The Emergence of EyePlay: A Survey of Eye Interaction in Games

Eduardo Velloso  
Microsoft Research Centre for Social NUI  
The University of Melbourne  
eduardo.veloso@unimelb.edu.au

Marcus Carter  
Microsoft Research Centre for Social NUI  
The University of Melbourne  
marcus.carter@unimelb.edu.au

ABSTRACT
As eye trackers become cheaper, smaller, more robust, and more available, they finally leave research labs and enter the home environment. In this context, gaming arises as a promising application domain for eye interaction. The goal of this survey is to categorise the different ways in which the eyes can be incorporated into games and play in general as a resource for future design. We reviewed the literature on the topic, as well as other game prototypes that employ the eyes. We compiled a list of eye-enabled game mechanics and derived a taxonomy that classifies them according to the eye movements they involve, the input type they provide, and the game mechanics that they implement. Based on our findings we articulate the value of gaming for future HCI gaze research and outline a research program around eye interaction in gaming.

Author Keywords
Survey; Eye tracking; Games

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
Marty McFly: (showing two boys how to play an arcade game in an idealised 80’s vision of 2015) I’ll show you, kid. I’m a crack shot at this.
Boy #1: You mean you have to use your hands?
Boy #2: That’s like a baby’s toy!

Back to the Future Part II (1989)

Gaming is an application domain traditionally receptive to novel input devices, driving their commercial adoption and proliferation. Until recently, eye-enabled games were restricted to research prototypes and student projects, as prices in the order of tens of thousands of dollars made purchasing eye-trackers prohibitive for the consumer market. However, more recent affordable trackers specifically designed for gaming are bringing this technology to the home environment. 2015 represents an important landmark in this timeline, with the release of Assassin’s Creed Rogue [68], the first natively gaze-enabled AAA game; and the SteelSeries Sentry, the first gaming-specific eye tracker. Emerging VR and AR devices are also increasingly incorporating eye-tracking, as exemplified by the upcoming FOVE VR headset.

Eye interaction research has been prolific in HCI and associated fields since the late 1970’s, but has typically focused on applications relating to health, user-experience evaluation, marketing, or accessibility. Thus, as we highlight in this paper, current literature has little contribution to its likely commercial success in gaming. Further, as is typical in HCI [10], the gaze-enabled game research that does exist often only utilizes the game element to demonstrate novel interaction techniques or to collect data, frequently without consideration of the game-specific demands or opportunities, and typically without deeper evaluations of the play experience.

In this paper we present a survey of 70 works that have interrogated gaze-enabled interaction in the context of games or play, categorizing them based on the 112 individual game mechanics we identify within them, their motivations, and relevance to game design and the games industry. This work contributes (1) a clear vocabulary for discussing and designing eye-based games and play, (2) a comprehensive and practical list of existing gaze-enabled game mechanics, and (3) a critique of gaze-in-games’ overemphasis on specific eye-movements and mechanics, and minimal engagement with opportunities for gaze to provide new game experiences.

Based on this comprehensive literature review, we propose the existence of an ‘EyePlay’ research program as a distinct and emerging field within HCI. Rather than utilizing games as an instrument for peripheral research, EyePlay research is concerned with game experiences that are only possible through eye-based interaction. In the discussion we articulate the boundaries and importance of this emerging research program and distinguish it from marketing, broader HCI and accessibility gaze research involving games.

RELATED WORK
Gaze is a term with many meanings in the literature. In the field of inter-personal communication, it often refers to mutual eye contact [2]. In feminist theory, media studies, and game studies, it relates to the concept of the male gaze [57]. Many post-modernist thinkers, such as Sartre, Foucault, and Lacan, presented theories based on the concept of gaze [46]. Conse-
We however disagree with their analysis of indicators based on Almeida et al. also overviewed the literature, dividing the works provide an excellent starting point for reviewing draw when designing eye-enabled games. These works offer a practical toolbox from which to offer game designers a practical toolbox from which to offer game designers a practical toolbox from which to offer game designs. While working at these deeper abstraction levels, we were able to extract a taxonomy of game mechanics that offer game designers a practical toolbox from which to draw when designing eye-enabled games.

Isokoski et al. were the first to overview the use of the eyes for game control, with a focus on assistive interfaces [37]. The authors begin by describing eye movements and their relation to gaming. They classify different works according to how they incorporate eye tracking: emulating mice and keyboards; using additional middleware; modifying the source code of existing games; and building a new game. However, as game engines, level editors, and other prototyping technologies enable the easy creation of new games from scratch, this taxonomy becomes less relevant. In fact, modern eye trackers even offer tools specifically directed at game development (e.g. the Tobii EyeX offers an SDK for the Unity game engine).

They also discuss possibilities for future implementations based on the requirements for different game genres. The authors propose that positive indicators for the use of eye tracking are ‘one-player mode’ and ‘turn-based gameplay’. Negative indicators are ‘online multiplayer’, ‘online real-time multiplayer’, ‘continuous position control’, ‘dissociation of focus of attention and control’, and ‘large number of commands’. They conclude with an analysis of four case studies. We however disagree with their analysis of indicators based on genre. Though a lot of the challenges identified by Isokoski et al. remain true for trying to retro-fit an existing game for eye tracking use, all the challenges posed by eye tracking can be cleverly solved and incorporated into new games with careful interaction design, regardless of the genre.

Sundstedt surveyed the literature on the topic in the form of notes for a SIGGRAPH course [73], which led to a book about the topic [74]. She describes different works on the topic and discusses three cases studies from her own research that investigate the combination of gaze with voice interaction. Almeida et al. also overviewed the literature, dividing the works into eye tracking as input for video games and visual attention studies in video games [1].

These works provide an excellent starting point for reviewing the related literature. However, they limit the scope of their analysis to the level of the papers surveyed. We take a substantially different approach, by delving into the eye movements, interaction techniques, and individual mechanics implemented in different games. While working at these deeper abstraction levels, we were able to extract a taxonomy of game mechanics that offer game designers a practical toolbox from which to draw when designing eye-enabled games.

![Figure 1. Research methodology: From different works, we extracted individual game mechanics and for each of them, we created a card. We then conducted three analyses, clustering the cards by eye movements involved in the mechanic, the type of input and the type of game mechanic.](image)

**METHODOLOGY**

EyePlay research is disparate, irregular and often in little conversation with its peers, consequently utilizing a wide variety of terms. Unlike other similar reviews [10, 44], the conference and disciplinary boundaries of this work are unclear, and a strict Boolean-type review would be inappropriate. We started with the initial set of works surveyed in previous reviews [1, 37, 73, 74]. We found more recent papers through Google Scholar and the ACM Digital Library by searching not only for works that are cited by these surveys but also the works that have cited them since. We repeated this process until it converged into the set that we examined. We also searched for videos on YouTube and magazine articles that demonstrated unpublished prototypes of eye-enabled games, and reviewed the official forums and documentation provided by eye-tracker manufacturers.

We compiled a total of 106 papers, videos, and articles. We discarded works that did not present a concrete implementation of a game mechanic controlled by the eyes (e.g. eye tracking for evaluation), coming down to a total of 70 items. For each of them, we annotated individual game mechanics, and created a Game Mechanic card, containing the reference, a summary of the game mechanic and an image that demonstrates it, for a total of 112 cards (see Figure 1). To categorize these mechanics we conducted a series of workshops, in which the authors started with an abstraction level (eye movements, input type, game mechanic) and used a affinity mapping to cluster them. First, we spread all cards on a large whiteboard placed horizontally on a tabletop. We then organised the cards by placing similar mechanics (according to the abstraction level) close to each other. As patterns started to emerge, we connected related mechanics with a marker, until we achieved a map of themes. In accordance with our methodology, these themes were articulated as memos that formed the basis of the theory development. As our lines of inquiry developed through the research, we returned to the papers and similarly summarized the papers to identify the motivations, methodologies and contributions of the paper to EyePlay research.

As a principal aim of this comprehensive literature review is to contribute a practical foundation for future design and research, we present a bottom-up approach for understanding eyes in games. At the lowest level, we look at the eye movements.
that form the building blocks of eye behaviour. These help us understand what are the capabilities and limitations of players’ eyes when using them for input. One level of abstraction above, we look at how these movements are interpreted in the software side, considering the different input types. Finally, we overview how these are incorporated into the game design as concrete game mechanics.

EYE MOVEMENTS

Our eyes are limited to specific types of movements. Explanations of how each movement works at a physiological level, and how to identify and track them can be found elsewhere [19, 34]. As these movements are rigid categories, we used them as a basic framework for understanding the different ways the eyes have been enabled in HCI games research.

Fixations: These are movements that stabilise the retina over a stationary object of interest so that its image falls on the fovea of the eyes [19]. As it is analogous to pointing, mechanics that use dwell time for selection leverage this kind of behaviour—by setting a threshold that outlasts the normal fixation length, it is possible to estimate intention. Measuring the length of fixations can also hint at players’ attention pattern and other cognitive processes, which can be leveraged for implicit interaction mechanics.

Saccades: These are the rapid movements used in repositioning the fovea as the eyes jump from one fixation to the next [19]. Mechanics that focus on saccades usually involve eye gestures [18], which can be used for gesture matching [8] or mode selection [39]. An advantage of working with saccades is that they can still be detected even with no registration information with the environment, making it well suited for ubiquitous games or trackers that only observe relative movement, such as electrooculography (EOG) ones [8].

Smooth Pursuits: These are the smooth movements that occur when tracking a moving object and keeping the gaze fixated on the target [19]. Whereas saccades are sudden and quick, pursuits are smooth and match the relative velocity of the target. Smooth pursuits cannot be faked—to exhibit this movement, the eyes must be following a moving target. This makes them particularly useful for identifying whether the user is actually following a moving target on the screen by comparing the relative trajectory of the target and of the gaze point. Vidal et al. found that a simple Pearson’s correlation coefficient thresholding algorithm can be used for such detection [91]. EyePlay applications include using these movements to calibrate the eye tracker unobtrusively [26, 66] and to select moving targets [24, 91].

Compensatory Eye Movements: These are involuntary smooth movements that occur when moving the head whilst keeping the eyes fixated at the same point in the visual field [19]. Until recently, the use of these movements for HCI was difficult, because eye trackers needed users’ heads to remain stationary (some even requiring a chin support) in order to track gaze accurately. However, modern trackers are now robust enough to compensate for head movements, and are able to detect that the user is looking at the same spot, even as they translate and rotate their head. Whereas we did not find examples of mechanics that leverage this behaviour, the combination of an eye tracker with the 3D head tracking capabilities of depth cameras can enable its incorporation into games. An simple example could be a virtual environment in which the player can direct nods and other head gestures at different NPCs, depending on which avatar she looks at.

Vergence: These movements focus the eyes at a distant target [19]. The further the target, the more parallel the two eyes will be. In HCI, they are particularly useful for transparent displays, as the system can detect whether the user is looking at the foreground (i.e. at the screen) or the background (i.e. through the screen) [92]. Even though, we did not find applications for vergence in games, this could be incorporated into games like Keyewai [3] and Relationship Tunnel Vision [14], which use a see-through, gaze-enabled display where which players sit across from each other. By observing vergence, the games would be able to detect whether the focus of attention is at the screen or at the other player.

Optokinetic Nystagmus: This is a conjugate eye movement comprised of smooth pursuits interspersed with saccades that enable us to perceive continuously an object that moves across the field of view [19]. These are the movements that happen when we observe a train passing by or we look through the window from inside a moving car. We did not find any direct use of these movements in games, but they nevertheless happen when playing games.

In summary, as eye-tracking technology has improved, more types of eye-movement are increasingly trackable. However, as our review demonstrates, re-articulating the different eye-movements available to game designers is necessary, as Eye-Play research has overwhelmingly focused on fixations. By highlighting other types of movements, we hope to uncover new opportunities for novel eye-based game mechanics.

INPUT TYPE

Modern eye trackers can capture a lot of information beyond the gaze point. From an application perspective, the input captured from a sequence of eye movements and gestures can be classified as discrete, continuous, or as a combination of them. From this perspective, we take an abstract approach to examine individual interaction techniques, illustrating them with example mechanics.

Discrete-Only

These are actions that do not incorporate the point-of-regard as input, but use other eye actions as a trigger. They are broadly categorised as Eye Gestures and Eyelid Gestures. Eye gestures are comprised of a sequence of saccades that correspond to different commands. In games, eye gestures can be used as an end in itself (e.g. in EyeMote, users must repeat the gestures indicated by the game to gain points [8]) or as a replacement for other command triggers. Examples of the latter include selecting different modes of operation [40] and triggering healing and attack spells [4]. Common problems with eye gestures include accidental activation due to the potential overlap with natural search patterns; limited screen real estate for gesture anchors; and saccadic fatigue, as intentional control of saccades is unnatural and tiresome [55].
**Eyelid gestures** include winks and blinks. Winks are voluntary closures of one of the eyes. Winks have been long proposed as an eyes-only trigger to replace mouse clicks [54], and they have been used in games as a trigger while users aim with a pistol [35]. However, not every user is able to wink voluntarily, with some people only being able to wink with one eye, and some not being able to wink at all [65]. Winking also transmits substantial nonverbal content, signaling collusion, shared secrecy, momentary intimacy, flirtation and trust [52, 56]. As a social gesture, Da Silva et al. implemented a game in which a little girl steals the player’s homework, which he can recover by winking at her [15]. Blinking is part of the natural maintenance of the eyes and occurs semi-automatically. Therefore, when used for interactive purposes, blinks must be held for longer to distinguish them from natural blinks.

**Continuous-Only**

These are mechanics that take as input the X and Y coordinates of the point-of-regard on the screen. We identified three ways to implement it: target pursuit, target avoidance, and always-on tracking. **Target pursuit** mechanics are those in which the player actively scans the virtual environments and triggers actions, events or commands by looking at specific objects or regions of the screen. These include triggering social and interactions and attacks with NPCs [3, 13, 78, 90, 95], triggering gaze-contingent objects [27, 70], looking at active regions to control avatar navigation [40, 63, 64, 71, 85, 87] and panning the camera by looking at the edges of the screen [12, 27, 37, 61].

**Target Avoidance** mechanics are those that require the player to look away from objects of interest or even the entirety of the screen. This can be implemented by requiring players to close their eyes (e.g. in *Invisible Eni* the player can close both eyes to make the avatar disappear and hide from enemies [23]) or by triggering negative actions when the user looks at certain areas of the screen (e.g. in *The Royal Corgi* this is implemented in multiple ways, including looking down to show respect, looking away to distract an annoying NPC, and looking away from an NPC to avoid engaging a conversation [90]). Because we are naturally drawn to moving and salient objects, this can offer players a challenging experience.

In **Always-On** tracking, the X and/or Y coordinates of the gaze point are interpreted continuously. This can be used explicitly to continuously control parameters, in one dimension (e.g. the horizontal position of a *Breakout* or *EyeGuitar* paddle [17, 88]) or more (e.g. absolute position of a character or control [14, 62], or speed and camera rotation [76]). Alternatively, it can be used implicitly by making inferences about players based on their visual attention patterns.

**Continuous + Discrete**

These are mechanics that use gaze for pointing and an additional modality for confirmation. The additional modality can be eye-based or not. If so, it is appropriate as an accessible input alternative for disabled users. Eye-based confirmation modalities include winks [35], blinks [16, 49, 69, 95] and gestures [37, 39, 69], as discussed before, as well as dwell time [4, 5, 6, 12, 16, 28, 29, 37, 40, 69, 77], smooth pursuits [91, 93], and voluntary changes in pupil size [23]. Other modalities for confirmation include voice [95], hand movement [97], and mechanical triggers, including buttons [67], mouse clicks [5, 6, 16, 28, 29, 36, 43, 48, 67], key presses [12] and gamepad controllers [35, 36].

Each confirmation modality has its advantages and disadvantages. We have already discussed winks, blinks and eye gestures above. Dwell time activation is known to suffer from the *Midas Touch* problem, i.e. unintentional target activation while users scan the environment [42]. Whereas increasing the dwell time thresholds can minimise this problem, it also slows down the interaction and leads to unnatural fixation behaviour. Smooth pursuits are somewhat more natural than selection by dwell time, as the eyes are naturally drawn to moving objects, reducing strain. However, they are also susceptible to unintentional activation and slow selection times. Selection by voluntary changes in pupil size are possible, but difficult. Such changes can be induced by physical activity, self-induced pain, positive emotions, negative emotions, cognitive tasks, focusing gaze and concentration [22]. Voice interaction has the advantage of offering multiple possible actions and make the player feel more immersed, but it can be slow and susceptible to mistakes in the recognition. For an in-depth treatment of voice interaction, see Harris [31]. Mechanical triggers require additional hardware, but offer an inexpensive, fast and intuitive way of triggering actions.

**GAME MECHANICS**

In this section, we categorise mechanics according to how they are employed as elements of game play. While some are similar and overlapping, our categorisations are meant to assist game designers and thus map to existing mechanics implemented with other modalities. We identified 5 categories which we explore in the following subsections, namely: Navigation, Aiming & Shooting, Selection & Commands, Implicit Interaction, and Visual Effects. Our classification should be seen as a snapshot of current implementations rather than a comprehensive overview of what can be accomplished with the eyes. In concretely categorising what has previously been implemented, we hope to inspire and uncover novel categories of mechanics that can be explored in future work, and assist researchers in avoiding common duplications.

**Navigation**

Navigation mechanics are those that involve moving the player’s avatar through the virtual environment. We identified five subcategories of navigation mechanics: 1:1 mapping, look to go there, virtual buttons, Cartesian spaces, and gradients.

**1:1 Mapping:** In this mechanic, the player’s avatar is positioned exactly where the gaze point is. Because of the jittery and sudden saccadic movements of the eyes, avatars controlled in this fashion tend to “teleport” around the screen. If using this pattern, designers must make sure that this behaviour is supported by corresponding game metaphors. This is usually implemented for non-humanoid avatars, such as in Dorr et al.’s *Breakout*, and in Vickers et al.’s *EyeGuitar* [17, 88]. In both cases, the player is substantiated in the game by a paddle
Look to go there: In this mechanic, the player looks to indicate the end position of the avatar (confirmed by dwell time [23] or other modality [67]) and using pathfinding techniques, the game guides the avatar to the desired position. In Invisible Eni, players fixate at a point to make the character go there [23]. Smith and Graham implemented this mechanic in the game Neverwinter Nights, a third-person RPG with a top-down view [67]. In first-person games this can be implemented by constraining the movement along a pre-defined path and navigating through it by dwelling at anchor points [12].

Virtual Buttons: In this mechanic, gazing at specific regions on the screen triggers a continuous action. This is often implemented by toggling a constant stream of keys, such as WASD. Previous works have explored a wide variety of designs, but there are few empirical comparisons between them, making it difficult to confidently assess what works and what does not. The first decision in the design of virtual buttons regards the number of buttons. Examples in the literature range from two [64, 86, 87] to twelve [71]. Designs with two buttons only enable the lateral rotation of the camera. This means that to move their avatars, players have to use a separate modality (e.g. voice commands as in Rabbit Run [64]) or use another mode of interaction that enables movement [38]. Additional buttons increase functionality of this mechanic, such as sidestepping, diagonal movement and vertical camera rotation [71].

The second decision regards the visual feedback, which can be achieved through opaque buttons, transparent buttons, or through no feedback. Opaque buttons, may communicate their functionality more clearly, but they occlude the content behind them. An example of such implementation can be found in Vickers et al.’s evaluation of the use of a gaze-controlled cursor as input for Second Life [84]. Their results show that because of the discrepancy between the locus of control (i.e. the widget) and the locus of attention (i.e. the environment), users would often get distracted and steer the avatar in the wrong direction. Further, because such controls occlude the game, they tend to be small, which can be a challenge for inaccurate trackers. Semi-transparent buttons allow users to see through the widget while still communicating their functionality explicitly [12, 39, 71]. The final alternative is to not display any controls on the screen. A recent trend in game interface design is to move away from persistent on-screen elements of head-up displays (HUDs) towards more immersive and realistic diegetic representations [96], strengthening the case for Virtual Buttons with no on-screen representation. The different designs that implement this pattern aim at leveraging the natural eye behaviours that players present when performing these actions with other controllers. Examples in the literature include controls for Second Life and World of Warcraft [40, 41, 86, 87].

The third decision relates to the shape of the buttons. We found in the literature three major kinds of shapes: triangular [64, 39], rectangular [12, 40, 41, 86, ?], and elliptical sections [71], with little report of the different usability or justifications of the shapes.

Cartesian space: In this mechanic, the whole screen acts as a Cartesian space where the XY coordinates of the gaze point control the movement. Nielsen et al. implemented this pattern to control a virtual aeroplane [63] (see Figure 3A). However, to fully control an aircraft in 3D space, a third degree-of-freedom is necessary. Hansen et al. assigned this extra degree-of-freedom to the keyboard, and compared different mappings of parameters to the different coordinates of the Cartesian space [30](see Figure 3B). Another possibility of adding additional degrees of freedom is to combine the Cartesian space with virtual buttons at the bottom of the screen, as Tall et al. did for steering a remote control toy car [76] (see Figure 3C).

Gradients: Similarly to Virtual Buttons, looking at different areas on the screen triggers movement and camera rotations (see Figure 3). The difference is that in gradient patterns, the specific point within the region being gazed at modulates the rate of change of the parameter under control (similarly to Cartesian spaces). Because users usually keep their gaze at the centre of the screen, the mapping usually flows from the centre towards the edges, with points further from the centre representing higher values (e.g. of rotation velocity). Stellmach et al. explored different designs for such navigation interfaces [71] (see Figure 3D,E).
Aiming & Shooting

Despite the despair of many concerned parents, shooting is still a widely popular game mechanic. As our eyes are naturally and quickly drawn to points of interest, many games have explored the possibility of augmenting performance by assigning the control of the weapon to players’ gaze. In most shooting games, the four parameters of interest are the character’s movement, the viewport control, the weapon’s aim (the crosshair), and the weapon’s trigger. In the previous section, we covered movement mappings to the eyes. In this section, we cover the viewport, crosshair, and trigger mappings. Table 1 shows how different shooting games mapped the eyes to the four shooting-related mechanics.

**Viewport:** In conventional first- and third-person shooting games, the movement, viewport, and crosshair are inherently coupled—the viewport determines both the direction of movement and the aim of the weapons, as the crosshair remains static at a fixed position (usually at the centre). In most games, the movement is assigned to one hand, and the viewport/direction of movement/crosshair to the other hand.

The viewport control works well with mice and joysticks, as these are *relative* input devices—they measure changes in position and orientation, respectively. This translates well in users’ mental models of viewport control: for the character to turn right, the player moves the mouse to the right. If she wants to turn a bit more, she moves the mouse a bit more. She can repeat this process until she has to clutch the mouse back to the starting position. An eye tracker, however, is an *absolute* input device. This means that it outputs coordinates in the screen frame of reference. Because of the jittery and sudden nature of eye movements, a 1:1 mapping of gaze to the viewport control is infeasible.

We found two types of implementation for controlling the viewport with the eyes. The first is to scroll to the viewport sideways when the user looks close to the edge of the screen. In this pattern, the closer to the edge of the screen the gaze point is, the faster the camera rotates in that direction. An early example of this feature can be found in the work of Gips in an adventure game for disabled children [27]. Castellina and Corno implemented this camera control technique among many others [12]. Isokoski et al. discuss the implementation of this technique for scrolling the camera sideways in the Chicken Run shooting game [37].

<table>
<thead>
<tr>
<th>Game Title</th>
<th>Mov.</th>
<th>Aim</th>
<th>Cam.</th>
<th>Trigger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feyerball Mage [82]</td>
<td>Feet</td>
<td>Eyes</td>
<td>Hands</td>
<td>Pinch</td>
</tr>
<tr>
<td>Quake II [67]</td>
<td>Keys</td>
<td>Eyes</td>
<td>Eyes</td>
<td>Keys</td>
</tr>
<tr>
<td>Half-Life 2 [43]</td>
<td>Keys</td>
<td>Eyes</td>
<td>Eyes</td>
<td>Mouse</td>
</tr>
<tr>
<td>Son of Nor</td>
<td>Keys</td>
<td>Eyes</td>
<td>Mouse</td>
<td>Keys</td>
</tr>
<tr>
<td>Isokoski et al., 2007 [36]</td>
<td>Keys</td>
<td>Eyes</td>
<td>Mouse</td>
<td>Keys</td>
</tr>
<tr>
<td>Half-Life 2[43]</td>
<td>Keys</td>
<td>Eyes</td>
<td>Mouse</td>
<td>Mouse</td>
</tr>
<tr>
<td>A.C. Rogue</td>
<td>Keys</td>
<td>Mouse</td>
<td>Eyes</td>
<td>Mouse</td>
</tr>
<tr>
<td>Battlefield 3 [81]</td>
<td>Keys</td>
<td>Mouse*</td>
<td>Mouse*</td>
<td>Mouse</td>
</tr>
<tr>
<td>Chicken Run [37]</td>
<td>Static</td>
<td>Eyes</td>
<td>Eyes</td>
<td>Look up</td>
</tr>
<tr>
<td>Killer Penguins [95]</td>
<td>Static</td>
<td>Eyes</td>
<td>Static</td>
<td>Blink</td>
</tr>
<tr>
<td>Eye Asteroids [77]</td>
<td>Static</td>
<td>Eyes</td>
<td>Static</td>
<td>Dwell</td>
</tr>
<tr>
<td>Sacrifice [43]</td>
<td>Static</td>
<td>Eyes</td>
<td>Static</td>
<td>Mouse</td>
</tr>
<tr>
<td>Leyba et al., 2004[48]</td>
<td>Static</td>
<td>Eyes</td>
<td>Static</td>
<td>Mouse</td>
</tr>
<tr>
<td>Lunar Command[67]</td>
<td>Static</td>
<td>Eyes</td>
<td>Static</td>
<td>Mouse</td>
</tr>
<tr>
<td>Hülsmann et al., 2011 [35]</td>
<td>Static</td>
<td>Eyes</td>
<td>Static</td>
<td>Pistol</td>
</tr>
<tr>
<td>Frog Game[91, 93]</td>
<td>Static</td>
<td>Eyes</td>
<td>Static</td>
<td>Pursuit</td>
</tr>
<tr>
<td>Killer Penguins[95]</td>
<td>Static</td>
<td>Eyes</td>
<td>Static</td>
<td>Voice</td>
</tr>
<tr>
<td>Hülsmann et al., 2011 [35]</td>
<td>Static</td>
<td>Eyes</td>
<td>Static</td>
<td>Wink</td>
</tr>
<tr>
<td>Hülsmann et al., 2011[35]</td>
<td>Static</td>
<td>Pistol</td>
<td>Static</td>
<td>Mouse</td>
</tr>
</tbody>
</table>

Table 1. Mappings for shooting mechanics. Games without a title are indicated by their authors. *Aim and camera are controlled by the mouse, but modulated by gaze.*

The second type of implementation is to continuously centre the viewport around the gaze point. This can mean that the camera is entirely controlled by gaze or the mouse acts as the primary control while gaze modulates the speed [81]. This mechanic was first proposed by Smith and Graham in their modification of *Quake 2*. In their game, these movement was controlled by WASD keys, and the camera constantly rotated towards the gaze point at a constant speed [67]. Isokoski et al. had a similar implementation for a FPS, but the camera only rotated towards the gaze point when the player pressed a button on the gamepad [36]. Further, the longer the player pressed this button, the faster was the rotation. In Ubisoft’s *Assassin’s Creed Rogue* gaze enables players to re-orient their cameras around the third-person avatar using gaze [68].

**Aiming:** The most common mapping of the eyes in shooting games is assigning the gaze point to control the aim of the weapon. This is a natural mapping, as even in conventional games we tend to look first before shooting the target. In terms of aiming, the crosshair can be controlled directly by the gaze position or controlled by a separate input device and modulated by gaze. As shown in Table 1, most surveyed games use the former implementation. The exception is Velloso et al.’s MAGIC techniques for *Battlefield 3*, in which (rather than having the eyes control the aim directly) the eyes modulated the speed of the aim and viewport, which were primarily controlled by the mouse [81].
**Trigger:** The final possible shooting-related mechanic that can be controlled by the eyes is to trigger the shots. Whereas any discrete technique could be used for this purpose, we found uses of eye gestures [37], blinks [95], winks [35], dwell time[77], and smooth pursuits [91, 93].

**Selection & Commands**
Selecting objects and issuing commands are mechanics existing in virtually every game. We found many diegetic (i.e. belonging to the game world) and non-diegetic instances of eye-based selection and commands. The input type for these mechanics can be discrete-only (e.g. making an eye gesture to invoke a menu) or discrete + continuous (e.g. looking to point and dwelling to select), so similarly to the shooting mechanics discussed above, the interaction techniques can be eyes-only or multimodal.

**Pick & Drop:** Many board, puzzle, and adventure games implemented mechanics that require the player to move objects around the game space. As this is a mechanic that can be fairly easy to implement, there are many examples in which picking and dropping objects with the eyes is used to demonstrate an eye tracker’s capabilities. Examples include card games [97], chess [69], and puzzles [5, 6, 29].

**Action Activation:** These are mechanics in which a given eye behaviour triggers a generic game action. This can range from throwing healing and attack spells with eye gestures [39] to activating some functionality of in-game objects [23]. In particular, it is interesting to think about how other modalities can influence the gameplay when combined with gaze for pointing, especially when the interaction technique matches in-game events. For example, in *Feyereball Mage*, players throw a fireball in the gaze direction by pinching in the air, an action that is repeated by the game avatar [82]. Previous work has shown how voice interaction in games can increase immersion and flow as players’ and characters’ voices overlap [9]. Analogously, activating actions with the eyes has great potential for merging players’ and characters’ gaze, consequently increasing immersion.

**In-Game Widgets:** As a supporting modality, the eyes offer an alternative way to interact with in-game widgets, such as keypads [4] and context menus [12, 78]. This offers a way for separating diegetic and non-diegetic interactions by input modality. For example, whereas primary interactions can remain mapped to conventional controllers, the eyes can support the hands in accomplishing secondary tasks.

**Implicit Interaction**
Because in eye-enabled applications, the eyes serve the dual purpose of not only providing input to the system, but also capturing information from the environment around us, they reveal a lot about our cognitive states and serve an important social role. Therefore, games can use gaze data implicitly, for example to affect social interactions with NPCs, to trigger environmental effects, and to adapt the game AI.

**Social Gaze:** Social aspects of gaze have been a huge element of gameplay even for non-digital games: a glance that gives away the poker player’s bluff; a killing wink in *Wink Murder*; a constant staring at Europe that can predict that player is about to make an attack in *Risk*. Vidal et al. surveyed the Psychology literature and identified eight gaze behaviours that could be implemented as game mechanics: desire for interaction, avoidance of interaction, apparent distraction, cultural disrespect, dominance test, gaze following trigger, signs of intent, and joint attention. In this work, we clustered the individual mechanics by similarity, finding three clusters of behaviours: Engagement/Disengagement, Dominance/Submission, and Shared Attention.

First, gaze can be used to engage and disengage in conversation with other characters. An example suggested by Tobii is that of a player that enters a tavern and looks at the bartender, who greets him when establishing eye contact [78]. Eye contact can also engage multiple NPC’s as in WAYLA, a game where the player looks at NPC couples to harmonise their relationship. This pattern can also be used to punish the player, by forcing them to engage in unwanted conversations. In *The Royal Corgi*, the player must avoid an annoying NPC by not looking at her [80]. An even earlier example is in the work of Da Silva et al., in which the player roams around a university lab and must avoid an evil student by looking down [15]. Another negative effect of engagement by eye contact is attracting the attention of enemies. For example, in *Keyewai*, players are stranded in an island, and looking at the native cannibals make them chase the player [3]. In the same way that eye contact can initiate interactions, looking away can terminate them. This effect can be as mild as simply finishing the conversation, but can also have repercussions for subsequent interactions with the NPC. For example, in *The Royal Corgi*, one of the NPCs demands all the player’s attention and will get offended if the player looks away [80].

A second popular effect of the eyes on NPCs is that of domination and submission. Humans and other animals perceive eye contact as a dominant behaviour. Taken to one extreme, this prolonged eye contact can indicate a lack of respect in certain cultures. Conversely, lack of eye contact can convey a submissive personality. In games, this pattern has been implemented in several ways. Tobii suggests an Agatha Christie-like scenario in which the player-detective looks at different suspects at a crime scene until prolonged eye contact with the killer makes him panic [78]. In *Shynosaurs*, the title creatures become intimidated and run away after the player looks at them for a certain time [89]. This mechanic can also involve multiple players. In *Keyewai*, both players must look at the cannibal to confuse and repel him [3]. Taken to an extreme, this mechanic can completely freeze an enemy—the so-called *Medusa Effect* [60]. Lankes et al. investigated this mechanic in depth in a platformer game, comparing the effects on the player experience depending on the relation between the player’s gaze and the avatar’s gaze [47].

The Dominance/Submission pattern has also been implemented in staring competitions with NPCs, which function as a dominance test. In *The Revenge of the Killer Penguins*, the player must stare at an NPC’s hypnotic eyes [95] and in *The Royal Corgi*, if the player wins the staring competition, the NPC proceeds to exhibit a submissive behaviour [90]. Con-
versely to dominant behaviour, it might be in the player’s best interest to also show submissive behaviour. In The Royal Corgi, to receive support from a certain NPC, the player must look down to show respect [90]. The behaviour required to avoid the evil student in Da Silva et al. can also be interpreted as submissive [15].

A third possible use of social gaze is to indicate shared attention, for example, when the player looks at an object or location during an interaction with an NPC. There are many ways in which this can be interpreted by the game. First, it can be seen as a sign of intent. The premise that guides the design of interaction techniques such as MAGIC pointing [98] is that gaze often precedes action. In a social sense, this can be used by NPCs to predict players future actions. For example, Tobii suggests a football game in which the goalkeeper uses the player’s gaze point to predict where he is going to shoot the ball [78]. An even more interesting use case happens when players can use their gaze behaviour to deceive NPCs. For example, to evade the annoying character in The Royal Corgi, players can look behind her. The character then gets distracted and looks away, giving the player an opportunity to escape the conversation [90].

Such pattern can add richness and complexity to conversations, by incorporating the context of the interaction. This can be used for positive or negative effects on the relationship with the NPC. An example of a positive effect is that of the conversation with the painter in The Royal Corgi. During this conversation, if the player dedicates her attention to this NPC’s favourite painting, she can gain his support [90]. Conversely, also in The Royal Corgi, one of the NPCs will get angry if the player stares at his wife during the conversation [90].

Responsive Environments: Many games offer mechanics that leverage users’ behaviours while exploring the game world to create magical effects. For example, Eagles Eyes, showed instructions when players looked at certain objects and displayed weather information when the player looked at the window [27]. In Starker and Bolt’s prototype, the game played synthesised speed depending on users’ visual attention patterns as they scanned the planet where the story of The Little Prince took place [70]. Dechant’s horror game Sophia triggers scary events as the player looks at different objects in a haunted mansion\(^1\).

Adaptive AI: A third way in which gaze can implicitly influence game play is having the game AI learn from users’ visual attention patterns and predict future actions and intentions. For example, Wetzel et al. adapted the AI’s strategy in Hex [94], Hillaire et al. predicted in which direction players were going to turn when exploring an environment [32], Munoz et al. predicted player’s actions in Super Mario Bros. [58], and Vesterby triggered different narrative branches depending on which character the user looked at the most [83].

Visual Effects
The last category of uses of eye tracking in games involves how the knowledge of the gaze point can enhance the presentation of the game’s visuals to the player. Such effects can range from subtle changes in how graphics are rendered to fully fledged game mechanics.

Depth-of-Field: This subtle effect renders the object being looked at sharply, while blurring objects at different distances. Hillaire et al. implemented gaze-contingent depth-of-field blur in Quake III [33] and Mantiuk et al. suggest that the use of players’ eyes for this effect does increase their preference over a control condition [51]. This effect improves the perception of 3D objects over flat images, but the quantitative depth information conveyed by this channel is limited [53].

Compensating Camera Motion: Another effect that Hillaire et al. implemented for Quake III was based on the focal point [33]. This effect simulates eye movements when walking, such as the vestibulo-ocular reflex, by compensating the camera motion normally created by walking FPS characters based on where the player is looking. The authors suggest that using gaze to compensate this motion is superior to other techniques, improving the sense of immersion and fun.

Hiding and Enhancing: Blinding players is an element of many games, ranging from children’s games, such as Blind Man’s Bluff to more recent exertion games, such as Reindeer & Wolves [25]. Knowledge of players’ gaze point can be used to create blindness by hiding objects and obscuring the field of view. A common pattern found in games is reducing the field of view with a flashlight metaphor. This mechanic effectively eliminates players peripheral vision, which makes it well suited for horror games [78]. It has also been implemented in Bala et al.’s Keyewai, and Velloso et al.’s StarGazing [3, 82]. Conversely, the gaze point can be used to enhance visuals. For example, when retro-fitting games to incorporate eye tracking input, interaction with small targets that would be trivial with a mouse, suddenly become a challenge. To address this problem, Istance et al. implemented a magnifier glass that zooms what is underneath, that can be dropped with gaze where fine grained positioning is required. Also, in graphically intensive games, visuals can be rendered from the gaze point outward, allowing faster frame rates and levels of details where they are required, crucial for applications such as VR [20, 59].

DISCUSSION
In this paper, we presented a survey of 70 works that have interrogated eye interaction in the context of games and play—a research agenda we refer to as EyePlay. From them, we distilled 112 individual game mechanics that employ the eyes in some fashion. This contributed a vocabulary for discussing and designing eye-enabled games, and a comprehensive and practical list of existing mechanics that serves as a resource for design and future work.

Our initial analysis was based on the (immutable) physiological nature of eye-movements, finding that research has overwhelmingly focused on fixations, and that little to no research has used more complex eye movements such as vergence and optokinetic nystagmus in gaze-game research. A further categorisation based on input type highlighted the relevance of eyelid gestures (winks and blinks) to EyePlay research, as well as the prevalence of work that has combined gaze with other forms of input to overcome the Midas Touch problem.

\(^1\)Sophia, https://youtu.be/CRJpxu5NQes
Subsequently, we identified and categorised individual game mechanics in the games surveyed according to how the eyes are used in the different games: navigation, aiming & shooting, selection & commands, implicit interaction, and visual effects. Of these, navigation and shooting were respectively the most common applications for gaze-enabled game mechanics in games research. In addition to highlighting commonly repeated applications of gaze in games, this summary provides a rich resource for design through highlighting typical pitfalls, drawbacks and successes.

In our discussion, we draw attention to the fact that even though game interaction design can be seen as a subset of HCI, designing interaction techniques for gaming follows a substantial set of rules and goals that distinguish EyePlay research from HCI research on gaze-based computer interaction. We demonstrate this through highlighting the many ways that HCI ‘problems’ have been implemented as resources for play in game design. Consequently, we conclude by arguing for the formulation of a formal EyePlay research agenda, as a distinct and emerging field within the CHI Play research community.

**Game Design Solutions to HCI Problems**

Even though game interaction design can be seen as a subset of HCI, designing interaction techniques for gaming follows a different set of rules and goals. Whereas in most domains, an interaction technique can be appropriately evaluated by measuring completion times and error rates, these metrics are not always adequate for evaluating games as they cannot capture social factors, the state of the user and other emotional, sensorial, cognitive, and behavioural factors that affect the game experience [80]. Moreover, specifically related to input design is the notion that mastering the input technology can be part of the play experience, and sometimes even the game’s core mechanic. Examples of the latter include the humorously difficult and frustrating QWOP, which challenges players to control individual joints of the characters legs, and There is Only One Level, which challenges players to discover and master unintuitive mappings.

Given that gaze-enabled interaction is harnessing a behaviour already in use—a fact that HCI research has explored extensively—challenging users’ natural reactions can create novel game mechanics. Conversely, sometimes leveraging them can create uninspiring play experiences. As the eyes are naturally drawn to moving targets, shooting with the eyes can be both a fun mechanic that leverages users’ natural intuitions. However, many games can create fun experiences that go against this natural behaviour. For example, they can force the user to use their peripheral vision to play the game, either by offering positive incentives (e.g. in Shynosaurs, players save the cuties by dragging them with the mouse while staring at the Shynosaurs to scare them away [89]) or negative incentives (e.g. in Virus Hunt, users must touch the viruses to kill them, but if they look at them, the infection spreads [82]).

However, a natural mapping is no guarantee of a positive game experience. For example, in Breakout, players must align the paddle with a bouncing ball to advance in the game. By assigning the control of the paddle to the eyes, Dorr et al. found that all that players had to do was to follow the ball on the screen to beat the game, which effectively removed all challenge [17]. In fact, after playing the game for two minutes, one of their participants asked when the game would actually start. This is not to say that the challenge to master the interface can be an excuse for poor mappings, but that careful balance between challenge and intuitiveness is paramount for a positive play experience. As a consequence, we look at the game genres that Isokoski et al. recommend against, not with discouragement, but as opportunities for ingenious game design solutions to these HCI problems [37].

The capacity of game design solutions to overcome an input device’s shortcomings offers fertile grounds for rethinking the challenges presented by many interaction techniques. In this paper, we consider the three common problems for eye-based interaction found in the literature; the Midas Touch, inaccuracies of the tracking, and the double role of the eyes for observation and control [72].

The Midas Touch is a well-known problem based on the difficulty of distinguishing between an intentional gaze interaction (for example, selecting a target with dwell) from natural eye behaviour. This leads to the unintentional activation of targets, due to the continuous tracking of the gaze point. A huge number of projects in HCI have addressed this problem, but games can take advantage of it in different ways. On one hand, the game might encourage players to scan the whole environment: in Starker and Bolt’s work, the more regions on the planet the player looked at, the more of the story she would uncover [70]. On the other hand, the game can penalise players for looking at certain areas, such as having an enemy chase them in case they look at them [3].

**Inaccuracies** in gaze tracking come in three forms. First, due to the jittery movement of the eyes, the estimated gaze point is less like a pixel and more like a thumb-sized region on the screen. Second, issues in the calibration can lead to an offset between the ground truth and the estimated gaze point. Third, depending on the tracker’s capabilities, a significant delay can be incurred in the estimation which can add perceivable lag in the interaction. Typical HCI solutions to these ‘problems’ include increasing the size of interactive objects, snapping user’s gaze to interactive areas, and ‘smoothing’ gaze feedback [50].

In game design, these inaccuracies can be dealt with in many ways. In games that require pointing, such as in adventure games, inaccuracies can be dealt with similarly to HCI. However, the dynamic nature of games creates an opportunity for correcting the calibration on-the-fly. For example, by correlating the movement of the eyes with the movement of in-game characters, an algorithm such as Pfeuffer et al.’s Pursuit Calibration, would be able to update the calibration dynamically [66]. Second, depending on the game this might not be a problem to begin with. For example, many passive eye-based mechanics can deal elegantly with ambiguity: in the social mechanic in which characters start a conversation with eye contact, a false activation due to inaccuracies could be attributed to the outgoing personality of the NPC, who starts conversations even though no eye contact has been established. Finally, inaccuracies can be leveraged within the game to increase the challenge. This is sometimes desirable in shooter
games, where perfect accuracy does not reflect the way in which guns shoot in the real world, and where the mastery of the input device is a core component of the game’s challenge.

When eye-tracking is introduced, the eyes serve a double-role for observation and input, introducing numerous challenges, from the Midas Touch to social and privacy concerns. This duality is very well illustrated by the attention dilemma explored in Shynosaurs: on one hand, the player must look at the cuties to place them in the pen, but, on the other hand, players must stare at the Shynosaurs to scare them away [89]. Similarly, the social implications of gaze-monitoring are explored in The Royal Corgi, where a gaze towards the Treasurer’s attractive wife can get the player in trouble [90].

While we note that these examples are hardly exhaustive, our goal is to highlight how gaze-interaction challenges which HCI has long perceived as shortcomings of the input device can be elegantly incorporated into the game to create novel experiences. Alongside the proliferation of gaze trackers and the NUI game interfaces, the different motivations, methods and contributions of gaze-enabled games research calls for the articulation and formulation of a novel research program.

EyePlay as a Research Program

This paper has so far demonstrated that there is a wide variety of research that has engaged with gaze in games, but that this body of research is often repetitive, disparate, irregular and often in little conversation with its peers. Often this work is motivated by, or situated within, broader HCI gaze research where games are used to evaluate an interaction technique, or gaze for accessibility research. As noted above, the transfer of new knowledge between these domains is not always direct. We argue that, as HCI games research is increasingly established as a distinct research discipline within HCI, and as gaze-tracking devices increase in commercial availability (and look to be embedded in the next generation of VR and AR devices), there is the emergence of a distinct research program within HCI games research, that of EyePlay Research.

EyePlay is thus a research program within HCI games research concerned with game experiences that are only possible through gaze-based interaction. Our call to formalizing the boundaries of this research area in this way is specifically intended to allow researchers to more clearly situate the motivations and goals of their work, avoid duplication and more clearly contribute to game design practice. Our emphasis on game experiences reflect the emergence of HCI games research as distinct from broader HCI work (consider CHI Play, established in 2014, the Games and Play subcommittee at CHI, and also see [10]), allowing for work which considers the creative and ludic potential of interfaces in games, while limiting EyePlay to those types of experiences only possible through gaze excludes the excellent and valuable work in research with eyes-only interaction as alternative input for disabled users which is often of little relevance or contribution to the gaming industry-at-large. Despite the breadth of work surveyed in this paper that has explored different means for gaze-only virtual world navigation, none have exceeded the capacity of the keyboard and mouse or game controller.

EyePlay is consequently motivated by improving and creating new game experiences, and the presentation of results that are accessible and relevant to game design practice. This is important because the majority of mechanics surveyed in this paper are aimed at desktop computer use, unsurprising considering both laboratory and consumer grade trackers are aimed at this context. As gaze-tracking becomes increasingly available in consumer devices (such as in the affordable Tobii Eyex, or MSI gaming laptop with built-in gaze tracking), EyePlay research has the potential to contribute meaningfully to game development.

Furthermore, this survey and our articulation of the EyePlay research program draws attention to significant gaps in the potential contribution of gaze-games research to industry. The emerging paradigms of computer interaction (natural user interfaces, virtual and/or augmented reality) are likely to vastly increase the ubiquity of gaze-tracking devices in non-desktop context, in which we found few examples of EyePlay research. Although likely remain too expensive for widespread adoption in the short term, head-mounted VR or AR devices presents prime opportunity for low-impact, high-accuracy inclusion of gaze-tracking, and many of the mechanics identified in our review could immediately be transplanted to these contexts. As games are likely to be a core driver of the adoption of these new computer interfaces, EyePlay presents a clear opportunity for HCI games research to have an impact on this area.

Similarly, even though the current eye tracking capabilities of mobile phones and tablets are substantially inferior to those of desktop eye trackers, recent models offer some rudimentary eye and head tracking capabilities for scrolling. However, several research works have investigated gaze interactions with small screens, including phones [21], tablets [79], and even smart watches [24]. Mobile gaming has completely revolutionised the gaming industry despite not providing input as precise as the mouse or gamepads, due to the size of users’ fingers. We believe that even rudimentary eye tracking capabilities can create compelling playful experiences if well designed and distinguishing EyePlay research from HCI, interaction design work—overwhelmingly focused on accuracy [11]—creates a context for this future work.

CONCLUSION

In this paper, informed by an extensive literature review we compiled and classified game mechanics that involve the eyes form three different perspectives. By doing so, we contribute a practical toolbox that game designers can draw from to incorporate in their games and to inspire new ones. Noting that gaze-enabled games research has been highly disparate and conducted over many years, this paper has provided a thorough overview of existing work in order to provide a clear resource for design and highlight common opportunities and pitfalls. In doing so, it becomes clear how the ludic nature of gaze in games calls for a distinct research agenda—EyePlay—which we articulate based on its methods, motivations, research context and contributions. In doing so, we note the existing limitations and gaps in gaze-enabled games research, and call for future work as gaze-interaction continues to proliferate commercially as well as into novel devices and domains.
REFERENCES


