Gaze-Controlled Gaming
Immersive and Difficult but not Cognitively Overloading

Abstract
A user study is described focusing on the cognitive and usability consequences of cueing visual attention in gaze-controlled gaming. Results show that such cueing influences performance and affects the subjective gaming experience. Accordingly, visual cues make the experience less enjoyable and less immersive. Interestingly, gaze-controlled gaming appears not to require additional cognitive effort. Taking into account previous findings, results suggest that applicability of overt cueing may depend on the game type. Our study shows the potential for optimization of interaction of gaze-controlled arcade games with visual cueing.

Author Keywords
gaze-controlled gaming, visual attention cueing, cognitive processes interference, gaming experience, cognitive load

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

Introduction and Background
Interest in gaze-controlled gaming is resurfacing with recent developments of eye tracking technology. Although the idea of controlling games solely through eye movement offers a potentially rich gaming experience, it also raises basic issues concerning players’ cognitive processes and performance. It
is especially important to consider perceived experience, cognitive function, and behavioral performance in gaze-controlled games in comparison with more traditional input devices such as the joystick, keyboard, or mouse.

Comparison of game interfaces controlled either with gaze input or a traditional computer mouse, has shown that gaze-gaming players outperform the others on problem-solving measures, commit fewer errors, are more immersed, and have a better user experience [?]. Moreover, research has already documented gaze as an input modality superior to the mouse in classic, well-known games such as ‘Breakout’ [?].

For large targets, performance of video game tasks, such as target acquisition and target tracking, is similar with gaze input as it is with the mouse, touch screen, and head tracking, but better with gaze input than with the joystick. Gaze input in game interaction, combined with a well-designed interface, has considerable potential [?].

Existing research suggests that players respond positively to eye tracking, although they have very different opinions about its appropriateness. Preference for either mouse or eye control remains largely inconclusive [?].

Review of research on visual cueing in the field of multimedia learning and driving safety shows that cueing is an effective method of directing attention [?] and optimizing information processing [?]. However, visual attention studies show that while cues can be effective for static graphics, they may lose their potency when used with animations [?].

The novelty of the present approach to gaze-controlled gaming is based on attentional cueing techniques used during gaming gaze control. Our use of overt cues is similar to McNamara et al.’s subtle gaze direction [?, ?] without the subtlety. That is, the cues are always visible and always present and do not disappear upon detection of gaze direction. We test the impact of these visual cues and hypothesize that gaze-controlled gaming with overt cues increases performance by lowering cognitive effort while increasing the subjective gaming experience.

Methodology
Participants. Twelve subjects (6 male and 6 female, aged $M = 30.5, SD = 4.06$) took part in the experiment voluntarily after signing a consent form. All participants reported they used the internet every day but they did not have any experience with gaze-controlled games.

Experimental design: dependent and independent variables. The experimental design was fully within-subjects with two fixed factors: game-control type (gaze-controlled with cues vs. gaze-controlled without cues vs. keyboard-controlled) and maze complexity (easy vs. hard). Dependent variables were:

- **gaming performance**: indicated by game completion time and success rate;
- **cognitive load**: indicated by pupil dilation, fixation duration and blink count;
- **visual attention distribution**: indicated by percentage of gaze on paths and the rest of the maze and number of saccades; and
- **gaming experience**: indicated by a questionnaire.

Procedure and research materials. During the experiment, participants played a simple arcade game. The goal of the game was to guide the game character through the maze. Each maze, consisting of 25 columns and 15 rows of square tiles, was built in such a way that the maze started in one corner of the screen and ended in the opposite corner, so that the player had to traverse each quadrant of the screen to finish the game. Each quadrant of the screen contained
a path segment of approximately equal length. Code for the game used in the study was written in Python and Pygame. Maze floor tiles and player images were free graphics from Daniel Cook.

The easy maze contained a simple path with no branching. There were 119 floor tiles from start to finish and the shortest path was 119 tiles long. The hard maze featured a non-trivial path, with 167 floor tiles in the whole maze, and with shortest path of 103 tiles long (see Figure 2).

There were three modes of game interaction:

1. **Gaze-controlled with no visual cues**: the player controlled the character by looking at the screen. Gaze position data was sent to a function that checked if the player’s gaze was above, below, to the right or to the left of the player character displayed on the screen and then moved the character by one tile in the corresponding direction if that was a valid move (not blocked by wall or end of screen) and waited 0.1 seconds for the next point of data.

2. **Gaze-controlled with visual cues**: same as the eye tracker version described above, but with visual cues in the form of arrows appearing on the screen pointing in the directions of possible moves.

3. **Keyboard-controlled**: the player controlled the game character by pressing cursor keys.

Each participant played all three versions of the game. The order of playing was counterbalanced and each version of the game was played consecutively twice, with easy and hard versions of the maze presented in random order. After playing each of the two games (easy and hard) with a particular control mode, participants rated the gaming experience.

A Gaming Experience Questionnaire was devised, focusing on evaluation of interaction with the game and consisted of four items: enjoyment, naturalness, difficulty and engagement, each evaluated on a 9-point Likert type scale. The lowest value of the scale was definitely disagree and highest was definitely agree. The questionnaire was adopted from Bednarik et al.

**Apparatus.** The games were presented on a 22-inch LCD computer monitor with 1680 × 1050 resolution and 60 Hz refresh rate, connected to a laptop. For standard game control, a PC keyboard was used. Eye movements were recorded at 120 Hz with an SMI RED 250 eye tracking system.

For better gaze control, participants were asked to keep their chin and forehead on a chin-rest (see Figure 1). SMI’s BeGaze v2 software was used for fixation and saccade detection with a dispersion-based event detection algorithm. The dispersion was 100 px, with minimum fixation duration set to 100 ms.

**Results**
Analyses of the eye movement data focused on classic eye movement indices and subjective game experience evaluation. Analyses were computed with R statistical software using mainly linear mixed models (LMMs) fit with repeated data. Fixed within-subject factors included game type (controlled with keyboard vs. cued gaze-controlled vs. gaze-controlled with no cues) and game complexity (easy vs. complex). Participant ID and trial order were treated as random factors. Effects of LMMs were calculated with analysis of variance (ANOVA) with type III sum of squares correction.

**Performance.** All participants were able to complete the game (steer the game character to the end of the maze). However, participants differed in time to completion.
Linear mixed model analysis was used to compare differences in completion times. Results revealed a significant main effect of game type, $F(2, 11) = 9.91, p < 0.01$, see Figure 3. Contrast analyses showed a statistically significant difference between cued gaze-controlled ($M = 45.32\, s$, SE = 23.37) and keyboard-controlled ($M = 33.43\, s$, SE = 24.59) modes, $t(11) = 3.02, p < 0.02$. The difference between the gaze-controlled game with no cues and keyboard was not significant ($p > 0.1$). Additional contrast analysis showed a significant difference between both gaze-controlled game types compared with keyboard control, $t(11) = 3.79, p < 0.01$.

Analysis also showed that neither main effect of game complexity ($F(1, 16) = 1.20, p > 0.1$) nor the interaction term ($F(2, 16) = 1.82, p > 0.1$) was statistically significant.

Eye movements analysis. A Chi-squared test on cross table was used to analyze visual attention distribution between paths and the rest of the maze in different game types. In the analysis we crossed frequencies of the binomial variable (gaze on path vs. gaze on other parts of maze) with game type (keyboard controlled vs. gaze-controlled with cues vs. gaze-controlled with no cues). Contrary to the hypothesis that participants playing the game with additional visual cues would spend more time gazing on paths, the Chi-squared test was not statistically significant, $\chi^2(2) = 0.23, p > 0.1$. In all conditions participants spent about 60% of time gazing on paths.

Analysis of saccade count revealed a statistically significant main effect of game type, $F(2, 11) = 11.53, p < 0.01$. Contrast analysis showed that the cued gaze-controlled game evoked significantly more saccades ($M = 128.00, SE = 8.15$) than the keyboard-controlled game ($M = 93.43, SE = 5.12$), $t(11) = 3.47, p < 0.01$. The difference between the gaze-controlled game with no cues ($M = 113.09, SE = 8.58$) and keyboard version was not significant ($p > 0.1$). This finding should not be surprising: with greater time to completion (see above), participants had more time to issue more eye saccades. The analysis also showed a main effect of game complexity, $F(1, 16) = 12.29, p < 0.01$, meaning that participants exhibited more saccades with the complex maze ($M = 122.80, SE = 8.52$) compared to the easy one ($M = 100.80, SE = 6.01$). However, the ANOVA interaction term was not significant, $F(2, 16) = 2.76, p = 0.094$, and contrasts analysis revealed that the difference between cued gaze-controlled and keyboard-controlled games for the complex maze was only at a statistical tendency level, $t(11) = 1.77, p = 0.09$, see Figure 4. No other contrasts were significant ($p > 0.01$).
Contrary to the hypothesis, no significant effects on eye movement indicators of cognitive load were found, including pupil dilation, $F(2, 16) = 1.16, p > 0.1$, average fixation time, $F(2, 16) < 1$, or blink count, $F(2, 16) = 1.11, p > 0.1$. This suggests no discernible effect of gaze-controlled or traditional keyboard game control on cognitive load.

Interpreting the observed number of saccades above, leads to two possible conflicting hypotheses regarding the distribution of visual attention during cued gaze-controlled game play:

1. first, participants were visually exploring the maze in order to find the optimal path, suggesting more exploratory eye movements with relatively larger deviations from the optimal maze path, or,

2. second, participants performed more local saccades moving their gaze back and forth from the game character to the arrow cues, suggesting relatively smaller deviations from the optimal maze path.

To test these hypotheses, we calculated the mean distance between the optimal path for each maze and the observed scanpaths. Each maze was divided into $6 \times 4$ AOI grids and the cumulative scanpath on such a grid was calculated for each participant. We then compared these scanpaths with the optimal path using the standardized Levenshtein distance $[?]$. The outcome vector was treated as a dependent variable ranging from 0 (no difference between individual scanpath and optimal path) to 1 (maximum difference between individual scanpath and optimal path).

Linear Mixed Models analysis on standardized Levenshtein distance revealed a statistically significant main effect of game version, $F(2, 16) = 4.75, p < 0.05$. Contrast comparisons show that distance in the gaze-controlled cued version was smaller ($M = 0.12, SE = 0.01$) than in the keyboard controlled game ($M = 0.16, SE = 0.01$). Moreover, the interaction effect between game version and complexity was statistically significant, $F(1, 23) = 3.86, p < 0.05$, see Figure 5. For complex games, participants' scanpaths were more similar to the optimal path in the cued gaze-controlled game ($M = 0.13, SE = 0.01$) than in the keyboard controlled game ($M = 0.19, SE = 0.02$).

This result supports the claim that participants in the cued gaze-controlled game performed more saccades but were forced by the mechanics of the game to keep their eyes on the path. It also suggests that this style of gameplay imposes greater requirements on the player and presumably influences their subjective gaming experience.
Figure 5: Interaction of game version and complexity on normalized Levenshtein distance of cumulative scanpaths from optimal maze path. Whiskers represent boundaries of confidence interval for $p < 0.05$.

**Game experience evaluation.** Four questionnaire items were given concerning interaction (enjoyment, immersion, difficulty, and naturalness) and were compared between game types (keyboard control, gaze control with cues and gaze control with no cues). LMM analysis was used.

Analyses revealed statistically significant effects of game type in terms of naturalness, $F(2, 22) = 12.49, p < 0.001$, difficulty, $F(2, 22) = 19.94, p < 0.001$, immersion, $F(2, 22) = 11.35, p < 0.001$, and enjoyment, with the latter at a statistical tendency level $F(2, 22) = 3.43, p = 0.051$, see Figure 6.

Figure 6: Effect of game version on subjective gaming experience. Whiskers represent boundaries of confidence interval for $p < 0.05$.

Contrast comparisons showed that players found the gaze-controlled interaction with no cues more difficult, $t(22) = 3.73, p < 0.01$, but at the same time more immersive than keyboard control, $t(22) = 3.85, p < 0.001$. On the other hand, gaze-controlled interaction with visual cues was evaluated negatively, as less enjoyable, $t(22) = 2.51, p < 0.05$, less natural, $t(22) = 3.15, p < 0.01$, and more difficult, $t(22) = 2.25, p < 0.05$ compared with keyboard control.

**Discussion and Conclusions**
Contrary to our hypotheses we did not observe an impact of gaze control on users’ cognitive effort. However, gaze-controlled gaming negatively influenced gaming performance and experience. Maze completion times in both gaze-controlled
conditions were longer than with keyboard control. A negative impact on performance was observed especially for the visually cued gaze-controlled condition.

An earlier study of subjective game experience showed that gaze-controlled games are more entertaining and engaging despite being more difficult [?]. Our results do not support this finding: gaze control (especially with cues) made the experience less enjoyable and less immersive. As suggested by others, perhaps not every game genre is suited for gaze control [?]. Our game was possibly too easy to provide sufficient entertainment derived from mastering the game’s difficulty. While natural eye movements are largely executed without conscious control, gaze-controlled gaming may impinge on cognitive processes related to the game itself, e.g., decision making. The lack of impact of game type on eye movement indicators of cognitive load suggests the game may have been not difficult enough to aggravate working memory resources.

Another possible explanation of the lowered game experience for the gaze-controlled games comes from the analysis of the saccade count and distances of scanpaths from the optimal maze path. These analyses revealed that the keyboard controlled version of the game was the only one in which players were scanning the visual field. Gaze-controlled games prevent such scanning since every gaze movement triggers player movement, e.g., an instance of the classical Midas Touch problem [?]. For a maze game, the ability to scan the visual field is an important factor influencing gameplay.

The above observation leads us to conclude that for successful overt gaze cueing, it would be beneficial to implement a mechanism which allows switching between two modes: gaze-controlled gaming and visual field scanning, similar in spirit to Snap-Clutch [?]. Figure 7 shows a mockup of the “visual four directional D-pad”, which we dub the V-pad—a system in which the movement of the game character is controlled only when gaze is within a given radius of the character’s position. When gaze falls outside this radius, the game switches to visual scanning mode.

Existing research on the topic of visuo-spatial cueing suggest that it may lack effectiveness [?] and that using arrows as external indicators may be sufficiently imprecise to effectively direct attention [?, ?, ?]. In our opinion, the implementation of the dual mode mechanism for gaze-controlled gaming will improve both the effectiveness of attention cuing as well as the subjective gaming experience.

Figure 7: V-Pad: a mockup of cued gaze-controlled game interaction that allows switching between gaze control and visual scanning.