Volume xx (200y), Number z, pp. 1–17

# Gaze-Aware Visualization: Design Considerations and Research Agenda

R. Jianu†[1], N. Silva[2,3], N. Rodrigues[4], T. Blascheck[4], T. Schreck[5], D. Weiskopf[4]

[1]City, University of London, UK
[2]Medical University of Graz, Austria
[3]Interdisciplinary Transformation University, Austria
[4]University of Stuttgart, Germany
[5]Graz University of Technology, Austria

**Abstract**

*Eye tracking provides a unique perspective on the inherently visual discourse between visualization systems and their users, and has recently become sufficiently precise and affordable to be integrated as regular input into workstations and virtual or augmented reality headsets alike. As such, real-time eye tracking can now contribute significantly toward the development of gaze-aware visualizations that infer and monitor users' needs to actively support their activities. To facilitate such systems we make three contributions. First, we structure and discuss design considerations for gaze-aware visualizations along four axes: measurable data; inferable data; opportunities for support; and limiting factors to beware. Second, we distill visualization research challenges that preclude such systems. Finally, we show via three usage scenarios how to apply these design considerations to imagine how existing systems can benefit from real-time eye tracking. We combined a structured literature analysis, a consideration of suitable places for eye-tracking integration in the typical visualization ecosystem, and design space modeling. Eye tracking has significant potential to improve the interactive visual analysis of data across many visualization domains. Our paper attempts to provide a comprehensive, general survey and conceptual discussion in this promising field, outlining the state-of-the-art and future research opportunities.*

**CCS Concepts**

• ***Human-centered computing*** → ***Visualization; Interaction paradigms; Interaction techniques;*** *Visualization design and evaluation methods;*

## 1. Introduction

Classic pointing devices and keyboards still serve the interaction needs of most visualization systems, but natural interaction modalities such as voice, gestures, and gaze have recently come into focus. Among these, gaze is promising and often applied in studies, perception research, and as a user interface. However, its application and potential to enhance interactive visual data analysis are underexplored. We contribute a structured design framework for gaze-aware visualizations and means to inform researchers who wish to develop such systems.

The visual analysis of large, multi-faceted datasets challenges analysts' perceptual and cognitive capacities as they navigate visually dense displays, create and maintain mental models of data interdependencies, and manage analytic processes. Adaptive and mixed-initiative visualizations are seen as a potential solution [EHR*14, Hor99, Opp17]. They seek to build an awareness of an analyst's goals and contribute proactively towards their resolution. Initial work on such systems relied mainly on active interaction and explicit input such as relevance feedback, interaction analysis, provenance data, or annotation. Lately, computers' growing ability to interpret natural language using machine learning has taken center stage in supporting more seamless human-computer interaction (HCI) and demonstrates the benefits of novel interaction modalities [MNES22, SSL*22].

High-quality eye tracking has become affordable and can be a rich source of information about visual attention and analytic intentions. Eye tracking can enrich efforts to develop adaptive visualizations by providing a window into users' visual and analytic behavior that is distinct from existing traditional modalities like keyboard and pointing devices. While a few visualization systems already explored the interactive use of eye tracking, for example, to support visual search, recommend unseen data patterns, or capture the analysis process [OAJ14, SSES17, SSS*16a, SSV*18] (see Figure 1 for an illustration), the potential of eye tracking to support adaptive visualizations remains largely untapped.

† e-mail: radu.jianu@city.ac.uk

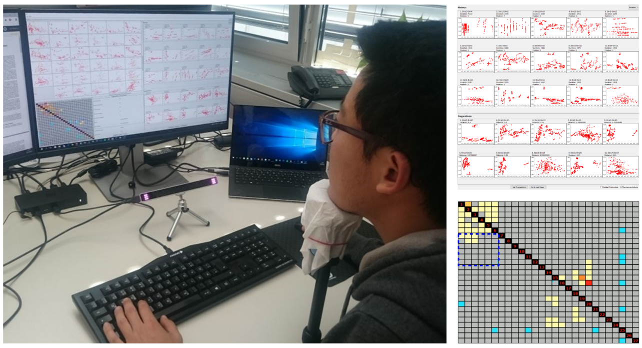

**Figure 1:** *Visual exploration of a scatter plot matrix aided by provenance-based recommendations [SSES17]. Via gaze, the system tracks the plots users view (left). Dissimilar plots are identified and recommended to encourage exploration of diverse data patterns. The history of both seen and recommended patterns are shown in sorted lists of plots (top-right). A color-coded matrix summarizes the view history and recommendations (bottom-right). Chin rest is used for evaluation only. Figure courtesy of Lin Shao.*

We contribute **(i)** a design framework for using gaze in adaptive visualizations, grounded in a comprehensive literature review (Figure 3 and Sections 4, 5, 6), **(ii)** research avenues to catalyze the use of gaze in the visualization domain (Sections 4, 5, 6), and **(iii)** three usage scenarios evidencing how the framework can inform how concrete visualization systems can benefit from eye tracking (Section 7). Our design framework follows a logical sequence of *measuring* behavior from gaze and other modalities, *inferring* about visual and analytic intent, then using this to enable *support* for operations, reasoning, and communication, while at the same time underlining challenges about inference accuracy, involved uncertainties, and goals for usefulness and adaptability.

Our work aims to help researchers and practitioners build gaze-aware visualizations by providing structured insight into the practical possibilities and challenges they ought to consider as well as catalyze visualization research on aspects that underlie such systems. More broadly, we provide a holistic view of how eye tracking can support visualization, with a particular focus on the development of multimodal, adaptive, and mixed-initiative systems. To date, relevant work has been dispersed across different research communities, including eye-tracking technology and applications (e. g., ACM ETRA), HCI (e. g., ACM CHI), and visualization (e. g., IEEE VIS). It was also unstructured, and not linked to current themes of visualization research (e. g., guidance, provenance, multimodal interaction). Applied work on interactive uses of eye tracking in visualization is also limited to a small set of technical solutions. Overall, our work complements current visualization research on guidance, provenance, and multimodal interaction, especially through natural language [CAS*18, CGM19, LSS*18, MNES22, PMCEA*22, SSL*22].

## 2. Background and related work

Our work is at the junction of adaptive, guiding, and mixed-initiative visualization; multi-modal interaction; and eye tracking.

### 2.1. Intelligent visualization systems

As data and the systems we use to interpret it become increasingly complex and multi-faceted, research into systems that can proactively support analysts and optimize their processes is underway. *Adaptive* systems and visualizations adjust themselves dynamically based on interactions, preferences, or changing data characteristics to provide tailored views, streamline interaction, and optimize data comprehension and analysis [AZZR07, CZK*22, Opp17]. *Guidance* involves providing analysts with intelligent support and recommendations during the data exploration process to assist them in decision-making and interpretation [CAS*18, CGM*17]. *Mixed-initiative* visualization emphasizes collaboration between analysts and computers, which are elevated to partners that contribute proactively to the analysis process [EHR*14, Hor99, MGG*23, SJB*21].

Much work in visualization has focused on adaptation, guidance, and mixed-initiative analytics, and we can only discuss a selection of it here. The Degree-of-Interest/Fisheye framework [Fur06] suggests defining a user-specific degree of interest function to describe the data and using this to adapt the level of data detail shown in a visualization. Brown et al. [BOZ*14] used features from the movement of a pointing device to train a model of user proficiency, which can distinguish between systematic and non-systematic search behavior as a basis for adaption. Cook et al. [CCI*15] explored how to integrate task-driven recommendations into visual analytics based on provided entities of interest. Other work proposed semantic interaction [End16] and adapting underlying data analysis models in response to interactions.

In all cases, such visualizations rely on real-time understanding of the analyst, their context, and their analytic objectives [AZZR07, KT17, MGG*23, OGW19, SJB*21]. These can be modeled from both active (e. g., keyboard, mouse, touch) and passive (e. g., camera, microphone, biophysical monitoring) input modalities.

### 2.2. Multimodal visualization

Multimodal visualization seeks to develop interaction experiences that let analysts focus on their data and promote flow by leveraging input devices and modalities that transcend the WIMP paradigm [EMJ*11, LIRC12, LSS*18]. A significant body of initial work focused on pen, touch, and gesture interactions [AEYN11, BLC*11, JLLS17, RK14]. More recently, interacting with visualization through natural language has received considerable attention [SSL*22]. Applications range from visualization generation and modification via spoken commands [AKG*16, NSS20, SLJL10] to multi-turn conversational analytics [MNES22, STD19]. The seamless combination of natural language with other modalities such as pen, touch, and gestural interaction is also effective [CMH*20, KR18, SS17, SLS20, SSS20].

Advances in machine learning and artificial intelligence are an enabling force behind these efforts. They often underlie the translation of typically messy, context-laden, and ambiguous multimodal data such as natural language utterances and free-form gestures into actionable interactions [KK23, PSMJ22, WCWQ21, WWS*21].

### 2.3. Eye tracking and applications in visualization

Eye tracking allows the mapping of the user's eye movements to fixations on a visible target, including locations on the display of a system or objects in a scene [HNA*11]. Eye trackers can capture fixations and saccades (rapid eye movements between fixations) at high spatial and temporal resolution. Fixations can then be mapped to foveated display elements. Eye trackers come in many forms, including devices for the desktop, head-mounted displays (HMDs), or software-based solutions for extracting gaze from webcams.

Eye-tracking data is relevant for many applications and studies. This is due to the *eye-mind hypothesis* [JC80]. It implies, in simple terms, that when people look at visual objects, they also attend to them. While the general validity of this is still under debate and other hypotheses on the relation between fixations and cognition exist [DS96, Pos80, WW08], many eye-tracking applications build on this presumed link between seeing and attending.

The most established use of eye tracking today is in psychometric studies, interface design, and usability evaluations [GW03, TGL*09]. Analyzing eye-tracking data allows one to assess interfaces and behavior patterns [BKR*17], evaluate visual analytics [KFBW16], understand how we read text combined with visualization [HKA*24], or bridge quantitative and qualitative research methods [Wei24].

Another use that receives increasing attention is the real-time use of eye tracking to support HCI, in particular, to facilitate adaptive interfaces. This development was driven by the advent of a new generation of high-accuracy, affordable eye trackers. Eye tracking is a particularly interesting indirect input channel for visualization interaction due to the possible link between gaze and cognitive processing [Bed05] and provides a likely useful complementary modality to natural language [ARZ*22, KNBV22]. In our previous work [SBR*19], we discussed how eye tracking may be integrated into the visual data analysis process by imagining how it may connect to the visual analytics process model of Keim et al. [KMS*08]. Additionally, we suggested ideas for adapting existing visual analytics applications. Similarly, Srinivasan et al. recently advocated that eye-tracking underlie the development of Attention Aware Visualizations (AAV) and proposed a set of design guidelines [SEB*24]. Eye tracking can also support collaborative learning and visualization by attention monitoring and guidance in environments equipped with augmented reality [RKP*24].

We now propose a larger encompassing design framework, based on the dimensions *measuring*, *inferring*, *supporting*, and a broad space of challenges affecting the design. Our work also touches on practical implementations that have used eye-tracking data to make inferences about users and their tasks to adapt visualizations accordingly. For example, a robust body of work, in particular by Conati et al., looked into leveraging gaze patterns to infer users' degree of visualization expertise and the visualization tasks they were pursuing [SCC13, TCSC13, CCTL15, CLRT17, LCT19, CLRT20]. Similarly, researchers looked at the links between gaze and momentary user states such as boredom and engagement [KGK*14, SSE*15], stress and cognitive load [SSE*15], and learning progress [TSG*14]. In more direct attempts to visual analysis, eye tracking was used to interact with and navigate hierarchy visualizations [SSS*16b], to retrieve or rank visual patterns of interest in

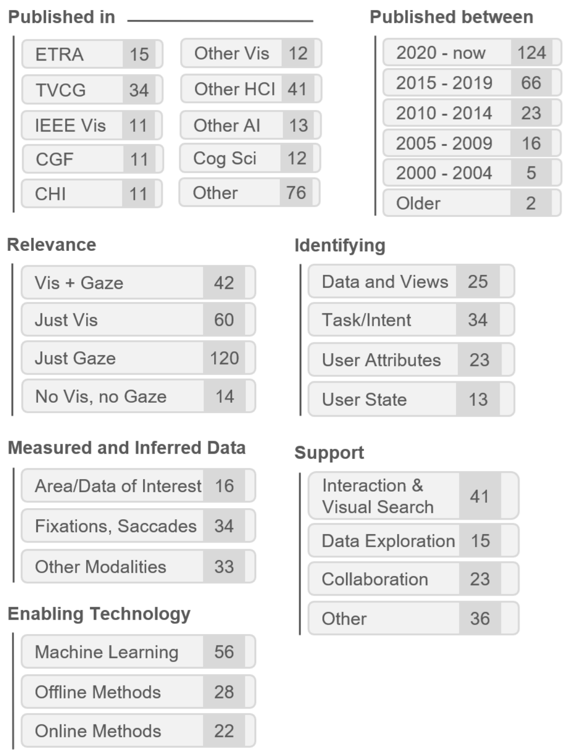


**Figure 2:** *Results of our literature analysis: **(Top)** About a quarter of the papers we reviewed were from the visualization domain (TVCG, IEEE Vis, CGF) and a quarter from publications in HCI (CHI). The rest were sourced from a range of domains including dedicated eye-tracking venues (ETRA), artificial intelligence, and cognitive science. Our analysis focused largely on recent papers, with a large majority from the past eight years. **(Bottom)** Our thematic coding reveals the most prevalent themes and topics from the considered papers. For example, 42 papers combined visualization and gaze in some way; machine learning was used to support inference in 56 papers; and 25 papers touched on the possibility of inferring data and views of interest from gaze.*

scatter plots [RSY*22, SSES17] and time series [SSV*18], to support data-reading in networks [OAJ14] and maps [GKG*18], and to navigate within and between visualization views [SLMS09]. Finally, in an attempt to facilitate the exploration of such work, Lalle et al. [LCT19] proposed a dedicated testbed platform for experimenting with gaze-adaptive visualizations.

These existing works focus on specific selected tasks and visualization techniques and do not provide a holistic view of how eye tracking fits within the broader visualization ecosystem.

## 3. Methodology

The framework shown in Figure 3 and described in Sections 4, 5, 6 is our main contribution. It serves as a checklist for researchers and practitioners who design gaze-aware visualizations, pointing them to opportunities and limitations that are important. It also suggests research avenues to catalyze the use of gaze in visualization. We relied on five activities to assemble the framework.

First, we analyzed existing literature on interactive uses of eye tracking, with an emphasis on visualization (Section 3.1). Second, we considered general facets of visualization ecosystems (e. g., data, visual encoding, interaction, users, devices) and imagined their potential synergies with real-time eye tracking (Section 3.2). This was necessary because existing literature on the use of real-time eye tracking in visualization is relatively limited. Third, we organized the resulting design considerations, challenges, and future research needs into a structured framework (Section 3.3).

Finally, we operationalized the framework for use in practice and demonstrate its ability to inform the design of gaze-aware visualizations via three usage scenarios (Section 7). We selected concrete applications to cover a diverse range of uses and technologies (e. g., business intelligence; law enforcement; desktop; mobile; virtual reality (VR)) and used our framework as a structured checklist to imagine how they could benefit from eye tracking.

### 3.1. Literature analysis

Our literature analysis aimed to uncover types of gaze support other researchers explored, as well as challenges and opportunities they identified, with emphasis on the potential to help visualization. It involved three activities: compiling a set of papers of interest; using thematic analysis to code their methods, opportunities, and challenges; and feeding those themes into design considerations.

To assemble relevant papers, we used Webster and Watson's snowballing strategy for literature reviews [WW02]. We first used our combined research experience on eye tracking and visualization to propose a preliminary list of papers broadly covering topics such as: interactive use of eye tracking in visualization; eye tracking as a means to infer users' abilities, goals, or tasks; multimodal interaction in visualization; models, principles, and examples of guidance and mixed-initiative visualisation; principles of recommendation systems and use of eye tracking in such systems. This initial list contained 71 papers. During their analysis and coding (described below), this list was expanded by a further 165 relevant papers that were either cited by ("backward snowballing") or inspired by the initial papers ("forward snowballing").

We chose this strategy over a traditional keyword-based search of publication databases because we observed that eye-tracking research is distributed across many communities (e. g., visualization, HCI, games, behavioral and social, cognitive, and vision sciences) and has variability in keyword conventions (see Figure 2).

Three of the authors then read the papers, quantified their relevance, and used thematic coding to identify common themes they addressed. One primary coder did so for all 236 papers while two secondary coders analyzed a subsample of 100 papers. The resulting codes were refined to remove overlaps, merge similar ones, describe the remaining ones using short sentences, and group them into broader themes. Example themes and codes include among others: measured data (fixations and saccades, area of interest (AOI)/ data of interest (DOI), other modalities); enabling technologies (offline, online, machine learning); identifying (user abilities, user state, tasks, data or views of interest). More predominant themes and codes are shown in Figure 2. The resulting codes and themes informed the design of our framework.

### 3.2. The ecosystem of visualization systems

Literature focusing explicitly on eye-tracking use in visualization is relatively limited. To obtain a more exhaustive view of opportunities for integrating gaze with visualization, we augmented our literature analysis by considering facets typical of visualization ecosystems and imagining their potential synergies with eye tracking. Blandford and Furniss' [BF05] framework for collaborative system design inspired this methodology, which advocates a structured consideration of all possible aspects of collaborative problem solving (e. g., physical layout, artifacts, information flows).

We used the facets described in our visual analytics model [SBR*19], one that was purposefully designed to support the integration of visual analytics and eye tracking: *System* (technological medium, support for other modalities); *Data* (types, sources); *Visualization* (marks, visual variables, techniques); *Interaction* (devices, style, techniques, intent); *Model*; *Analyst* (knowledge, goals, tasks); *Context* (domain particularities, situatedness).

We then considered how eye tracking could support each facet by referring to support in the literature. For example, when considering users' goals and tasks, we relied on visualization research on task taxonomies and models [AS04, BM13, LTM17] to reason how eye tracking can support tasks of different scopes; when imagining how eye-tracking data can help flesh out the analysis process we considered work on capturing provenance and its uses [BCC*05, DHRL*12, GZ09, RESC15, XAJK*15]; or, when considering users, we were inspired by Conati et al.'s [CCTL15, CLRT17, SCC13, TCSC13] work on using eye tracking to identify characteristics and cognitive states. Overall, this stage allowed us to consider potential eye-tracking opportunities not yet explored and ground our framework in a wider research landscape.

### 3.3. Organizing the framework

We organized our framework around four main themes. *Measured* data refers to data that eye trackers or other modalities and contexts readily provide (Section 4). *Inferred* data is then obtained by interpreting measured data computationally to learn about users and their interests, tasks, intent, and analysis process (Section 5). *Support* opportunities encompass the ways in which visualization systems can leverage such information to support users in ways previously not possible (Section 6). Finally, *Beware* factors are orthogonal to the previous three themes. They arise from limitations in eye-tracking technology and the sometimes inscrutable ways in which human perception and cognition work, and may constrain support possibilities. Each theme contains a set of numbered factors (M1-3, I1-6, S1-7, B1-9), which are described and linked to existing research and grouped further based on higher-order concepts they relate to (e. g., display and data; user and process).

Additionally, we suggest research avenues to tackle limitations and support applications that our framework exposes. These can serve as *catalysts* to the widespread use of gaze-aware visualizations. We focus primarily on catalysts that are within the remit of the visualization community, such as the accurate mapping of gaze to visualized data; the inference of visualization-specific tasks; or the evaluation of gaze-enabled visualizations. We defer a few other research challenges to research communities specialized in pursuing them. For example, the need to measure gaze reliably and transparently across a wider range of technological mediums is already actively and effectively explored by the eye tracking community.

To arrive at our final framework organization, we drew inspiration from Blandford and Furniss' [BF05] DiCoT framework to group granular factors into broader themes, and from Brehmer and Munzner's [BM13] task typology to arrange them into a three-tiered structure. However, we note that our framework evolved iteratively and other organizations were considered. Most notably, we initially attempted to structure design considerations around the facets of our earlier VA model [SBR*19]. However, while doing so we realized that we were often referring to data and inferences eye trackers provide, support opportunities, and limitations. Elements of those themes recurred in different facets of our framework, which resulted in overlaps. These shortcomings led us to ultimately reorganize the framework into the one outlined above and detailed in the following three sections.

## 4. Framework: Measure

The foundation of our framework is measurable data—information that commercial eye trackers and other input devices readily capture. This raw data serves as the basis for further interpretation, providing essential input for computational methods that infer deeper insights into user behavior and intent.

### 4.1. Gaze and other modalities

**M1 Gaze:** Most eye trackers report gaze-streams (2D or 3D coordinates supplied at time intervals that depend on the eye tracker's sampling frequency), fixations, saccades, blink events, and lower-level eye-properties such as pupil size. An extensive review of such eye-tracking measures is given by Poole and Ball [PB06]. Such data is derived by eye trackers from imagery and the algorithms and thresholds used in this process impact the quality of reported eye-tracking data [HNM12, SG00]. However, discussing these is beyond the intended scope of our work, because we primarily mean to inform visualization researchers who would likely use off-the-shelf eye trackers. We consider data that commercial eye trackers provide as the starting point of our discussion.

**M2 Enriching other modalities:** Eye-tracking data can be combined with other modalities such as directed manual interactions and gestures, natural language, pulse rate, electrodermal response, and data from electrocardiography (ECG) or electroencephalography (EEG) to help systems infer user intent and state. Such possibilities are detailed further below (S2).

**M3 Enriching context:** Other information about context (e.g., time, location, domain) may also be combined with eye-tracking data to support adaptive applications [CG07, LQS*03, PB05].

### 4.2. Beware measuring inaccuracies

Technical limitations and context determine how precisely, accurately, and quickly an eye tracker can find where a user is looking.

**B1 Technology:** Unlike most manual input devices, eye tracking can only report small screen areas that are viewed, rather than precise pixels. Eye trackers generally report accuracy (i. e., difference between true and reported gaze) of $< 0.5°$ of the field of vision, which at a typical viewing distance, translates to a screen area of about 0.5–1 cm in diameter [DMH*18, FWT*17]. The data that eye trackers report is typically delayed by 5–10 ms from when gaze events occurred, due to computations needed to interpret where the user is looking. This is distinct from an eye tracker's sampling frequency, which is often between 60 and 250 Hz, but may be as high as 1200 Hz. Additionally, gaze events need to be parsed from gaze streams using different algorithms and thresholds. Together, these three factors determine how quickly and reliably a system can identify and respond to a gaze event [SG00].

**B2 Context:** Accuracy varies based on lighting conditions, whether the user is wearing glasses or make-up, and even based on which screen area is viewed. Furthermore, eye tracking in mobile contexts is less accurate due to limitations in mobile eye trackers and the variability in their context of use. Feit et al. [FWT*17] provide an empirical account of what accuracy to expect.

### 4.3. Research catalysts

Whether and how gaze can be used effectively in practice depends on its measured accuracy. Furthermore, the adoption of the technology is contingent on its costs, form factors, and integration in diverse visualization hardware (e. g., desktop, augmented, and virtual reality). Over the last decade, the eye-tracking community has grown considerably and shown remarkable ability to improve the technology. Visualization research can build on these advances.

## 5. Framework: Infer

Building on measurable data, inference involves computational techniques that extract higher-level insights about users, their tasks, and cognitive states. By interpreting gaze data in combination with other inputs, we can uncover patterns that inform system adaptation and facilitate richer, more responsive user interactions.

### 5.1. Data and display

**I1 Foveated GUI elements:** Eye-tracking data can be related in real-time to a visualization's GUI controls (e. g., panels, toolbars, buttons) to reveal which of them a user is looking at [Duc02].

**I2 Foveated visual marks:** Similarly, eye-tracking data can be related to visualizations and the visual marks they contain, i. e., the points, curves, and shapes [AJ17, BSHW14, OAJ14, SEB*24].

**I3 Foveated data:** Via visual marks, a user's gaze can be matched to data that those marks depict facilitating an account of foveated data items or data attributes over time [AJ17, SSV*18, TSTR12].

### 5.2. User and intent

**I4 Tasks and intent:** Eye-tracking data can be used to understand the users' tasks. Researchers were so far able to detect task boundaries [BK06]; low-level interactions (e. g., drag-and-drop) [KS18]; low-level visualization or data-reading tasks [CCTL15, OAJ14]; domain-independent tasks such as reading or image interpretation [HSW*13, IB04]; and tasks that are

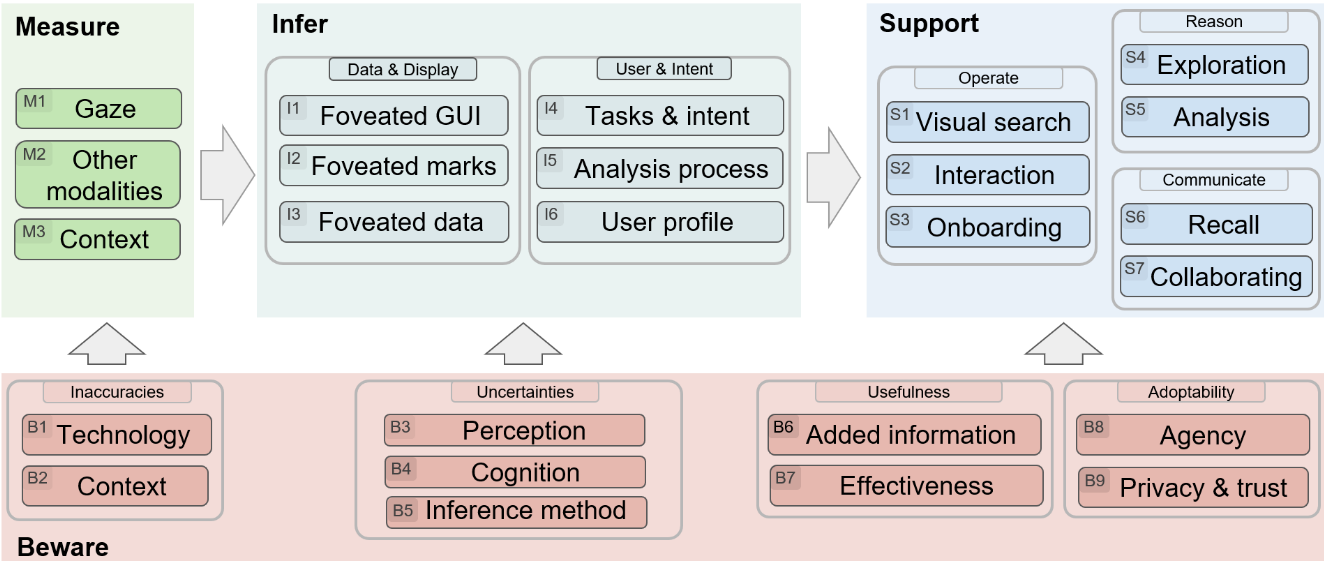


**Figure 3:** *Design considerations for gaze-aware visualizations.* (Measure) *Eye trackers provide eye-tracking data in real-time.* (Infer) *Eye-tracking data can be combined with data describing context and other modalities to support inference of higher order information such as visual marks or data that users look at, their intended tasks and analysis goals, and their user profile and state.* (Support) *This allows us to support aspects of the visual data analysis process ranging from interaction to reasoning and communication.* (Beware) *However, to ensure usefulness and adaptability, a range of limitations in technology and the information we can derive from gaze about perception and cognition should be considered.*

application-specific [CAD*11]. In some cases, this can be done from data describing fixations and saccades (M1) [GC15], but tracking foveated marks and data may help detect more complex analysis tasks and intent [AJ17, AJ15].

**I5 Analysis process:** Provenance was studied extensively by the visualization community [DJS*09, KCD*09, RESC15], but the benefits of integrating eye-tracking data into it were not. Eye-tracking data may allow to capture and describe analysis processes more faithfully by providing clues about the mental processes that prompt observed manual interactions [SSS*16b, SSES17].

**I6 User profile:** Eye-tracking data can reveal a range of individual user characteristics such as degree of expertise [BPW17, CCTL15, TCSC13], cognitive style (e. g., field-dependent vs. field-independent, locus of control) [GC15, RKB*17], cognitive abilities (e. g., visual and verbal memory, perceptual speed) [CCTL15, TCSC13], and cognitive impairment or disease [LMZ*11]. It can expose momentary user states such as boredom and engagement [KGK*14, PPG*19, SSE*15], stress and cognitive load [DKK*18, PPG*19, SSE*15], and learning progress [TSG*14]. Lallé et al. [LCC16, LCT19] reviewed such research.

### 5.3. Beware inference uncertainties

The way our perception and cognition work means inferred data typically comes with a degree of uncertainty.

**B3 Perception:** Measuring inaccuracies (B1) are compounded by how perception works. First, foveated vision spans about 5° of the field of view, around its center. At a typical viewing distance, people can accurately view screen areas of 2–4 cm around reported fixations. If multiple shapes intersect a foveated area it is difficult to know which is the focus of the user's attention.

Second, users can see, interpret, and register shapes using their parafoveal and peripheral vision, albeit less accurately. Kim et al. [KDX*12] found that the decision-making of visualization analysts relies on information perceived in this way and recent research increasingly emphasizes the broader role of peripheral vision in visual tasks [DBV21, LBX22, Ros16, WLDM23]. Such information eye trackers cannot capture, meaning we cannot know with absolute certainty what users have and have not seen.

**B4 Cognition:** Inferring interest and intent from eye-tracking data alone is challenging. For example, there are uncertainties in how we should interpret long fixations or the absence thereof. Long fixations can mean both that content is relevant to the user but also that it is difficult to parse visually [BPW17]. Furthermore, while a long fixation is often a sign of interest, learning, or novelty [EBH20], it is less clear what it means for a fixated item to be viewed once and never again. The element may have been seen, assessed, and deemed irrelevant, or it may have been remembered and factored into the decision process [ZCHK16].

**B5 Inference method:** Inferring higher-order information from eye-tracking data includes uncertainty. First, mapping gaze to visual marks and data accurately (I2, I3) depends on the visual encoding, clutter, and overlap [AJ17, OAJ14]. For example, it is easier to map gaze to point-like glyphs in uncluttered visualizations than to curves and complex shapes in dense visualizations [OAJ14]. Three-dimensional visualizations are difficult to track due to $z$-axis layering [ESDM18, MSSB16]. Second, inferring tasks, goals, and user profiles (I4, I6) has various degrees of effectiveness. For example, Gingerich and Conati [GC15] classified a series of visualization tasks with an accuracy of over 90% but only achieved around 60% on aspects of working memory and perceptual speed.

### 5.4. Research catalysts

**Mapping gaze to visualization data:** The premise of I3 is that if we capture the effect of the visualization pipeline on data (e. g., visual encoding, transformation, interaction effects), then we would know what data is mapped where on the display and could perform an *inverse mapping* from a visual stimulus (and implicitly eye-tracking coordinates) to data. Jianu et al. refer to this as Data-of-Interest mapping while Srinivasan et al. as Data-Aware eye tracking [JA18, SEB*24].

Doing so in practice is not trivial. Visualization pipelines are often non-injective: a single point in the output image may map back to multiple data features. This problem, which arises, for example, when visual primitives overlap, introduces mapping uncertainties. These are further compounded by the inherent uncertainty that comes with gaze information (see Sections 4 and 5). Inverse mappings need to model such uncertainties, reduce them if possible, and propagate them back to data elements. Three lines of research could meet these challenges:

- *Quantifying the quality of gaze-to-data mapping that can be expected in different types of visualization*, especially when factoring in gaze-inherent uncertainties. This would likely need to be tackled by a combination of experimentation and use of perceptual theories and models.
- *Improving the accuracy of gaze-to-data mapping* may be achieved through use of additional task information and other modalities. Visualizations are typically not viewed and used at random. Manual interactions, gaze, and the data they refer to are coordinated as part of coherent tasks. Inverse mappings that incorporate such information may help reduce uncertainty.
- *Integration of gaze-to-data mapping in visualization systems and libraries* is challenging as full access to visualization pipelines is needed. Lalle et al. [LCT19] proposed a dedicated testbed platform for experimenting with gaze-adaptive visualizations. Another approach could be to extend rendering libraries like Vega [SMWH16] to support an effortless and accurate mapping of gaze to data elements.

**Improved intent recognition:** We showed before that it is possible to recognize low-level activities and tasks from gaze data. However, most existing work is restricted to simple tasks and targets general HCI scenarios. We need theory and practical tools that connect gaze to visualization-specific workflows:

- *Visualization intent inferred at multiple levels of granularity* (e. g., interaction, task, goal, process). One potential research avenue is to look at concrete visualizations (e.g., scatterplots, networks, maps) and the tasks they afford. The fact that such common visualizations are well understood, provides a strong basis for this research. Another is to look at whether gaze patterns can reveal more general visualization tasks documented in task and interaction frameworks [BM13, YKSJ07].
- *Modeling the relation between gazing behavior and visualization intent* and using such models to infer intent from users' gazes. Computational models of cognition such as EMMA [Sal00], ACT-R [ABB*04], or EPIC [KM97] model complex user behavior from low-level building-blocks such as visual or manual interactions. They could provide the basis for further work that is specific to visualization intent.
- *Leveraging foveated data as a task descriptor*. Existing work has primarily leveraged raw gaze measures (e. g., fixations duration, saccade length) to classify general tasks that users engage in. Using the inverse mapping described previously and including foveated data as a task descriptor could significantly contribute to more accurate and descriptive intent inference.
- *Synergistic use of foveated data and other modalities* have potential to reduce uncertainties and provide a more complete picture of what users are trying to do. Combining multiple modalities (e. g., natural language, interactions and gestures, gaze) was shown to lead to superior inferences of user intent and facilitate fluid interactions that align more closely with how people communicate naturally [CMH*20, HTSCH14, KDEB*22, KNBV22, SLS20, SS17, SSS20]. However, more research is needed in translating such findings to visualization specific contexts.
- *Boosting inference with machine learning*. Classification algorithms can help identify intent from messy natural input modalities such as gaze, gestures, or language but rely on finding appropriate training data to be learned from. A combination of unsupervised learning methods, theories emerging from the research proposed above, and manual annotations will need to be employed to build such datasets.
- *Developing libraries for seamless intent inference from messy multi-modal data* can reduce the overhead of inferring intent in real-life visualization systems.

## 6. Framework: Support

The insights gained through inference can, in turn, support visualization analysts by enabling systems to respond dynamically to user behavior. We explore how these measured and inferred signals create new opportunities while also addressing the challenges of integrating gaze-aware interactions into visualization systems. We focus our discussion primarily on opportunities and challenges that derive from the use of eye-tracking data.

### 6.1. Operate

**S1 Visual search:** Duchowski et al. [DCM04] recognized the potential of eye tracking to support visual search early on, but the idea gains a new dimension in visualization, given its rich encodings, semantics, and tasks. By tracking foveated visual marks and data (I2, I3), and intended tasks (I4), systems can help users find information within visual displays.

This can be done by increasing the visual saliency of relevant marks and data or by de-emphasizing context. Such changes can be immediate and short-lived to support a momentary data reading task, or gradual and permanent to support users' common tasks. Examples of the first type of intervention include work by Okoe et al. [OAJ14], who support users of node-link diagrams during their visual tasks by detecting when they trace edges (I2) and temporarily increasing their saliency within a split-second; and by

Göbel et al. [GKG*18], who help their users by visually emphasizing parts of a legend that relate to the currently viewed geographic map elements (I2). Adaptation could follow a better understanding of a user's cognitive characteristics or expertise (I6) because research shows that visualization effectiveness is linked to users' individual cognitive types [OCZC15, ZK09]. Conati et al. [CCTL15, CLRT17, CLRT20] show across multiple experiments that users' personal characteristics, as inferred through eye tracking, can predict their performance on visualization tasks.

Eye tracking can also support foveated rendering, whereby quality is deliberately degraded outside foveated regions of the display to reduce computation needs and increase rendering speed. Foveated rendering was explored both conceptually and in practice [DCM04, Duc18] and shown to be effective when needing to render high-detail graphics in real time. It might bring benefit to visual analytic systems that struggle to show their data at interactive rates, either because the data is too large (e. g., unaggregated big data) or because the representations are too complex (e. g., immersive 3D analytics).

**S2 Interaction:** Duchowski [Duc18] distinguishes two ways to support interaction using gaze: actively and passively. The former involves using gaze for explicit control by triggering interactions on foveated GUI elements, visual marks, and data (I1, I2, I3). For example, Streit et al. [SLMS09] allow users to navigate within and between visualization views using gaze control. Figure 4 shows that their visualization includes a lens that scales up visual regions when they are focused. Silva et al. [SSS*16b] use gaze to control the navigation of visualizations of hierarchical data. More broadly, Velloso et al. [VC16] propose a taxonomy of gaze-based controls for gaming, some of which can provide inspiration for interactions in visualization. However, a widespread view in the HCI community is that using gaze for explicit control is not effective for two reasons: users do not typically associate eyes with actuation; and overloading vision with control makes it difficult to distinguish between seeing and controlling [Duc02, Duc18, HMR05].

Alternatively, a system can continuously monitor a user's gaze, tasks, and preferences and adapt unobtrusively to support them. This implicit form of gaze interaction, labeled by Duchowski [Duc18] as passive and by Hyrskykari et al. [HMR05] as gaze-attentive, may align better with the eyes' natural role as a perceptual organ [Duc18]. While examples of gaze-attentive interfaces exist in HCI [SSV03], the visualization community has explored adaptation in response to manual interactions alone [GW09].

A further possibility is to combine gaze with other modalities such as manual input, language and utterances. Several studies found eye tracking to significantly aid the interpretation of voice commands by disambiguating references to interface elements [HTSCH14, KDEB*22, KNBV22, PC08]. Conati et al. [CLRT20] showed that a combined tracking of both gaze and manual interactions leads to accurate inference of user intent. These findings align with a growing body of research showing that natural language interaction, in particular, becomes more fluid and precise when coupled with other modalities [CMH*20, KNBV22, SLS20, SSS20]. Silva et al. [SSE*15] combined eye tracking with a range of other sensors (e. g., ECG) to accurately track the state (e. g., stress) of air-traffic controllers. Khedher et al. [KJF19] showed that combining EEG with eye tracking allows to identify struggling students in e-learning settings more accurately than using a single modality.

**S3 Onboarding:** Eye-tracking data can be used to detect user confusion (I6) and pinpoint its possible causes (I1, I2, I4). Visualization systems could use such information to guide novice users, either explicitly by providing suggestions and hints, or implicitly by making helpful GUI features or marks more salient and guide visual attention [SSES17, SSV*18]. Moreover, such guidance could be tailored to the expertise of users by tracking the evolution from novices to experts [TSG*14]. Finally, gazing behavior of experts could be captured and used to train novices [GKR13].

### 6.2. Reason

**S4 Exploration:** Systems can help users explore large datasets by using gaze histories to infer their data interests (I3), then recommend undiscovered data with similar properties. For example, Silva et al. [SSV*18] use gaze to model time-series that their users are interested in, then recommend views containing similar data, while Jianu [Jia18] allows a point-cloud to gradually evolve by replacing data points that do not match the user's interests, as inferred from gaze histories (I3), with new ones that do. Rodrigues et al. [RSY*22] monitor users' gaze to understand what makes a cluster interesting, then recommend unseen clusters with similar attributes. Beyond visualization, eye tracking was used in similar ways to improve the ability of recommender systems to keep track of preferences and recommend items of interest [PSS*05].

Alternatively, analysts can be nudged to consider more diverse perspectives by recommending data that is dissimilar to that already explored. For example, Shao et al. [SSES17] track the plots that users view (I3) while exploring scatterplot matrices. They then suggest novel, dissimilar views to ensure broader data coverage. Similarly, Chegini et al. [CASS19] apply eye tracking in a parallel coordinates visualization to suggest unexplored dimensions.

Such approaches build on a user's own interaction history to recommend new data, a technique known as content filtering [SK09]. The alternative, collaborative filtering, can also be envisioned: multi-user systems could track data preferences of many users, combine these with users' profile information (e. g., domain, analytic interests), and recommend data other users viewed with similar profiles. We were unable to find examples of gaze-driven collaborative filtering.

**S5 Analysis:** Tracing the analysis process as it unfolds (I5) creates the premise for the system to support it. The visualization literature broadly suggests two types of support. The first seeks to improve the process itself by, for example, ensuring broader coverage of the space of possible analyses or avoiding decision-making biases [BKE16, WSE19]. The second seeks to support the user by offloading tasks and analyses to the computer [CCI*15, ERT*17]. We were unable to find examples that leverage eye-tracking data but visualization systems that rely on manual interactions exemplify the point. *CzSaw* [KCD*09] allows users to visualize and reflect upon their analysis process to consider alternative analysis paths, reinterpret their problem, and learn, refine, and reuse their analysis patterns. Sarvghad et al. [STM16] use provenance to improve data coverage in multidimensional data analysis. Wall et al. [WSE19] propose a design space for analytics systems that detect, quantify, and mitigate cognitive biases. Finally, ongoing work on mixed-initiative systems shows how computation can be offloaded to the computer if the user's intentions are either inferred [CCI*15, SJB*21, WMA*15] or expressed through natural language [MNES22].

### 6.3. Communicate

**S6 Recall:** Recall can be interpreted as communication from one's past to current self. Using gaze to augment provenance data (I5) could support long-term memory by allowing users to restore their thought processes between analysis sessions or revisit past analyses. Giannopoulos et al. [GKR13] exemplify this by visualizing gaze histories (I5) on maps to repeat previously completed tasks. Srinivasan et al. recently propose options for "revisualization" through eye-tracking attention maps [SEB*24].

**S7 Collaboration:** In addition to supporting collaborative filtering (S4), sharing gaze and provenance data explicitly can enhance communication [DG16]. In synchronous settings, eye-tracking data can complement other channels such as speech [Ken19]. In asynchronous settings, eye-tracking data can help a collaborator understand what another has done and how it was done, also known as analysis *handoff* [ZGI*17]. Several software systems exist that allow users to share eye-tracking data to improve their collaboration [DG16, DG18, NSA*19, NVA*17]. Beyond eye-tracking data, Xu et al.'s [XJL08] *Chart Constellations* use sophisticated computation to arrange views users explored to support handoff in collaboration. VizTrail [BCC*05] tracks the progression of exploratory visual analysis of scientific data as it unfolds, while GraphTrails [DHRL*12] captures and shows pathways taken during exploration of network data to support recall and communication. Gratzl et al. [GLG*16] allowed users to turn provenance data into 'vistories' to help recollection and communication. Less well explored so far are scenarios in which collaborators do not share the same roles within their teams. For example, instructors may observe their students' gaze patterns to monitor their learning process (e. g., work by Yao et al. [YBD*18]).

### 6.4. Beware usefulness of eye tracking

**B6 Added information:** Preference and intent inference is possible from manual interactions alone, and the added value of eye-tracking data ought to be considered carefully. For highly interactive visualizations, manual interactions may provide sufficient information for inference and adaptation even without the addition of eye-tracking data. For example, Conati et al. [CLRT20] found that intent is inferred most accurately when combining eye tracking and manual interaction data, but that using manual interactions alone also provides satisfactory results. Additionally, for small displays, eye tracking might only discriminate between a handful of foveated screen areas, and, depending on the context of use, the accuracy might be too low for actionable inferences.

**B7 Effectiveness:** Many factors that impact effectiveness are tied to the type of delivered support rather than to eye tracking itself. For example, recommending data that aligns with users' interests (S4) may lead to 'filter bubbles' that isolate users from diverse content [NHH*14]. Such challenges should be considered during the design of gaze-aware adaptations, but describing them here is outside the scope of this paper. Instead, we warn against effectiveness issues specific to eye tracking and which derive from two factors:

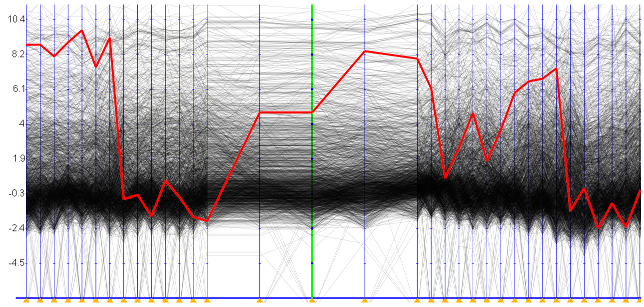


**Figure 4:** *The parallel coordinates plot by Streit et al. [SLMS09] uses screen space efficiently by extending foveated elements (green axis) and compressing the peripheral areas (sides). Cropped and reprinted from [SLMS09].*

First, eye-tracking data and the inferences we can make from it have a high level of uncertainty (B1, B3, B4, B5) and systems that react too quickly and too often to potentially erroneous inferences may seem erratic and unpredictable [KGA*17]. Decisions of whether and how to adapt an interface should be taken into consideration. For example, erroneous adaptations should be easily reversible and not disrupt the user's flow, or adaptations should be gradual and easily ignored if not useful. Recent work proposes design guidelines for effective delivery of proactive computer interventions [KKR23, ZWG*22].

Second, users do not control their gaze in the same way they do with manual interactions [Duc18]. Eye movements are often driven by the visual stimulus and the information requirements of high-level cognitive goals but not by an intent to actuate (see S2). This impacts the types of adaptations a system ought to implement and how it should deliver them to maximize effectiveness.

Finally, a problem not unique to gaze adaptation but worth drawing attention to is the ability of the system to produce responses sufficiently quickly to be helpful. Locating gaze, inferring intent from it, and deciding on an appropriate response may take time, depending on the complexity of required computations. This time should be balanced against the duration of the supported activity. Support for short visual tasks (e. g., visual scanning, foveated rendering) will likely need to be delivered within a split second to be perceived as real-time. More complex interactive tasks, which themselves take multiple seconds, may tolerate support delivered within a second or two, while lengthy analysis processes such as exploration or complex data searches take even longer. The timing of interventions is discussed by Cenada et al. [CAGM21].

### 6.5. Beware adoptability

**B8 Agency:** Users need to feel in control of the analysis process rather than controlled by the system. This is not a problem unique to gaze-adaptive support and requires the right balance between automation and interest steering. Adaptation processes that are transparent may support a feeling of agency [SSES17, SSV*18].

**B9 Privacy and trust:** Users need to trust the support indications the system gives and the process that led to them. Unpredictable, ineffective, and obscure adaptations would undermine a user's trust in the system. Users also need to trust that their eye-tracking data remains private. Eye-tracking data are a rich source of information about a user's preferences, interests, diseases, and cognitive profiles [PPD*20, SSE*15, SSV*18]. Moreover, gaze is largely involuntary. This means eye-tracking data should be under stricter privacy provisions than manual interaction data. This includes user consent but also encryption of eye-tracking data for most, if not all, applications and regular security audits. The absence of such provisions is likely to reduce adoption considerably.

### 6.6. Research catalysts

**Designing adaptive visualizations:** We showed above how eye-tracking could enhance different visualization workflows. While grounded in established research, many possibilities we described are conjectured and yet to be explored:

- *What* support flavors are possible in visualization systems? Ties with current efforts within the visualization community to build user-adaptive, mixed-initiative, and multimodal systems - especially with the advent of Large Language Models (LLMs) - should be strengthened, by providing an emphasis on how eye tracking can uniquely contribute to them.
- *Where and when* the use of eye-tracking is beneficial? We show in Section 7 that some visualizations allow more accurate inverse mappings from gaze to data or inference of tasks, have a wider range of workflows needing support, and can be used in more gaze-amenable contexts than others. Such aspects need to be considered against eye-tracking particularities, especially those related to uncertainty and the ability to reveal visual behavior that other modalities cannot. Developing an empirical understanding of when gaze can deliver gains is necessary.
- *How* support should be delivered? Given the breadth of possible support and uncertainties in eye-tracking data and inferences from it, we need to look beyond designs based on notifications and recommendation lists. We need support that is delivered in subtle ways, that is helpful, transparent, unobtrusive, and predictable. This may include brief and subtle increases in the saliency of visual marks, attention guidance, gradual visual reconfiguration, or multimodal feedback. Different designs may also be more suited for different supported activities. For example, small visual tasks may require adaptations that are automatic, momentary, and unobtrusive, while recommendations for more disruptive actions, such as suggestions for different visualizations, could be delivered as explicit, opt-in recommendations. Some of these questions have recently also been raised by Srinivasan et al. in their call for Attentiona Aware Visualizations [SEB*24]. Ultimately, such research should lead to an empirically derived and tested taxonomy of support delivery options. Recent work has started to recognize the importance of nuanced support delivery and can serve as a starting point [CAGM21, KKR23, ZWG*22].

**Seamless, multimodal human-computer communication:** Rapid advances in machine learning and natural language interaction prompt the need to "recalibrate the roles of humans and machines as teammates" [WFMD21]. This is recognized by recent research spanning the visualization domain [MGG*23, OGW19, PMCEA*22, SJB*21] and beyond [NSA*19]. Li et al. recently ourlined a preliminary roadmap for LLMs as Visual Data Analysts [LAS24]. However, to support a symbiotic partnership, both humans and computers need to establish and maintain awareness of each other's goals, intents, and workflows [OGW19, SJB*21].

This is facilitated by new tools that allow moving from the granular control of WIMP interfaces to more natural interactions that are high-throughput and effortless. Through rapid advances in LLMs, natural language emerges as a modality that lets users articulate analytic goals in human-like ways and engage in a discourse with the computer to facilitate transparent interaction [MNES22, NSS20]. However, just like people rely on context and multiple modalities to communicate with each other (e. g., using referential gestures and gazes), research increasingly finds that combining multiple natural modalities can lead to significantly superior computational inferences of user intent [CMH*20, HTSCH14, KDEB*22, KNBV22, SS17, SLS20]. While combining natural language with manual interactions and gestures has been explored, integrating gaze input has received less attention from the visualization community.

**Evaluating gaze-aware visualizations** to explore and demonstrate their effectiveness is orthogonal to research listed above. Visualization evaluation is difficult in general and is even harder for adaptive systems. So far, evaluations of single components of the adaptation pipeline (e. g., restricted task inference, single type of adaptation) in highly controlled settings are prevalent. However, these need to be augmented by holistic assessments of the benefits of adaptation in real use. This is challenging because observed effects depend on individual differences between users, context of use, and the interplay between the multiple parts of the larger adaptive system. Ultimately, both theoretical work on evaluation methodologies that can bridge between controlled, component-based vs. realistic, holistic assessments, as well as practical work on actual evaluations, is necessary.

**Social context and adoption:** Finally, before gaze-aware visualizations are adopted widely we expect answers to questions such as: How can we produce gaze-aware visualizations that actually are adopted? Are users willing to accept such systems? How can we build trust in such systems? In particular, how can we guarantee an appropriate level of privacy protection, given that eye-tracking data provides access to highly personal information related to cognitive processes? Many such questions link to other disciplines and benefit from collaboration with other fields.

## 7. Design framework in action: usage scenarios

We demonstrate how our framework can serve as a structured design checklist for researchers and practitioners who wish to develop gaze-aware visualization by using it to analyze how eye tracking could help three concrete systems—Tableau [Tab19], Mobile VA for Law Enforcement [RMA*14], and FiberClay [HRD*19].

To facilitate this analysis, we first operationalized our framework into the worksheet exemplified in Figure 5. We did so by reframing its key factors aspects into a checklist of short, concrete questions grouped into six blocks that match the structure of the framework: measure, infer data and display, infer

| Measure | Inaccuracies |
|---|---|
| ☑ **M1:** What gaze data can you measure (e.g. fixations, saccades, pupil size, blinks)?<br>☑ **M2:** What other modality could you measure (e.g., voice, body pose, face, ECG, EKG)? Can you imagine synergies with gaze-data?<br>☒ **M3:** What data about context could you measure (e.g., time, location, weather)? Can you imagine synergies with gaze-data? | ☒ **B1**: Are there limitations imposed by technology?<br>☒ Are you tracking a small screen? The ET may only discriminate between a handful of broad regions that users view.<br>☒ Is ET technology too expensive for users of your application?<br>☒ Are ET specs (e.g., accuracy, frequency, latency) not enough for the envisioned inferences and support?<br>☒ Do users lack the technology to record other modalities (consider aquisition costs)?<br>☑ **B2**: Is the context of use limiting what you can measure?<br>☒ Are lighting conditions too variable for accurate gaze tracking?<br>☒ Is the positioning of my screen relative to the ET too variable?<br>☑ Is context of use preventing expression of other modalities (e.g., users can't talk/gesture in their workplace)?<br>☑ Is context data uninformative (e.g., location doesn't change in a desktop application)? |
| Rationale: ***Tableau is primarily for every-day desktop use with steady lighting and on large screens: gaze data could be measured with cheap add-on eye trackers intended for interaction and performance would suffice for a broad range of support. Voice could be collected easily too but users may not want to use it in an office setting; so, it could be an option but should not a prerequisite. Given limited variations in context in a desktop setting, keeping track of context data may not be helpful.*** | |
| **Infer Display and Data ....** | **Uncertainties ...** |
| **Infer User and Intent ....** | **Uncertainties ...** |
| **Support Operate ...** | **Usefulness ... / Adoptability ...** |
| **Support Reason ...** | **Usefulness ... / Adoptability ...** |
| **Support Communicate ...** | **Usefulness ... / Adoptability ...** |

**Figure 5:** *Design framework in action: Using the framework as a checklist of opportunities and challenges structures the process of considering how a system can benefit from eye tracking. The image exemplifies the first step of this process (Measure) for Tableau, and lists the remaining ones.*

use and intent, support interaction, reasoning and communication. Applying the worksheet to inform the design of a gaze-aware visualization then involves following the checklist step-by-step and considering how each question applies to the visualization's context. Our framework and worksheet are not meant to provide strong prescriptive guidelines or definite yes/no answers, but rather prompt thoughtful, nuanced consideration of relevant aspects of gaze-adaption in the given context. The framework itself and literature it cites should inform this thought process, which can then be captured in the worksheet as a series of descriptive 'design rationales' for each framework block (see Figure 5).

We exemplify this process for our three chosen systems and show how, for each one, it can result in concrete ideas about new data that could be collected or inferred, the adaptions these could afford, and a realistic assessment of their potential utility in the face of likely challenges. In the interest of space, we only summarize these analyses below. We provide the worksheets and complete notes for all three systems as a persistent online resource at `https://doi.org/10.5281/zenodo.5665612`.

### 7.1. Tableau — business intelligence analytics

Tableau [Tab19] is an application for visualization, data analysis, and business intelligence. It is intended for broad use by analysts with diverse levels of expertise and domains. Tableau relies on a model that integrates data tables tightly with varied visualizations (e. g., bar, pie, and line charts, box and scatter plots, maps). Analysts can interactively author visualizations by choosing how visual marks are mapped to data items and attributes.

Eye-tracking data could be measured (M1) relatively accurately with cheap add-on eye trackers (B1) because Tableau is typically used in desktop environments with large screens and stable lighting (B2), and should be used in combination with data describing manual interactions (M2). Tableau could identify analysts' visual marks, data dimensions, and measures of interest (I2, I3). Tableau uses relatively standard visualizations with minimal clutter and overlap, which together with optimal tracking conditions, creates the premise for robust inferences (B3, B5). Elements of Tableau's rich GUI (e. g., buttons, panels) could also be tracked (I1).

These measurements and inferences would provide insight into analysts' tasks (I4), which may be visual (e. g., scanning measures along a dimension; following a leader-line to a visual element) or operational (e. g., adding a column, using a formula to combine data dimensions). The fact that Tableau implements standard visualizations with relatively well-understood uses and a consistent interaction paradigm may allow this to be done robustly (B4, B5). They would also add detail to broader goals and analytic processes (I5). For example, they may help Tableau track analysts' interests in data (e. g., which data dimensions are foveated often?) and visual patterns (e. g., are analysts typically foveating outliers, clusters?). When combined with manual interaction data, this could better explain what analysts are trying to do (e. g., "find a set of dimensions along which measures cluster", "are there correlations between particular data dimensions?"). Finally, they would help Tableau establish an analyst's profile (I6). This can include the evolution of their data and visual interests over time, their cognitive profile (e. g., locus of control), and momentary states (e. g., flow vs. confusion).

Such inferences would pave the way for a range of support. First, quick visual tasks could be supported momentarily when inferred, for example, by increasing the saliency of their targets (S1). Visual displays could be optimized to match an analyst's profile, for example, by emphasizing data or visual patterns of interest or recommending suitable visualizations (S1). Second, operational tasks could be streamlined by providing proactive details on demand about foveated visuals and offering auto-complete suggestions (S2). Third, a better understanding of analysts' broader goals and preferences would allow Tableau to suggest more meaningful recommendations for data subsets, visualizations, or computations to support their analysis (S2,S5). Finally, adding gaze-based inferences to provenance histories can tell richer visualization stories [GLG*16] and allow to revisit analyses to support recall (S6), scrutinize analyses (S5), and communicate them to others (S7).

To be effective, support should account for inherent uncertainties in eye tracking and inference should be delivered unobtrusively, as well as keep the analyst in charge (B7, B9, B8). Support for low-level tasks (e. g., visual, operational) could be delivered as momentary adaptations (e. g., increased saliency, guided attention) that are both sufficiently salient to help but also sufficiently subtle to be ignored. More meaningful or disruptive adaptations, such as recommending a new visualization or computation, could be presented as explicit opt-in recommendations.

The fact that Tableau analysts engage with the tool for extended periods of time to conduct complex analyses makes tracking and supporting the analysis process and user profile especially promising (B6, B7). While manual interactions are probably sufficient for broadly capturing what analysts are doing (e. g., cycling through different visualizations or combinations of dimensions), eye tracking can reveal more about the reasons behind their actions (e. g., looking for dimensions with outlier values, B6).

### 7.2. Mobile visual analytics for law enforcement

Razip et al. [RMA*14] developed a VA system to assist law enforcement. Officers, crime analysts, and detectives can use the system on mobile devices in the field and on desktops in their offices. It includes a map with glyph overlays, a time-series chart, a heatmap, a clock view, and a GUI composed of a few buttons and sliders.

Desk-working analysts could benefit from support similar to that described in the Tableau usage scenario (Section 7.1), with the difference that this system is significantly less complex. This makes it easier to augment with eye tracking (e. g., fewer visualizations to track; lower number of tasks to differentiate between), but also relatively straightforward to use as is. The value that complex gaze-adaptive support would add is uncertain (B6).

Mobile contexts are variable, leading to less accurate eye-tracking data (B2) and a need for more elaborate tracking solutions (B1), while tracking small screens (e. g., mobile phones) is likely to reveal only coarse information (B6). On the plus side, voice could be collected too and context information (e. g., time, location, weather) is likely to be valuable (M2, M3).

Reliably tracking individual visual elements an analyst looks at (e. g., glyphs, heatmap cells, buttons) would be impossible, but we could find out when analysts foveate entire visualizations or GUI panels (I1, I2). From such coarse information, it would be difficult to differentiate between granular tasks (I4). Furthermore, officers on the ground would likely conduct predominantly short, query-like tasks rather than engage in in-depth analyses (I5). As such, the complex adaptations discussed for Tableau are unlikely to be both possible and necessary here (B6, B7).

Instead, the main benefit from eye tracking may be to combine gaze with voice commands and context information to support interaction (S2) when an analyst's hands are unavailable. Specifically, eye-tracking data can reveal useful clues about what the analyst refers to when issuing verbal commands. For example, the application could interpret the sentence "When are these (gazed) crimes happening here (device location)?" as "When do burglaries occur at the intersection of Universe Boulevard and Galaxy Drive?"

### 7.3. FiberClay — visual analytics in virtual reality

FiberClay is a VR application for the visualization of trajectories [HRD*19]. It uses an HMD and hand-held controllers [HRD*19]. Head posture orientates the scene camera, while the controllers serve to move and scale the view. The same physical controllers also generate virtual rays that analysts use to select entire ranges or individual trajectories within the data.

Eye trackers could be integrated into the HMD and measure gaze (M1) accurately without interference from foreign light sources (B2). This could provide input for foveated rendering (I2) to improve performance and reduce simulator sickness (S1). FiberClay could use gaze rays to intersect with the shown data to infer high interest (I3) or to detect the current analytical task (I4). It could then recommend visualization types that better fit the data (S4, S5).

Gaze rays from both eyes intersect within the visualized data, providing an intuitive method for 3D pointing (I3). The application could then adjust the rendering or move the viewpoint to avoid occlusion of the target data (S4). This might be of special interest in collaborative data analysis, to share the portion of data that is currently the topic of conversation (S7). Sharing the current gaze position in real time might be confusing for the collaborators and would need to be triggered via a button on the hand-held controllers to avoid pitfalls related to effectiveness (B7, B9). Additionally, recent research [ÖRB*20] indicates that the accuracy of 3D gaze tracking has to improve further before it can substitute hand-held controllers (B1). However, using the head pose from the HMD's developer framework could provide an intuitive means for an analyst to specify the depth of data of interest or for selection (S2): they would move their head forward and backward.

Knowledge of the foveated visual marks (I2) could also be combined with adaptive optics. These would constantly adjust the focal distance of viewed objects to avoid the vergence-accommodation conflict. This could also be used with user profiles or automated detection to correct refractive errors (I6) to effectively replace glasses in the confined space of the VR headset (S1).

## 8. Conclusion

Following an analysis of eye-tracking literature pertinent to data visualization ecosystems, we arrived at a design framework for gaze-aware visualization. The elements of this framework are organized as *measures* that capture eye-tracking data, as *inferring* derived information for the data display and the user, and as *support* for improved systems. These opportunities are accompanied by limiting factors we have to *beware* of and avenues for future visualization research to alleviate them. We have showcased how the design considerations can be used in practice via three usage scenarios. Overall, we see great opportunity for the visualization community to advance gaze-aware visualizations by including real-time gaze information.

## Acknowledgments

NS is supported by the Austrian Science Fund – FWF (KLI811). NR and DW acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 251654672 – TRR 161, TB by the European Social Fund and the Ministry of Science, Research, and Arts Baden-Württemberg.

## References

[ABB*04] ANDERSON J. R., BOTHELL D., BYRNE M. D., DOUGLASS S., LEBIERE C., QIN Y.: An integrated theory of the mind. *Psychol. Rev. 111*, 4 (2004), 1036. 7

[AEYN11] ANDREWS C., ENDERT A., YOST B., NORTH C.: Information visualization on large, high-resolution displays: Issues, challenges, and opportunities. *Inf. Vis. 10*, 4 (2011), 341–355. 2

[AJ15] ALAM S. S., JIANU R.: What are they doing?: Real-time analysis of eye-tracking data. IEEE VIS Poster Compendium, 2015. 6

[AJ17] ALAM S. S., JIANU R.: Analyzing eye-tracking information in visualization and data space: from where on the screen to what on the screen. *IEEE Trans. Vis. Comput. Graph. 23*, 5 (2017), 1492–1505. 5, 6

[AKG*16] AURISANO J., KUMAR A., GONZALEZ A., LEIGH J., DIEUGENIO B., JOHNSON A.: Articulate2: Toward a conversational interface for visual data exploration. In *IEEE Vis.* (2016), vol. 8. 2

[ARZ*22] ALAÇAM Ö., RUPPERT E., ZARRIESS S., MALHOTRA G., BIEMANN C.: Modeling referential gaze in task-oriented settings of varying referential complexity. In *Findings Assoc.Comp. Ling.* (2022), pp. 197–210. 3

[AS04] AMAR R., STASKO J.: A knowledge task-based framework for design and evaluation of information visualizations. In *Proc. IEEE Inf. Vis.* (2004), pp. 143–150. 4

[AZZR07] ALVAREZ-CORTES V., ZAYAS-PEREZ B. E., ZARATE-SILVA V. H., RAMIREZ URESTI J. A.: Curr. trends adaptive user interf: Challenges applic. In *Electronics, Robotics and Automotive Mechanics Conference* (2007), pp. 312–317. 2

[BCC*05] BAVOIL L., CALLAHAN S. P., CROSSNO P. J., FREIRE J., SCHEIDEGGER C. E., SILVA C. T., VO H. T.: Vistrails: Enabling interactive multiple-view visualizations. In *IEEE Vis.* (2005), IEEE, pp. 135–142. 4, 9

[Bed05] BEDNARIK R.: Potentials of eye-movement tracking in adaptive systems. In *Proc. Empir. Eval. Adapt. Sys.* (2005), pp. 1–8. 3

[BF05] BLANDFORD A., FURNISS D.: DiCoT: a methodology for applying distributed cognition to the design of teamworking systems. In *Int. Design Specific. Verific. Interact. Sys.* (2005), Springer, pp. 26–38. 4, 5

[BK06] BAILEY B. P., KONSTAN J. A.: On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Comput. Hum. Behavior 22*, 4 (2006), 685–708. 5

[BKE16] BLASCHECK T., KOCH S., ERTL T.: Logging interactions to learn about visual data coverage. *Logging Interact. Vis. & Vis. Interact. Logs* (2016). 8

[BKR*17] BLASCHECK T., KURZHALS K., RASCHKE M., BURCH M., WEISKOPF D., ERTL T.: Visualization of eye tracking data: A taxonomy and survey. *Comput. Graph. Forum 36*, 8 (2017), 260–284. 3

[BLC*11] BROWNE J., LEE B., CARPENDALE S., RICHE N., SHERWOOD T.: Data analysis on interactive whiteboards through sketch-based interaction. In *Proc. ACM Intl. Conf. Interactive Tabletops & Surf.* (2011), pp. 154–157. 2

[BM13] BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. *IEEE Trans. Vis. Comput. Graph. 19*, 12 (2013), 2376–2385. 4, 5, 7

[BOZ*14] BROWN E. T., OTTLEY A., ZHAO H., LIN Q., SOUVENIR R., ENDERT A., CHANG R.: Finding Waldo: Learning about users from their interactions. *IEEE Trans. Vis. Comput. Graph. 20*, 12 (2014), 1663–1672. 2

[BPW17] BORYS M., PLECHAWSKA-WÓJCIK M.: Eye-tracking metrics in perception and visual attention research. *Europ. J. Med. Techn. 3* (2017), 11–23. 6

[BSHW14] BERNHARD M., STAVRAKIS E., HECHER M., WIMMER M.: Gaze-to-object mapping during visual search in 3d virtual environments. *ACM Trans. Appl. Percep. 11*, 3 (2014), 1–17. 5

[CAD*11] COURTEMANCHE F., AÏMEUR E., DUFRESNE A., NAJJAR M., MPONDO F.: Activity recognition using eye-gaze movements and traditional interactions. *Interact. Comp. 23*, 3 (2011), 202–213. 6

[CAGM21] CENEDA D., ARLEO A., GSCHWANDTNER T., MIKSCH S.: Show me your face: towards an automated method to provide timely guidance in visual analytics. *IEEE Trans. Vis. Comput. Graph. 28*, 12 (2021), 4570–4581. 9, 10

[CAS*18] COLLINS C., ANDRIENKO N., SCHRECK T., YANG J., CHOO J., ENGELKE U., JENA A., DWYER T.: Guidance in the human-machine analytics process. *Vis. Informatics 2*, 3 (2018), 166–180. 2

[CASS19] CHEGINI M., ANDREWS K., SCHRECK T., SOURIN A.: Eye-tracking based adaptive parallel coordinates. SIGGRAPH Asia Posters, 2019. 8

[CCI*15] COOK K., CRAMER N., ISRAEL D., WOLVERTON M., BRUCE J., BURTNER R., ENDERT A.: Mixed-initiative visual analytics using task-driven recommendations. In *Proc. IEEE Vis. Analyt. Sci. Techn.* (2015), IEEE, pp. 9–16. 2, 8, 9

[CCTL15] CONATI C., CARENINI G., TOKER D., LALLÉ S.: Towards user-adaptive information visualization. In *Proc. AAAI Artif. Intel.* (2015). 3, 4, 5, 6, 8

[CG07] CUTRELL E., GUAN Z.: What are you looking for?: An eye-tracking study of information usage in web search. In *Proc. SIGCHI Conf. Human Fact. Comp. Syst.* (San Jose, California, USA, 2007), ACM Press, pp. 407–416. 5

[CGM*17] CENEDA D., GSCHWANDTNER T., MAY T., MIKSCH S., SCHULZ H.-J., STREIT M., TOMINSKI C.: Characterizing guidance in visual analytics. *IEEE Trans. Vis. Comput. Graph. 23*, 1 (2017), 111–120. 2

[CGM19] CENEDA D., GSCHWANDTNER T., MIKSCH S.: A review of guidance approaches in visual data analysis: A multifocal perspective. *Comput. Graph. Forum 38*, 3 (2019), 861–879. 2

[CLRT17] CONATI C., LALLÉ S., RAHMAN M. A., TOKER D.: Further results on predicting cognitive abilities for adaptive visualizations. In *Int. Joint Conf. Artif. Intell.* (2017), pp. 1568–1574. 3, 4, 8

[CLRT20] CONATI C., LALLÉ S., RAHMAN M. A., TOKER D.: Comparing and combining interaction data and eye-tracking data for the real-time prediction of user cognitive abilities in visualization tasks. *ACM Trans. Interactive Intell. Sys. 10*, 2 (2020), 1–41. 3, 8, 9

[CMH*20] CHOWDHURY I., MOEID A., HOQUE E., KABIR M. A., HOSSAIN M. S., ISLAM M. M.: Miva: Multimodal interactions for facilitating visual analysis with multiple coordinated views. In *Intl. Conf. Inf. Vis.* (2020), IEEE, pp. 714–717. 2, 7, 8, 10

[CZK*22] CHIOSSI F., ZAGERMANN J., KAROLUS J., RODRIGUES N., BALESTRUCCI P., WEISKOPF D., EHINGER B. V., FEUCHTNER T. M., REITERER H., CHUANG L. L., ERNST M. O., BULLING A., MAYER S., SCHMIDT A.: Adapting visualizations and interfaces to the user. *it Inf. Technol. 64*, 4-5 (2022), 133–143. 2

[DBV21] David E. J., Beitner J., Võ M. L.-H.: The importance of peripheral vision when searching 3d real-world scenes: A gaze-contingent study in virtual reality. *J. Vision 21*, 7 (2021), 3–3. 6

[DCM04] Duchowski A. T., Cournia N., Murphy H.: Gaze-contingent displays: A review. *CyberPsych. & Behavior 7*, 6 (2004), 621–634. 7, 8

[DG16] D'Angelo S., Gergle D.: Gazed and confused: Understanding and designing shared gaze for remote collaboration. In *Proc. ACM Hum. Fact. Comput. Sys.* (2016), ACM, pp. 2492—-2496. 9

[DG18] D'Angelo S., Gergle D.: An eye for design: Gaze visualizations for remote collaborative work. In *Proc. ACM Hum. Fact. Comput. Sys.* (2018), ACM, pp. 1–12. 9

[DHRL*12] Dunne C., Henry Riche N., Lee B., Metoyer R., Robertson G.: Graphtrail: Analyzing large multivariate, heterogeneous networks while supporting exploration history. In *Proc. ACM Hum. Factors Comput. Sys.* (2012), pp. 1663–1672. 4, 9

[DJS*09] Dou W., Jeong D. H., Stukes F., Ribarsky W., Lipford H. R., Chang R.: Recovering reasoning processes from user interactions. *IEEE Comput. Graph. Appl. 29*, 3 (2009), 52–61. 6

[DKK*18] Duchowski A. T., Krejtz K., Krejtz I., Biele C., Niedzielska A., Kiefer P., Raubal M., Giannopoulos I.: The index of pupillary activity: Measuring cognitive load vis-à-vis task difficulty with pupil oscillation. In *Proc. ACM Hum. Fact. Comput. Sys.* (2018), pp. 1–13. 6

[DMH*18] Dalrymple K. A., Manner M. D., Harmelink K. A., Teska E. P., Elison J. T.: An examination of recording accuracy and precision from eye tracking data from toddlerhood to adulthood. *Frontiers Psychol. 9* (2018). 5

[DS96] Deubel H., Schneider W. X.: Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vis. Res. 36*, 12 (1996), 1827–1837. 3

[Duc02] Duchowski A. T.: A breadth-first survey of eye-tracking applications. *Behavior Res. Meth., Instrum., & Comput. 34*, 4 (2002), 455–470. 5, 8

[Duc18] Duchowski A. T.: Gaze-based interaction: A 30 year retrospective. *Computers & Graphics 73* (2018), 59–69. 8, 9

[EBH20] Ernst D., Becker S., Horstmann G.: Novelty competes with saliency for attention. *Vis. Res. 168* (2020), 42–52. 6

[EHR*14] Endert A., Hossain S., Ramakrishnan N., North C., Fiaux P., Andrews C.: The human is the loop: new directions for visual analytics. *J. Intel. Inf. Sys. 43*, 3 (2014), 411–435. 1, 2

[EMJ*11] Elmqvist N., Moere A. V., Jetter H.-C., Cernea D., Reiterer H., Jankun-Kelly T.: Fluid interaction for information visualization. *Inf. Vis. 10*, 4 (2011), 327–340. 2

[End16] Endert A.: *Semantic Interaction for Visual Analytics: Inferring Analytical Reasoning for Model Steering*. Morgan & Claypool, 2016. 2

[ERT*17] Endert A., Ribarsky W., Turkay C., Wong B. W., Nabney I., Blanco I. D., Rossi F.: The state of the art in integrating machine learning into visual analytics. *Comput. Graph. Forum 36*, 8 (2017), 458–486. 8

[ESDM18] Elmadjian C., Shukla P., Diaz-Tula A., Morimoto C.: 3D gaze estimation in the scene volume with a head-mounted eye tracker. In *Proc. Work. Communic. Gaze Interac.* (2018), ACM, pp. 1–9. 7

[Fur06] Furnas G. W.: A fisheye follow-up: further reflections on focus + context. In *Proc. ACM Hum. Fact. Comput. Sys.* (2006), pp. 999–1008. 2

[FWT*17] Feit A. M., Williams S., Toledo A., Paradiso A., Kulkarni H., Kane S., Morris M. R.: Toward everyday gaze input: Accuracy and precision of eye tracking and implications for design. In *Proc. ACM Hum. Factors Comput. Sys.* (2017), pp. 1118–1130. 5

[GC15] Gingerich M. J., Conati C.: Constructing models of user and task characteristics from eye gaze data for user-adaptive information highlighting. In *Proc. AAAI Artif. Intel.* (2015). 6, 7

[GKG*18] Göbel F., Kiefer P., Giannopoulos I., Duchowski A., Raubal M.: Improving map reading with gaze-adaptive legends. In *Proc. ACM Eye Track. Res. Appl.* (2018), ACM, pp. 29:1–29:9. 3, 8

[GKR13] Giannopoulos I., Kiefer P., Raubal M.: The influence of gaze history visualization on map interaction sequences and cognitive maps. In *Proc. ACM Intern. Wor. MapInteraction - MapInteract* (2013), ACM. 8, 9

[GLG*16] Gratzl S., Lex A., Gehlenborg N., Cosgrove N., Streit M.: From visual exploration to storytelling and back again. *Comput. Graph. Forum 35*, 3 (2016), 491–500. 9, 12

[GW03] Goldberg J., Wichansky A.: *Eye tracking in usability evaluation: A practitioner's guide*. Elsevier, 2003. 3

[GW09] Gotz D., Wen Z.: Behavior-driven visualization recommendation. In *Proc. Intel. User Interf.* (2009), pp. 315–324. 8

[GZ09] Gotz D., Zhou M. X.: Characterizing users' visual analytic activity for insight provenance. *Inf. Vis. 8*, 1 (2009), 42–55. 4

[HKA*24] Huth F., Koch M., Awad-Mohammed M., Weiskopf D., Kurzhals K.: Eye tracking on text reading with visual enhancements. In *Proc. ACM Eye Track. Res. Appl.* (2024), ACM, pp. 80:1–80:7. 3

[HMR05] Hyrskykari A., Majaranta P., Räihä K.-J.: From gaze control to attentive interfaces. In *Proc. Hum.-Comput. Interact. Intern.* (2005). 8

[HNA*11] Holmqvist K., Nyström M., Andersson R., Dewhurst R., Jarodzka H., Van de Weijer J.: *Eye tracking: a comprehensive guide to methods and measures*. Oxford University Press, 2011. 3

[HNM12] Holmqvist K., Nyström M., Mulvey F.: Eye tracker data quality: what it is and how to measure it. In *Proc. ACM Eye Track. Res. Appl.* (2012), pp. 45–52. 5

[Hor99] Horvitz E.: Principles of mixed-initiative user interfaces. In *Proc. SIGCHI Conf. Human Fact. Comp. Syst.* (1999), pp. 159–166. 1, 2

[HRD*19] Hurter C., Riche N. H., Drucker S. M., Cordeil M., Alligier R., Vuillemot R.: Fiberclay: sculpting three dimensional trajectories to reveal structural insights. *IEEE Trans. Vis. Comput. Graph. 25*, 1 (2019), 704–714. 10, 12

[HSW*13] Henderson J. M., Shinkareva S. V., Wang J., Luke S. G., Olejarczyk J.: Predicting cognitive state from eye movements. *PLOS one 8*, 5 (2013). 5

[HTSCH14] Hakkani-Tür D., Slaney M., Celikyilmaz A., Heck L.: Eye gaze for spoken language understanding in multi-modal conversational interactions. In *Proc. Intl. Conf. Multimod. Interac.* (2014), pp. 263–266. 7, 8, 10

[IB04] Iqbal S. T., Bailey B. P.: Using eye gaze patterns to identify user tasks. In *Grace Hopper Celebr. Women Comput.* (2004), pp. 5–10. 5

[JA18] Jianu R., Alam S. S.: A data model and task space for data of interest (doi) eye-tracking analyses. *IEEE Trans. Vis. Comput. Graph. 24*, 3 (2018), 1232–1245. 7

[JC80] Just M., Carpenter P.: A theory of reading: from eye fixations to comprehension. *Psychol. Rev. 87*, 4 (1980), 329–354. 3

[Jia18] Jianu R.: Gaze-enabled data recommendations in visualization: First considerations. In *AVI Wor. Multimod. Interact. Data Vis.* (2018). 8

[JLLS17] Jo J., L'Yi S., Lee B., Seo J.: Touchpivot: blending wimp & post-wimp interfaces for data exploration on tablet devices. In *Proc. SIGCHI Conf. Human Fact. Comp. Syst.* (2017), pp. 2660–2671. 2

[KCD*09] Kadivar N., Chen V., Dunsmuir D., Lee E., Qian C., Dill J., Shaw C., Woodbury R.: Capturing and supporting the analysis process. In *Proc. IEEE Vis. Analyt. Sci. Techn.* (2009), IEEE, pp. 131–138. 6, 8

[KDEB*22] KUMAR A., DI EUGENIO B., BHATTACHARYA A., AURISANO J., JOHNSON A.: Reference resolution and context change in multimodal situated dialogue for exploring data visualizations. *arXiv preprint arXiv:2209.02215* (2022). 7, 8, 10

[KDX*12] KIM S.-H., DONG Z., XIAN H., UPATISING B., YI J. S.: Does an eye tracker tell the truth about visualizations?: findings while investigating visualizations for decision making. *IEEE Trans. Vis. Comput. Graph. 18*, 12 (2012), 2421–2430. 6

[Ken19] KENTTÄLÄ K.: *Interactive Voice Response System and Eye-tracking Interface in Assistive Technology for Disabled*. PhD thesis, University of Oulu, 2019. 9

[KFBW16] KURZHALS K., FISHER B., BURCH M., WEISKOPF D.: Eye tracking evaluation of visual analytics. *Inf. Vis. 15*, 4 (2016), 340–358. 3

[KGA*17] KIEFER P., GIANNOPOULOS I., ANAGNOSTOPOULOS V. A., SCHÖNING J., RAUBAL M.: Controllability matters: The user experience of adaptive maps. *GeoInformatica 21*, 3 (2017), 619–641. 9

[KGK*14] KIEFER P., GIANNOPOULOS I., KREMER D., SCHLIEDER C., RAUBAL M.: Starting to get bored: an outdoor eye tracking study of tourists exploring a city panorama. In *Proc. ACM Eye Track. Res. Appl.* (2014), pp. 315–318. 3, 6

[KJF19] KHEDHER A. B., JRAIDI I., FRASSON C.: Predicting learners' performance using eeg and eye tracking features. In *Thirty-Sec. Int. Flairs Conf.* (2019). 8

[KK23] KESKIN M., KETTUNEN P.: Potential of eye-tracking for interactive geovisual exploration aided by machine learning. *Intl. J. Cartography* (2023), 1–23. 2

[KKR23] KWOK T. C., KIEFER P., RAUBAL M.: Unobtrusive interaction: a systematic literature review and expert survey. *Human–Computer Interact.* (2023), 1–37. 9, 10

[KM97] KIERAS D. E., MEYER D. E.: An overview of the epic architecture for cognition and performance with application to human-computer interaction. *Human–Comp. Inter. 12*, 4 (1997), 391–438. 7

[KMS*08] KEIM D., MANSMANN F., SCHNEIDEWIND J., THOMAS J., ZIEGLER H.: Visual analytics: scope and challenges. In *Vis. Data Mining*. Springer, 2008, pp. 76–90. 3

[KNBV22] KHAN A. A., NEWN J., BAILEY J., VELLOSO E.: Integrating gaze and speech for enabling implicit interactions. In *Proc. SIGCHI Conf. Human Fact. Comp. Syst.* (2022), pp. 1–14. 3, 7, 8, 10

[KR18] KASSEL J.-F., ROHS M.: Valletto: A multimodal interface for ubiquitous visual analytics. In *Extend. Abstr. 2018 CHI Conf. Human Fact. Comp. Syst.* (2018), pp. 1–6. 2

[KS18] KARAMAN Ç. Ç., SEZGIN T. M.: Gaze-based predictive user interfaces: Visualizing user intentions in the presence of uncertainty. *Int. J. Hum.-Comput. Stud. 111* (2018), 78–91. 5

[KT17] KROGSÆTER M., THOMAS C. G.: Adaptivity: System-initiated individualization. In *Adaptive User Support*. CRC Press, 2017, pp. 67–96. 2

[LAS24] LI H., APPLEBY G., SUH A.: A preliminary roadmap for llms as visual data analysis assistants. In *NLVIZ Work: Exploring Res. Opp. Nat. Lang., Text & Data Vis.* (2024). 10

[LBX22] LLERAS A., BUETTI S., XU Z. J.: Incorporating the properties of peripheral vision into theories of visual search. *Nat. Rev. Psychology 1*, 10 (2022), 590–604. 6

[LCC16] LALLÉ S., CONATI C., CARENINI G.: Prediction of individual learning curves across information visualizations. *User Model. User-Adap. Interact. 26*, 4 (2016), 307–345. 6

[LCT19] LALLÉ S., CONATI C., TOKER D.: A gaze-based experimenter platform for designing and evaluating adaptive interventions in information visualizations. In *Proc. ACM Eye Track. Res. Appl.* (2019), pp. 1–3. 3, 6, 7

[LIRC12] LEE B., ISENBERG P., RICHE N. H., CARPENDALE S.: Beyond mouse and keyboard: Expanding design considerations for information visualization interactions. *IEEE Trans. Vis. Comput. Graph. 18*, 12 (2012), 2689–2698. 2

[LMZ*11] LAGUN D., MANZANARES C., ZOLA S. M., BUFFALO E. A., AGICHTEIN E.: Detecting cognitive impairment by eye movement analysis using automatic classification algorithms. *J. Neurosci. Methods 201*, 1 (2011), 196–203. 6

[LQS*03] LIN J., QUAN D., SINHA V., BAKSHI K., HUYNH D., KATZ B., KARGER D. R.: The role of context in question answering systems. In *CHI '03 Extended Abstr. Human Fact. Comp.Syst.* (2003), ACM, pp. 1006–1007. 5

[LSS*18] LEE B., SRINIVASAN A., STASKO J. T., TORY M., SETLUR V.: Multimodal interaction for data visualization. In *Proc. Adv. Vis. Interf.* (2018), pp. 11:1–11:3. 2

[LTM17] LAM H., TORY M., MUNZNER T.: Bridging from goals to tasks with design study analysis reports. *IEEE Trans. Vis. Comput. Graph. 24*, 1 (2017), 435–445. 4

[MGG*23] MONADJEMI S., GUO M., GOTZ D., GARNETT R., OTTLEY A.: Human-computer collaboration for visual analytics: an agent-based framework. *arXiv preprint arXiv:2304.09415* (2023). 2, 10

[MNES22] MITRA R., NARECHANIA A., ENDERT A., STASKO J.: Facilitating conversational interaction in natural language interfaces for visualization. In *IEEE Vis. & Vis.Analytics* (2022), IEEE, pp. 6–10. 1, 2, 9, 10

[MSSB16] MANSOURYAR M., STEIL J., SUGANO Y., BULLING A.: 3D gaze estimation from 2D pupil positions on monocular head-mounted eye trackers. In *Proc. ACM Eye Track. Res. Appl.* (2016), ACM, pp. 197–200. 7

[NHH*14] NGUYEN T. T., HUI P.-M., HARPER F. M., TERVEEN L., KONSTAN J. A.: Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proc. Int'l Conf. on WWW* (2014), pp. 677–686. 9

[NSA*19] NEWN J., SINGH R., ALLISON F., MADUMAL P., VELLOSO E., VETERE F.: Designing interactions with intention-aware gaze-enabled artificial agents. In *Proc. IFIP Hum.-Comp. Interact.* (2019), Springer, pp. 255–281. 9, 10

[NSS20] NARECHANIA A., SRINIVASAN A., STASKO J.: Nl4dv: A toolkit for generating analytic specifications for data visualization from natural language queries. *IEEE Trans. Vis. Comput. Graph. 27*, 2 (2020), 369–379. 2, 10

[NVA*17] NEWN J., VELLOSO E., ALLISON F., ABDELRAHMAN Y., VETERE F.: Evaluating real-time gaze representations to infer intentions in competitive turn-based strategy games. In *Proc. Comput.-Hum. Interact. Play* (2017), ACM, pp. 541—552. 9

[OAJ14] OKOE M., ALAM S. S., JIANU R.: A gaze-enabled graph visualization to improve graph reading tasks. *Comput. Graph. Forum 33* (2014). 1, 3, 5, 6, 7

[OCZC15] OTTLEY A., CROUSER R. J., ZIEMKIEWICZ C., CHANG R.: Manipulating and controlling for personality effects on visualization tasks. *Inf. Vis. 14*, 3 (2015), 223–233. 8

[OGW19] OTTLEY A., GARNETT R., WAN R.: Follow the clicks: Learning and anticipating mouse interactions during exploratory data analysis. *Comput. Graph. Forum 38*, 3 (2019), 41–52. 2, 10

[Opp17] OPPERMANN R.: *Adaptive user support: ergonomic design of manually and automatically adaptable software*. Routledge, 2017. 1, 2

[ÖRB*20] ÖNEY S. Z., RODRIGUES N., BECHER M., ERTL T., REINA G., SEDLMAIR M., WEISKOPF D.: Evaluation of gaze depth estimation from eye tracking in augmented reality. In *Proc. ACM Eye Track. Res. Appl.* (2020), ACM, pp. 9–16. 12

[PB05] POOLE A., BALL L. J.: Eye tracking in human-computer interaction and usability research: Current status and future. In *Prospects", Chapter in C. Ghaoui (Ed.): Encyclopedia of Human-Computer Interaction. Pennsylvania: Idea Group, Inc* (2005). 5

[PB06] POOLE A., BALL L. J.: Eye tracking in HCI and usability research. In *Ency. Hum. Comput. Interact.* IGI Global, 2006, pp. 211–219. 5

[PC08] PRASOV Z., CHAI J. Y.: What's in a gaze, role of eye-gaze in reference resolution in multimodal conversational interfaces. In *Proc. Intel. User Interf.* (2008), ACM, pp. 20—-29. 8

[PMCEA*22] PÉREZ-MESSINA I., CENEDA D., EL-ASSADY M., MIKSCH S., SPERRLE F.: A typology of guidance tasks in mixed-initiative visual analytics environments. *Comput. Graph. Forum 41*, 3 (2022), 465–476. 2, 10

[Pos80] POSNER M. I.: Orienting of attention. *Quarterly Jour. Experim. Psychol. 32*, 1 (1980), 3–25. 3

[PPD*20] PALETTA L., PSZEIDA M., DINI A., RUSSEGGER S., SCHUESSLER S., JOS A., SCHUSTER E., STEINER J., FELLNER M.: MIRA–a gaze-based serious game for continuous estimation of alzheimer's mental state. In *Proc. ACM Eye Track. Res. Appl.* (2020), pp. 1–3. 10

[PPG*19] PALETTA L., PSZEIDA M., GANSTER H., FUHRMANN F., WEISS W., LADSTÄTTER S., DINI A., MURG S., MAYER H., BRIJACAK I., REITERER B.: Gaze-based human factors measurements for the evaluation of intuitive human-robot collaboration in real-time. In *IEEE Emerging Techn. Factory Autom.* (2019), pp. 1528–1531. 6

[PSMJ22] PATHIRANA P., SENARATH S., MEEDENIYA D., JAYARATHNA S.: Eye gaze estimation: A survey on deep learning-based approaches. *Expert Syst. Applic. 199* (2022), 116894. 2

[PSS*05] PUOLAMÄKI K., SALOJÄRVI J., SAVIA E., SIMOLA J., KASKI S.: Combining eye movements and collaborative filtering for proactive information retrieval. In *Proc. ACM Res. Dev. Inf. Retrieval* (2005), pp. 146–153. 8

[RESC15] RAGAN E. D., ENDERT A., SANYAL J., CHEN J.: Characterizing provenance in visualization and data analysis: an publisheral framework of provenance types and purposes. *IEEE Trans. Vis. Comput. Graph. 22*, 1 (2015), 31–40. 4, 6

[RK14] RZESZOTARSKI J. M., KITTUR A.: Kinetica: Naturalistic multi-touch data visualization. In *Proc. SIGCHI Conf. Human Fact. Comp. Syst.* (2014), pp. 897–906. 2

[RKB*17] RAPTIS G. E., KATSINI C., BELK M., FIDAS C., SAMARAS G., AVOURIS N.: Using eye gaze data and visual activities to infer human cognitive styles: method and feasibility studies. In *Proc. User Model. Adapt. Personaliz.* (2017), pp. 164–173. 6

[RKP*24] RAU T., KOCH M., PATHMANATHAN N., ISENBERG T., WEISKOPF D., SEDLMAIR M., KÖHN A., KURZHALS K.: Understanding collaborative learning of molecular structures in AR with eye tracking. *IEEE Comput. Graph. Appl. 44*, 6 (2024), 39–51. 3

[RMA*14] RAZIP A. M., MALIK A., AFZAL S., POTRAWSKI M., MACIEJEWSKI R., JANG Y., ELMQVIST N., EBERT D. S.: A mobile visual analytics approach for law enforcement situation awareness. In *Proc. IEEE Pac. Vis. Symp.* (2014), IEEE, pp. 169–176. 10, 12

[Ros16] ROSENHOLTZ R.: Capabilities and limitations of peripheral vision. *Annual Rev. Vision Sci.*, 2 (2016), 437–57. 6

[RSY*22] RODRIGUES N., SHAO L., YAN J. J., SCHRECK T., WEISKOPF D.: Eye gaze on scatterplot: Concept and first results of recommendations for exploration of sploms using implicit data selection. In *Proc. ACM Eye Track. Res. Appl.* (2022), pp. 1–7. 3, 8

[Sal00] SALVUCCI D. D.: A model of eye movements and visual attention. In *Proc. Int. Conf. Cognitive Modeling* (2000), pp. 252–259. 7

[SBR*19] SILVA N., BLASCHECK T., RODRIGUES N., WEISKOPF D., RAUBAL M., SCHRECK T.: Eye tracking support for visual analytics systems: theoretical foundations, opportunities, and research themes. In *Proc. ACM Eye Track. Res. Appl.* (2019), ACM, pp. 131–139. 3, 4, 5

[SCC13] STEICHEN B., CARENINI G., CONATI C.: User-adaptive information visualization: using eye gaze data to infer visualization tasks and user cognitive abilities. In *Proc. Intel. User Interf.* (2013), ACM, pp. 317–328. 3, 4

[SEB*24] SRINIVASAN A., ELLEMOSE J., BUTCHER P. W., RITSOS P. D., ELMQVIST N.: Attention-aware visualization: Tracking and responding to user perception over time. *arXiv preprint arXiv:2404.10732* (2024). 3, 5, 7, 9, 10

[SG00] SALVUCCI D. D., GOLDBERG J. H.: Identifying fixations and saccades in eye-tracking protocols. In *Proc. ACM Eye Track. Res. Appl.* (2000), pp. 71–78. 5

[SJB*21] SPERRLE F., JEITLER A., BERNARD J., KEIM D., EL-ASSADY M.: Co-adaptive visual data analysis and guidance processes. *Computers & Graphics 100* (2021), 93–105. 2, 9, 10

[SK09] SU X., KHOSHGOFTAAR T. M.: A survey of collaborative filtering techniques. *Adv. Artif.* (2009). 8

[SLJL10] SUN Y., LEIGH J., JOHNSON A., LEE S.: Articulate: A semi-automated model for translating natural language queries into meaningful visualizations. In *Intl. Symp. Smart Graphics* (2010), Springer, pp. 184–195. 2

[SLMS09] STREIT M., LEX A., MÜLLER H., SCHMALSTIEG D.: Gaze-based focus adaption in an information visualization system. In *Proc. IADIS Comput. Graph., Vis., Comput. Vision, and Image Proc.* (2009), pp. 303–307. 3, 8, 9

[SLS20] SRINIVASAN A., LEE B., STASKO J.: Interweaving multimodal interaction with flexible unit visualizations for data exploration. *IEEE Trans. Vis. Comput. Graph. 27*, 8 (2020), 3519–3533. 2, 7, 8, 10

[SMWH16] SATYANARAYAN A., MORITZ D., WONGSUPHASAWAT K., HEER J.: Vega-lite: A grammar of interactive graphics. *IEEE Trans. Vis. Comput. Graph. 23*, 1 (2016), 341–350. 7

[SS17] SRINIVASAN A., STASKO J.: Orko: Facilitating multimodal interaction for visual exploration and analysis of networks. *IEEE Trans. Vis. Comput. Graph. 24*, 1 (2017), 511–521. 2, 7, 10

[SSE*15] SILVA N., SETTGAST V., EGGELING E., ULLRICH T., SCHRECK T., FELLNER D.: Increasing fault tolerance in operational centres using human sensing technologies: approach and initial results. In *Eur. Proj. Space Comput. Vision, Graph., Optics, Photonics*. SciTePress, 2015, pp. 25–49. 3, 6, 8, 10

[SSES17] SHAO L., SILVA N., EGGELING E., SCHRECK T.: Visual exploration of large scatter plot matrices by pattern recommendation based on eye tracking. In *Proc. ACM Explorat. Search Interact. Data Analyt.* (2017), ACM, pp. 9–16. 1, 2, 3, 6, 8, 9

[SSL*22] SHEN L., SHEN E., LUO Y., YANG X., HU X., ZHANG X., TAI Z., WANG J.: Towards natural language interfaces for data visualization: A survey. *IEEE Trans. Vis. Comput. Graph.* (2022). 1, 2

[SSS*16a] SILVA N., SHAO L., SCHRECK T., EGGELING E., FELLNER D. W.: Sense.me - open source framework for the exploration and visualization of eye tracking data. In *IEEE VIS Poster Compend.* (2016). 1

[SSS*16b] SILVA N., SHAO L., SCHRECK T., EGGELING E., FELLNER D. W.: Visual exploration of hierarchical data using degree-of-interest controlled by eye-tracking. In *Proc. Forum Media Techn. & All Around Audio Symp.* (2016), pp. 82–89. 3, 6, 8

[SSS20] SAKTHEESWARAN A., SRINIVASAN A., STASKO J.: Touch? speech? or touch and speech? investigating multimodal interaction for visual network exploration and analysis. *IEEE Trans. Vis. Comput. Graph. 26*, 6 (2020), 2168–2179. 2, 7, 8

[SSV03] SHELL J. S., SELKER T., VERTEGAAL R.: Interacting with groups of computers. *Commun. ACM 46*, 3 (2003), 40–46. 8

[SSV*18] SILVA N., SCHRECK T., VEAS E. E., SABOL V., EGGELING E., FELLNER D. W.: Leveraging eye-gaze and time-series features to predict user interests and build a recommendation model for visual analysis. In *Proc. ACM Eye Track. Res. Appl.* (2018), pp. 13:1–13:9. 1, 3, 5, 8, 9, 10

[STD19] SETLUR V., TORY M., DJALALI A.: Inferencing underspecified natural language utterances in visual analysis. In *Proc. Intl. Conf. Intellig. User Interf.* (2019), pp. 40–51. 2

[STM16] SARVGHAD A., TORY M., MAHYAR N.: Visualizing dimension coverage to support exploratory analysis. *IEEE Trans. Vis. Comput. Graph. 23*, 1 (2016), 21–30. 8

[Tab19] TABLEAU: Tableau software. https://www.tableau.com/, 2019. [Online; accessed 08-Jan-2019]. 10, 11

[TCSC13] TOKER D., CONATI C., STEICHEN B., CARENINI G.: Individual user characteristics and information visualization: connecting the dots through eye tracking. In *Proc. ACM Hum. Factors Comput. Sys.* (2013), pp. 295–304. 3, 4, 6

[TGL*09] TSIANOS N., GERMANAKOS P., LEKKAS Z., MOURLAS C., SAMARAS G.: Eye-tracking users' behavior in relation to cognitive style within an e-learning environment. In *Proc. IEEE Adv. Learning Techn.* (2009), pp. 329–333. 3

[TSG*14] TOKER D., STEICHEN B., GINGERICH M., CONATI C., CARENINI G.: Towards facilitating user skill acquisition: identifying untrained visualization users through eye tracking. In *Proc. Intel. User Interf.* (2014), pp. 105–114. 3, 6, 8

[TSTR12] TOLEDO A., SOOKHANAPHIBARN K., THAWONMAS R., RINALDO F.: Personalized recommendation in interactive visual analysis of stacked graphs. *ISRN Artif. Intel. 2012* (2012). 5

[VC16] VELLOSO E., CARTER M.: The emergence of eyeplay: a survey of eye interaction in games. In *Proc. Comput.-Hum. Interact. Play* (2016), pp. 171–185. 8

[WCWQ21] WANG Q., CHEN Z., WANG Y., QU H.: A survey on ml4vis: Applying machine learning advances to data visualization. *IEEE Trans. Vis. Comput. Graph. 28*, 12 (2021), 5134–5153. 2

[Wei24] WEISKOPF D.: Bridging quantitative and qualitative methods for visualization research: A data/semantics perspective in light of advanced AI. In *Proc. IEEE Eval. Beyond – Methodol. Approach. Vis.* (2024), IEEE, pp. 119–128. 3

[WFMD21] WENSKOVITCH J., FALLON C., MILLER K., DASGUPTA A.: Beyond visual analytics: Human-machine teaming for ai-driven data sensemaking. In *IEEE Work. TRust & EXpertise VA* (2021), IEEE, pp. 40–44. 10

[WLDM23] WANG F. S., LOHMEYER Q., DUCHOWSKI A., MEBOLDT M.: Gaze is more than just a point: Rethinking visual attention analysis using peripheral vision-based gaze mapping. In *Proc. ACM Eye Track. Res. Appl.* (2023), pp. 1–7. 6

[WMA*15] WONGSUPHASAWAT K., MORITZ D., ANAND A., MACKINLAY J., HOWE B., HEER J.: Voyager: Exploratory analysis via faceted browsing of visualization recommendations. *IEEE Trans. Vis. Comput. Graph. 22*, 1 (2015), 649–658. 9

[WSE19] WALL E., STASKO J., ENDERT A.: Toward a design space for mitigating cognitive bias in vis. In *IEEE Vis. Conf.* (2019), IEEE, pp. 111–115. 8

[WW02] WEBSTER J., WATSON R. T.: Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly* (2002), xiii–xxiii. 4

[WW08] WRIGHT R., WARD L.: *Orienting of Attention*. Oxford University Press, 2008. 3

[WWS*21] WU A., WANG Y., SHU X., MORITZ D., CUI W., ZHANG H., ZHANG D., QU H.: Ai4vis: Survey on artificial intelligence approaches for data visualization. *IEEE Trans. Vis. Comput. Graph.* (2021). 2

[XAJK*15] XU K., ATTFIELD S., JANKUN-KELLY T., WHEAT A., NGUYEN P. H., SELVARAJ N.: Analytic provenance for sensemaking: A research agenda. *IEEE Comput. Graph. Appl. 35*, 3 (2015), 56–64. 4

[XJL08] XU S., JIANG H., LAU F.: Personalized online document, image and video recommendation via commodity eye-tracking. In *Proc. ACM Recom. Sys.* (2008), ACM, pp. 83–90. 9

[YBD*18] YAO N., BREWER J., D'ANGELO S., HORN M., GERGLE D.: Visualizing gaze information from multiple students to support remote instruction. In *Ext. Abstr. Conf. Hum. Factors Comput. Sys.* (2018), pp. 1–6. 9

[YKSJ07] YI J. S., KANG Y., STASKO J., JACKO J.: Toward a deeper understanding of the role of interaction in information visualization. *IEEE Trans. Vis. Comput. Graph. 13*, 6 (2007), 1224–1231. 7

[ZCHK16] ZHAO Q., CHANG S., HARPER F. M., KONSTAN J. A.: Gaze prediction for recommender systems. In *Proc. ACM Recom. Sys.* (2016), ACM, pp. 131–138. 6

[ZGI*17] ZHAO J., GLUECK M., ISENBERG P., CHEVALIER F., KHAN A.: Supporting handoff in asynchronous collaborative sensemaking using knowledge-transfer graphs. *IEEE Trans. Vis. Comput. Graph. 24*, 1 (2017), 340–350. 9

[ZK09] ZIEMKIEWICZ C., KOSARA R.: Preconceptions and individual differences in understanding visual metaphors. *Comput. Graph. Forum 28*, 3 (2009), 911–918. 8

[ZWG*22] ZHOU Z., WANG W., GUO M., WANG Y., GOTZ D.: A design space for surfacing content recommendations in visual analytic platforms. *IEEE Trans. Vis. Comput. Graph.* (2022). 9, 10