

Evaluating Real-Time Gaze Representations to Infer Intentions in Competitive Turn-Based Strategy Games

Joshua Newn¹, Eduardo Velloso¹, Fraser Allison¹, Yomna Abdelrahman², Frank Vetere¹

¹Microsoft Research Centre for Social NUI, University of Melbourne, Australia

²Institute for Visualization and Interactive Systems, University of Stuttgart, Germany

¹{newnj,velloso,fallison,f.vetere}@unimelb.edu.au, ²yomna.abdelrahman@vis.uni-stuttgart.de

ABSTRACT

In this paper, we investigate nine different visual representations of gaze in a competitive digital game setting. We evaluate the ability of spectators to infer a player's intentions in the game for each visual representation. Our results show that spectators have a remarkable ability to infer intent accurately using all nine visualizations, but that visualizations with certain characteristics were more comprehensible and more readily revealed the player's intent. The real-time *Heatmap* visualization was the most highly preferred by participants and the most effective in revealing intent, due to its ability to balance real-time gaze information with a persistent summary of recent gaze behaviour. Our findings show that eye-tracking visualization can enable playful interactions in competitive games based on players' ability to interpret opponents' attention and intention through gaze information.

ACM Classification Keywords

H.5.2. Information Interfaces and Presentation: User Interfaces: Input devices and strategies

Author Keywords

Eye tracking; Gaze; Nonverbal leakage; Shared gaze; Gaze awareness; Competitive Gaming; Intent Prediction

INTRODUCTION

Players' nonverbal signals are an important element in many competitive board and card games. By monitoring an opponent's body language, facial expressions and gaze direction, a player can infer their next moves or call their bluffs. Gaze is a strong signal of a person's area of attention and interest [9], and can give away unintended information—what psychology researchers call 'nonverbal leakage' [8]. Gaze is equally important in collaborative settings, as a cue for collaborators to monitor each other's attention and coordinate their actions [14]. Despite its importance in co-located settings, shared gaze awareness is typically absent in remote multiplayer games. In

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

CHI PLAY '17, October 15-18, 2017, Amsterdam, Netherlands

© 2017 ACM. ISBN 978-1-4503-4898-0/17/10...\$15.00

DOI: <https://doi.org/10.1145/3116595.3116624>



Figure 1: An opponent's gaze presented as a real-time Heatmap over *Ticket to Ride* game showing interest in the east coast of the North America. This is an animated figure and it is best viewed in Adobe Reader.

this paper, we demonstrate an opportunity to bring shared gaze awareness into this setting by comparing nine different ways of visualising gaze within a digital multiplayer version of a popular board game—*Ticket to Ride*¹.

The advantages in observing a partner's gaze in a *collaborative* task are clear, leading researchers to build multiple systems where users' gaze points are streamed over a network to improve coordination [3]. However, the opportunities for gaze in *competitive* settings are more nuanced. Recent studies have explored these possibilities by using visualizations of players' gaze to open up novel gameplay experiences and increase social presence [16, 20]. Normally in remote multiplayer games, a player's only clues for predicting their opponents' current and future plans are the in-game actions they have previously taken; players do not have access to the bodily cues (such as gaze) that enrich the experience of games played around a table or in a shared space. When given information about an opponent's gaze in a networked game, a player can build hypotheses about what their opponent is thinking through observation of the areas or game elements that receive the most visual attention. This creates further opportunities for a player to deliberately manipulate their own gaze behaviour to deceive an opponent about their plans and goals. This contrast between the tendency for observed gaze to reveal intent and the ability of a player to use their eye movements deceptively makes shared gaze awareness a rich resource for game design.

¹<https://www.daysofwonder.com/online/en/t2r/>

However, studies on shared gaze visualizations for various applications have found that the introduction of shared gaze awareness can be *disruptive*, often being referred to as ‘distracting’ and ‘confusing’ [6, 16]. We hypothesise that this is caused by two primary problems with gaze visualization for digital interfaces. First, the way in which gaze has been visually represented in past works may be inadequate or ill-suited for the task and setting. This is expected, as the design of visual representations in graphical user interfaces is highly dependent on their context of use. This problem is exacerbated by an inherent problem with gaze: gaze is ‘always on’, which makes designing gaze representations difficult as the technology cannot distinguish between focused scrutiny and more absent-minded looking.

Second, humans are not accustomed to interpreting visual representations of gaze, as the focal point of gaze is ‘invisible’ in normal interpersonal interaction. Gaze information can be ‘noisy’ due to the jittery movements and the way the eyes work (fixation-saccade-fixation cycle), and this is reflected in the representations. Gaze data is difficult to interpret and often misunderstood [6], as even trained professionals in eye-tracking evaluations make use of a range of gaze visualizations to make sense of the same collected data. Further, as gaze is fast-moving and never entirely still, any real-time visualization overlaid on a user interface adds a highly distracting element that may interfere with the viewer’s own thought process.

To effectively use gaze awareness in a game, the designer must understand how the choice of gaze visualization will affect players’ understanding and enjoyment. Inspired by how players observe the gaze of one another in tabletop boardgames, we selected a popular board game—*Ticket to Ride*—for the purposes of our research. Using recorded clips from eye-tracked gameplay, we assessed nine gaze representations in three different stages of the game (early, mid, late). This paper serves as an initial novel groundwork for determining the characteristics of gaze representations that allow human players to make inferences about other players’ intentions and into derivation of potential insights into their strategies (goals and plans).

RELATED WORK

In this section, we summarise related works that require a user to interpret gaze representations of another, which falls into two key areas: research on *shared gaze*, and works that compare and interpret the use of gaze representations. We have chosen to exclude works related to eye-tracking evaluations in which gaze data is interpreted by trained professionals, as our interest is specifically on the ability of non-specialist users to interpret visual gaze representations.

Shared Gaze

Shared gaze systems are those that give users the ability to see the gaze of a partner, typically to improve communication and coordination between pairs in collaborative problem-solving tasks. Here, the eyes are typically not directly visible, and the addition of a gaze visualization provides a complementary layer of nonverbal communication. Knowing where someone is looking can provide rich information, from inference about their intentions to clues about their current cognitive activity (based on behaviours such as scanning, focused

interest in an object, and repeated comparison of different objects) [24]. With this knowledge, much research in HCI (Human-Computer Interaction) has used gaze in collaborative settings to foster coordination. For instance, Stein and Brennan [24] found an improvement in problem-solving task performance when another person’s gaze was used as a cue. Gaze has been shown to be effective in co-located settings, e.g. when multiple users perform a visual search task on a large display [29]. A common approach is to present a user’s gaze as a cursor overlaid on the shared visual display space, to create an awareness of that user’s focus area [5, 24].

Gaze has shown promise when used implicitly with other modalities in remote communication. Qvarfordt and Zhai applied gaze in a dialogue system, where a remote assistant is able to detect the interest of a remote user through their gaze patterns while conversing with them (remote user is unaware that their gaze is being tracked) [22]. Tracking gaze behaviour can serve as a natural deictic pointer and thus reveal interest in a natural way. Brennan et al. have demonstrated that shared gaze can be more efficient than speech for the rapid communication of spatial information (spatial referencing) [3].

Despite the numerous benefits mentioned, gaze as an input suffers from inherent problems is that gaze is ‘always on’, related to the ‘Midas Touch’ problem that leads to accidental triggering of gaze-enabled interface widgets [12]. To combat this, gaze is often paired with another modality, such as voice interaction; however, this can also lead to confusion when there is a mismatch between gaze and voice inputs [6]. However, in competitive settings, this ambiguity can work as a resource for game design, revealing a glimpse into players’ intentions without giving it away entirely.

Gaze Representation

How gaze information is represented strongly influences gaze interaction [4]. We reviewed two recent works in HCI that compare and contrast gaze representations to gather design implications for our study. Zhang et al. evaluate gaze representations for use in co-located collaboration, where two users share a screen and two gaze points are represented [29]. From a review of prior works, the authors selected four different gaze visualization formats for testing—*Cursor*, *Trajectory*, *Highlight* and *Spotlight*—of which participants preferred highlight and spotlight. The authors determined that attaining a balance between *visibility* and *distraction* was one of the biggest challenges when designing and selecting gaze visualizations.

In another work, Li et al. designed and evaluated two gaze representations aimed at improving the coordination between pairs of users [17]: *zoom focus* and *gaze trail*. Both representations used fixations processed from the gaze stream; the former aimed to show the precise location of a fixation, while the latter aimed to show the most recent gaze locations. *Zoom focus* was found to be cognitively demanding and *gaze trail* was found to have excessive information. The authors conclude that when designing gaze representation, one should consider *distraction* and *visual information*.

Games Context and Summary

Gaze is an emerging input in computer games, with a focus on using a player’s own gaze as a direct control mechanism to en-

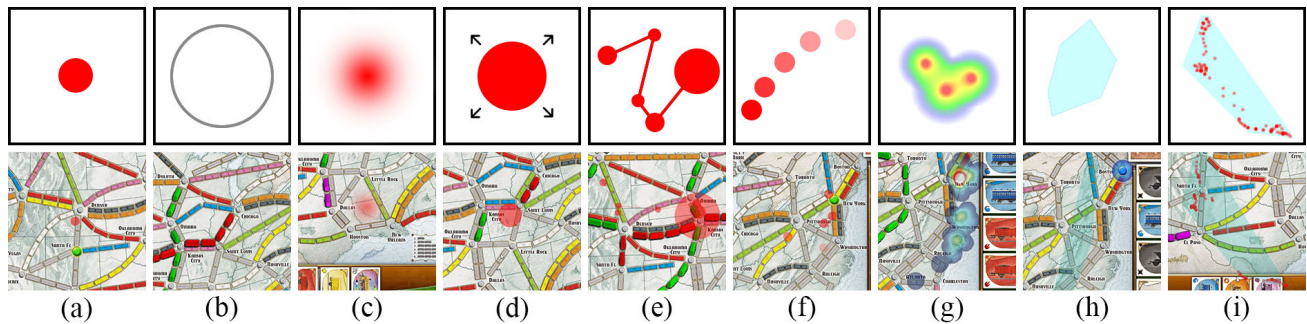


Figure 2: Top: The nine real-time gaze representations selected. Bottom: How each representation appears over the game *Ticket to Ride*. L-R: (a) Dot, (b) Cursor, (c) Spotlight, (d) Fixation, (e) Scanpath, (f) Fixation Trail, (g) Heatmap, (h) Convex Hull and (i) Bee Swarm.

able novel playful experiences [25, 26]. There is growing interest in implementing gaze as a mechanic for multiplayer games, in both cooperative [19, 21] and competitive [16, 20] settings. While the use of gaze in cooperative gameplay can directly leverage the findings of ‘shared gaze’ literature (e.g. [21]), competitive gameplay presents a different design challenge, as players have opposing objectives and are more likely to see different information on screen. We seek to find a gaze representation for use in competitive games that maintains the game balance, giving neither player a disproportionate edge over their opponent while not being disruptive to gameplay.

The existing literature makes clear that there is a need to balance *visibility*, *distraction* and *visual information*. Furthermore, we note that the problem with a majority of gaze representations is that they are designed to be used in post-processing of gaze data. Therefore, there is a need to design and evaluate novel gaze representations, such as that presented by Li et al. [17], to cater for settings where gaze input is not commonly used. Before we can design novel gaze representations, we need to understand the intrinsic characteristics of existing representations in relation to the context of use. In the next section, we present a variety of visual gaze representations that we have selected for use in our study.

REAL-TIME GAZE REPRESENTATIONS

Figure 2 shows the nine visual gaze representations which were evaluated in this study. For each representation, we provide an account of its current applications, its characteristics and the parameters for our use in the study. Most of the representations are common in existing in eye-tracking software, such as Tobii Pro Studio². These representations are primarily used for gaze data analysis and to replay collected gaze data. We also considered a wide selection of representations from works that: (1) survey gaze representations (e.g. [2, 18]), (2) evaluate gaze representations (e.g. [29]) and (3) use gaze for interaction (e.g. [21]). In making our selection, we required the representations to be clearly distinguishable from one another. The selection was based on the characteristics *conspicuousness* and *visual information* (see Table 1). We chose the term *conspicuousness* (rather than ‘visibility’) to highlight the degree to which the representations can be used to display and draw attention, beyond just how well the viewer is able to see the representation. *Visual information* refers to the amount of information the representation provides. This is

²<http://www.tobii.com>

Representation	Conspicuousness	Visual Information	Category
Dot	Low	High	Point
Cursor	Low	Low	Point
Spotlight	Low	Moderate	Point
Fixation	High	Low	Trajectory
Scanpath	High	Moderate	Trajectory
Fixation Trail	Moderate	Moderate	Trajectory
Heatmap	Moderate	High	AOI
Convex Hull	Moderate	Low	AOI
Bee Swarm	High	High	AOI

Table 1: Real-time gaze representations in relation to conspicuousness and visual information.

measured by the amount of aggregated gaze data used to generate the representation in addition to its overall size. We further sub-categorised the selected representations by their characteristics into three equally-sized subgroups: (1) Point-based, (2) Trajectory-based and (3) Area of Interest (AOI)-based.

Point-Based Representations

For this category, we first selected two types of gaze representations previously examined by Zhang et al. [29]: *cursor* and *spotlight*. We then added the commonly-used *dot* representation as a contrast to the first two in terms of precision and visual information. The point-based representations are mapped directly to the real-time gaze coordinates from the eye tracker. To reduce the jittery motions that naturally arise from eye movements, we applied a weighted moving average filter [27] on top of the eye tracker’s already lightly filtered data to smooth out the motion. This gave all three representations a hovering effect when a user’s eyes moved across the screen and reduced distracting jitter during gaze fixations.

Dot: This representation appears as a translucent red dot, akin to a laser pointer beam or, at larger sizes, a torch beam (Fig. 2a). It has commonly been adopted for use in gaze-based interaction as a deictic pointer and has been demonstrated to be an effective referential pointer to facilitate coordination in remote collaboration [1]. This representation allows a viewer to see gaze with high precision but can be potentially distracting. Our implementation used a 0.5 opacity with a radius of 30-px.

Cursor: This representation is a large ring with a transparent centre (Figure 2b). Its use has been demonstrated by Zhang et al. [29] and in *GazeArchers* [21], a game in which two players use gaze and touch simultaneously on a shared surface. Further, Tobii divisions has implemented a similar representation

in many of their recent products (Tobii Gaming³ *Gaze Trace*, Tobii Pro *Gaze Share*⁴). Tobii Pro also uses this representation for their wearable eye tracker controller to view the gaze of a user in the real world in real-time. This representation is equivalent to the size of the foveal vision which is about the size of a thumb at arm's length, making explicit the confidence range of the gaze estimation. We matched the size of this representation to default size of Tobii's *Gaze Trace* (150-pixel radius) and coloured the outline in grey to give this a low visibility. This representation provides more visual information than the previous representation (*Dot*) as it allows users to see what is enclosed in the ring as opposed to a specific point.

Spotlight: Figure 2c depicts this representation where the gaze point is shown as a red circle with high opacity in the centre and decreasing opacity towards the edges. This was the most widely preferred representation in Zhang et al.'s evaluation [29]. It was noted to be more subtle and less distracting due to the reduced overall visibility and conspicuousness. We modified the parameters to suit our study, reducing its overall size to a radius of 90 pixels (30 pixels at the highest opacity) and changing the colour to red, as white blended into the light background of the map area of the *Ticket to Ride* game board.

Trajectory-Based Representations

The representations in this category use the sensitive fixation stream provided by the eye tracker's SDK to reveal the sequence of points at which the player has recently looked, leaving behind a short history of the player's gaze. We initially drew from professional eye tracking usability studies that use animated representations to replay gaze data. In our review, we found similar representations being used in eye tracking research. For example, Qvarfordt and Zhai use a scanpath representation to infer dynamic eye movements in their study [22]. We also found proposed representations aimed to enhance remote collaboration, demonstrating the importance of assessing representations that fall into this category [17]. A red fill is used for representations in this category, with each fixation presented as a translucent ellipse.

Fixation: This representation aims to illustrate points of focus by displaying a single, unmoving fixation point as a red dot (Figure 2d). When a player maintains their fixation, the dot remains small in size. If the distance to the next fixation is within a threshold, the dot becomes larger. In real-time, this representation gives the appearance of small eruptions occurring on the interface that indicates whether a player is sustaining attention in an area (large dots) or flicking around between multiple areas (small dots). This representation was inspired by the 'Live Viewer' feature in the Tobii software, which Higuch et al. implemented for use in a remote collaborative scenario [11]. Our adaptation removes the elements of direction (tail, fade) used in other implementations to potentially reduce the mental load. Therefore, participants see a single point that appears then disappears instead of following a point.

Scanpath: The scanpath representation shows the length, duration and direction of gaze by using a sequence of fixations connected by a line (Fig. 2e). Variation of this representation

³<https://www.tobiigaming.com>

⁴<https://tobiicloud.com/share/>

has long been used to visualise gaze patterns in both psychology, eye tracking usability studies and in HCI. This representation has been used unidirectionally such as in Qvarfordt and Zhai's study to show the interest of a remote unaware user, or bi-directionally such as in Jermann and Nussli to improve communication in a remote pair-programming scenario [13, 22]. We implemented this representation based on Goldberg and Helfman [10]. Many approaches have been presented to overcome the problem of visual clutter such as by bundling close fixations into a single point. Another approach is to use time duration such as in previous works (e.g. 5 seconds [28]). Instead, we opted to limit the number of fixation bundles displayed to 5, reduce the visibility of the lines connected the fixation bundle and removed other elements such the numbering of each fixation bundle to reduce visual information.

Fixation Trail: This representation was originally proposed by Li et al. to illustrate gaze paths as opposed to a single fixation point [17]. In our study, we renamed this representation to *Fixation Trail* (originally *Gaze Trail*) as it uses fixation data and leaves a trail of fixations (see Figure 2f). Li et al. believe that the availability of prolonged gaze information allowed participants to make references to multiple locations (e.g. contrasting two different fixation clusters). This representation is different to *Scanpath* in that we fixed the size of the fixation to a 30-px radius. Only 5 fixation points remain fully visible while the remainder fades over time.

AOI-Based Representations

In this category, we selected gaze representations that allow an observer to view the Areas of Interest (AOI) at which the user has looked as a whole rather than a specific point. Representations in this category use fixations aggregated over time. We can control how much visual information is displayed by adjusting its time window. Our implementation uses 120 fixation points (2 seconds) so as not to occlude the screen (increase visibility, reduce visual information). We selected three representations with varying levels of visual information based on their popularity in eye tracking research.

Heatmap: Heatmaps represent the distribution of gaze points by a colour gradient (Figure 2g). We included this representation as it has commonly been used to visualise data in many domains and is not limited to eye tracking. In eye tracking usability studies it is also used for aggregating the gaze data from multiple participants. While the representation requires the aggregate participant data, it is possible to display a dynamic heatmap of a single tracked user [7]. Our implementation calculates the density of fixations within a 2-second window and displays the heatmap accordingly. The window size in our implementation attempts to balance visual information, allowing several 'hot spots' to be shown but do not remain for too long that it starts to occlude the interface.

Convex Hull: This representation is also known as *Cluster* (Tobii Pro) which display areas with a high concentration of gaze points as polygons to show areas of interest (Figure 2h). Our implementation uses the fixation data stream in which we use to render a convex hull polygon in real-time using a clustering algorithm. As the stream continuously updates the convex hull, this creates a shifting polygon that shows the

area in which the area is interested in. When the player looks rapidly between two points, the polygon stretches towards the points, but when the focus on a small area is high, the representation becomes small. This representation aims to provide minimal visual information of the area in which the participant has shown interest in.

Bee Swarm: Bee Swarm has been typically used to replay gaze points of several subjects against dynamic content such as videos, allowing efficient comparison of recordings from several subjects. Our variation aims to emulate this representation for use with a single user by displaying all fixation points in the last 2 seconds in which we enclose in a convex hull (Figure 2i). Each point was 5 pixels in radius with a red translucent fill. This visualization allows the player not only to see where the area in which the opponent was looking but the also points within that area. When the user focuses in a particular area, the points will become denser in the area, revealing an interest in the area. This representation has the highest amount of visual information as it combines characteristics from both AOI- and point-based representations.

RESEARCH DESIGN

In this study, we aim to understand the following research questions derived from our review of the literature:

[RQ1] How well are human subjects able to infer intent from a range of gaze representations?

[RQ2] What characteristics of gaze representations enable humans to perform better in predicting intentions?

[RQ3] Do gaze representations have an influence on predicting intent in a competitive gameplay setting?

To answer these questions, we adopted a spectator (‘third person’) approach. This means that participants other than the players watched clips of gameplay sessions after the fact, with different clips featuring different gaze representations. We chose to make our participants spectators rather than active players in order to simplify the task they were asked to perform, as they could focus on interpreting the gaze data without worrying about managing their own strategy. This decision was informed by the common finding in prior works that live gaze data can be highly distracting from a task at hand [17, 20, 29]. As this was the first study to assess people’s ability to infer intent from gaze data of a remote opponent in a game setting, we decided to begin with a more concentrated test of interpretive ability, before introducing the distractions of live gameplay in a future study.

We recorded our data in two stages. In the first, we recruited 8 players to record 4 games of our selected game—*Ticket to Ride* (described below) in which one player’s gaze was shared. In the second, a separate group of 27 participants watched short selected clips (< 1 minute) of the recorded gameplay from the perspective of the player who could see the real-time gaze representations. These participants provided an assessment of each clip, which we used in our analysis. A secondary benefit of this two-stage approach was that we were able to conduct a larger number of gaze visualization assessments in a shorter period of time than would have been possible if all participants had had to actively complete multiple games. This enabled a



Figure 3: Participant assessing a clip in a lounge room setting.

larger sample size for this foundational work, which we will build on in future studies involving active players.

Application

We developed a custom networked system that overlays real-time gaze visualizations over any application (see through, click through). We used this application to stream gaze data between two laptops, each equipped with a Tobii EyeX eye tracker, in different rooms. We streamed two types of gaze data concurrently from the SDK: the lightly filtered gaze data and the sensitive fixation data. Our application is capable of displaying the gaze of one or both users simultaneously, meaning that we have the ability to switch on the gaze of one user to be visualised by themselves or by another user. Moreover, we can easily switch the gaze between representations through a control window on an extended screen.

Game and Stages

In *Ticket to Ride*, each player competes to claim train routes between cities across a map of North America, seeking to connect certain pairs of cities (‘tickets’) in a continuous path of linked routes. Only one player can claim each route, and players do not know which cities their opponent is seeking to connect. Therefore, players must place their routes carefully, staying alert to the risk that their opponent will claim a route before them and block the most convenient path between their goal cities. Keeping information hidden plays a large role in *Ticket to Ride*, as a player can gain significant advantages over their opponents by correctly guessing which routes they are planning to take. Hence, *Ticket to Ride* works as an excellent use case for our study for several reasons. First, it is a widely-played board game, already familiar to many, with a digital game version that is almost identical to the board game; and because it has relatively simple rules involving clear goals and actions, which can be learned in under 15 minutes by interested players who have not played it previously. Second, as a tabletop game, it demands constant attention, requiring regular thought processing and frequent changes to short- and long-term strategies. This affords us to obtain a range of behaviours to display. Third, Lankes et al. have determined that turn-based games are better-suited to gaze transfer applications (as opposed to real-time first person shooters) [16]. This is because such games give players the opportunity to alternate between playing their own moves and watching their opponents’ behaviour as they take their turn. Fourth, it is a map-based game, in which the players’ focus and strategic calculations are closely concentrated on the geography of the

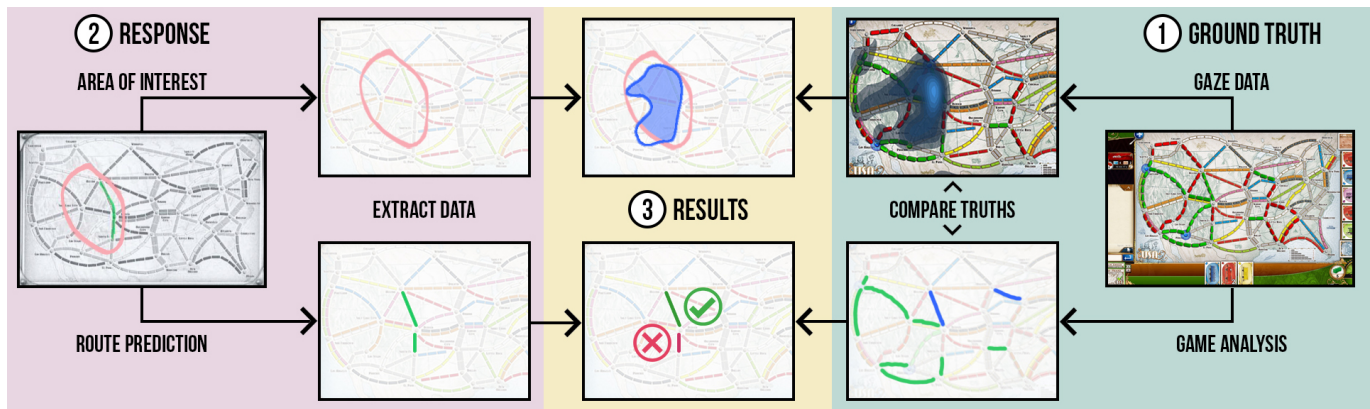


Figure 4: Method for analysis for Area of Interest (AOI) Estimation and Route Prediction.

map; gaze awareness and eye movement studies with map-based applications have been shown to be highly successful, as the static map gives users a point of reference [15, 22, 29].

As a playthrough of *Ticket to Ride* progresses, the map becomes increasingly populated with trains, and the strategic information available to the players becomes consequently richer while their choices of unclaimed routes become more constrained. Each player starts with a stock of 45 trains, and the game ends when any player's stock of unplaced trains falls below three. We use this train stock as a metric for game progression to divide the recorded games into three stages: early (45-31 trains), middle (31-16 trains), late (15-0 trains). The early stage is typically given over to planning, and placing trains are minimised as players try not to give away their goals. The middle stage sees greater competition for routes, and is where players often find themselves clashing over the most efficient paths between their respective goal cities. The late stage involves players either slotting in the last remaining routes that they need, seeking to block their opponent, or looking for other routes they can pick up extra points. We include an even selection of representations of each stage in our study, as we hypothesised that the usefulness and relevance of gaze visualizations might change as the game progresses.

Gaze Visualization Recording

We recruited 8 players (4 pairs) to record the necessary gameplay clips between a *naive player* (whose gaze was being recorded) and an *aware player* (for whom their opponent's gaze was visualised). Both players were placed in separate rooms with an experimenter and given the opportunity to play the in-game tutorial for up to 10 minutes to understand the rules and mechanics of the game. Once the *naive player* was confident with the game, they were asked to perform the default calibration provided by the eye tracker. The *naive player* does not know that they are being watched to record and observe 'natural gaze'. Meanwhile, the *aware player* was briefed that different gaze visualizations would be shown to them and that their opponent was not aware that their gaze was being shown. Each pair played three games, and the gaze representation was changed at every stage of the game. We recorded the displays using *Open Broadcaster Software (OBS) Studio*⁵ and logged the gaze data with our custom application.

⁵<https://obsproject.com/download>

Following the recordings, we reviewed the videos and noted instances in which the naive player exhibited various behaviours, such as planning, checking game counters or tickets, and looking for routes to claim on the map. From the videos, we selected three short clips for each of the nine gaze visualization formats: one in early game, one in middle game and one in late game. This gave us a total of 27 clips, exhibiting a variety of behaviours, and balanced between stages to control for the influence of game stage. Moreover, as different visualizations were used in different stages of different games, we are able to randomise the clips further. The length of the clips ranged between 20 and 60 seconds. All clips were saved without audio, to remove any potential sound clues such as the recorded player speaking aloud about their opponents' plans. As expected, the *aware players* experienced a level of distraction causing them to be drawn away from their own strategy and that some visualizations were preferred over others. However, we disregarded the feedback to avoid any early speculations.

STUDY

For the second stage of our study, we recruited 27 participants from the University of Melbourne. There were 11 male and 16 female participants, aged 18 to 44 years (mean=25.6), and included undergraduate students, graduate students and non-faculty staff. Only two had previously played the board game version of *Ticket to Ride*, but 17 indicated they had played a game of the same type in the past week. Participants were compensated with a \$5 coffee voucher for their time. We encouraged participants to watch a short video⁶ that describes the rules and objectives as part of the recruitment process.

Procedure

Upon arrival, participants were greeted and seated comfortably in a lounge room setting, facing a 48-inch (1080p) television display connected to a PC as depicted in Figure 3. They completed a consent form and a simple demographic questionnaire and were briefed on the rules of *Ticket to Ride*. The facilitator showed them an example clip of gameplay footage and with it, explained how they should assess each clip on a paper survey, with the measures outlined under 'Measures and Analysis'. Here, the facilitator ensured that the participants understood the rules of the game and what was required of them before starting the study. Each session ran between 30 to 60 minutes.

⁶<https://www.youtube.com/watch?v=uDODCHcj6Sg>

Representation	Early	Mid	Late	Overall
Heatmap	.44(.78)	.68(.63)	.73 (.75)	.62(.72)
Fixation Trail	.48(.41)	.66(.72)	.48 (.75)	.54(.62)
Dot	.59(.49)	.37(.55)	.67(.36)	.54(.47)
Cursor	.46(.73)	.68(.68)	.44(.28)	.53(.56)
Fixation	.63(.72)	.51(.87)	.41(.80)	.52(.80)
Scanpath	.40(.65)	.50(.79)	.65(.70)	.52(.71)
Bee Swarm	.59(.58)	.36(.63)	.43(.72)	.46(.64)
Convex Hull	.40(.62)	.17(.44)	.23 (.44)	.27(.50)
Spotlight	.24(.37)	.25(.43)	.28 (.65)	.26(.48)

Table 2: AOI Estimation Results. Precision(Recall).

We employed a repeated-measures design: each participant was shown all 27 (3 stages × 9 visualizations) clips, each of which was considered one study ‘condition’, representing one gaze visualization format and one game stage. Participants were divided into three sets (A, B and C), and the order of the clips was randomised for each set.

Participants were given a maximum of 10 seconds to analyse the game and board state before the clip was played. In almost all cases, participants indicated that this was enough time for them to understand the game state, either with a verbal acknowledgement or a non-verbal indication such as a nod of the head. This time limit was imposed to prevent the participant from basing their inferences too heavily on reasoning about the layout of the game board, rather than the visualizations of the gaze data, and in order to more closely approximate a live gameplay setting in which the gaze information would be processed and responded to quickly. The facilitator encouraged participants to *think aloud* about what they were seeing, and to ask for any clarifications about the game if required. After the participant had viewed and assessed all 27 clips, they were finally asked to rank the nine gaze visualization formats from most preferred to least preferred. In summary, we collected a total of 27 participants × 9 visualizations × 3 stages = 729 responses for each of our dependent measures.

Measures and Analysis

In response to each clip, participants were asked to record the following five pieces of information: (1) *Area of Interest Estimation*, using a map of the game board they circled the area (or multiple areas) in which they believed the gaze-tracked player had been interested. (2) *Route Prediction*, on the same map they highlighted any specific routes that they thought the gaze-tracked player was planning to build upon. They were also encouraged to write a short interpretation of what the gaze-tracked player was thinking during the clip or to explain their interpretation of the gaze behaviour out. (3) They rated how informative they found the gaze visualization in the clip on a 7-point scale ranging from *Not informative at all* to *Extremely informative*. (4) They rated how well they felt they could predict the gaze-tracked player’s intentions in the clip on a 10-point rating scale. Participants were permitted to skip any of these questions when they felt unable to give a response, for example, if they could not predict any routes upon which the gaze-tracked player was considering to build. Finally, we collected (5) *Qualitative Responses* at the end of the session, participants were asked to rank the representations in order of preference and explain the reasoning behind their choices.

Representation	Early	Mid	Late	Overall
Bee Swarm	.70(.37)	.67(.43)	.43 (.28)	.60(.36)
Cursor	.69(.85)	.54(.37)	.53 (.19)	.59(.47)
Heatmap	.57(.31)	.63(.29)	.50(.63)	.57(.41)
Scanpath	.40(.46)	.45(.26)	.63(.25)	.50(.32)
Spotlight	.50(.22)	.49(.26)	.25(.14)	.41(.20)
Fixation Trail	.44(.14)	.50(.17)	.26(.26)	.40(.19)
Dot	.59(.37)	.37(.14)	.10(.04)	.35(.18)
Fixation	.37(.44)	.26(.14)	.27(.21)	.30(.26)
Convex Hull	.25(.11)	.27(.10)	.74(.02)	.20(.08)

Table 3: Route Prediction Results. Precision(Recall).

Area of Interest Estimation and Route Prediction

To determine how closely each participant guessed the area of interest of the naive player, we compared the intersection of each response against the ground truth for the respective condition. Participants recorded their responses by circling the corresponding area in the map. To be able to analyse this data digitally, we scanned the highlighted pages and extracted the contours of the highlights using a custom-built OpenCV computer vision application. To establish a ground truth, we generated 2D density maps from the gaze data in each clip using the R statistical package. We then refined the AOIs against the manual analysis of the clips, discarding outliers that did not correspond to the routes the player eventually built. We then compared the intersection of the two regions, computing the precision (area of intersection divided by the total area circled by the participant) and the recall (area of the intersection divided by the area of the ground truth contour) of the participant’s estimate (see Figure 4, top path).

Similarly to the analysis of the AOI Estimation, we scanned the highlighted maps and extracted the highlighted routes using our OpenCV application (see Figure 4, bottom path). We established the ground truth by recording the routes that the player eventually built after the clip. We also recorded routes that the player was considering to build but never did. We inferred such routes based on evidence from their hidden goal cards, subsequent game actions, and in-game comments. As with the AOI Estimation, we operationalise their performance primarily based on precision and recall, but also considered their harmonic mean i.e. $F\text{-Score}$ ($F = 2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$). The details of these performance metrics are explained by Rijsbergen [23].

Quantitative and Qualitative Responses

For our third and fourth measure, we will use descriptive statistics to analyse how informative as well as the predictability of the gaze representation. To gain richer insights beyond the quantitative performance data, we collected subjective perceptions in three ways. First, we transcribed all verbal comments as they thought out loud when watching the clips. Second, in the questionnaires for the clips, participants answered an open-ended question that asked them to predict what the player was thinking. Third, we transcribed the interview about their order of preference at the end of the study. We were particularly interested in comments beyond the areas of interest and routes, including potential strategies, such as building a route to block the opponent, planning a route, verifying whether they have enough resources to build a certain route.

Representation	Early	Mid	Late	Overall
Spotlight	5.1 (1.0)	5.1 (1.1)	4.8 (1.0)	5.0 (1.0)
Heatmap	4.8 (1.0)	5.0 (0.8)	4.9 (1.0)	4.9 (0.9)
Cursor	4.9 (1.3)	4.8 (1.4)	4.6 (1.3)	4.8 (1.3)
Dot	5.1 (1.6)	4.8 (1.1)	4.1 (1.7)	4.7 (1.5)
Fixation	4.7 (1.0)	4.5 (1.2)	4.6 (1.1)	4.6 (1.1)
Scanpath	4.7 (1.0)	4.5 (1.2)	4.8 (1.0)	4.6 (1.1)
Bee Swarm	4.6 (1.1)	4.6 (1.3)	4.4 (1.4)	4.5 (1.3)
Fixation Trail	4.4 (1.2)	4.7 (1.0)	4.3 (1.0)	4.5 (1.1)
Convex Hull	3.2 (1.2)	3.6 (1.4)	3.0 (1.5)	3.3 (1.4)

Table 4: Ratings of Informativeness Results 1 Not Informative at all - 7 Extremely Informative). Score (Standard Deviation)

RESULTS

We report the results of the nine visualizations at each of the three game stages and their overalls. Five primary metrics are reported: the percentage overlap of the AOI indicated by the participant against the AOI detected by the software; the performance between routes predicted by the participant and routes plausibly considered by the naive player; participant ratings of the visualizations' informativeness; participant self-assessments of how well they can predict what the naive player is planning to do; and participant rankings of gaze visualization preference. Lastly, we compared the scores from their performance in our first two metrics against the last three.

Area of Interest Estimation

Table 2 shows the precision and recall for each representation, ordered by precision. The order of the performance of the visualizations in terms of precision were the *Heatmap* (.62), followed by the *Fixation Trail* (.54), and finally the *Dot* (.54). This means that these are the best representations for selecting areas of interest within the ground truth. In terms of recall, the *Fixation* (.80) was the best performing, followed by the *Heatmap* (.72) and the *Scanpath* (.71). This indicates that these were the representations in which the selected area captured a larger area within the area of interest. If we combine both metrics using the *F-Score*, we find that the overall best representation was the *Heatmap* (.66), followed by the *Fixation* (.63) and the *Scanpath* (.63). To test whether the stage of the game had an effect on the precision and recall, we conducted a repeated-measures ANOVA for each representation. Except for *Spotlight*, we found that the stage had a significant effect on the stage on the precision and recall. Precision significantly increased in the *Heatmap* and *Fixation Trail* as the game progressed, and decreased in the remaining conditions.

Route Prediction

Table 3 shows the computed precision and recall results for each visualization ordered by precision. The best performing visualizations when it comes to route prediction in terms of *F-Score* are *Cursor* (.52), *Heatmap* (.48), and *Bee Swarm* (.45). *Convex Hull* was ranked significantly lower in both precision and recall (.20 and .08 respectively) and this was primarily from the late stage precision score (.02). The other low score noted, also in the late stage, is the *Dot* representation recall score (.04). We tested the effects of the gameplay stage on both precision and recall for each representation using repeated-measures ANOVA. For precision, only the *Scanpath* representation showed significance ($p = 0.014$) in which a

Representation	Early	Mid	Late	Overall
Heatmap	6.3 (2.0)	6.7 (1.9)	7.1 (1.7)	6.7 (1.9)
Cursor	7.0 (1.7)	6.1 (2.0)	6.6 (2.0)	6.6 (1.9)
Scanpath	6.0 (1.5)	6.0 (2.2)	6.6 (1.6)	6.2 (1.8)
Fixation	6.9 (1.7)	5.4 (2.1)	6.3 (1.5)	6.2 (1.9)
Spotlight	6.2 (1.7)	6.5 (1.5)	5.6 (1.9)	6.1 (1.8)
Bee Swarm	5.9 (2.1)	6.2 (2.0)	5.8 (2.3)	6.0 (2.1)
Fixation Trail	6.1 (1.7)	6.2 (1.5)	5.6 (2.2)	6.0 (2.0)
Dot	6.9 (2.3)	5.7 (1.8)	3.8 (2.0)	5.5 (2.4)
Convex Hull	4.0 (1.8)	4.7 (2.0)	3.9 (2.1)	4.2 (2.0)

Table 5: Predictive Ability Results. Score (Standard Deviation)

Bonferroni post hoc test showed a significant difference between the mid and late stage ($p = 0.0069$). Here, we note that the *Scanpath* representation is the only representation to have increase in precision from the mid stage ($M = 45.5$) to the late stage ($M = 63.4$). A majority of representations, 6 out of 9 (*Bee Swarm*, *Heatmap*, *Spotlight*, *Fixation Trail*, *Scanpath* and *Convex Hull*), remained precise for the early and mid stages but significantly becomes less precise in the late stage. Both *Fixation* and *Cursor* precision dropped significantly from early ($M = 37.3$, $M = 69.1$) to mid game ($M = 27.2$, $M = 54.0$) but remained consistent in the late game. The precision of the *Dot* representation dropped linearly from early to late stage. There was no significant effect on recall on the gameplay stages.

Ratings of Informativeness

The responses follows a normal distribution (*Pearson chi-square normality test*, $p < 0.05$), in which calculated the mean scores and standard deviation for each visualization for each stage, and overalls (see Table 4). We test the effects of stage on their rating on informativeness using repeated-measures ANOVA and found a significance in the mid and late stage ($p = 0.0247$). We tabulated the results and ranked them by overall rating accordingly. We found that *Spotlight* (5.0), *Heatmap* (4.9) and *Cursor* (4.8) received the highest rating while *Convex Hull* was rated significantly lower (3.3). *Heatmap* had the lowest deviation (0.9) among the visualizations showing an agreement among participants on its informativeness.

Self-Assessment of Predictive Ability

We ran the same statistical tests used in the previous measure and found that the responses also follows a normal distribution. We calculated the mean scores and standard deviation of each visualization for each stage and overalls, and ranked them by overall rating (see Table 5). We proceeded to test the effects of stage on the responses using repeated-measures ANOVA and found significance in the early and late stage ($p = 0.331$). The results shows that *Heatmap* (6.7), *Cursor* (6.6) and *Scanpath* (6.2) received the highest rating while *Dot* (5.5) and *Convex Hull* (4.2) was rated significantly lower than the others.

Preference Rankings

Figure 5 shows the ranking data for each representation by all participants, showing that the *Convex Hull* was ranked the lowest ($M = 2.0$) while *Heatmap* was the highest ranked ($M = 6.7$). *Cursor* and *Scanpath* also was ranked highly. We further analysed the ranked data using the Friedman test in which we found a significance ($p < 0.05$). The post hoc test showed that *Convex Hull* representation was ranked significantly lower than

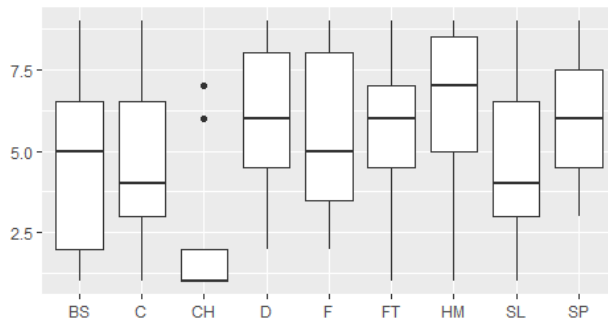


Figure 5: Boxplot of Preference Ratings given for each visualization. (BS) Bee Swarm, (C) Circle, (CH) Convex Hull, (D) Dot, (F) Fixation, (FT) Fixation Trail, (HM) Heatmap, (SL) Spotlight and (SP) Scanpath.

all other representations. We ran the Friedman test once again but this time removing the *Convex Hull* representation from the test. We found that there were no significance differences between the other rest ($p = 0.07547$). We proceeded to analyse the interviews to gain further insights into these preferences.

We received many positive insights into why *Heatmap* was ranked the highest and preferred by a majority of participants. Participants that gave the *Heatmap* a high score mentioned its ‘persistence’ i.e. the visualization gave them a moment to make an inference before it fades away (e.g. “*I like that it stayed on the screen for awhile*” [P3]). A small number of participants (e.g. P7) noted that they were familiar with the visualization having seen or used them in other domains, noting its inherent usefulness. When participants were asked why they preferred *Heatmap* over the others, a majority they mentioned that they had difficulty following them. Two contributing factors for this were (1) speed (“*too quick*” [P16], “*moved too much and too quickly*” [P11]) and (2) lack of conspicuity (e.g. “*more subtle ones are harder to keep track of*” [P19], “*it was hard to see*” [P25]). Some participants mentioned that the *Scanpath* visualization blended into the game making it hard to differentiate (e.g. P20) while others found it useful (e.g. “*I like it because it gives you a track.*” [P11]). Conversely, some participants did rate visualization that falls under point-based category highly. Participant 22 provided his explanation, stating that he preferred the point-based representations as he was able to process the visual information quickly and described trajectory- or AOI-based representations as “a lot going on”. Participants often compared *Convex Hull* and *Bee Swarm* and found that the ‘red dots’ made it better (e.g. P17). *Bee Swarm*’s ability to have persistence yet show movement at the same time it is probably a reason why it scored well in the route prediction measure. Overall, participants provided rich feedback, noting that AOI-based representations were good for predicting overall strategy while point-based representations appear to be better moment-to-moment attention.

DISCUSSION

Our study explored the characteristics of nine visual representations of gaze to determine which was most effective at conveying intent to an observer by exploring relationship between the characteristics of representations and human ability to infer intent from them. This section discusses the results from the five metrics as well as the post study interviews.

Representation Performance and Preference

Once again *Convex Hull* was largely indecipherable to participants, with the lowest scores for route prediction and close to the lowest scores for AOI estimation. However, the other AOI-based visualizations, *Bee Swarm* and *Heatmap*, had among the best results for route prediction. Unlike *Convex Hull*, these visualizations can clearly display both the broad area of sustained interest and the focal point of the player’s moment-to-moment attention. The *Heatmap* shows multiple hot spots and recent zones in which the player has looked. Interestingly, *Bee Swarm* performed well for route prediction but relatively poorly in AOI estimation measure. Our version of the *Bee Swarm* is encompassed by a convex hull display, which allows viewers to see at a glance the general region of the player’s gaze while their more exact moment-to-moment attention is captured by a large number of rapidly-propagating red dots. We believe that participants used the convex hull to narrow their focus, enabling them to see the routes more clearly; avoiding a visual search for the next sequence.

Conversely, *Fixation Trail* and *Dot* performed strongly for AOI estimation, but relatively poor for route prediction, suggesting that these visualizations were effective at conveying general information but were not fine-grained enough for inferring specific intent. We believe that trajectory-based representations caused participants to divide their attention between two areas of interest and this was specifically mentioned by Participant 22. None of the point-based representations performed particularly well. We believe by the time participants start to make a connection between objects, the visual has moved somewhere else, causing them to start making a new inference elsewhere.

The *Heatmap* gaze representation overall scored highly in all five measures, and this is primarily attributed to its ability to provide a high level of visual information by using averaged information. Moreover, *Heatmap* takes a small amount of time to ‘heat up’ and cool down, allowing the player to know areas which are gaining interest and losing interest. We earlier noted that *Scanpath* performance increased from the early-mid to late stage of the game which may be reflective of the nature of the game. Towards the late stage players tend to track the missing paths which reduces the search space and enhances the performance of the *Scanpath* gaze representation. Further, with the increased availability of other cues to select from as the game progresses, a sustained gaze (presenting sustained interest) at a location is likely to signal with greater confidence the likelihood of an intention to action. To some degree, we expected the *Heatmap* to have performed the best being a well-used representation. On the other hand, the *Convex Hull* was notably disliked, being given the lowest or second-lowest ranking by all but four participants.

Preferences were mixed for the rest, with a wide variation in scores indicating that different participants valued different characteristics. The trajectory-based representations were all ranked relatively highly; we hypothesise that this is because trajectory-based representations naturally calls attention to the connections between different points of interest as the player moves their gaze between them without overloading the user and therefore remained neutral in our results.

Interplay Between Game Cues and Overall Scores

As expected, we found that participants make use of in-game and behavioural cues to predict routes such as by observing what cards the naive player looked at or picked. From the game analysis, there was 17 showed cards being picked and many participants drew conclusions by matching the route seen and card colour which is a probable cause of differences in scores between the clips. Across all visualizations, prediction accuracy was lower in the late game than in the early and mid game. We consider several factors that may account for this trend, with different implications for the use of gaze visualizations in other contexts. First, the visual clutter on the screen increases as the game progresses, so that in the late game there are substantially more train routes that may potentially distract the viewer from the gaze representation. Second, the number of available routes is lower in the late game, which tends to require players to draw more convoluted routes; the combination of zigzagging existing routes and fewer possibilities may increase the difficulty of recognising an obvious path that the player is aiming for. Third, the behaviour of the player is likely to be different, as they transition from long-term planning in the earlier stages to opportunistic route-grabbing and path-blocking in the late stage; this behaviour may be more difficult for a viewer to interpret. We believe that it is plausible that all three of these factors play a role, but based on our own observation of the game recordings, we expect that the third factor to be the most significant as gaze behaviour clearly shows a different pattern late in the game. This is exhibited in our results such as in the case of the *Dot* representation where the naive player was looking for was considering a route to build between a few potential routes.

Humans naturally rely on explicit actions as they confirm their thoughts, such as in real life. Regardless of representation shown, we are impressed at the ability of participants to put themselves into the mind of the player with the introduction of displayed gaze. In the same way, spectators rely on these actions, and therefore participants employ a variety of cues in making inferences. However, we observed in our study that some participants better than other in keeping track of the elements in the game, allowing them to make better inferences.

Design Implications

Our results present various implications towards the digital games community. First and foremost, we demonstrated the possibilities of using different types of representations to display mental processes (e.g. intention) of players in real-time, which has direct implications towards *game developers*. When using gaze explicitly for use in games, game developers must be wary of how it is actually used in its execution. For example, when we tested the *Convex Hull* representation, it worked well allowing two developers to reveal what the other was looking at. However, during our evaluation, *Convex Hull* obtained consistently low scores revealing a flaw in our implementation when converting a static visualisation to real-time. Then again, this consistently low score gave us a reference point and confidence in the validity of our data. The design of gaze representations should further consider the amount it reveals without giving it away as it would give a player a disadvantaged. Imagine playing a game and your op-

ponent is constantly one step ahead of you. For the evaluation of games in real-time with gaze visualisation (implicit use), representations such as the *Heatmap* can be used to reveal intentions of players from natural gaze behaviour, and therefore provide insights into whether players understand the task in hand or is making correct strategic decisions. Second, a range of selectable gaze visualizations can implemented in live game streaming platforms such as Twitch, where *audiences* to create sensory excitement and anticipation. For example, audiences can anticipate intentions of live players before the actual move, providing a new experience. Lastly, *players* themselves can benefit from visualisations the gaze of their opponent in the form of an in-game ‘power up’ to create new gameplay experiences, especially in similar competitive contexts.

CONTRIBUTIONS, LIMITATIONS AND FUTURE WORK

This work presents three main contributions. First, we provide an understanding of the role of gaze and its importance to game designers in choosing and designing the right visualisation. Second, an assessment of how people interpret visual gaze visualisations, which has not been done before in this context. The findings also highlight the effect of the different techniques on the inferring ability. Third, we contribute the assessment of AOI-based representations in the context of gameplay. Limitations wise, we used clips in our study from the entire game session which might not be comprehensive to show the goals of the whole game but enough to determine short term intention. Moreover, as each visualization was tested in only three different clips, it is possible that one outlier clip could affect the overall result for that visualization; nevertheless, we have taken this into account in our reporting of the findings. We also acknowledged that we only used an example of a strategic game to demonstrate gaze. However, we believe our findings apply to similar games.

Our future work focuses on *players*, where we give one player the ability to predict intentions using a real-time *Heatmap*, and with room for deceptive or misinterpreted behaviours to happen. We have also considered testing the viability of our findings with a different game (e.g. on with a less cluttered board) to provide further insights into the generalisation of our results. Furthermore, our work has further implications in gaze-aware artificial intelligence, and we have begun putting together a planning-based AI that predicts players potential intentions based on the recorded games from this study.

CONCLUSION

This paper contributes the groundwork for determining the characteristics of gaze representations that allow human players to make inferences about other players’ intentions and to derive insights into potential strategies. This has been achieved by measuring the performance of human in their ability to predict intentions using a range of visual gaze representations. Our results showed us that human subjects have a remarkable ability to infer intent with the introduction of gaze. We conclude that representations that show both characteristics of point- and AOI- based representations would work well in our setting. Our work ultimately contributes gaze input as a viable implicit modality in multiplayer competitive computer games and with broader implications towards the games community.

ACKNOWLEDGEMENTS

We acknowledge the Australian Commonwealth Government and the Microsoft Research Centre for Social NUI for their support on this project. We further acknowledge Zaher Joukhadar for major improvements made on the custom built software.

REFERENCES

1. Deepak Akkil, Jobin Mathew James, Poika Isokoski, and Jari Kangas. 2016. GazeTorch: Enabling Gaze Awareness in Collaborative Physical Tasks. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 1151–1158. DOI : <http://dx.doi.org/10.1145/2851581.2892459>
2. T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl. 2014. State-of-the-Art of Visualization for Eye Tracking Data. In *EuroVis - STARS*, R. Borgo, R. Maciejewski, and I. Viola (Eds.). The Eurographics Association. DOI : <http://dx.doi.org/10.2312/eurovisstar.20141173>
3. Susan E. Brennan, Xin Chen, Christopher A. Dickinson, Mark B. Neider, and Gregory J. Zelinsky. 2008. Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition* 106, 3 (2008), 1465 – 1477. DOI : <http://dx.doi.org/10.1016/j.cognition.2007.05.012>
4. Marcus Carter, Joshua Newn, Eduardo Velloso, and Frank Vetere. 2015. Remote Gaze and Gesture Tracking on the Microsoft Kinect: Investigating the Role of Feedback. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction (OzCHI '15)*. ACM, New York, NY, USA, 167–176. DOI : <http://dx.doi.org/10.1145/2838739.2838778>
5. Mauro Cherubini, Marc-Antoine Nüssli, and Pierre Dillenbourg. 2008. Deixis and Gaze in Collaborative Work at a Distance (over a Shared Map): A Computational Model to Detect Misunderstandings. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications (ETRA '08)*. ACM, New York, NY, USA, 173–180. DOI : <http://dx.doi.org/10.1145/1344471.1344515>
6. Sarah D'Angelo and Darren Gergle. 2016. Gazed and Confused: Understanding and Designing Shared Gaze for Remote Collaboration. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2492–2496. DOI : <http://dx.doi.org/10.1145/2858036.2858499>
7. Andrew T. Duchowski, Margaux M. Price, Miriah Meyer, and Pilar Orero. 2012. Aggregate Gaze Visualization with Real-time Heatmaps. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*. ACM, New York, NY, USA, 13–20. DOI : <http://dx.doi.org/10.1145/2168556.2168558>
8. Paul Ekman and Wallace V. Friesen. 1969. Nonverbal Leakage and Clues to Deception. *Psychiatry* 32, 1 (1969), 88–106. DOI : <http://dx.doi.org/10.1080/00332747.1969.11023575>
9. Alexandra Frischen, Andrew P Bayliss, and Steven P Tipper. 2007. Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin* 133, 4 (2007), 694.
10. Joseph H. Goldberg and Jonathan I. Helfman. 2010. Visual Scanpath Representation. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications (ETRA '10)*. ACM, New York, NY, USA, 203–210. DOI : <http://dx.doi.org/10.1145/1743666.1743717>
11. Keita Higuch, Ryo Yonetani, and Yoichi Sato. 2016. Can Eye Help You?: Effects of Visualizing Eye Fixations on Remote Collaboration Scenarios for Physical Tasks. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5180–5190. DOI : <http://dx.doi.org/10.1145/2858036.2858438>
12. Robert J. K. Jacob. 1991. The Use of Eye Movements in Human-computer Interaction Techniques: What You Look at is What You Get. *ACM Trans. Inf. Syst.* 9, 2 (April 1991), 152–169. DOI : <http://dx.doi.org/10.1145/123078.128728>
13. Patrick Jermann and Marc-Antoine Nüssli. 2012. Effects of Sharing Text Selections on Gaze Cross-recurrence and Interaction Quality in a Pair Programming Task. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 1125–1134. DOI : <http://dx.doi.org/10.1145/2145204.2145371>
14. Adam Kendon. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica* 26 (1967), 22 – 63. DOI : [http://dx.doi.org/10.1016/0001-6918\(67\)90005-4](http://dx.doi.org/10.1016/0001-6918(67)90005-4)
15. Peter Kiefer, Ioannis Giannopoulos, and Martin Raubal. 2013. Using Eye Movements to Recognize Activities on Cartographic Maps. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL '13)*. ACM, New York, NY, USA, 488–491. DOI : <http://dx.doi.org/10.1145/2525314.2525467>
16. Michael Lankes, Bernhard Maurer, and Barbara Stiglbauer. 2016. An Eye for an Eye: Gaze Input in Competitive Online Games and Its Effects on Social Presence. In *Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology (ACE2016)*. ACM, New York, NY, USA, Article 17, 9 pages. DOI : <http://dx.doi.org/10.1145/3001773.3001774>
17. Jerry Li, Mia Manavalan, Sarah D'Angelo, and Darren Gergle. 2016. Designing Shared Gaze Awareness for Remote Collaboration. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion (CSCW '16 Companion)*. ACM, New York, NY, USA, 325–328. DOI : <http://dx.doi.org/10.1145/2818052.2869097>

18. Paivi Majaranta, Hirotaka Aoki, Mick Donegan, Dan Witzner Hansen, and John Paulin Hansen. 2011. *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies* (1st ed.). Information Science Reference - Imprint of: IGI Publishing, Hershey, PA.
19. Bernhard Maurer, Michael Lankes, Barbara Stiglbauer, and Manfred Tscheligi. 2016. *EyeCo: Effects of Shared Gaze on Social Presence in an Online Cooperative Game*. Springer International Publishing, Cham, 102–114. DOI: http://dx.doi.org/10.1007/978-3-319-46100-7_9
20. Joshua Newn, Eduardo Velloso, Marcus Carter, and Frank Vetere. 2016. Exploring the Effects of Gaze Awareness on Multiplayer Gameplay. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (CHI PLAY Companion '16)*. ACM, New York, NY, USA, 239–244. DOI: <http://dx.doi.org/10.1145/2968120.2987740>
21. Ken Pfeuffer, Jason Alexander, and Hans Gellersen. 2016. GazeArchers: Playing with Individual and Shared Attention in a Two-player Look&Shoot Tabletop Game. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia (MUM '16)*. ACM, New York, NY, USA, 213–216. DOI: <http://dx.doi.org/10.1145/3012709.3012717>
22. Pernilla Qvarfordt and Shumin Zhai. 2005. Conversing with the User Based on Eye-gaze Patterns. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*. ACM, New York, NY, USA, 221–230. DOI: <http://dx.doi.org/10.1145/1054972.1055004>
23. C. J. Van Rijsbergen. 1979. *Information Retrieval* (2nd ed.). Butterworth-Heinemann, Newton, MA, USA.
24. Randy Stein and Susan E. Brennan. 2004. Another Person's Eye Gaze As a Cue in Solving Programming Problems. In *Proceedings of the 6th International Conference on Multimodal Interfaces (ICMI '04)*. ACM, New York, NY, USA, 9–15. DOI: <http://dx.doi.org/10.1145/1027933.1027936>
25. Jayson Turner, Eduardo Velloso, Hans Gellersen, and Veronica Sundstedt. 2014. EyePlay: Applications for Gaze in Games. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-human Interaction in Play (CHI PLAY '14)*. ACM, New York, NY, USA, 465–468. DOI: <http://dx.doi.org/10.1145/2658537.2659016>
26. Eduardo Velloso and Marcus Carter. 2016. The Emergence of EyePlay: A Survey of Eye Interaction in Games. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '16)*. ACM, New York, NY, USA, 171–185. DOI: <http://dx.doi.org/10.1145/2967934.2968084>
27. Oleg Špakov. 2012. Comparison of Eye Movement Filters Used in HCI. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*. ACM, New York, NY, USA, 281–284. DOI: <http://dx.doi.org/10.1145/2168556.2168616>
28. Nadir Weibel, Adam Fouse, Colleen Emmenegger, Sara Kimmich, and Edwin Hutchins. 2012. Let's Look at the Cockpit: Exploring Mobile Eye-tracking for Observational Research on the Flight Deck. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*. ACM, New York, NY, USA, 107–114. DOI: <http://dx.doi.org/10.1145/2168556.2168573>
29. Yanxia Zhang, Ken Pfeuffer, Ming Ki Chong, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2017. Look Together: Using Gaze for Assisting Co-located Collaborative Search. *Personal Ubiquitous Comput.* 21, 1 (Feb. 2017), 173–186. DOI: <http://dx.doi.org/10.1007/s00779-016-0969-x>