Individual User Characteristics and Information Visualization: Connecting the Dots through Eye Tracking

Dereck Toker, Cristina Conati, Ben Steichen, Giuseppe Carenini
Department of Computer Science
University of British Columbia, Vancouver, Canada
{dtoker, conati, steichen, carenini}@cs.ubc.ca

ABSTRACT
There is increasing evidence that users' characteristics such as cognitive abilities and personality have an impact on the effectiveness of information visualization techniques. This paper investigates the relationship between such characteristics and fine-grained user attention patterns. In particular, we present results from an eye tracking user study involving bar graphs and radar graphs, showing that a user's cognitive abilities such as perceptual speed and verbal working memory have a significant impact on gaze behavior, both in general and in relation to task difficulty and visualization type. These results are discussed in view of our long-term goal of designing information visualisation systems that can dynamically adapt to individual user characteristics.

Author Keywords
User characteristics, Information Visualization, Eye Tracking, Adaptive Information Visualization.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
User Studies; Human Factors; Design; Measurement.

INTRODUCTION

Information visualization (Infovis for short) aims to assist users in exploring, managing, and understanding the ever-growing amount of digital information. While visualizations have gained increasingly in terms of general usage and usability, they have traditionally followed a one-size-fits-all model, typically ignoring user differences. However, recent research has shown that individual differences can indeed have a significant impact on task effectiveness and user satisfaction during Infovis usage. For example, personality traits have been found to impact a user's performance with different Infovis designs [31, 15]. Velez et al. [30] found that a user's abilities for spatial reasoning (e.g., spatial orientation) were correlated with visualization comprehension. Similarly, Conati & Maclaren [3] and Toker et al. [28] found that cognitive abilities such as perceptual speed, visual/verbal working memory and expertise can impact user performance or subjective preference with a given visualization.

CHI 2013, April 27–May 2, 2013, Paris, France.

These studies indicate that it is important to investigate the possibility of user-adaptive information visualization systems, namely, Infovis that can dynamically adapt to individual differences. User-adaptive interaction has been shown to be effective in a variety of applications such as web search, desktop assistance, and e-learning [18], but it is largely unexplored in information visualization. Notable exceptions are [13, 14], which monitor a user's interaction data to detect and adapt to suboptimal usage patterns. In contrast, our research goal is to investigate how to detect and adapt to longer-term user cognitive abilities, which have been shown to be relevant for effective information visualization processing.

While Conati & Maclaren [3] investigated the impact of these cognitive abilities on overall user performance with different visualizations, the research presented in this paper aims to gain a more fine-grained understanding of the impact that these cognitive abilities have on visualization processing. One of the most informative (and sometimes the only available) sources of real-time information on visualization processing is a user's gaze data, because visual scanning and elaboration are fundamental components of working with a visualization (they are in fact the only components for non-interactive visualizations). Therefore, in this paper we aim to determine if and how features in user gaze behavior are impacted by different user characteristics. In particular, we aim to answer the following questions:

1) Do individual user characteristics influence a user's gaze behavior in a way that is detectable by state of the art eye tracking?
2) If yes, (a) which gaze features are influenced by which user characteristics? (b) Is the effect modulated by task context (e.g., task difficulty), and visualization type (e.g., bar vs. radar graph)?

Answering these questions can provide a better understanding of how specific user characteristics influence the processing of both information visualizations in general, as well as different visualization types (e.g., bar graphs), and individual visualization components (e.g., a graph's legend). Moreover, these answers can help drive the design of user-adaptive visualizations. In particular, by finding that user characteristics influence a user's gaze behavior in a way that is detectable via eye tracking, we can consider exploring eye tracking as a source of real-time information to provide adaptive interventions targeting these characteristics. For instance, if an Infovis system could detect from gaze data that the current user has low perceptual speed (an element that, as we will see later in the paper, may negatively impact the user's interaction with a visualization), it could then generate interventions to facilitate visualization processing, e.g., through highlighting or by explanatory material.

In exploring these research questions, this paper makes the following contributions. First, we present a statistical analysis using Mixed-Models [8] to investigate how a user's gaze behavior relates to user characteristics, task difficulty, and visualization type. We argue that a Mixed Model analysis is the most suitable statistical model to leverage at best the generally noisy eye tracking data. Secondly, we present a novel definition of task difficulty, derived through applying Principal Component Analysis to a selection of both objective and subjective performance measures. Thirdly, the results from the analysis show that user characteristics indeed have a significant influence on user gaze behavior, and that this influence is detectable through a variety of eye tracking metrics. We discuss these results in detail, as well as how they can inform the design of user-adaptive information visualizations

RELATED WORK

The use of eye tracking has long been established in Psychology as a means for analyzing user attention patterns in information processing tasks [20]. Research in this field has also investigated the impact of individual user differences on reading and search tasks [23].

Researchers in human-computer interaction and information visualization have also started to use eye tracking technology to investigate trends and differences in user attention patterns and cognitive/decision processing. This research has typically focused on either identifying differences in gaze patterns for different visualizations [12], task types [17, 26], and activities within a task [4], or on explaining differences in user accuracy between alternative visualization interfaces [22]. While these studies provide valuable insights on how different tasks and/or activities affect a user's gaze behaviors, they have traditionally ignored individual differences among study participants.

Recent research, however, has shown that user differences can have a significant impact on a user's performance during Infovis tasks. For example, Ziemkiewicz et al. [31], and Green and Fisher [15], looked at the influence of personality traits, showing that locus of control impacts performance across different visualizations. Cognitive measures such as perceptual speed and visual memory have been shown to influence a user's ability to complete a task effectively [3, 30]. These results were extended in a recent study [28] showing that perceptual speed and visual/verbal working memory influence not only task performance, but also a user's subjective preference for two different visualizations. While each of these studies clearly indicate that user differences should be considered in Infovis, they do not explain why or how these differences impact visualization processing, nor do they examine how this impact could be detected in real time. In this paper, we address these issues by providing a detailed analysis of how a set of user characteristics (including three cognitive measures and visualization expertise) influences a variety of eye gaze features during visualization processing.

To the best of our knowledge, there is no established comprehensive theory connecting eye gaze patterns and individual user traits that could guide our investigation of gaze patterns during visualization processing. Previous work has empirically identified relationships between eye gaze and individual user differences in attention-related tasks [21] (e.g., measuring a user's susceptibility to distraction), which is not directly relevant to our focus. Our research adds to this body of empirical work by providing detailed evidence of how individual differences affect a user's gaze patterns during Infovis usage. The closest to our research is work by Tai et al. [26] and Tang et al. [27], who focused on a single, domain-specific user trait (task-domain expertise), showing that domain experts and novices display different gaze behaviors. The scope of our work is broader, since we investigate a comprehensive array of user characteristics including cognitive abilities and *visualization expertise*, (i.e., expertise on specific visualizations). These characteristics are domain-independent, thus our results are more general across different Infovis tasks. Furthermore, we perform a detailed analysis of gaze data in order to link different characteristics to standard Infovis components (e.g., legends).

One approach to analyze eye tracking data is to apply data mining techniques, such as Hidden Markov Models [4], Scan-Path clustering [11], or specifically defined unsupervised algorithms [5, 16]. While data mining methods can quickly identify clusters of similar attention patterns during visualization tasks, the results they return are often difficult to interpret, since unsupervised algorithms are typically applied as black-boxes. By contrast, although traditional human-guided statistical analyses can be more time-consuming, its findings tend to be more transparent and easier to interpret. Our paper presents such a human-guided analysis of how user gaze behavior relates to user, task, and visualization characteristics. In particular, the paper provides fine-grained insights on how a set of user characteristics interact with different visualization types, components, and task difficulty to impact gaze patterns.

USER STUDY

In this section, we describe the study that we conducted to investigate the relationship among user characteristics, task difficulty and gaze patterns while using different visualizations. As case studies, we considered two basic visualization techniques: bar graphs (Figure 1, top) and radar graphs (Figure 1, bottom). Bar graphs were chosen because they are one of the most popular and effective visualization techniques. We chose radar graphs because, although they are often considered inferior to bar graphs on common information seeking tasks [7], they are widely used for multivariate data. Furthermore, there are indications that radar graphs may be just as effective as bar graphs for more complex tasks [28].

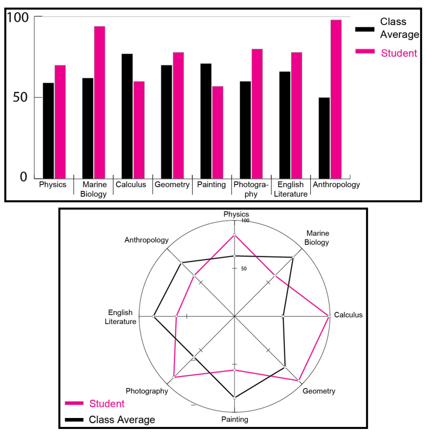

Figure 1. Sample bar (top) and radar graph (bottom)

User characteristics explored in the study

The user characteristics that we investigate in this study consist of two measures of prior *visualization expertise*, one for each of the two visualizations, as well as three cognitive abilities: *perceptual speed* (a measure of speed when performing perceptual tasks), *verbal working memory* (*verbal WM* - a measure of storage and manipulation capacity of verbal information), and *visual working memory* (*visual WM* - a measure of storage and manipulation capacity of visual and spatial information). Visualization expertise was chosen because we hypothesized that users with different levels of expertise might exhibit different gaze behaviors. For our study, participants self-reported their expertise by expressing their agreement with the statement "I am an expert in using radar(bar) graphs," on a Likert-scale from 1 to 5. Perceptual speed and visual WM were selected because they were among the perceptual abilities explored by Velez et al. [30], as well as among the set that Conati and Maclaren [3] found to impact user performance with radar graphs and a Multiscale Dimension Visualizer (MDV). We chose verbal WM because we hypothesized that it may affect performance in processing textual components of a visualization (e.g., labels/ legends).

Experiment tasks

Participants were asked to perform a set of tasks related to evaluating student performance in eight different courses. This task domain was chosen because it did not require any specific background knowledge, thus domain expertise was not a factor in our study. The tasks were based on a set of low-level analysis tasks that Amar et al. [1] identified as largely capturing people's activities with visualizations. The tasks were chosen so that each of our two target visualizations would be suitable to support them. A first battery of tasks involved 10 questions comparing the performance of one student with the class average for eight courses (e.g., *"In how many courses is Maria below the class average?")*. A second battery of tasks involved 4 questions comparing the performances of two different students with respect to the class average, e.g., *"Find the courses in which Andrea is below the class average and Diana is above it"*.

Procedure

Thirty-five subjects (18 female), ranging in age from 19 to 35, participated in the experiment. Participants were recruited via advertising at our university, with the aim of collecting a heterogeneous pool of participants with suitable variability in the target characteristics. Ten participants were CS students, while the rest came from a variety of backgrounds, including microbiology, economics, classical archaeology, and film production. The experiment was a within-subjects study, designed and pilot-tested to fit in a single session lasting at most one hour. Participants began by completing tests for the three cognitive measures: a computer-based OSPAN test for verbal WM [29] (lasting between 7 and 12 minutes), a computer-based test for visual WM [10] (10 minutes long), and a paper-based P-3 test for perceptual speed [6] (3 minutes long). The experiment was conducted on a Pentium 4, 3.2GHz, with 2GB of RAM and a Tobii T120 eye tracker as the main display. Tobii T120 is a remote eye tracker embedded in a 17" display, providing unobtrusive eye tracking. After undergoing a calibration phase for the eye tracker, each participant performed the 14 tasks described in the previous section twice, once with each of the two target visualizations. The presentation order with respect to visualization type was fully counterbalanced across subjects. Each task consisted of presenting the participant with a radar/bar graph displaying the relevant data, along with a textual question. Participants would then select their answer from a set of available options, and click OK to advance to the next task. Before seeing the next task, participants were shown a screen asking them to rate their confidence in their answer on a Likert scale from 1 to 5. The experimental software was fully automated and coded in Python.

DATA ANALYSIS

Independent Measures

The independent measures for our study consisted of the collected cognitive abilities and expertise measures (continuous), visualization type (categorical: bar vs. radar), and task difficulty (continuous - values described in the next section). Table 1 presents summary statistics on the user characteristics data collected from the study. The rather large variances for most measures indicate that we succeeded in collecting a diverse pool of participants. We also verified our results with other studies involving the same tests and found similar results from other populations.

Measure	N	Min	Max	Mean	Std. Dev
Percep. Speed	35	54.00	96.00	85.70	11.64
Visual WM	35	0.3	5.4	2.72	1.36
Verbal WM	35	2	6	4.58	1.10
Bar Expert	35	2	5	4.15	0.87
Radar Expert	35	1	5	2.12	1.22

Table 1. User characteristics collected from our study

Task difficulty

Defining tasks as being easy or difficult a priori is challenging, since difficulty depends upon user expertise and perceptual abilities, which were varied on purpose in our study. We therefore defined task difficulty a posteriori, based on four different measures (two objective and two subjective) aggregated using a principal component analysis [8]. Because there was a ceiling effect on task correctness, our first objective measure of task difficulty is *task completion time* (assuming that, in general, more time is needed for more difficult tasks). However, longer completion times may also simply be an indication of a task being longer while not necessarily being more difficult. Therefore our second objective measure of difficulty is the *standard deviation of completion time* for each task, across all users. A high value of this metric indicates a high variability among users' completion times, an indicator that the task may be difficult or confusing for some users.

Our two chosen subjective measures of task difficulty are based on the users' reported confidence of their performance, which was elicited after each task. The first subjective measure is the *average confidence* reported by users on each task. Intuitively, less difficult tasks would have higher values for this average. However, we also want to take into account that some users may tend to be more confident overall than other users. Therefore, our second subjective measure is the *average deviation of confidence* for each task across all users and is computed as follows. For each user, we look at their average confidence across their tasks. Then, for each task, we compute the *deviation of confidence* as the difference between the user's reported confidence for that task and the user's average confidence across tasks. Finally, for each task, we average the deviation of confidence across all users. This average indicates for which tasks users were giving confidence ratings that were above or below their typical input.

In order to combine the four variables above, we performed a *Principal Component Analysis (PCA)*. PCA is a form of dimension reduction that allows one to identify and combine groups of inter-related variables into components more suitable for data analysis [8]. A PCA on our four measures of task difficulty resulted in one output component. Bartlett's test of sphericity (x^2 = 73.35, df = 6, p < .001) indicated that the principal component analysis was appropriate. Kaiser's sampling adequacy was 0.55 and all variables showed a communality > 0.52 which was above the acceptable limit of 0.5[1] [8]. The component we generated had an eigenvalue over Kaiser's criterion of 1 and explained 62.22% of the variance. In sum, we use the output component generated by this PCA (i.e., dimensional reduction) as the measure of task difficulty that we will investigate in our analysis.

Dependent Measures: Eye Tracking Features

Eye tracking measures

An eye tracker captures gaze information in terms of *fixations* (i.e., maintaining gaze at one point on the screen) and *saccades* (i.e., a quick movement of gaze from one fixation point to another), which can then be analyzed to derive a viewer's attention patterns. In this paper, we use a large set of basic eye tracking features described by Goldberg and Helfman [11] as the building blocks for comprehensive gaze processing. These features are built by calculating a variety of statistics upon the basic eye tracking measures that are described in Table 2.

Measure	Description
Fixation rate	Rate of eye fixations per milliseconds
Number of Fixations	Number of eye fixations detected during an interval of interest
Fixation Duration	Time duration of an individual fixation
Saccade Length	Distance between the two fixations delimiting the saccade (d in Figure 2)
Relative Saccade Angles	The angle between the two consecutive saccades (angle y in Figure 2)
Absolute Saccade Angles	The angle between a saccade and the horizontal (angle x in Figure 2)

Table 2. Description of basic eye tracking measures

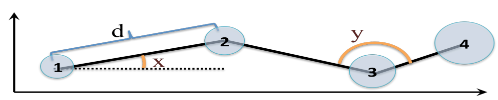

Figure 2. Saccade based eye measures

Among the measures described in Table 2, fixation rate, number of fixations, and fixation duration are widely used. In addition, we included saccade length, relative saccade angle, and absolute saccade angle, as suggested by Goldberg and Helfman [11], because these measures are useful to summarize trends in user attention patterns within a specific interaction window (e.g., if the user's gaze seems to follow a planned sequence as opposed to being scattered).

The gaze features for our analysis are obtained by computing statistics such as sum, average, and standard deviation over the measures shown in Table 2, at two levels of granularity. At the *Task Level*, features are computed over each task as a whole (Table 3). At the *AOI level,*

[1] All subsequent PCA results reported meet the required criteria, and for simplicity we only report the value of Bartlett's test.

features are computed based on gaze activity within a specific region of the screen, or *Area Of Interest* (Table 4), including transitions between pairs of defined AOIs (five in our analysis, as described in the next section). To limit our analysis to a reasonable number of features, at the AOI level, we opted to calculate only proportionate features and did not include features related to path angles (note that each AOI feature increases complexity by a factor of 5). In total, we included 49 different features in our analysis (14 Task-level and 35 AOI-level), computed by processing raw data from the Tobii using customized Python scripts[2].

Total Fixation rate
Total Number of Fixations
Sum, Mean and Std. Dev. of Fixation Durations
Sum, Mean and Std. Dev. of Saccade Length
Sum, Mean and Std. Dev. of Relative Saccade Angles
Sum, Mean and Std. Dev. of Absolute Saccade Angle

Table 3. Task-level eye tracking features

Proportion of Fixation Durations
Proportion of Total Number of Fixations
Number of Transitions from this AOI to each other AOI (5 separate measures for each AOI)

Table 4. AOI-level eye tracking features

Areas of interest (AOI)

A total of five AOIs were defined for each of the two visualizations. These regions were selected in order to capture the distinctive and typical components of these two information visualizations. Figure 3 and 4 show how these AOIs map onto bar graph and radar graph visualizations.

- *High Area:* covers the upper half of the data elements of each visualization. This area is the graphical portion of an Infovis that contains the relevant data values. On the bar graph, it corresponds to a rectangle over the top half of the vertical bars (see Figure 3); for the radar graph, it corresponds to the combined area of the 8 trapezoidal regions covering the data points (see Figure 4).
- *Low Area:* covers the lower half of the data elements for each visualization.
- *Labels Area:* covers all the data labels in each graph.
- *Question Text Area:* covers the text describing the task to be performed.
- *Legend Area:* covers the legend showing the mapping between each student and the color of the visualization elements that represent her performance.

The selection of these five AOIs is the result of a trade-off between having detailed information on user attention by measuring very specific areas that are salient for task execution, versus keeping the number of AOIs manageable for data interpretation and analysis. Because each added AOI increases transition analysis by n², we opted to include fewer AOIs (e.g., we don't have an AOI for each bar in the bar graph or radial element in the radar graph), while still capturing distinctive areas that can be considered general components of many Infovis.

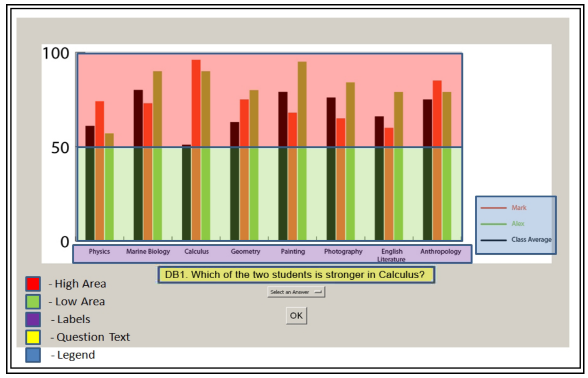

Figure 3. The five AOI regions defined over Bar Graph

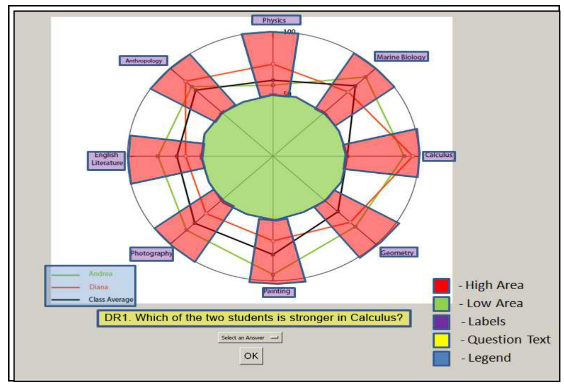

Figure 4. Five AOI regions defined over Radar Graph

Feature Reduction: Principal Component Analysis

To account for correlations among measures, we used three PCAs on our initial set of 49 gaze features. We grouped the gaze features into three non-overlapping families according to how the measures were intuitively related, namely (i) task-level features (e.g., *fixation rate*); (ii) AOI proportionate features (e.g., *proportion of fixation durations* in each AOI) and (iii) AOI transitions (e.g., *number of transitions* from one AOI to another). One PCA was performed on each of these three families, which allows us to discuss results in terms of high-level related gaze components rather than many low-level features.

The task level family consisted of 14 gaze features. Five components were generated[3] using PCA ($x^2 = 15434.49$, df = 91, $p < .001$), explaining 86.69% of the variance. Table 5 shows the breakdown of the original 14 features into the five components. Note that, for most components, it is quite easy to identify intuitive commonalities among its features, as reflected in the components' names (e.g., all

[2] These scripts are part of EMDAT (Eye Movement Data Analysis Toolkit), an open-source toolkit developed by our group.

[3] The number of components generated is always determined by using the Catell scree test, or if this test results in an ambiguous scree plot, we then use Kaiser's criterion to select only components with eigenvalues greater than 1 [7].

features of component 1 are based on sums, all features in component 2 relate to fixation measures, etc.). The same is true for the components resulting from PCA on the other two families of features.

Component Name	Task-Level Measures
Sum Measures	*num. fixations, sum rel. path angles, sum abs. path angles, sum path distance, sum fixation duration*
Fixation Measures	*mean fixation duration, std. dev. fixation duration, fixation rate***
Path Distance	*mean path distance, std. dev. path distance*
Std.Dev. Path Angles	*mean rel. path angles, std. dev. rel. path angles, std. dev. abs. path angles*
Mean Abs. Path	*mean abs. path angles*

Table 5. Generated components for the task-level family **fixation-rate is the only measure that inversely correlates with the other members of its component

The AOI proportionate family consisted of 10 gaze features. Five components (see Table 6) were produced from the PCA (x^2= 5706.32, df = 300, p < .001) and explained 97.13% of the variance. The AOI transitions family consisted of 25 features and the PCA generated four components (see Table 7), (x^2 = 8506.86, df = 45, p < .001), which explained 45.24% of the variance. Note that PCA proved to be especially useful for reducing the many AOI transition features to a small set of meaningful components, each including features mostly related to a specific AOI.

Component Name	AOI Proportionate Measures
Low prop.	*low prop. num. fixations, low prop. time*
Label prop.	*labels prop. num. fixations, labels prop. time*
Legend prop.	*legend prop. num. fixations, legend prop. time*
Text prop.	*text prop. num. fixations, text prop. time*
High prop.	*high prop. num. fixations, high prop. time*

Table 6. The components for AOI Proportionate Measures

Component Name	AOI Transition Measures
Legend Transitions	*legend to legend, legend to high, high to legend, text to legend, legend to text, legend to low, text to text, low to legend, legend to labels, labels to legend*
Low Transitions	*high to low, low to high, text to low, low to low, low to text*
Label Transitions	*labels to labels, labels to low, low to labels, text to labels, labels to text*
High Transitions	*high to labels, labels to high, high to text, high to high, text to high*

Table 7. Components for the AOI Transitions Family

Mixed Model Analysis

Since the study data involved repeated measures (e.g., each subject performed the same task type with each of the two different visualizations), a suitable means for analysis is a *Mixed Model* [8]. Mixed models can handle both repeated measures as well as the mix of categorical and continuous independent measures that we consider. An alternative model commonly used for repeated-measures analysis is a General Linear Model Repeated Measures analysis (GLM for short) [8]. GLM, however, is less suitable than a Mixed Model for eye tracking analysis, because it is less resilient to missing data. This issue is due to the fact that GLM requires data to be in wide format, where all repeated measures (trials) for each participant are listed in one data entry row. When there is an invalid trial, GLM is forced to discard the entire data for that participant. This can be costly in an experiment with several invalid trials, as is often the case when using unobtrusive eye trackers that do not constrain subjects' movements. By contrast, a Mixed Model uses data in long format, listing each trial as a different data entry, and discarded invalid trials do not interfere with valid ones. Thus, a Mixed Model analysis is able to leverage at best potentially noisy eye tracking data.

For each of our three families of gaze features (i.e., task level, AOI Proportionate, and AOI transition) we ran a mixed model over each of the generated PCA components within that family[4]. Each mixed model was a 2 (visualization type) by 2 (visualization order) model, with the user characteristics and task difficulty as the model's covariates. We report statistical significance at the 0.05 level. Effect sizes of our results are reported small for r = 0.1, medium for r = 0.3, and large as r = 0.5 [9]. In the next section, we report the most salient results of the analysis. When going over the results involving directionality, the reader should keep in mind that our dependent measures are PCA components, each consisting of a single value that represents a much larger collection of underlying measures. Each component is generated by (i) calculating the weighted values of its underlying members; (ii) aggregating and scaling these values into one number typically ranging from -1 to +1. If an underlying member is positively correlated to its corresponding component the directionality will be the same, otherwise it will be opposite.

RESULTS

In this section, we present results that provide answers to our original research questions: do individual user characteristics influence a user's eye gaze behavior in a way that is detectable by state of the art eye trackers? If so, which gaze features are influenced by which particular user characteristics, and is the effect modulated by task and visualization type? The analysis results are discussed per user characteristic.

Perceptual Speed - Main Effects

We found main effects of perceptual speed on three PCA components (see Table 8).

One main effect was at the task level (first row in Table 8), showing that High perceptual speed users had lower values

[4] Mixed Models are univariate analyses (ANOVA), thus do not support having more than one depended measure per model. We adjusted our models for family-wise error by applying the Bonferroni adjustment to each family of results, according to the number of components within that family.

of Fixation Measures than low perceptual speed users. An analysis of the underlying members of this component shows that users with high perceptual speed had a higher fixation-rate than low perceptual speed users, indicating that they were able to scan the screen more quickly. They also had lower average and standard deviation of fixation durations, i.e. shorter and more consistently timed fixations. These combined findings closely match the definition of perceptual speed, and are interesting because they show that individual differences for this cognitive ability may be captured via eye tracking measures that are not related to information on specific elements of the visualization.

Family	Component	F-Ratio	Effect Size	Sig. Value
Task level	Fixation Measures	F(1,27) = 8.9	r = 0.37	p = 0.03
AOI	Legend Proportion	F(1,21) = 25.2	r = 0.21	p < 0.001
	Legend Transitions	F(1,26) = 10.25	r = 0.16	p = 0.016

Table 8. Main effects of perceptual speed

The other two main effects of perceptual speed are at the AOI level (see Table 8), showing that this cognitive ability also affects eye gaze measures relating to specific visualizations elements. The main effects are on the two components *Legend Proportion* and *Legend Transitions*: low perceptual speed users spent more of their time in the legend AOI and transitioned to it more often than high perceptual speed users. This result indicates that users with low perceptual speed took more time to process/store legend-related information and looked at the legend more frequently (possibly because they tended to forget the contained information).

Perceptual Speed - Interactions

We found significant interactions of perceptual speed with both task difficulty and visualization type.

Family	Component	F-Ratio	Effect Size	Sig. Value
AOI	Legend Transitions	F(1,686)=6.85	r = 0.10	p < 0.05
	Label Transitions	F(1,676)=7.97	r = 0.11	p = 0.02

Table 9. Interaction Effects for Perceptual Speed and Task Difficulty.

Interactions with Task difficulty. There are significant interactions of task difficulty and perceptual speed on both the *Legend Transitions* and *Label Transitions* components (see Table 9). For *Legend Transitions,* all users generate more legend-related transitions with difficult tasks than with easy tasks (see Figure 5), likely due to the fact that an increased difficulty increases cognitive load and causes users to forget some of the information in the legend. This effect, however, is higher for low perceptual speed users.

For *Label Transitions* (Figure 6), all users show more label-related transitions for easy tasks, but the difference is much higher for low perceptual speed users. This effect is not as intuitive as the one found on *Legend Transitions*, but, irrespective of what causes users to have more label-related transitions for easy tasks, it seems to affect low perceptual speed users the most.

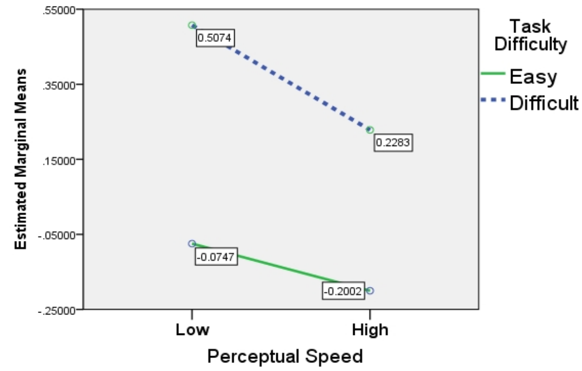

Figure 5. Interaction between Perceptual Speed and Task Difficulty on AOI Legend Transitions

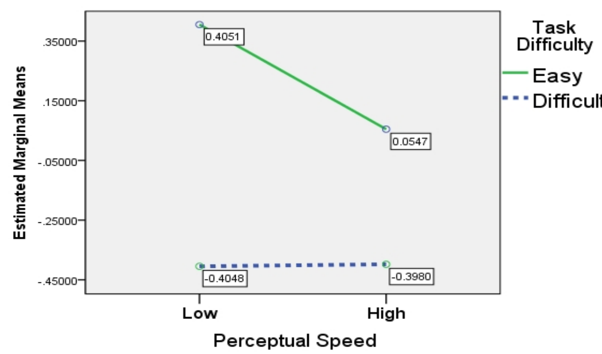

Figure 6. Interaction between Perceptual Speed and Task Difficulty on AOI Label Transitions

Interactions with visualization type. There was a significant interaction effect between perceptual speed and visualization type in terms of the High AOI Transitions component (F(1,680)=22.2, r=0.18, p < 0.001) (see Fig. 7).

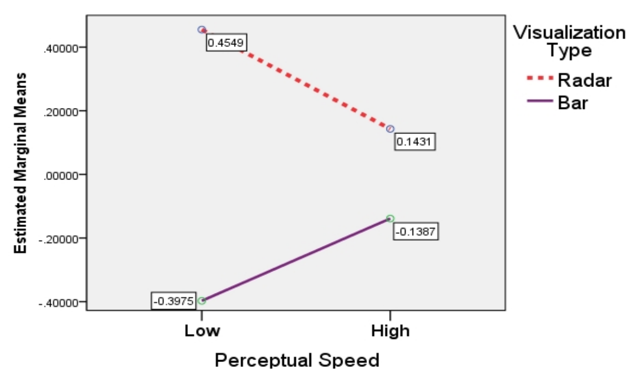

Figure 7. Interaction between perceptual speed and visualization type for *AOI High Transitions*

All users showed more High AOI related transitions with the radar graph than with the bar graphs, but the difference is much higher for low perceptual speed users. Given that the High AOI is the graphical portion of an Infovis that contains the relevant data values, this effect indicates that low perceptual speed users are more affected by different ways of visualizing data.

Verbal WM - Main Effects

There are two main effects of verbal WM: one on the *Text Proportion* component and one on *Standard Deviation of Path Angles* (see Table 10).

Text Proportion relates to the most textual element in our visualizations, namely the question text. An analysis of the members of this component shows that the proportionate amount of time spent on the Text AOI and the number of fixations in this area are lower for users with high verbal WM. This effect indicates that high verbal WM users refer to the task question less often than their low verbal WM counterparts, which is consistent with the definition of verbal WM as a measure of storage and manipulation capacity of verbal information. This result is interesting because it shows that differences in users' verbal WM can be directly captured by eye tracking features related to the primary textual elements of a visualization.

Family	Component	F-Ratio	Effect Size	Sig. Value
AOI	Text Proportion	$F(1,28) = 7.24$	$r = 0.36$	$p = 0.04$
Task Level	Std.Dev. PathAngles	$F(1,25) = 8.06$	$r = 0.32$	$P < 0.05$

Table 10. Main effects of Verbal WM

Std. Dev. Path Angles: this component essentially captures the consistency of a user's gaze patterns during a visualization task, because it is built upon features related to measuring the deviation of angles between subsequent saccades. Users with low verbal WM had higher values for *Std. Dev. Path Angles* than users with high verbal WM. When these values are higher, it indicates that a user is frequently looking across different areas of the screen, rather than following more planned or consistent path directions. Therefore, the finding that users with low verbal WM had higher values for *Std. Dev. Path Angles* is consistent with the finding that low verbal WM users referred back to the question text more often.

Bar and Radar Graph Expertise

There are two non-significant main effects of both bar graph and radar graph expertise, which we discuss because of their large effect sizes.

There was a main effect of *Bar Graph Expertise* on the AOI Label proportion component, ($F(1,21) = 6.042$, $r = 0.80$, $p = 0.1$), showing that users with high bar expertise spent a greater proportion of their time looking at labels compared to non-experts. Similarly, there was a main effect of *Radar Graph Expertise* on the AOI Legend proportion component ($F(1,21) = 5.732$, $r = 0.78$, $p = 0.129$), with radar experts spending less time looking at the legend when compared to non-experts. The discrepancy between strong effect sizes for these two expertise-related measures and the lack of statistical significance is likely due to limited statistical power. The power for the effect of Bar Graph expertise on the AOI Label is 0.67, and the power for the effect of Radar Graph expertise on AOI Legend is 0.64. A commonly recommended value of power is 0.8 [8], and we would have to add 17% (or 6) more users to reach this value.

It may seem surprising that we did not find stronger influences of visualization expertise on gaze patterns. This result, however, is consistent with findings in [28], which showed that bar and radar graph expertise may only have significant effects on user visualization preference, but not on performance. These findings suggest that there might not be easily detectable differences in the visualization processing behaviors of experts and novices, as defined by our self-rated measures of expertise.

Visual WM

We found no effects worth reporting for visual WM. This lack of findings may be due to the fact that the study tasks were relatively easy and that the visualizations were static in nature. It is thus likely that users did not require to reach their maximum visual memory capacity, especially since they could easily get an overview of the whole graph in a single look. Moreover, individual tasks were independent of each other, thus users were not required to store any successive visual information (one of the functions affected by visual WM).

SUMMARY OF FINDINGS & DISCUSSION

The goal of our study was to investigate 1) if user characteristics impact gaze patterns during visualization processing, and if the impact can be detected though eye-tracking; 2) which gaze measures are influenced by which user characteristics, as well as if/how the influence is mediated by task difficulty and visualization type. In this paper, we chose to focus on the five characteristics listed in Table 11 and, as shown in this table, we found a number of effects (either statistically significant or with large effect sizes) on various gaze measures.

Perceptual speed is the cognitive measure with the highest number of effects. This finding provides encouraging evidence that this cognitive ability could be reliably detected in real time using gaze information. This result is particularly important for our long-term goal of designing user-adaptive visualizations, especially in light of previous studies, which showed that low perceptual speed can negatively affect task performance, in terms of both accuracy [3] as well as task completion time [28]. We have shown that perceptual speed influences AOI-specific gaze measures relating to the legend, labels and High AOI. These findings suggest that adaptive interventions could be particularly useful if they support the access and/or processing of such AOIs for low perceptual speed users. In addition, the interaction effects we found for perceptual speed suggest that task difficulty and visualization type should be taken into account, if known, when providing adaptive interventions. For instance, we found that low perceptual speed users tended to access a visualization legend more than high perceptual speed users, suggesting that they should be specifically supported in terms of legend processing. However, we also found that this effect is exacerbated in the presence of difficult tasks. Thus, while it may not be worthwhile disrupting a low speed user with a legend-related intervention for tasks known to be easy, it may be important to do so as task difficulty increases.

User Characteristic	Eye tracking measure component
Perceptual Speed	*Fixation Measures* (main effect) *Legend Proportion* (main effect) *Legend Transitions* (main & interaction effect) *Label Transitions* (interaction effect) *High AOI Transitions* (interaction effect)
Verbal WM	*Std. Dev. Path Angles* (main effect) *Text Proportion* (main effect)
Bar Expertise	*Label Proportion* (p> 0.05, but large effect size)
Radar Expertise	*Legend* (p> 0.05, but large effect size)
Visual WM	*None*

Table 11. Overall Results

The results on *verbal WM* indicate, intuitively, that this cognitive ability affects eye-tracking features related to the main textual element of a visualization, and thus may be detectable in real time by tracking these features. In our experiment, the textual element was the question text, but in other settings this could be the visualization caption or the portion of text in which the visualization is embedded (e.g. possibly providing verbal descriptions of the displayed data). In terms of adaptation, it is plausible that users with low verbal WM may benefit if textual elements of a visualization were given more emphasis than the purely graphical elements. However, because we do not have information on whether verbal WM affects performance during Infovis processing, it remains a topic for future research to investigate if and how adaptive interventions would impact visualization effectiveness for users with different levels of verbal WM.

We discussed two non-significant main effects of the expertise-related user characteristics because of their large effect sizes. *Bar expertise* had a large effect on label access, while *radar expertise* had a large effect on legend access. These results may indicate that non-experts could benefit from adaptive interventions that guide them to access these elements in a way that is more similar to experts. However, we need to run further studies with more reliable, objective measures of expertise (the ones used in this study were self-reported) before we can make a more informed decision on how to provide adaptive support for novice users.

In summary, we have identified a set of user abilities that have a strong impact on gaze measures related to specific AOIs of a visualization, and discussed how adaptive interventions driven by these abilities and targeting such AOIs may improve a user's experience with a given visualization. While our study has only investigated two simple visualization techniques, several results may be generalized to a wider array of visualization designs, since they involve AOIs that are common to most types of visualizations (such as a graph's labels or legend). In fact, the majority of our results are effects that are actually independent of visualization design. Similarly, while the study has focused on an artificial data set involving student grades (in order to eliminate domain expertise as a study factor), the actual tasks were derived from an established set of general, low-level analysis tasks for information visualization [1]. Lastly, while this work has focused on an analysis for the purpose of adaptive information visualization, similar user studies could be performed in other areas of HCI (e.g., desktop interfaces), to determine whether the influence of individual user differences can also be detected in those scenarios.

CONCLUSION AND FUTURE WORK

We presented research aimed at investigating the relationship between a set of user cognitive and expertise measures, task difficulty, and user attention patterns when using different visualization techniques. Our analysis reveals that some of the tested user characteristics do have a significant influence on user gaze behavior, and that this influence is detectable through a variety of eye tracking metrics. Based on these findings, we provided general suggestions for adaptive visualization design in relation to components that are common to most types of visualizations, for example suggesting that low perceptual speed users may need support in processing legends. Our results may therefore be of interest when designing systems for specific user groups that are known to have high/low cognitive abilities (e.g., older adults and people with autism are known to have lower values for perceptual speed).

We see the analysis presented here as a first step towards understanding the complex relationships between user traits, visualizations, and gaze patterns. However, additional studies are necessary to investigate these relationships at the level of more basic Infovis properties such as color, size, and shape. Similarly, studies should be run to investigate these relationships in more complex visualizations such as time series, networks, as well as interactive visualizations. Along these lines, we are currently applying the experimental design described in this paper to investigate the impact of user traits on different versions of a complex interactive visualizations involving multiple, aligned bar charts [2] for preference elicitation. Because of the added complexity, we expect the impact of user characteristics, task difficulty, and visualization type to be even more pronounced than in the current study. The next step of our research is to show that the relevant user characteristics can be detected in real-time to drive adaptive interventions benefiting users with those characteristics. We are currently investigating a variety of machine learning techniques to perform this real-time inference task, and we already have encouraging results [25]. Lastly, we are in the process of running a user study to test different ways of providing adaptive interventions, both in general, and in relation to individual user differences.

REFERENCES

1. Amar, R.A., Eagan J., & Stasko, J.T. Low-Level Components of Analytic Activity in Information Visualization. In 16th IEEE Vis. Conf., (2005), 15-21.
2. Bautista, J., & Carenini, G. An integrated task-based framework for the design and evaluation of visualizations to support preferential choice. Proc. of the conf. on Adv. Vis. Interf. (AVI '06), (2006), 217-224.
3. Conati, C., & Maclaren, H. Exploring the Role of Individual Differences in Information Visualization. In Proc. of the working conf. on Adv. visual interfaces (AVI '08). ACM, New York, NY, USA, (2008), 199-20.
4. Courtemanche, F., Aïmeur, E., Dufresne, A., Najjar, M., & Eboa, F.H.M. Activity recognition using eye-gaze movements and traditional interactions. In Proceedings of In-teracting with Computers, (2011), 202-213.
5. Eivazi, S., Bednarik, R. Predicting Problem-Solving Behavior and Performance Levels from Visual Attention Data. In the proceedings of 2nd Workshop on Eye Gaze in Intelligent Human Machine Interaction at IUI 2011, (2011), 9-16.
6. Ekstrom, R., French, J., Harman, H. and Dermen, D, Manual from Kit of Factor-References Cognitive Tests. Educational Testing Service (1976): Princeton, NJ.
7. Few, S. Keep Radar Graphs Below the Radar - Far Below: Information Management Magazine, (2005).
8. Field, A. Discovering Statistics Using SPSS, Third Edition. Sage Publications, London, (2009).
9. Field & Hole, How to Design and Report Experiments. Sage Publications, (2003).
10. Fukuda, K., & Vogel, E.K. Human variation in overriding attentional capture. Journal of Neuroscience, (2009), 8726-8733.
11. Goldberg, J.H., & Helfman, J.I. Comparing Information Graphics: A Critical Look at Eye Tracking, BELIV ’10, (2010), 182-195.
12. Goldberg, J.H., & Helfman, J.I. Eye tracking for visualization evaluation: Reading values on linear versus radial graphs. Inf. Vis. 10(3), (2011), 182-195.
13. Gotz D., & Wen, Z.. Behavior-Driven Visualization Recommendation. Int. Conf. on Intelligent User Interfaces, (2009), 315-324.
14. Grawemeyer, B. Evaluation of ERST – an external representation selection tutor. In Proc. of the 4th int. conf. on Diagram. Repr. and Inference, (2006), 154-167.
15. Green, T.M., & Fisher, B. Impact of personality factors on interface interaction and development of user profiles: Next steps in the Personal Equation of Interaction. Inf. Visualization, 11(2), (2012), 1-17.
16. Grindinger, T., Duchowski, A., & Sawyer, M. Group-wise similarity and classification of aggregate scanpaths. In Proceedings of the 2010 Symposium on Eye-Tracking Research 38, (2010), 101-104.
17. Iqbal, S.T., & Bailey, B.P. Using eye gaze patterns to identify user tasks. In: The Grace Hopper Celebration of Women in Computing, (2004).
18. Jameson, A. (2008). Adaptive Interfaces and Agents. The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications (2nd ed.), pp. 433–458.
19. Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. In the eyes of the beholder: How experts and novices interpret dynamic stimuli. Learning and Instruction. 20, (2010), 146-154.
20. Keith, R. Eye movements and cognitive processes in reading, visual search, and scene perception. Eye Movement Research Mechanisms, Processes, and Applications. North-Holland Press, (1995).
21. Kruschke, J. K., Kappenman, E. S., & Hetrick, W. P. Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, (2005), 830–845.
22. Plumlee, M.D., & Ware, C. Zooming versus multiple window interfaces: Cognitive costs of visual comparisons. ACM Trans. Comput.-Hum. Interact, 13, 2, (2006), 179-209.
23. Rayner, K. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin 124, (1998), 372-422.
24. Simola, J., Salojarvi, J., & Kojo, I. Using hidden Markov model to uncover processing states from eye movements in information search tasks. Cognitive Systems Research. 9, (2008), 237-251.
25. Steichen, B., Carenini, G., & Conati, C. User-Adaptive Information Visualization - Using eye gaze data to infer visualization tasks and user cognitive abilities. Int. Conf. on Intelligent User Interfaces, (2013), (to appear).
26. Tai, R.H., Loehr, J.F., & Brigham, F.J. An Exploration of the Use of Eye-Gaze Tracking to Study Problem-Solving on Standardized Science Assessments. International Journal of Research & Method in Education 29, 2, (2006), 185-208.
27. Tang, H., Topczewski, J.J., Topczewski, A.M. & Pienta, N.J. Permutation test for groups of scanpaths using normalized Levenshtein distances and application in NMR questions. In Proc. of the Symposium on Eye Tracking Research and Applications, (2012), 169-172.
28. Toker, D., Conati, C., Carenini, G., & Haraty, M. Towards Adaptive Information Visualization: On the Influence of User Characteristics. In Proc. of the 19th international conference on user modeling, adaptation, and personalization, UMAP, (2012), 274-285.
29. Turner, M. L., & Engle, R.W. Is working memory capacity task dependent? Journal of Memory and Language, 28(2), (1989), 127-154.
30. Velez, M.C., Silver, D., & Tremaine, M. Understanding visualization through spatial ability differences. Proceedings of Visualization, (2005), 511-518.
31. Ziemkiewicz, C., et al. How Locus of Control Influences Compatibility with Visualization Style. Proc. IEEE VAST, (2011), 81-90.