Visual Perception of International Traffic Signs: Influence of e-Learning and Culture on Eye Movements

Gergely Rakoczi
Vienna University of Technology
Inst. for Design and Assessment of Technology
Vienna, Austria
+43 1 58801 41544
gergely.rakoczi@tuwien.ac.at

Andrew Duchowski
Clemson University
School of Computing
Clemson
South Carolina, USA
+1 864 6567677
duchowski@clemson.edu

Helena Casas-Tost
Universitat Autònoma de Barcelona
Traducció i Interpretació
Barcelona, Spain
+34 93 581 3382
helena.casas@uab.cat

Margit Pohl
Vienna University of Technology
Inst. for Design and Assessment of Technology
Vienna, Austria
+43 1 58801 18753
margit.pohl@tuwien.ac.at

ABSTRACT
Various eye movement metrics were recorded during the visual perception of international traffic signs embedded within an e-learning course designed to familiarize participants with foreign signage. Goals of the were to gauge differences in task types, sign origin, and ethnicity (American, Chinese, and Austrian) as well as effectiveness of the e-learning teaching materials in terms of prior preparation. Results, in contrast to other studies, suggest that teaching materials had no overall effect on either eye movement metrics nor on task success rates. Instead, sign origin had the strongest effect on gaze, as foreign signs in mixed presentation with domestic signs, elicited a larger number of fixations with longer mean fixation durations, highest regression rates, and lower performance scores. Possible effects of ethnicity were also noted: Americans showed lower mean fixation durations over the entire experiment, independent of test conditions, with Chinese participants fixating faster on (correct) road signs than the other ethnic groups.

Categories and Subject Descriptors
H.1.2 [**Computer Applications**]: Social and Behavioural Sciences – *psychology*.

General Terms
Visual Perception, e-Learning.

Keywords
Eye Tracking, e-Learning, Visual Behaviour, Learning Instructions, Visual Perception, Signage and Culture.

1. INTRODUCTION
Visual perception of traffic signs can be measured in two ways: while driving, e.g., on the road or in driving simulators, or when learning their meaning, e.g., with driver education software. The latter is commonly assisted by online teaching systems, where increasingly e-learning is used. This paper describes an experiment that focuses on the visual perception of traffic signs that are embedded in an e-learning course especially designed for drivers to learn to distinguish foreign and domestic signage.

Eye tracking research in the context of driving is almost exclusively concentrated on the investigation of attentional allocation when performing the act of driving. There is a good deal of eye tracking research supporting drivers during recognition of traffic signs in different driving (often simulated) environments [12] or investigating the conspicuity and capability of traffic sign recall when driving [8]. However, there is hardly any research on eye movement behaviour during the visual perception of traffic signs within a (theoretical) learning scenario, which is also crucial for driving, especially for novice drivers. This paper investigates an e-learning system designed for learning international traffic signs, e.g., prior to traveling abroad and/or studying for an international driving permit. It should also be pointed out that e-learning is becoming increasingly important for traffic safety courses and for obtaining a driver's license.

Beyond perception of traffic signs themselves, the second focus of this paper is related to the recording of eye movements when using the e-learning system. First, the influence of teaching materials displayed within the e-learning system prior to online tests is evaluated. There are numerous reports on the effect of task instructions on eye movement patterns. Examples include effects on reading [11] as well as during visual perception of scenes or websites, shifting people's visual attention to different regions [14]. Holmqvist, Nyström, Andersson, Dewhurst and Jarodzka report effects of shortening participants' attention levels, or producing changes in other eye movement metrics, which can be traced to task instructions [7]. In this paper, teaching materials are used to investigate two levels of learners' preparation. This assumption is closely related to research on learners' level of preparation. There are various reports of eye tracking studies investigating differences between novices or experts [1]. In this paper, we examine the effect of familiarity of domestic signs to the (presumable) unfamiliarity of foreign signs.

Finally, cultural aspects are also considered. Related work suggests cultural differences play a part in general scene perception [3]. Essentially, Easterners and Westerners use different visual strategies when perceiving fore- and back-ground objects of certain scenes. Americans, for example, tend to fixate faster and longer on objects in the foreground extracting more visual details than Chinese, whereas Easterners tend to move their eyes in a holistic manner over the screen. However, one limitation of this prior work was that tasks focused on free-viewing scene exploration, here, users' visual tasks are much more specific. Focusing on hemispheric dominance in terms of reading Chinese or English text, visual processing of English words shows left-hemispheric dominance as English words are processed in a sequential-analytic manner [6]. In contrast, Chinese visuals are processed holistically indicating dominance of the right-hemisphere, since hemispheric cortical areas are active during processing Chinese script [17]. In general the amount of information in Chinese characters is higher than in English words and therefore Chinese learn to pay more attention to the entire configuration

ETSA '13,Cape Town, South Africa, August 29–31, 2013.
Copyright 2013 ACM 978-1-4503-2110-5/00/0010….$15.00

of the characters when distinguishing these from others, while Westerners (readers of alphabetic languages) pay more attention to single letters (especially the initial letter) when recognizing words.

2. BACKGROUND

Prior work on the visual perception of traffic signs in an e-learning context is limited. Using a questionnaire, one study examined five cognitive design aspects of Chinese traffic signs such as familiarity, concreteness, simplicity, meaningfulness and semantic closeness [13]. Findings showed that the frequency of encounter (of the road signs' iconography) significantly influenced semantic closeness. Simply put, the more people see specific traffic signs the better they know their meanings. Another outcome of the study showed that the more representational the sign icons are (direct visualizations of objects) the better people could derive meaning and could more easily link the traffic signs to their function. However, within this study no eye movements were collected.

Another study related to perception of traffic signs tested their specific characteristics [18]. A complex model of visual attention was used to predict drivers' eye fixations. This study reported good results for letters and symbols depicted in road signs. However, the model could not sufficiently predict attention to complex traffic signs.

The main data gathering method of the present experiment is eye tracking, a methodology that records eye movements of participants when looking at a computer screen. Eye movements themselves are defined as changes in the direction of a participant's gaze falling on screen objects or Areas of Interest (AOIs). The theoretical foundation for using eye tracking data is based on the assumption that there is a relationship between one's eye movements and corresponding cognitive processes (e.g., visual attention) [10]. Eye movements may be regarded as valuable indicators of specific cognitive functions, e.g., attention, interest, etc., which may occur during learning [2]. Eye tracking in the investigation of e-learning is not common, although there are a few eye tracking examples in the literature [15]. General limitations of eye tracking results have also been reported [9]. As there are no accepted standards for interpretation of eye movement data, analysis has to be conducted carefully [7]. Accordingly, data analysis and interpretation of eye movements, especially when generalizing from results, requires expertise and experience [14].

3. DESCRIPTION OF THE STUDY

3.1 Research Goals

This paper evaluates specific eye tracking metrics (see Section 3.5) on the basis of three main research goals. The first goal of this paper is the *analysis of selected eye movements of learners when viewing traffic signs within an e-learning scenario*. The selection criteria for eye movement metrics are based on the fact that these parameters represent widely accepted and commonly applied metrics in eye-tracking research. Therefore these metrics benefit traceability and comparison with similar investigations in various fields of research. Furthermore, as indicated earlier there is no related work in this specific context and therefore this paper serves as a unique contribution to the field of e-learning and driving education. The second research goal is focusing on the *evaluation of two levels of learner's preparation*. Simply put the effect of absent or given teaching materials on eye movements is considered. Related eye tracking literature shows (see Sections 1 and 3.5) that there can be significant differences in visual perception for different levels of knowledge. However, in the field of computer supported learning there is no such investigation, therefore this paper tries to close this gap. Third, *cultural differences expressed by eye movements of various ethnic groups* are investigated. Because there are specific differences in scene perception that can be traced back directly to ethnicity (see Section 1), this paper evaluates if similar effects on eye movements can be measured during computer supported learning.

3.2 Procedure

The traffic sign experiment was designed as a single-user test scenario following a 3×2×2 mixed factorial design. The study's design mixes two between-subjects factors, namely ethnicity (at 3 levels: Austrian, Chinese and American participants) and teaching materials (at 2 levels: present or absent) with one within-subjects factor termed sign origin (at 2 levels: foreign or domestic traffic signs). The study design was blocked by tasks, of which there were three types (search, matching, and true/false tasks) each being homogeneous or mixed in terms of sign origin. To investigate differences within these test blocks, pairwise comparisons using t-tests with pooled SD and Bonferroni correction were obtained for each condition. The structure of this experiment is somewhat complex—Figure 1 provides an overview depicting all experimental conditions within a schematic diagram.

The investigator was a non-participating observer sitting behind participants during the entire test. The experiment was self-paced and participants had no time restrictions. To minimize

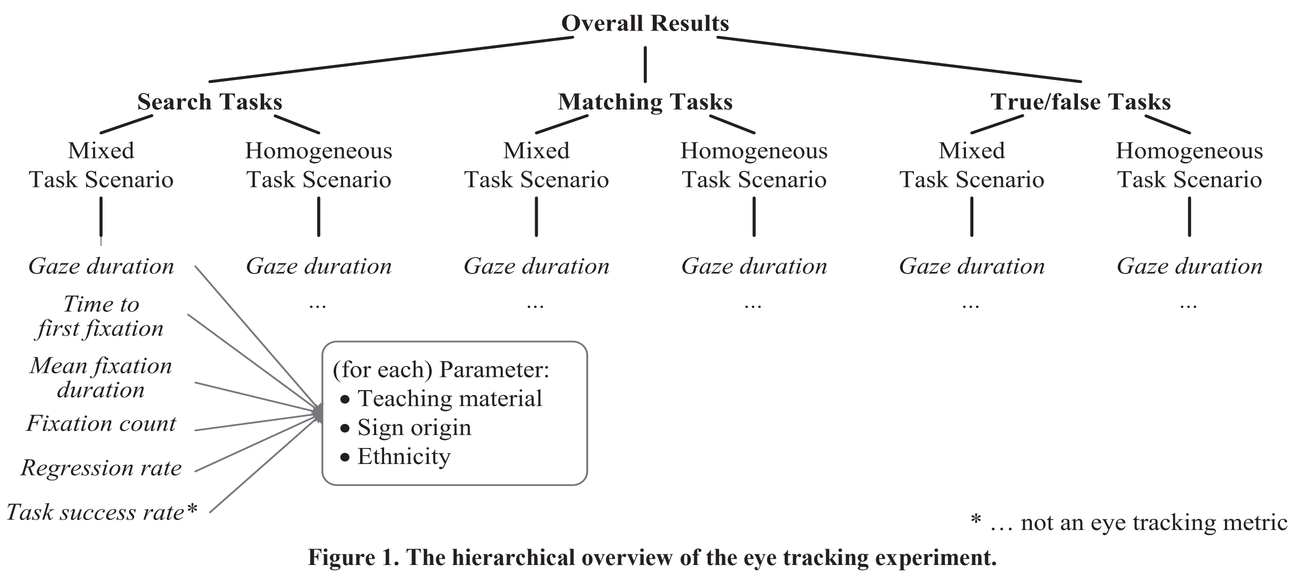


Figure 1. The hierarchical overview of the eye tracking experiment.

effects of intervening personal factors on work accuracy, participants were asked to solve the tasks as quickly as possible. Immediately after the eye tracking session a questionnaire about their driving experience, level of knowledge on traffic signs, and participants' personal opinions on the teaching lessons' in terms of complexity, difficulty, and usefulness was presented to the subjects. The questionnaire was designed with 7-point Likert scales. At the end of the experiment a retrospective interview was carried out, where users' experiences were jointly reflected with the investigator for about 30 minutes.

3.3 Stimulus

The stimuli of this study were international traffic signs from Austria, China and the United States of America embedded into an online course of TUWEL—the e-learning environment of Vienna University of Technology. It is based on Moodle (version 2.2.1), an open-source community-based e-learning and content management system. In order to minimize language barrier problems, the course language, as well as the e-learning environment (e.g., navigation menus and buttons), were translated to the official languages of the three countries. To be able to compare the eye tracking recordings among the different ethnic groups, the layout of the course as well as of its elements were displayed identically to all participants.

The landing page of the e-learning course (Figure 2) consisted of two topics. In topic 1 the links to the two teaching lessons were shown. Both teaching lessons related to the topic of the investigation served as online instruction for the oncoming tasks. The second teaching lesson (Figure 3) contained three paragraphs, one for each of the three countries. These paragraphs included short descriptions of the main differences of the traffic signs, pointing out differences in shape or colouring. Each paragraph was illustrated jointly by two visuals serving as representative traffic signs for the specific country. The intent was to educate learners about foreign signs which were assumed to be unfamiliar.

Figure 2. The landing page of the e-learning course

Figure 3. The second teaching lesson displayed to members of the treatment group

In topic 2 the links to the main tasks of the experiment were shown. According to Nielsen and Pernice [14], "tasks determine looks", prompting investigation of different e-learning task types to ascertain a more complete picture of learners' visual behaviour. The experiment investigated three different e-learning task types, namely search tasks, matching tasks, and true/false tasks (Figure 4a-c). The application of three different task types ensures that the investigation is valid on a broader basis in the context of learning. The reason for this is that in the practice of e-learning courses design mostly a mixture of different task types are applied. Therefore this investigation has to consider this important aspect as well. Within the search task participants had to find the corresponding traffic sign indicated within the task instructions. Within the matching task, users had to match different traffic signs to short textual descriptions. Within the true/false-task, users had to make a true/false decision about a textual expression describing the road signs.

To mitigate experimental learning and fatigue effects, the order of the tasks was counterbalanced via Latin square ordering and the questions within the tasks were displayed randomly. In total, participants had to complete 35 questions. One question was eliminated due to a design error. Users could freely decide the sequence to process the tasks, however only 2 out of 36 users did not follow the predefined order.

3.4 Apparatus

Data collection for this study was carried out in two laboratories. The data collection for the Austrian participants took place in the video lab at Vienna University of Technology in Austria (Institute of Software Technology and Interactive Systems, Interactive media systems group). The recording of the eye movements of American and Chinese participants (as depicted in Figure 5) was carried out in the eye tracking lab at Clemson University in South Carolina, USA (School of Computing).

(a) Search Task

(b) Matching Task

(c) True/false Task

Figure 4. Examples for the three e-learning task types.

Similar technological platforms were used at both locations, specifically Tobii eye trackers of the Tobii ET-1750 family. In Austria the Tobii X50 stand-alone tracker was used in combination with an IBM 20-inch TFT monitor, whereas in the USA the Tobii ET-1750 was used, where the eye tracker is integrated in a 17-inch TFT display. The resolution of the screen was set in both labs to 1280x1024 pixels. Both eye trackers are binocular and are equipped with an infrared camera to capture the coronal reflection of the eye at a sampling rate of 50 Hz. No chin rest was used in either lab. For identification of fixations from raw gaze data, the average of both left and right gaze points was used. Fixations as well as saccades were detected with filters defined by Tobii. The I-VT, or velocity-based filter was used [16], with velocity threshold set to 50 degrees, and duration threshold set to 100ms. For gaze analysis and statistical evaluation Tobii Studio (version 2.2.8), R (build 2.11.1), as well as Microsoft Excel (version 2007) were employed.

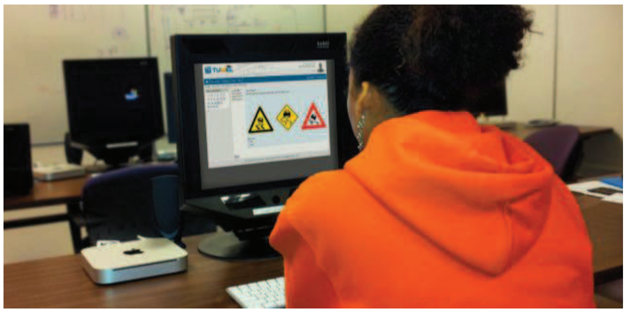


Figure 5. The eye tracker and the research setting at the American laboratory.

3.5 Eye Tracking Metrics

The following eye tracking metrics are used in order to track participants' visual perception of the given stimulus.

3.5.1 Gaze duration

Gaze duration, measured in seconds, stands for the total time between the first and the last fixation on any region within a task. Longer gaze durations indicate higher complexity of the visual stimulus [7]. Also, high cognitive effort during task solving may provoke longer gaze durations [14]. Therefore, for tasks depicting foreign signs and for members of the control group (no instruction given) longer total gaze duration is expected.

3.5.2 Time to first fixation

Time to first fixation is a latency measure also referred to as entry time in an AOI, measuring the time in seconds from the onset of stimulus until the AOI is first intersected by the viewer's gaze. In general, shorter values reflect higher localization efficiency which may be interpreted as influence on visual attention. That is, elements evoking shorter time to first fixations can be referred as more salient with higher visual trigger potential. In terms of computer interaction, users with less experience show longer entry times. Out-of-context objects may also attract early fixations. This metric is expected to be shorter for foreign signs (as they might be perceived as out-of-context) and for members of the treatment group fixating correct images earlier due to higher level of preparation [7].

3.5.3 Mean fixation duration

Mean fixation duration describes the average of fixation durations for participants on a specific AOI. Longer fixation durations suggest more effortful cognitive processing of the visual stimulus, i.e., longer fixation durations may indicate difficulty in information extraction [5]. Unexpected, out-of-context objects, less frequently encountered elements, or unclear/blurred parts of images may generate longer fixations. In contrast, there is evidence that non-demanding elements can push participants towards daydreaming, provoking longer fixations, suggesting fixation durations indicate shallow information processing or low levels of concentration [7]. Furthermore, stress, impatience, or higher levels of expertise may also shorten fixation durations. For this study, longer fixation durations are anticipated in the mixed traffic sign tasks. Members of the treatment group received initial teaching instruction, hence this prior exposure to unfamiliar signs would lead to the expectation of shorter mean fixation durations than those of the control group.

3.5.4 Fixation count

Fixation count indicates the total number of valid fixations within an AOI, ignoring fixation durations. This metric defines semantically informative areas within the stimulus. For task-solving processes the fixation count indicates regions where higher cognitive activity has occurred. Simply put, fixation count identifies areas of general importance [4]. However, for search tasks, fixation count is negatively correlated, as low values identify either task simplicity or high expertise. Also, in the context of search tasks, high fixation counts indicate difficulty in interpreting the stimulus [7]. For matching and true/false tasks, where foreign signs are involved, higher values were expected, whereas on domestic signs lower fixation counts were expected. For search tasks the opposite effect was expected. The influence of teaching materials should provoke faster task solving, with lower predicted fixation counts.

3.5.5 Regression rate

Regression rate describes the number of times participants re-visited previously fixated traffic signs. This metric has to be calculated relative to the number of fixations on traffic signs within each task. Eye movement revisits on AOIs usually indicate higher cognitive load related to deeper information extraction. Regression often suggests difficulty of visual processing of the object, or reveals more intense visual comparison identifying competing elements within the stimulus. Regressions may also be elicited in order to refresh short-term working memory [7]. For the present study the regression rate, calculated as the average number of re-fixations (all but the first fixation) over all AOIs in the stimulus, is expected to be lower for domestic signs. Furthermore, this metric is expected to be higher for the mixed task type scenarios and lower for members of the treatment group.

3.6 Participants

For this eye tracking study 36 participants were recruited including three ethnic groups representing three different continents, namely Europe (Austria), North America (USA) and Asia (China). There were 12 students from each country aged between 22 and 34 (M=26.3 years). The Austrian volunteers' average age was higher (M_{AT}=28.2 years) than those of Chinese and American participants (M_{CN}=25.2) and (M_{US}=25.4 years). To ensure gender balance for each country 6 males and 6 females were recruited. The average years of driving experience was 7.4 years. To be accepted to the experiment volunteers had to have a driving license (category A and/or B) or a driving permit for at least 2 years in their home countries. Participants were screened for any major visual deficiencies or any other remarkable constraints concerning their physical condition. Accordingly, two volunteers were excused, as the eye tracker could not be calibrated to them. As Chinese participants had to be recruited in the USA, they had to meet additional requirements. Only those Chinese volunteers participated in the study, whose residency did not exceed one year in the USA and had none (or minimal) driving experience on US roads. To keep ethnicity and gender balanced, 3 males and 3 females of each country were assigned to each of the control (no instruction) and treatment (instruction) groups.

4. RESULTS

Due to the experiment's complex structure and semantically different task types' outcomes, results are partially broken down into four parts. In Section 4.1 overall results are described. In Section 4.2 results of all task types are given, for both conditions (mixed and homogeneous scenarios).

4.1 Overall Results

For g*aze duration* no significant effects could be found of the influence of teaching materials ($F(1,34)=0.24$, $p=0.62$, n.s.) nor of ethnicity ($F(2,33)=0.11$, $p=0.89$, n.s.). In contrast, sign origin yielded significant differences ($F(1,71)=19.49$, $p<0.01$), as mixed representations of traffic signs and foreign road signs generated longer gaze durations.

Time to first fixation: No significant effects were observed of teaching materials ($F(1,34)=0.72$, $p=0.39$, n.s.) nor of sign origin ($F(1,71)=0.42$, $p=0.65$, n.s.). Interestingly, different results were detected for different ethnic groups ($F(2,33)=9.26$, $p<0.01$). Pairwise t-tests revealed Chinese participants fixated images twice as quickly as Austrian participants.(see Table 1).

Analysing *mean fixation duration* no differences were found for teaching materials ($F(1,34)=0.06$, $p=0.81$, n.s.). However, results revealed statistical tendencies for the ethnic groups ($F(2,33)=2.64$, $p=0.08$), as Americans' mean fixation durations were shorter than those of the Chinese or Austrians. The effect of sign origin was significant ($F(1,71)=7.23$, $p<0.01$), resulting in longer durations over foreign signs than over domestic road signs in the mixed scenario.

Focusing on *fixation count* neither teaching materials ($F(1,34)=0.26$, $p=0.60$, n.s.) nor ethnicity ($F(2,33)=0.46$, $p=0.63$, n.s.) showed significant effects. In contrast, sign origin

Table 3. Quantitative results for all eye tracking metrics in each conditions of the experiment.

	Gaze duration (time in sec)			**Task success rate** (in percentage %)			**Mean fixation duration** (time in sec)		
Overall	13.9			78			0.374		
Ethnicity (Austria /USA /China)	13.8	13.8	14.2	84	74	76	0.387	0.342	0.393
Sign origin (domestic /foreign)	9.3	11.1		85	78		0.369	0.385	
Teaching material (yes/no)	13.8	14.2		79	77		0.371	0.376	
Scenario (mixed /homogeneous)	13.6	16.9		72	83		0.369	0.378	
Search tasks (mixed)	9.8			72			0.367		
Search tasks (homogeneous)	9.5			89			0.360		
Matching tasks (mixed)	19.4			75			0.356		
Matching tasks (homogeneous)	29.3			95			0.364		
True/false tasks (mixed)	11.8			69			0.419		
True/false tasks (homogeneous)	11.9			65			0.409		
	Fixation count (integer)			**Regression rate** (integer)			**Time to first fixation** (time in sec)		
Overall	34.4			3.4			1.56		
Ethnicity (Austria /USA /China)	34.6	35.3	33.2	3.2	3.4	3.7	1.99	1.54	1.16
Sign origin (domestic /foreign)	22.7	26.3		2.9	3.4		1.65	1.61	
Teaching material (yes/no)	33.9	34.8		3.4	3.5		1.49	1.63	
Scenario (mixed /homogeneous)	41.6	33.4		3.7	3.3		1.62	1.58	
Search tasks (mixed)	24.4			2.8			1.49		
Search tasks (homogeneous)	24.1			2.6			1.45		
Matching tasks (mixed)	74.4			5.5			1.44		
Matching tasks (homogeneous)	49.4			3.2			1.44		
True/false tasks (mixed)	26			2.9			1.93		
True/false tasks (homogeneous)	26.8			4.1			1.87		

($F(1,71)=57.9$, $p<0.01$) revealed significant differences, as foreign signs in the mixed scenario elicited more fixations.

Regression rate: There was no significant effect of teaching materials ($F(1,34)=0.37$, $p=0.54$, n.s.) nor of ethnicity ($F(2,33)=1.1$, $p=0.34$, n.s.). However, strong statistical evidence was found for the effect of sign origin ($F(1,71)=64.8$, $p<0.01$), as both foreign signs and signs within the mixed scenario were re-visited more often.

For *task success rate* no differences were detected between the treatment and the control groups ($F(1,34)=0.38$, $p=0.55$, n.s.). However, a weak statistical effect of ethnicity was seen ($F(2,33)=4.6$, $p<0.05$). Austrians performed best, followed by the Chinese and Americans. Task scores were significantly lower when foreign signs were shown in the mixed scenario ($F(1,71)=28.7$, $p<0.01$).

4.2 Results for Search Tasks

Mixed Task Scenario. *Gaze duration* yielded no significant differences between presence of teaching materials ($F(1,34)=1.25$, $p=0.27$, n.s.). In contrast, participants' ethnicity revealed marginally significant differences ($F(2,33)=3.93$, $p<0.05$), as Austrians' gaze duration (M_{AT}=9.03s) was lower than those of the Chinese (M_{CN}=10.89s). For *time to first fixation* neither teaching materials ($F(1,34)=2.29$, $p=0.14$, n.s.) nor ethnicity ($F(1,34)=0.98$, $p=0.38$, n.s.) had statistical influence. For *mean fixation duration* there are no significant effects of ethnicity ($F(2,33)=2.15$, $p=0.13$, n.s.) nor of teaching materials ($F(1,34)=0.09$, $p=0.76$, n.s.). Similarly, *fixation counts* did not differ for teaching materials ($F(1,34)=1.14$, $p=0.29$, n.s.) or ethnic groups ($F(2,33)=1.82$, $p=0.18$, n.s.). The metric *regression rate* is not statistically significant for ethnicity ($F(1,34)=1.28$, $p=0.29$, n.s.) nor teaching materials ($F(1,34)=0.91$, $p=0.34$, n.s.). For *task success rate* no significance was detected either for teaching materials $F(1,34)=0.09$, $p=0.77$, n.s.) nor for ethnicity ($F(2,33)=0.85$, $p=0.43$, n.s.).

Homogeneous Task Scenario. For *gaze duration* teaching materials had no significant effect ($F(1,34)=0.25$, $p=0.62$, n.s.). In contrast, sign origin ($F(1,35)=30.71$, $p<0.01$) and ethnicity ($F(2,33)=5.55$, $p<0.01$) appeared to be significant. Results indicate that domestic signs and Austrian ethnicity elicited lower gaze durations during search. The results for *time to first fixation* revealed no significance for teaching materials ($F(1,34)=0.13$, $p=0.71$, n.s.) nor for sign origin ($F(1,35)=0.42$, $p=0.51$, n.s.). However, ethnicity seemed to lead to different outcomes ($F(2,33)=10.01$, $p<0.01$). Austrians (M_{AT}=1.955s) needed more than twice as long to fixate on the correct sign than both the Chinese and Americans (M_{US}=1.292s and M_{CN}=1.115s). The metric *mean fixation duration* was not significantly different for ethnicity ($F(2,33)=2.29$, $p=0.11$, n.s.) nor for teaching materials ($F(1,34)=0.01$, $p=0.89$, n.s.). However, ANOVA revealed a strong effect of sign origin ($F(1,35)=18.71$, $p<0.01$). Mean fixation durations were significantly higher when foreign traffic signs were fixated. The metric *fixation count* was marginally different given different ethnicity ($F(2,33)=3.98$, $p<0.05$), as Austrians exhibited fewer fixations for viewing traffic signs (M_{AT}=21.38) compared to both Chinese (M_{CN}=24.92) and Americans (M_{US}=25.87). In contrast, sign origin ($F(1,35)=20.80$, $p<0.01$) revealed high significance, with teaching materials imparting no significant effect ($F(1,34)=0.05$, $p=0.82$, n.s.). For *regression rate* there were no significant differences due to teaching materials ($F(1,34)=0.10$, $p=0.75$, n.s.), however there is a marginally significant effect of ethnicity ($F(2,33)=3.92$, $p<0.05$). Austrian participants produced fewer regressions on average (M_{AT}=2.394) than Americans (M_{US}=2.687) or the Chinese (M_{CN}=2.943). Strong statistical evidence was found for sign origin ($F(1,35)=21.61$, $p<0.01$), as perception of foreign traffic signs elicited higher regression rates. Focusing on *task success rate* there was no significant effect of teaching materials ($F(1,34)=1.24$, $p=0.27$, n.s.). ANOVA revealed marginally significant effects of sign origin ($F(1,35)=5.37$, $p<0.05$) and ethnicity ($F(2,33)=4.65$, $p<0.05$). To sum up, scores were generally higher for Austrians participants when viewing domestic road signs.

4.3 Results for Matching Tasks

Mixed Task Scenario. Focusing on *gaze duration* no significant differences were revealed for teaching materials ($F(1,34)=0.04$, $p=0.83$, n.s.) nor for ethnicity ($F(2,33)=0.38$, $p=0.68$, n.s.). For *time to first fixation* ANOVA showed no effect of teaching materials ($F(1,34)=0.46$, $p=0.50$, n.s.), however ethnicity indicated significance ($F(2,33)=11.42$, $p<0.01$), as again the Chinese (M_{CN}=0.808s) were twice as fast as the other ethnic groups (M_{AT}=2.175s and M_{US}=1.358s). ANOVA revealed a marginally significant effect of ethnicity on participants' *mean fixation durations* ($F(2,33)=4.71$, $p<0.05$), as Americans showed lowest mean fixation durations (M_{US}=0.319s). In contrast, teaching materials did not produce any effect ($F(1,34)=0.05$, $p=0.81$, n.s.). For *fixation count* there was no significant difference for the influence of teaching materials ($F(1,34)=0.03$, $p=0.86$, n.s.) nor for ethnicity ($F(2,33)=2.42$, $p=0.11$, n.s.). There were no significant differences for *regression rates* on teaching materials ($F(1,34)=0.13$, $p=0.71$, n.s.) nor for ethnicity ($F(2,33)=0.15$, $p=0.86$, n.s.). In contrast, teaching materials showed marginally significant differences for *task success rate* ($F(1,34)=6.90$, $p<0.05$), as those participants receiving prior preparation performed better (M_{no}=68% and M_{yes}=83%). Ethnicity also yielded a statistically significant effect ($F(2,33)=7.01$, $p<0.01$), as Austrians achieved highest scores (M_{AT}=85%), followed by the Chinese (M_{CN}=82%) and Americans (M_{US}=60%).

Homogeneous Task Scenario. Results revealed no statistical significance on *gaze duration* for teaching materials ($F(1,34)=1.83$, $p=0.18$, n.s.) nor for ethnicity ($F(2,33)=0.27$, $p=0.76$, n.s.). Focusing on *time to first fixation* teaching materials ($F(1,34)=0.03$, $p=0.85$, n.s.) revealed no differences. However, there was a marginally significant effect of ethnicity ($F(2,33)=5.27$, $p<0.05$), as the Chinese (M_{CN}=0.841s) were twice as fast to first fixations as the others (M_{AT}=2.084s and M_{US}=1.408s). In case of *mean fixation duration* both conditions—teaching materials ($F(1,34)=0.00$, $p=1.00$, n.s.) and ethnicity ($F(2,33)=2.24$, $p=0.23$, n.s.)—showed no significance. For *fixation count* there was no significant effect of teaching materials ($F(1,34)=1.64$, $p=0.21$, n.s.) nor of ethnicity ($F(2,33)=0.58$, $p=0.56$, n.s.). Similarly, analysis revealed no effects on *regression rate* of teaching materials ($F(1,34)=1.74$, $p=0.19$, n.s.) nor of ethnicity ($F(2,33)=2.76$, $p=0.079$, n.s.). Focusing on *task success rate* none of the factors showed significant differences with ANOVA producing non-significant results for ethnicity ($F(2,33)=1.66$, $p=0.21$, n.s.) and for teaching materials ($F(1,34)=0.23$, $p=0.63$, n.s.).

4.4 Results for True/false Tasks

Mixed Task Scenario. For *gaze duration* eye tracking revealed a highly significant effect of ethnicity ($F(2,69)=6.89$, $p<0.01$), as the Chinese (M_{CN}=9.696s) needed less time compared to Americans (M_{US}=13.95s). In contrast, ANOVA indicated no significant differences for teaching materials ($F(1,70)=0.01$, $p=0.94$, n.s.). The results for *time to first fixation* revealed no significance for teaching materials ($F(1,70)=0.13$, $p=0.71$, n.s.), however, there was strong influence of ethnicity ($F(2,69)=7.42$, $p<0.01$), as the Chinese (M_{CN}=0.942s) needed considerably less time for their first fixation on (correct) images than the other ethnic groups (M_{AT}=2.393s and M_{US}=2.485s). ANOVA yielded no effect of teaching materials on *mean fixation duration*

($F(1,70)=0.10$, $p=0.75$, n.s.), however marginal significance for ethnicity was seen ($F(2,69)=3.23$, $p<0.05$), as Americans showed the lowest mean fixation durations ($M_{US}=0.387$s). For *fixation count* there were no significant differences found for teaching materials ($F(1,70)=0.44$, $p=0.50$, n.s.), however a highly significant effect of ethnicity was observed ($F(2,69)=11.06$, $p<0.01$). The Chinese deposited fewer fixations ($M_{CN}=21.38$), than Americans ($M_{US}=31.79$). Similarly, there was also a marginally significant effect for *regression rate* among the ethnic groups ($F(2,69)=3.19$, $p<0.05$). Chinese participants produced the lowest regression rates ($M_{CN}=2.475$) compared to Americans and Austrians ($M_{AT}=3.121$ and $M_{US}=3.129$). Teaching materials had no effect ($F(1,70)=0.01$, $p=0.90$, n.s.). Finally, ANOVA revealed no significance on *task success* rate of teaching materials ($F(1,70)=1.06$, $p=0.30$, n.s.) nor of ethnicity ($F(2,69)=1.85$, $p=0.16$, n.s.).

Homogeneous Task Scenario. ANOVA showed no significant differences in *gaze duration* for teaching materials ($F(1,35)=0.07$, $p=0.79$, n.s.), sign origin ($F(1,35)=1.91$, $p=0.18$, n.s.) or ethnicity ($F(2,33)=0.16$, $p=0.84$, n.s.). Neither did it yield any effect on *time to first fixation* of teaching materials ($F(1,34)=0.08$, $p=0.56$, n.s.) or sign origin ($F(1,35)=0.10$, $p=0.74$, n.s.). Only the differences among ethnic groups indicated some marginal significance ($F(2,33)=4.07$, $p<0.05$), as Chinese participants' time to first fixations ($M_{CN}=1.423$s) was lower than those of the other groups ($M_{AT}=2.2$s and $M_{US}=2.018$s). For *mean fixation duration* neither teaching materials ($F(1,34)=0.24$, $p=0.62$, n.s.) nor sign origin ($F(1,35)=0.17$, $p=0.68$, n.s.) nor ethnicity ($F(2,33)=2.34$, $p=0.11$, n.s.) showed statistical effects. Results showed no significant effect on *fixation count* of ethnicity ($F(2,33)=0.17$, $p=0.84$, n.s.) nor of teaching materials ($F(1,34)=0.01$, $p=0.92$, n.s.). Only sign origin yielded some statistical tendencies ($F(1,35)=3.39$, $p=0.07$, n.s.), as counts on domestic traffic signs were lower compared to foreign road signs. Results for *regression rate* revealed no effect of teaching materials ($F(1,34)=0.11$, $p=0.73$, n.s.) nor of ethnicity ($F(2,33)=0.78$, $p=0.47$, n.s.). In contrast, sign origin revealed a marginally significant effect ($F(1,35)=4.33$, $p<0.05$), as domestic traffic signs were perceived with lower regression rates than foreign signs. Focusing on *task success rate* teaching materials ($F(1,34)=0.57$, $p=0.45$, n.s.) and sign origin ($F(1,35)=1.96$, $p=0.17$, n.s.) had no influence. Only ethnicity showed a highly significant effect ($F(2,33)=7.53$, $p<0.01$), as Austrians scored highest ($M_{AT}=75\%$), followed by the Chinese ($M_{CN}=69.4\%$) and Americans ($M_{US}=52.7\%$).

5. DISCUSSION

Results showed there was very little effect of the teaching material on participants' eye movements. No significant influence of prior preparation could be found for either task types nor for any of the eye tracking metrics. One reason for this may be that the quality of the teaching materials was low. However, questionnaire results did not support this. Teaching materials were perceived as simple, useful, and easy to understand. This opinion was shared among all participants. Eye movements on the teaching lessons revealed that the online instructions had been read intently by all ethnic groups: fixation clusters covered all areas, where important knowledge was described. Qualitatively, the proportion of fixations appeared higher on descriptions of foreign traffic signs suggesting that treatment group participants did not skip the instructions, and focused on all crucial parts of the text (Figure 6). Note, that text blocks describing foreign traffic signs were investigated more intensely than descriptions of domestic signs by all participants of all three ethnic groups. Another reason for lack of effect of teaching material may be that participants' level of knowledge of traffic signs prior to the study was already high, as the average driving experience was seven years. Perhaps no new knowledge was garnered from the teaching materials, rather, cognitive heuristics were used to solve the tasks. This assumption is supported by participants' feedback during the retrospective interview, when participants reported solving tasks by applying strategies of exclusion.

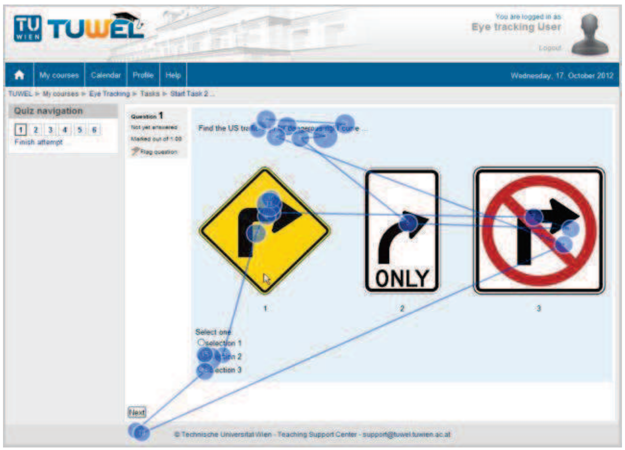
Figure 6. Textual descriptions of foreign traffic signs were perceived more intensely than those of the domestic signs.

Sign origin had a strong overall effect for almost all eye tracking metrics and all task types. The strongest overall statistical effect was measured for search tasks, the weakest for true/false tasks. In general, foreign signs elicited higher gaze durations, higher fixation counts, higher mean fixation durations, higher regression rates and lower performance scores than domestic signs.

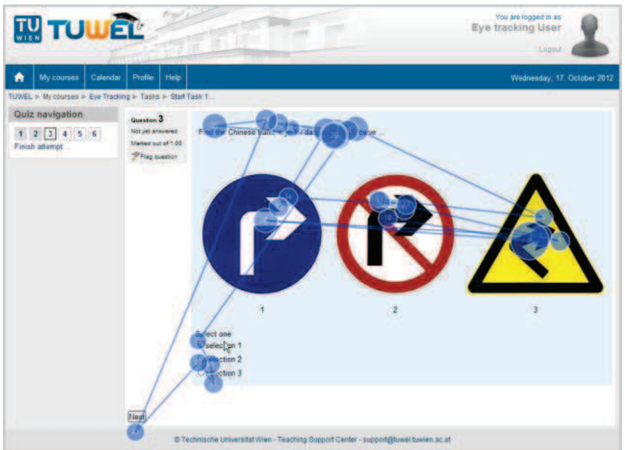
As an example, gaze plots depicted in Figure 7 illustrate higher regression rates on foreign traffic signs (see Fig.7b). Interestingly, time to first fixation is lower for foreign traffic signs. This may be explained by the thesis that out-of-context or unexpected visual elements attract visual attention faster. When focusing on mixed scenarios, where domestic and foreign signs were displayed simultaneously, gaze durations, mean fixation durations, and performance scores were lower than for the homogeneous task design. The other metrics seem to be negatively correlated as fixation counts, regression rates, and times to first fixation were higher within the mixed task scenario. These results indicate that participants were visually challenged most when fixating both familiar and unfamiliar objects at the same time.

When focusing on the interpretation of the eye tracking metrics in terms of task type, some interesting outcomes can be reported. Gaze duration for matching tasks was higher than for other task types. This is supported by participants' feedback indicating higher complexity of this task type. The reason for this may be related to the higher number of visual elements that had to be processed. This effect could also be seen for fixation counts, where matching tasks were perceived with higher numbers of fixations.

Mean fixation duration was about 370 milliseconds over all the eye tracking sessions indicating similar processing of visual stimulus among all conditions. However, there was one exception that could be linked to ethnicity. Statistical significance was found for different mean fixation durations. Americans' mean fixation durations were lowest for every task type and every scenario, as depicted in Figure 8. This appears to corroborate the finding that Americans try to fixate faster on foreground objects in order to retrieve more visual detail [3].



(a) Domestic traffic signs



(b) Foreign traffic signs

Figure 7. American participants had fewer regressions on domestic traffic signs (a) than on foreign – in this case Chinese - road signs (b).

However, as described in Section 3.5 lower mean fixation durations may also indicate impatience or higher level of expertise. Americans' highest fixation counts among all ethnic groups may suggest their willingness to spend more time on tasks negating the hypothesis of impatience. Since Austrians scored highest on performance and were the most experienced drivers, the impact of expertise can also be excluded.

Interestingly, regression rates were similar over the entire test. Exceptions included Austrians looking at foreign traffic signs, exhibiting low regression rates. True/false tasks (homogeneous scenario) as well as matching tasks (mixed scenario) provoked higher revisits. The latter had the highest regression rate of all. Results support the hypothesis of this eye tracking metric indicating deeper information extraction, since participants appeared to revisit signs more in this task to identify differences among competing domestic and foreign traffic signs.

The eye tracking metric time to first fixation yielded some interesting results, as Chinese participants fixated almost twice as fast on (correct) images than Austrians and one and a half times as fast as Americans (Table 1). This out-come might be related to cultural differences, as Chinese process Chinese scripts more holistically using left-hemispherical cortical areas. This might imply that Chinese users apply visual strategies in order to fixate faster on all elements of the entire test in order to understand the context faster, whereas Americans and Austrians apply visual skills to provoke more analytic processing by perceiving the visual configuration more sequentially. Results may also be interpreted as Chinese being more likely to be distracted by figurative visuals or preferring to devote more visual attention on images in general. Note that Chinese possessed similar computer literacy to other ethnic groups, excluding the supposition that novelty of the interface was a factor.

To sum up, recommendations for designing e-learning courses for driver education include the observation that matching tasks may be more visually demanding than true/false and search tasks. The more foreign the signs are and the greater the mixture of domestic and international signs, the more complex the learning scenario is likely to be. Teaching materials may not have any effect if learners were already familiar with the task. In the case of designing online education software for different ethnic groups, it should be taken into account that visual perception and specific eye movement characteristics may vary, e.g., mean fixation duration or time to first fixation.

Limitations of this study include 'Americanization' of Chinese participants, as the eye tracking study was conducted in the USA and not in China. For the study only those Chinese participants were accepted who resided in the USA no longer than 6 months. Still, questionnaire results reveal that even these recruits showed good knowledge of American traffic signs. However, Chinese participants' average driving experience (M_{CN}=3.7 years) was significantly lower than those of Austrians and Americans (M_{AT}=10.2 and M_{US}=8.3years).

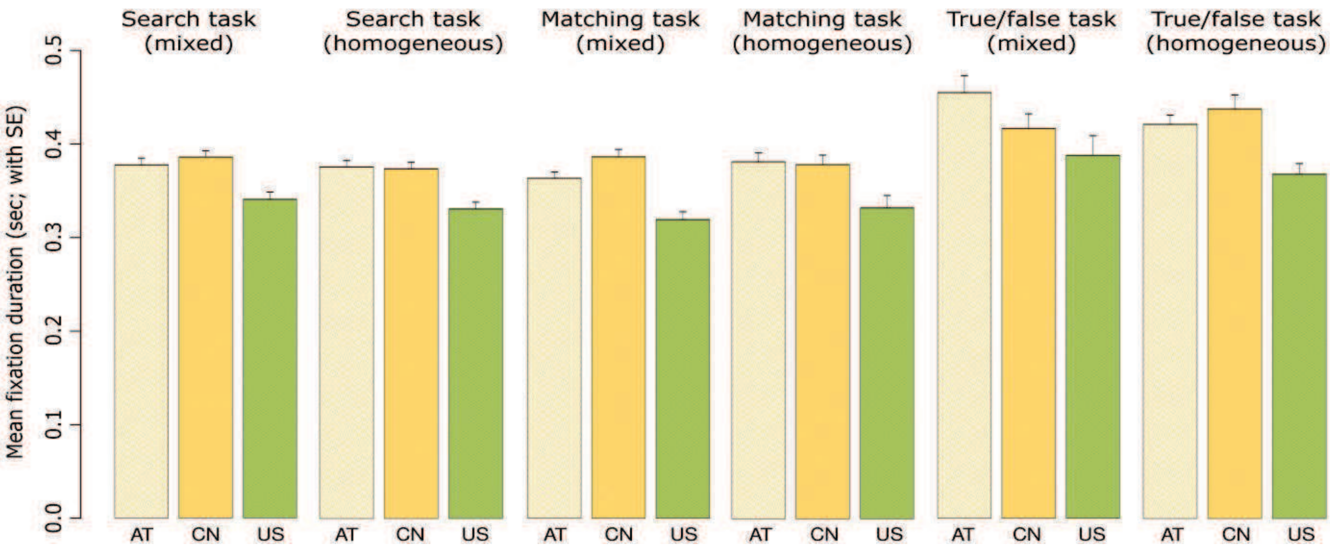


Figure 8. For all task types and for both conditions American participants had lowest mean fixation durations indicating faster information retrieval.

Participants' language preferences seemed to have a marginal impact on their overall eye movement distribution. Although the e-learning platform's layout was identical for all ethnic groups, the length of the text significantly varied in case of different languages thereby changing the lengths of instructions and those of the textual matching elements. Chinese text appeared to be the shortest, German text the longest.

6. CONCLUSION

Eye movements were captured during the visual perception of international traffic signs, embedded into an e-learning platform. The influence of different task types, sign origin, teaching materials, and ethnicity were analysed. In contrast to other studies, no statistically significant effect of prior preparation was found. Teaching materials seemed to have no overall effect on either the eye tracking metrics or on the task success rate. Sign origin produced the most significant effect on eye movements. Foreign signs and mixed presentation of domestic and foreign traffic signs provoked longer gaze durations, higher mean fixation durations, higher fixation rates, higher regression rates and lower scores. Other findings showed that there were two potentially significant cultural influences. The first was seen for Americans, who exhibited lower mean fixation durations overall, independent of test conditions. The second was observed for the Chinese who fixated fastest on (correct) road signs. However, more research is needed to clarify the reasons for these cultural differences. The investigation also yielded some interesting results in terms of learning design as well as usability improvements of the e-learning environment. These may help improve the e-learning platform's user interface and to design more visually effective e-learning tasks. Finally, our observations may contribute to better perception of traffic signs in driving education systems, which may benefit higher awareness and safety in driving.

7. REFERENCES

[1] Amadieu, F., Van Gog, T., Paas, F., Tricot, A., Mariné, C. 2009. Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction* 19, 5, 376-386.

[2] Bente, G. 2005. Erfassung und Analyse des Blickverhaltens. In *Lehrbuch Medienpsychologie*, Bente, G., Mangold, R., Vorderer, P., Eds. Hogrefe, Vienna, 297-324.

[3] Chua, H.F., Boland, J.E., Nisbett, R.E. 2005. Cultural variation in eye movements during scene perception. In *Proceedings of the National Academy of Sciences of the United States of America* 102, 35, 12629-12633.

[4] Ehmke, C., Wilson, S. 2007. Identifying web usability problems from eye-tracking data. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI*. British Computer Society, Swinton, 1, 119-128.

[5] Goldberg, J.H., Kotval, X.P. 1999. Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics* 24, 6, 631-645.

[6] Hellige, J.B., Cox, P.J. 1976. Effects of concurrent verbal memory on recognition of stimuli from the left and right visual fields. *Journal of Experimental Psychology: Human Perception and Performance* 2, 210-221.

[7] Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. 2011. *Eye Tracking: A Comprehensive Guide to Methods and Measures.* Oxford University Press, Oxford.

[8] Inman, V.W. 2012. Conspicuity of traffic signs assessed by eye tracking and immediate recall. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 56, 2251-2255.

[9] Jacob, R.J.K., Karn, K.S. 2003. Eye Tracking in Human Computer Interaction and Usability Research: Ready to Deliver the Promises. In *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, Hyönä, J., Radach, R., Deubel, H., Eds. Elsevier Science, Amsterdam, 573-605.

[10] Just, M.A., Carpenter, P.A. 1976. Eye fixations and cognitive processes. *Cognitive Psychology* 8, 441-480

[11] Kaakinen, J.K., Hyönä, J. 2010. Task effects on eye movements during reading. *Journal of experimental psychology. Learning, memory, and cognition* 36, 6, 1561-6.

[12] Ko, M., Higgins, L., Chrysler, S.T., Lord, D. 2010. *Effect of driving environment on drivers' eye movements: Reanalyzing previously collected eye-tracker data.* Technical Report. 89th Annual Meeting of Transportation Research Board.

[13] Ng, A.W.Y., Chan, A.H.S. 2007. Cognitive design features on traffic signs. *Engineering Letters* 14, 13-18.

[14] Nielsen, J., Pernice, K. 2010. *Eyetracking Web Usability. Voices That Matter.* New Riders, Berkeley.

[15] Rakoczi, G. 2012. Eye Tracking in Forschung und Lehre: Möglichkeiten und Grenzen eines vielversprechenden Erkenntnismittels. In *Digitale Medien*, Reichl, F., Csanyi, G., Steiner, A., Eds. Waxmann Verlag, Münster, 87-98.

[16] Salvucci, D.D., Goldberg, J.H. 2000. Identifying Fixations and Saccades in Eye-Tracking Protocols. In *Eye Tracking Research & Applications (ETRA) Symposium*. ACM, Palm Beach Gardens, 71-78.

[17] Tan, L.H., Liu, H.L., Perfetti, C.A., Spinks, J.A., Fox, P.T., Gao, J.H. 2001. The neural system underlying Chinese logograph reading. *NeuroImage* 13, 836-846.

[18] Yan, H., Guan, H., Xu, Z. 2011. Incorporating Saliency Map into Prediction of Drivers' Eye Fixations on Traffic Signs. In *11th International Conference of Chinese Transportation Professionals*. American Society of Civil Engineers, Reston, 1627-1639.