

Hertzum, M., Holmegaard, K., & c, H. (2009). Scrutinizing usability evaluation: Does thinking aloud affect behavior and mental workload? *Behavior & Information Technology*, *28*. https://doi.org/10.1080/01449290701773842

Jang, G., & Kim, S. (2019). Investigating the effect of a raised cycle track, physical separation, land use and number of pedestrian on cyclists' gaze behavior. *Journal of Architecture and Urbanism*, *43*(1), Article 1. https://doi.org/10.3846/jau.2019.3786

Jashami, H., Cobb, D., Sinkus, I., Liu, Y., McCormack, E., Goodchild, A., & Hurwitz, D. (2023). Evaluation of bicyclist physiological response and visual attention in commercial vehicle loading zones. *Journal of Safety Research*. https://doi.org/10.1016/j.jsr.2023.11.018

Jiang, K., Yang, Z., Feng, Z., Sze, N. N., Yu, Z., Huang, Z., & Chen, J. (2021). Effects of using mobile phones while cycling: A study from the perspectives of manipulation and visual strategies. *Transportation Research Part F: Traffic Psychology and Behavior*, *83*, 291–303. https://doi.org/10.1016/j.trf.2021.10.010

King, A. J., Bol, N., Cummins, R. G., & John, K. K. (2019). Improving Visual Behavior Research in Communication Science: An Overview, Review, and Reporting Recommendations for Using Eye-Tracking Methods. *Communication Methods and Measures*, *13*(3), 149–177. https://doi.org/10.1080/19312458.2018.1558194

Kircher, K., & Ahlström, C. (2020). Attentional requirements on cyclists and drivers in urban intersections. *Transportation Research Part F: Traffic Psychology and Behavior*, *68*, 105–117. https://doi.org/10.1016/j.trf.2019.12.008

Kircher, K., & Ahlström, C. (2023). Children and youngster's gaze behavior when cycling in familiar environments. *Journal of Cycling and Micromobility Research*, *1*, 100006. https://doi.org/10.1016/j.jcmr.2023.100006

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2023). *Segment Anything* (arXiv:2304.02643). arXiv. https://doi.org/10.48550/arXiv.2304.02643

Lamble, D., Laakso, M., & Summala, H. (1999). Detection thresholds in car following situations and peripheral vision: Implications for positioning of visually demanding in-car displays. *Ergonomics*, *42*(6), 807–815. https://doi.org/10.1080/001401399185306

Lim, T., Kalra, A., Thompson, J., Caldwell Odgers, J., & Beck, B. (2022). Physiological measures of bicyclists' subjective experiences: A scoping review. *Transportation Research Part F: Traffic Psychology and Behavior*, *90*, 365–381. https://doi.org/10.1016/j.trf.2022.09.007

Lu, Z., & Pesarakli, H. (2023). Seeing Is Believing: Using Eye-Tracking Devices in Environmental Research. *HERD: Health Environments Research & Design Journal*, *16*(1), 15–52. https://doi.org/10.1177/19375867221130806

Mantuano, A., Bernardi, S., & Rupi, F. (2017). Cyclist gaze behavior in urban space: An eye-tracking experiment on the bicycle network of Bologna. *Case Studies on Transport Policy*, *5*(2), 408–416. https://doi.org/10.1016/j.cstp.2016.06.001

Mekuria, M. C., Furth, P. G., & Nixon, H. (2012). *Low-Stress Bicycling and Network Connectivity*. San Jose, CA: Mineta Transportation Institute.

Nygårdhs, S., Ahlström, C., Ihlström, J., & Kircher, K. (2018). Bicyclists' adaptation strategies when interacting with text messages in urban environments. *Cognition, Technology & Work*, *20*(3), 377–388. https://doi.org/10.1007/s10111-018-0478-y

Onkhar, V., & Dodou, D. (n.d.). *Evaluating the Tobii Pro Glasses 2 and 3 in static and dynamic conditions*.

Pashkevich, A., Burghardt, T. E., Puławska-Obiedowska, S., & Šucha, M. (2022). Visual attention and speeds of pedestrians, cyclists, and electric scooter riders when using shared road – a field eye tracker experiment. *Case Studies on Transport Policy*, *10*(1), 549–558. https://doi.org/10.1016/j.cstp.2022.01.015

Pedrotti, M., Mirzaei, M. A., Tedesco, A., Chardonnet, J.-R., Mérienne, F., Benedetto, S., & Baccino, T. (2014). Automatic Stress Classification With Pupil Diameter Analysis.

International Journal of Human–Computer Interaction, *30*(3), 220–236. https://doi.org/10.1080/10447318.2013.848320

Pfeifer, C., Leinen, P., Puhl, J., & Panzer, S. (2023). Visual behavior and road traffic hazard situations when using a bike computer on a racing bike: An eye movement study. *Applied Ergonomics*, *112*, 104070. https://doi.org/10.1016/j.apergo.2023.104070

Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Rupi & Krizek. (2019). Visual Eye Gaze While Cycling: Analyzing Eye Tracking at Signalized Intersections in Urban Conditions. *Sustainability*, *11*(21), 6089. https://doi.org/10.3390/su11216089

Ryerson, M. S., Long, C. S., Fichman, M., Davidson, J. H., Scudder, K. N., Kim, M., Katti, R., Poon, G., & Harris, M. D. (2021). Evaluating cyclist biometrics to develop urban transportation safety metrics. *Accident Analysis & Prevention*, *159*, 106287. https://doi.org/10.1016/j.aap.2021.106287

Scott-Deeter, L., Hurwitz, D., Russo, B., Smaglik, E., & Kothuri, S. (2023). Assessing the impact of three intersection treatments on bicyclist safety using a bicycling simulator. *Accident Analysis & Prevention*, *179*, 106877. https://doi.org/10.1016/j.aap.2022.106877

Shiferaw, B., Downey, L., & Crewther, D. (2019). A review of gaze entropy as a measure of visual scanning efficiency. *Neuroscience & Biobehavioral Reviews*, *96*, 353–366. https://doi.org/10.1016/j.neubiorev.2018.12.007

Siegle, G. J., Ichikawa, N., & Steinhauer, S. (2008). Blink before and after you think: Blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology*, *45*(5), 679–687. https://doi.org/10.1111/j.1469-8986.2008.00681.x

Stelling-Konczak, A., Vlakveld, W. P., van Gent, P., Commandeur, J. J. F., van Wee, B., & Hagenzieker, M. (2018). A study in real traffic examining glance behavior of teenage cyclists when listening to music: Results and ethical considerations. *Transportation Research Part F: Traffic Psychology and Behavior*, *55*, 47–57. https://doi.org/10.1016/j.trf.2018.02.031

Strange, B. A., & Dolan, R. J. (2006). Anterior medial temporal lobe in human cognition: Memory for fear and the unexpected. *Cognitive Neuropsychiatry*, *11*(3), 198–218. https://doi.org/10.1080/13546800500305096

van Paridon, K. N., Leivers, H. K., Robertson, P. J., & Timmis, M. A. (2019). Visual search behavior in young cyclists: A naturalistic experiment. *Transportation Research Part F: Traffic Psychology and Behavior*, *67*, 217–229. https://doi.org/10.1016/j.trf.2019.10.014

van Paridon, K., Timmis, M. A., & Sadeghi Esfahlani, S. (2021). Development and Evaluation of a Virtual Environment to Assess Cycling Hazard Perception Skills. *Sensors*, *21*(16), 5499. https://doi.org/10.3390/s21165499

Vansteenkiste, P., Cardon, G., D'Hondt, E., Philippaerts, R., & Lenoir, M. (2013). The visual control of bicycle steering: The effects of speed and path width. *Accident Analysis & Prevention*, *51*, 222–227. https://doi.org/10.1016/j.aap.2012.11.025

Vansteenkiste, P., Cardon, G., & Lenoir, M. (2015). Visual guidance during bicycle steering through narrow lanes: A study in children. *Accident Analysis & Prevention*, *78*, 8–13. https://doi.org/10.1016/j.aap.2015.02.010

Vansteenkiste, P., Cardon, G., Philippaerts, R., & Lenoir, M. (2015). Measuring dwell time percentage from head-mounted eye-tracking data – comparison of a frame-by-frame and a fixation-by-fixation analysis. *Ergonomics*, *58*(5), 712–721. https://doi.org/10.1080/00140139.2014.990524

Vansteenkiste, P., Hamme, D. V., Veelaert, P., Philippaerts, R., Cardon, G., & Lenoir, M. (2014). Cycling around a Curve: The Effect of Cycling Speed on Steering and Gaze Behavior. *PLOS ONE*, *9*(7), e102792. https://doi.org/10.1371/journal.pone.0102792

Vansteenkiste, P., Zeuwts, L., Cardon, G., Philippaerts, R., & Lenoir, M. (2014). The implications of low quality bicycle paths on gaze behavior of cyclists: A field test. *Transportation Research Part F: Traffic Psychology and Behavior*, *23*, 81–87. https://doi.org/10.1016/j.trf.2013.12.019

Vansteenkiste, P., Zeuwts, L., van Maarseveen, M., Cardon, G., Savelsbergh, G., & Lenoir, M. (2017). The implications of low quality bicycle paths on the gaze behavior of young learner cyclists. *Transportation Research Part F: Traffic Psychology and Behavior*, *48*, 52–60. https://doi.org/10.1016/j.trf.2017.04.013

Velichkovsky, B. M., Rothert, A., Kopf, M., Dornhöfer, S. M., & Joos, M. (2002). Towards an express-diagnostics for level of processing and hazard perception. *Transportation Research Part F: Traffic Psychology and Behavior*, *5*(2), 145–156. https://doi.org/10.1016/S1369-8478(02)00013-X

Velichkovsky, B. M., Rothert, A., Miniotas, D., Dornhöfer, S. M., Joos, M., & Pannasch, S. (n.d.). *Visual Fixations as a Rapid Indicator of Hazard Perception*.

von Stülpnagel, R. (2020). Gaze behavior during urban cycling: Effects of subjective risk perception and vista space properties. *Transportation Research Part F: Traffic Psychology and Behavior*, *75*, 222–238. https://doi.org/10.1016/j.trf.2020.10.007

Zeuwts, L. H. R. H., Iliano, E., Smith, M., Deconinck, F., & Lenoir, M. (2021). Mental fatigue delays visual search behavior in young cyclists when negotiating complex traffic situations: A study in virtual reality. *Accident Analysis & Prevention*, *161*, 106387. https://doi.org/10.1016/j.aap.2021.106387

Zeuwts, L. H. R. H., Vanhuele, R., Vansteenkiste, P., Deconinck, F. J. A., & Lenoir, M. (2023). Using an immersive virtual reality bicycle simulator to evaluate hazard detection and anticipation of overt and covert traffic situations in young bicyclists. *Virtual Reality*, *27*(2), 1507–1527. https://doi.org/10.1007/s10055-023-00746-7

Zeuwts, L., Vansteenkiste, P., Deconinck, F., van Maarseveen, M., Savelsbergh, G., Cardon, G., & Lenoir, M. (2016). Is gaze behavior in a laboratory context similar to that in real-life? A study in bicyclists. *Transportation Research Part F: Traffic Psychology and Behavior*, *43*, 131–140. https://doi.org/10.1016/j.trf.2016.10.010

Zhao, G., Orlosky, J., Gabbard, J., & Kiyokawa, K. (2023). HazARdSnap: Gazed-based Augmentation Delivery for Safe Information Access while Cycling. *IEEE Transactions on Visualization and Computer Graphics*, 1–10. https://doi.org/10.1109/TVCG.2023.3333336

Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). *Pyramid Scene Parsing Network*. 2881–2890. https://openaccess.thecvf.com/content_cvpr_2017/html/Zhao_Pyramid_Scene_Parsing_CVPR_2017_paper.html