

An empirical comparison of first-person shooter information displays: HUDs, diegetic displays, and spatial representations



Margaree Peacocke¹, Robert J. Teather^{*,2}, Jacques Carette³, I. Scott MacKenzie⁴,
Victoria McArthur⁵

G-ScalE Lab, Dept. of Computing & Software, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4L8, Canada

ARTICLE INFO

Keywords:

First-person shooter
Video games
Information displays
Diegetic
Head-up displays
User interfaces

ABSTRACT

We present four experiments comparing player performance between several information displays used in first-person shooter (FPS) games. Broadly, these information displays included heads-up displays (HUDs), and alternatives such as spatial representations, and diegetic (in-game) indicators. Each experiment isolated a specific task common to FPS games: (1) monitoring ammunition, (2) monitoring health, (3) matching the weapon to the situation, and (4) navigating the environment. Correspondingly, each experiment studied a different information type, specifically ammunition (ammo) levels, health levels, current weapon, and navigation aids, while comparing HUDs to alternatives. The goal was to expose player performance differences between different classes of displays, and types of information displays (e.g., numeric, iconic, etc.). Results suggest that no one display type – HUDs or alternatives – are universally best; each performed well, depending on the type of information. For ammo, player performance was best with diegetic/spatial displays; for health information, players performed significantly better with a HUD. For weapon displays, results were best when showing a redundant HUD icon and a diegetic/spatial display (the actual weapon). Finally, for navigation, a spatial “navigation line” (showing the path) was best, but HUD-based mini-maps offered competitive player performance. We discuss implications for the design of first-person shooter games.

1. Introduction

In first-person shooter (FPS) games, the player sees the world from the eyes of a gun-wielding avatar, completing missions and shooting enemies. FPS games are wildly popular. The NPD Group reports that 3 of the top 10 bestselling games of 2015 were FPS games [33]. They are also highly profitable. For example, Activision’s *Call of Duty: Modern Warfare 3* earned \$400 million within 24 h of release and \$1 billion within 16 days [14]. Player engagement is crucial to their success.

FPS games are interesting platforms for human-computer interaction (HCI) research. Due to the genre’s success and large user base, improvements to FPS user interfaces (UIs) affects many users. Consequently, there is a large body of research on FPS games [21,18,10,20,2,15,8,12,32,36]. Previous work in this area generally focuses on input-related issues, for example, aiming [21,32] or input devices [19,36]. While these input-related tasks are undoubtedly

critical in FPS UIs, information displays within FPS games have been comparatively underexplored. We thus focus on the output-related task of effectively displaying and conveying game information to the player.

Feedback is long recognized as crucial in user interface design [27,28]. When displaying game information, “feedback is crucial for player learning and satisfaction with the game” [28]. Most games, until recently, have traditionally used heads-up displays (HUDs) to present the player with relevant information. These HUDs usually display critical information (e.g., health or ammo levels) as numbers, icons, or meters displayed around the edge of the screen. Schaffer [30] argues that since HUDs on the periphery of the display occupy little game space, they are not likely to distract from gameplay. Fig. 1 depicts example HUDs from *Call of Duty: Strike Team*, *Tom Clancy’s Rainbow Six: Vegas*, and *Call of Duty: Ghosts*.

Game designers increasingly seek to produce more *immersive* experiences. Immersion occurs when players “voluntarily adopt the game

* Corresponding author at: 230 G Azrieli Pavilion, School of Information Technology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada.

E-mail address: Robert.Teather@carleton.ca (R.J. Teather).

¹ Present address: D2L, 151 Charles St. W. Suite 400, Waterloo, ON N2G 1H6, Canada.

² Present address: School of Information Technology, Carleton University, 1125 Colonel By Dr, Ottawa, ON K1S 5B6, Canada.

³ Present address: Dept. of Computing & Software, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4L8, Canada.

⁴ Present address: Department of Electrical Engineering and Computer Science, York University, 4700 Keele St, Toronto, ON M3J 1P3, Canada.

⁵ Present address: Institute of Communication, Culture, Information and Technology, University of Toronto Mississauga, 3359 Mississauga Road, Mississauga, ON L5L 1C6, Canada.



Fig. 1. Example HUD displays. (a) *Call of Duty: Strike Team*, depicting controls (soft buttons, left-side), health (variation of bar), and ammunition as a number and bar (b) *Tom Clancy's Rainbow Six: Vegas*, depicting ammunition numerically (c) *Call of Duty: Ghosts* depicting ammunition both numerically and as a bar/meter.

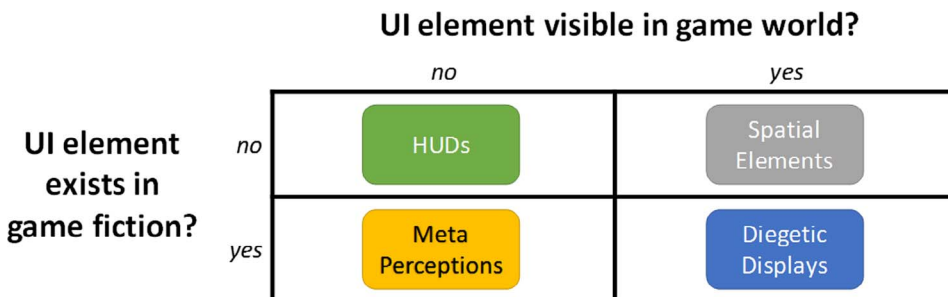


Fig. 2. A classification of game UI elements, based on diegesis (if the UI exists in the fictional game world) and spatial orientation (if the UI element is visualized as part of the 3D game space). Figure based on that of Fagerholt and Lorentzon ([10], p. 51), but simplified to include only display types found in our experiments.

world as a primary world and reason from the character’s point of view” ([10], p. 69). HUDs may compromise immersion as they are not part of the game world, but rather overlaid on it. A comparatively recent trend in games is the use of alternative displays that may better preserve, or even enhance, player immersion over HUDs [34]. The alternatives include geometric/spatial representations, meta perceptions, and diegetic elements [10]. Fagerholt and Lorentzon categorized game displays according to whether they are presented in the game world (spatiality) and whether they are part of the game fiction (diegesis). This suggests a two-dimensional model, as seen in Fig. 2.

Traditional HUDs fall within the upper left quadrant (Fig. 2, since they are neither part of the fictional game world nor are rendered (graphically) in the 3D game space. Conversely, diegetic displays are rendered in the game world, but are also a part of the game fiction. The term *diegetic*, borrowed from film, refers to elements within the game world that are part of the game world. In film, any element that can be perceived by the characters of the story is said to be diegetic [3]. By contrast, elements such as the musical score or subtitles are non-diegetic, since they are not perceived by the characters. By extension, in gaming, a diegetic UI element is part of the game world and is visible and/or audible to the characters in the game world while simultaneously providing information to the player [10]. Information displayed in the game and recognized in the game fiction is considered diegetic. In contrast, both spatial representations and meta representations are typically non-diegetic, and involve non-traditional UI elements or on-screen placements. Spatial representations are elements within the game's 3D space. Meta perceptions are not displayed in the game world. Instead, they are displayed in a fashion similar to HUDs, despite representing something in the game fiction. Common examples include “blood spatter” or cracked glass when the player takes damage. We collectively refer to these non-HUD displays as “alternative” displays (i.e., alternatives to HUDs).

These classes of displays are equivalent in the *type* of information they present the player; their comparative and objective *effectiveness* in presenting information is the primary motivation for our research. For example, an ammunition counter could be displayed as a component of a weapon, making the information seemingly visible to both the player and avatar in a singular format. In this case, the ammunition counter would be visible within the game space and is part of the game fiction, and hence would be considered a diegetic display. Alternatively, the same information could be presented as a label, counter, or icons on the

HUD. However, it is unclear which display type provides the player the most efficient way to quickly intake the information. For example, although HUDs often present information in a clear and concise way, this is not necessarily universal. Consider ammunition presented as a number compared to a meter (e.g., Fig. 1c). The number presents the information more clearly. Moreover, in addition to potential immersion benefits, alternative displays may also offer player *performance* advantages over HUDs. The information presented by diegetic displays, for instance, can be displayed centrally, decreasing the need to glance at the screen edges. Developers of several FPS (as well as third-person shooter) games have employed diegetic displays, typically to enhance immersion [17]. Some of the better-known examples include *Metro 2033* and *Dead Space*, although many games use variants of these. See Fig. 3 for some common examples.

Understanding the relative performance considerations of alternative displays may be important in the design of small-screen games, e.g., on mobile devices, now a bigger market than either the console or PC markets [24]. These small screens drive us to ask about *efficient* use of screen real estate. Alternative displays may help since they require less screen real-estate than HUDs – they are embedded in the game, rather than covering parts of it. When screen real estate is at a premium, developers must consider the most effective way to convey information to the player. Consider, for example, porting a PC game that makes heavy use of HUDs to a small-screen device. If diegetic options offer comparable user performance, then changing the HUDs to diegetic displays on the mobile app may be preferable to cluttering the interface with HUDs. However, to make such a decision, the performance differences between these display types must first be understood.

Our research focuses on the effectiveness – in terms of quantitative user performance – of in-game information displays. There is little quantitative research on the performance offered by diegetic displays. Most work in this area is qualitative [10,20,12] or focuses on immersion [2,16]. While past research makes a convincing case that diegetic displays are more coherent with an escapist philosophy of game enjoyment, it is unclear if this view is consistent with the more utilitarian view of performance-based enjoyment. We note that player enjoyment, performance, and immersion are not necessarily aligned. For example, players may *enjoy* using a diegetic display, even if it offers demonstrably worse performance than an equivalent HUD. That said, if a display is sufficiently ineffective, it may negatively impact player experience – much like any other bad user interface. We thus argue that

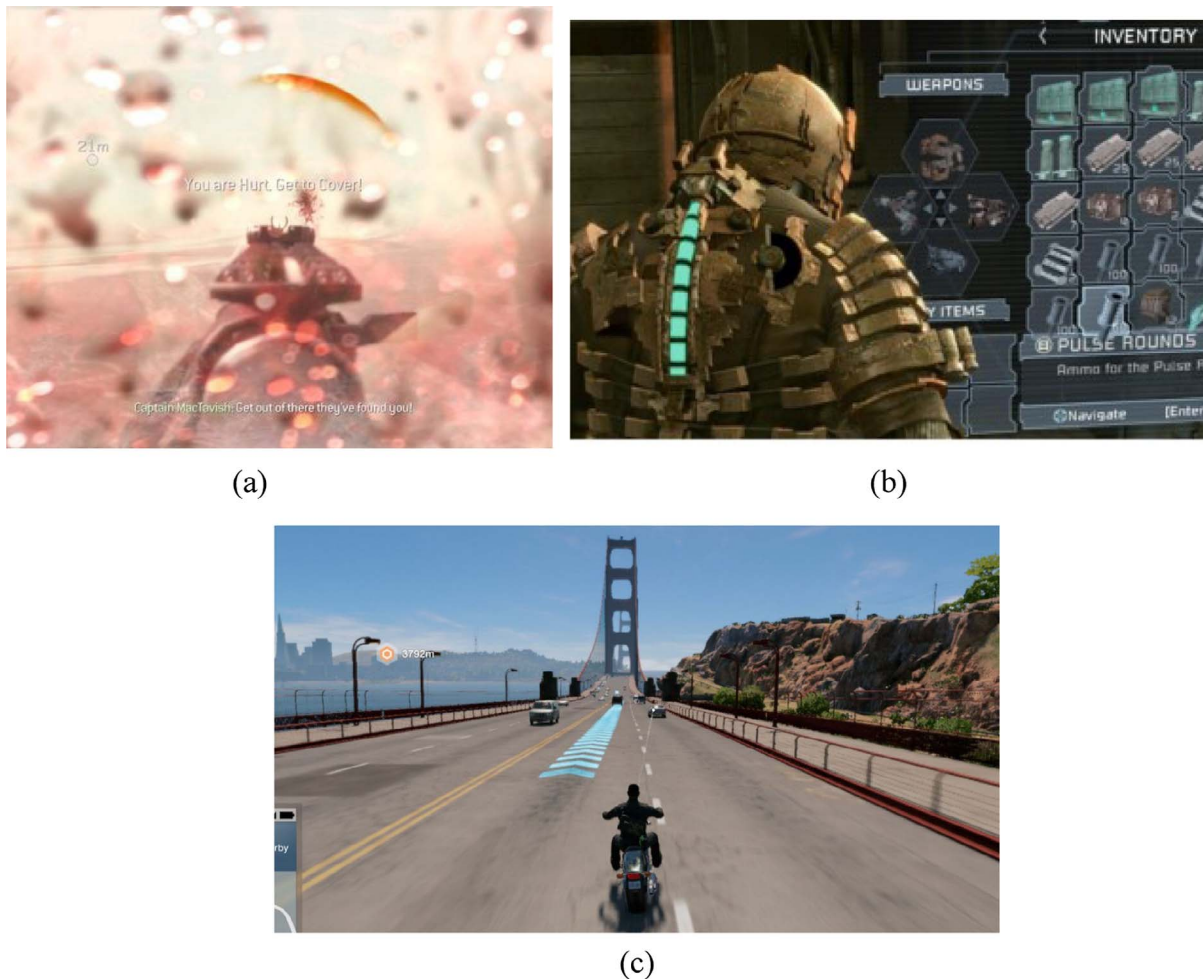


Fig. 3. In-game displays. (a) *Call of Duty: Modern Warfare 2* makes use of a meta-perception (blood splatter) to represent health information. (b) *Dead Space* displays the health meter (cyan bar mounted on player's back) diegetically. The in-game inventory is also presented like an augmented reality display floating in front of the player. (c) *Watch Dogs* uses a spatial navigation display which is visually similar to *Dead Space*'s diegetic navigation aid. Unlike *Dead Space*, this navigation aid is not part of the game fiction (it is visible to the player, not the character) and hence is not considered diegetic.

research in this area must also examine performance trade-offs between the different display types. Although we are also interested in the relationship between player experience and display type, our current work focuses exclusively on player performance.

In this context, our primary research question is thus empirical: How do various information displays affect in-game success with the “micro-tasks” that make up a game experience? Our work presents what is, to our knowledge, the first experimental comparison of the performance offered by HUDs compared to alternatives such as diegetic displays or spatial representations. We first present an analysis of recent FPS games to identify the most important information displayed during gameplay. This analysis and previous work [29] identified four types of information common to modern FPS games: health, ammunition level, the player's current weapon, and navigation aid. This led us to conduct four experiments comparing the information displays commonly used for each information type. These experiments are not intended as a series; rather, they represent a cross-section of tasks common to most FPS games, isolated to offer greater experimental control. In each experiment, we included both HUD and alternative display options. Our “diegetic” display options do not technically qualify as such, as defined by Fagerholt and Lorentzon [10]. After all, our experimental platform – a custom-developed game – presents no in-game fiction. In the absence of in-game fiction, diegetic displays and spatial representations essentially collapse into the same class of display, as do HUDs and meta-

perceptions. In terms of game mechanics, and in isolation from enjoyment or immersion, diegetic displays and spatial representations should offer comparable player performance. Consequently, the displays included in our study align with commercial games, as found in our initial analysis. They include HUDs, meta-perceptions, spatial representations, and diegetic displays. Our analysis further revealed that there is often a mix of HUDs and alternative displays within the same game. This variety is reflected in the options studied herein. Our study used a custom-developed FPS game, which offers better experimental control than commercial games [25,31], and avoids participant biases towards existing games [10]. The displays selected are discussed further in the Methodology section for each experiment.

We solicited participants who regularly play FPS games, since skilled gamers can quickly assess their status, while novice players cannot [8]. Hence expert gamers should be skilled enough to elicit differences between the conditions studied. In contrast, novice participants require training to get to this level of skill, and thus may not expose differences between the experimental conditions. From an experiment design point of view, this decision makes sense. Novice participants introduce a greater degree of variability in performance measures. Statistically, this means novice participants are less likely to produce statistically significant differences, despite potentially large differences in the conditions studied [22].

2. Related work

There is considerable research on FPS games in the HCI literature [19,26,32,36]. The general themes are aiming [32], targeting in the context of pointing [21,23,36], or metrics for empirical evaluation of FPS UIs [19]. Other work focused on developing or evaluating input devices [18,26,15] or immersion [5,10,2,16]. To date, *performance* comparisons of HUD and alternative displays have received little attention.

Research on FPS information displays is primarily qualitative [10,20,12]. Past results indicate that participants support the use of diegetic displays to enhance immersion, *as long as it does not impact performance*. Similarly, if information is clearly communicated, players tend not to care how it is displayed [20,12]. Applicable game design heuristics also exist. Federoff states that “the interface should be as non-intrusive as possible” [11] and that “a player should always be able to identify their score/status in a game” [11]. We argue that empirical studies on the effectiveness of these displays are needed to complement existing qualitative work and to improve design heuristics.

While there is little empirical work on player *performance* with diegetic and non-diegetic game elements, several studies have focused on the immersive qualities of diegetic displays. For example, Babu [2] compared immersion levels in two games with diegetic displays (*Metro 2033* and *Dead Space*) and two games with HUDs (*Bioshock* and *Resident Evil 5*). Immersion was assessed through self-reporting on a 5-point Likert scale and was not significantly different between display types. Participants instead suggested that graphics and storyline had a stronger impact on immersion. Recent work by Iacovides et al. [16] revealed that diegetic displays can indeed enhance immersion, but the effect is stronger for expert gamers.

Galloway [13] introduced the terms diegetic and non-diegetic to the study of video games. The terms originated in literary and film theory. He defines game diegesis as “the game’s total world of narrative action” [13], and non-diegetic as “gamic elements that are inside the total gamic apparatus yet outside the portion of the apparatus that constitutes a pretend world of character and story” [13]. He concludes that HUDs are non-diegetic elements.

Fagerholt and Lorentzon [10] also explore diegesis, developing a descriptive model categorizing FPS UI elements on whether the element exists in the fictional world and is a part of the game space. They recommend considering the game’s fiction when deciding if information should be displayed diegetically, arguing that game coherence is paramount. For example, diegetic options make sense in a game like *Dead Space*, as the futuristic setting allows designers to cast diegetic displays as future technologies such as augmented reality displays or holograms. Ultimately, Fagerholt and Lorentzon suggest using diegetic displays whenever appropriate and where game cohesion can be retained. However, the merit of this suggestion is questionable in the absence of empirical results assessing the potential performance impact of such a design choice. For FPS games, we believe that player performance is ultimately the most important factor.

Fragoso [12] conducted a qualitative study of the effects of diegetic displays on player immersion. Participants played EA’s *Battlefield 3*, which is considered more immersive than other games due to the minimal use of a HUD. For example, the game employs “blood splatter” as health status – as the player takes damage, the screen becomes increasingly occluded by blood – rather than using more common health bars or numeric displays. Participants reported that the lack of meaningful feedback was disruptive, due to the vagueness of the displays. The authors conclude that effective feedback is more important than realism. They further report that HUDs were less disruptive than their diegetic counterparts. These sentiments are echoed by Llanos and Jørgensen [20] who report that while players liked the aesthetic of diegetic displays, they preferred clear communication. However, they also note that players were annoyed when excessive information is displayed on HUDs. More recent work by Caroux et al. [6] suggest that

differences in player experience between HUDs and diegetic options are dependent on player expertise and game genre.

Zammitto [35] conducted a visual analysis of Valve’s *Half Life 2* to assess if visualization design principles were applied to presenting game information. She notes that the game applied two principles to the HUD ammunition display: silhouette and colour coding. These were implemented by (i) showing a bullet icon when the player should reload their weapon and (ii) changing the ammunition indicator from yellow to red when ammunition was low. Red is appropriate because it connotes “danger” in Western cultures. A similar approach was used in the game’s health indicators. Overall, Zammitto concluded that information visualization is not well used in video games.

Bowman et al. [4] share this sentiment, noting that because data visualization in games is new, it is relatively underutilized. They analyzed visualization in games and proposed a design framework. Their “primary purpose” dimension classifies critical game information as Status, noting that “visual representations are often chosen in lieu of a simple number ... because the game designers feel that visualization is more immersive and easier to read quickly” ([4], p. 1961). They recommend considering the target audience to ensure that “the visualization is in spirit with the game’s atmosphere and integrated within the game” ([4], p. 1962). This is consistent with the recommendation of Fagerholt and Lorentzon [10]. The consensus is that FPS players value cohesion in games and that proper data visualizations improve players’ situational awareness.

3. Analysis of current games

Before presenting our experiments comparing several information display types, we first explain how we determined which classes of information were most critical (i.e., ammunition, health, current weapon, navigation aid) and specifically which display types are most commonly used in modern shooter games (i.e., diegetic displays, HUDs, etc.). We analyzed several recent and popular shooter games (not exclusively FPS) across multiple platforms. The games were chosen because of sales and awards, and because they are available for large and small screen platforms. As discussed earlier, the latter necessitates UI changes between display sizes, which may favour the use of in-game displays (e.g., diegetic, spatial). The purpose of this analysis was to learn what types of information were consistently displayed in shooter games, and what types of displays were most commonly used – not only in terms of the display class (e.g., HUDs vs. diegetic), but also their presentation (e.g., numbers vs. icons, etc.). Our intent was to focus on the most critical information, to inform the design of our experiments. The games analyzed included Activision’s *Call of Duty: Strike Team*, *Call of Duty: Black Ops*, *Call of Duty: Ghosts*, Ubisoft’s *Tom Clancy’s Rainbow Six: Vegas*, Bioware’s *Mass Effect 3*, *Mass Effect Infiltrator* and EA’s *Dead Space*. The analysis involved playing these games, watching gameplay videos, and reading publicly available game reviews.

We found that four types of information were common to every game analyzed. These included player health, ammunition level, current weapon, and navigation aid. We thus conclude that these are the most important information displays in FPS games. The display methods used for health, ammo, weapon, and navigation are shown in Table 1. Games using alternative displays are shaded, with the alternative (non-HUD) option set in boldface. Note that navigation aid is mostly displayed as a mini-map in multiplayer mode, but as a navigation arrow in single-player campaign mode. Our analysis focused exclusively on single-player campaign modes.

Numeric displays show a numeric count, typically on a HUD (see Fig. 1b). They are useful for displaying “amounts of things for which you would normally use digits in the real world” [1], such as ammunition. Numeric displays are especially useful for large quantities.

Bars (see Fig. 1c) are also useful for large quantities [1]. These are often presented as a meter that is full at the maximum quantity, and empties as appropriate. The primary benefit is that bars can be

Table 1

Analysis of current game displays for health, remaining ammunition, and current weapon. Alternative (non-HUD) options are set in boldface font. Displays with at least one alternative (e.g., diegetic, spatial, etc.) option are highlighted.

Game	Platform (Year)	Health Display	Ammo Display	Weapon Display	Navigation Aid
Call of Duty: Strike Team	iOS (2013)	Bar	Icons-on-HUD + Number-on-HUD	Icon + In Front	Arrow
Call of Duty: Black Ops	PC (2010)	Blood Spatter	Number-on-HUD	Name + In Front	Arrow
Call of Duty: Black Ops	Nintendo DS (2010)	Blood Spatter	Number-on-HUD	Name	Mini-map
Call of Duty: Ghosts	PC (2013)	Blood Spatter	Number-on-HUD + Bar-on-HUD	In Front	Arrow
Tom Clancy's Rainbow Six: Vegas	Sony PSP (2007)	Bar	Icons-on-HUD	Name + Icon + In Front	Arrow
Tom Clancy's Rainbow Six: Vegas	PC (2006)	Blood Spatter	Number-on-HUD	Name + In Front	Arrow
Mass Effect 3	PC (2012)	Bar	Number-on-HUD + Bar-on-HUD	Icon + In Front	Arrow
Mass Effect Infiltrator	iOS (2012)	Bar	Bar-on-HUD	Icon + In Front	Arrow
Dead Space	PC (2008)	Bar in game	Number-in-game	In Front	Arrow/Line
Dead Space	iOS (2011)	Bar in game	Number-in-game	In Front	Arrow/Line
Metro 2033	PC (2010)	Blood Spatter	Icons-in-game	In Front	Compass
Halo 4	Xbox 360 (2012)	Bar	Number-in-game + Number-on-HUD + Icons-on-HUD	In Front + Icon	Mini-map



Fig. 4. Experimental setup and hardware apparatus. (a) A participant performing the experiment. (b) Annotated Xbox One controller.

interpreted at a glance.

Icon bars (Fig. 1c), or “small multiples” [1], are best for small-integer numeric data. Icons are thus useful for indicating the quantities of around five items or less. Players have difficulty taking in greater than five items at a glance, and thus have greater difficulty remembering the number. However, Adams suggests using graphical indicators rather than text or numbers because they are easier to read at a glance [1]. Our analysis indicates that bar and numeric displays are commonly used together. This offers players the ability to both read at a glance and receive more detailed information as desired.

Our analysis reveals some consistency in alternative health displays (favouring “blood spatter” – a meta perception) and weapon displays (typically diegetically displaying the weapon in front of the player, often with redundant HUD-based icons showing the current weapon). However, there is little consistency in ammunition displays. As seen in Table 1, there is great variety in the presentation of ammo HUDs, and even some diegetic options. Since alternative displays are relatively new (compared to HUD-based displays), design standards have yet to evolve and it is important to develop best practices early. EA’s *Dead Space* (see Fig. 3b) has been praised for its lack of a HUD, relying instead on diegetic displays. In *Dead Space*, ammo is displayed using a numeric count positioned directly above the weapon, and health as a bar physically mounted on the player’s back, which is coherent with its futuristic theming.

4. Common methodology

The following Sections 5–8 present four user studies designed to evaluate different UI display options in an instrumented FPS game. The studies were designed to compare different UI options (traditional HUDs vs. alternative displays) for ammunition, health, weapons, and navigation respectively. The design of each HUD and alternative UI display for each experiment was motivated by the analysis of commercial FPS games presented in the previous section. Section 4 provides details common to the four experiments here. Later sections provide experiment-specific details.

4.1. Apparatus

The experiments were conducted on a 3.4 GHz quad-core i7-based PC, with 8 GB of RAM running Windows 7. A 75-in. Samsung *Series 7 7100 Smart TV* (1920 × 1080 pixel resolution) was used for the display. The display ran in game mode to minimize latency. Participants were seated approximately 15 ft. from the screen, which allowed viewing of the entire display without excessive gaze shifts. The setup is shown in Fig. 4a.

A custom game was developed using Unity Technologies’ *Unity 4.5* engine. The game presented a first-person view to the player, consistent with the appearance of a typical FPS game. It supported typical controls offered by FPS games, such as turning, moving, and shooting. However,



Fig. 5. The custom FPS game (showing the ammo experiment task).

since we used experimental software and not a “true” game, there was no fiction, and no in-game objectives beyond those set for each experiment (e.g., shoot enemies, navigate a maze, etc.). Moreover, certain controls were enabled/disabled as necessary in each of the four experiments. For example, movement was disabled in three of the four experiments. Further details of the game appear in the following experiment Sections 5–8. Participants used a Microsoft *Xbox One* controller to play the game. See Fig. 4b. Viewpoint rotation/aiming was controlled by the right analog stick in all experiments. In experiments that included shooting, the right trigger was used to shoot. The game (during the ammo experiment) is depicted in Fig. 5.

4.2. Procedure

Upon arrival, participants were greeted and the purpose of the experiment was explained. Participants gave informed consent before proceeding. They were instructed on the mechanics of each display and the controls. They were then allowed to begin. Upon finishing all trials, participants completed a questionnaire asking about prior experience with FPS games and soliciting subjective feedback.

5. Experiment 1: Ammunition displays

The purpose of this experiment was to assess user performance differences in a simulated ammo monitoring task. Participants were presented with five different ammunition displays – including HUD and spatial/diegetic displays – and tasked with monitoring ammo while shooting enemies.

5.1. Participants

Twenty paid participants (16 male) took part in the study. Ages ranged from 18 to 38 years (*mean* 22.35, *SD* 4.31). Half reported that their preferred system was a console and the other half reported PC. All participants were regular gamers, playing between 1 and 10 h per week. Sixteen participants reported playing FPS games every week.

5.2. Apparatus

The player’s ammunition level was displayed using one of the five ammunition displays shown in Fig. 6. The ammunition displays included bar-on-HUD (BH), number-on-HUD (NH), icons-on-HUD (IH), number-in-game (NG), and icons-in-game (IG). Each presented the same information, but visualized it differently. The three HUD-based options were similar to those described earlier, presenting ammo as either a number, a “meter”, or as a set of icons displayed at the lower right corner of the display. The alternative options (NG and IG) were

modeled after diegetic displays found in games like Microsoft Studios’ *Halo 4* and *Dead Space*. In these games, futuristic weapons include an ammo numeric counter built into the gun, or bullets visualized *through* the gun (e.g., *Metro 2033*).

The game was set in a simulated warehouse. There were 25 enemy soldiers initially positioned in a semi-circle around the player (Fig. 5). The enemies walked slowly towards the player. The player had a rifle which fired one shot per trigger press. Enemies died when shot, and disappeared. The right trigger button shot, the x-button reloaded. Reloading was only possible upon running out of ammunition (i.e., using all shots in the clip).

The software automatically recorded the number of clips used, hits, misses, enemies remaining, shots before reload, and time before reload. For each shot, the time was recorded along with the remaining ammunition and whether the shot hit or missed an enemy.

5.3. Procedure

Participants were instructed to play the game to the best of their ability, shooting all enemy soldiers as quickly and accurately as possible. They were informed that they had unlimited ammunition, but each clip only had a certain number of shots. Consequently, participants had to reload when out of ammo. They were then instructed on the controls and were allowed to begin. A trial ended when all enemies were killed.

On starting the trial and after reloading, each clip had a random number of rounds. The number of rounds ranged from 7 to 16 (decided once per trial). Using a random number of shots per clip required participants pay greater attention to their ammunition level. This helped prevent participants from mentally tracking ammunition, and thus was expected to help elicit differences between the test conditions. Upon running out of ammunition, participants manually reloaded (and could not reload prior to running out). This task was representative of real games: Ammunition level becomes crucial when it is low in a battle situation. The task requires participants to be highly aware of their ammunition level.

Participants completed 15 trials for each of the five ammunition displays, completing 75 trials in total. After each trial participants could take a break before continuing. Each trial took between 30 and 45 s. In total, the experiment took approximately 1 h.

Upon finishing all trials, participants completed a questionnaire about prior experience with FPS games. The questionnaire also asked participants about their perceived effectiveness of each ammunition display.

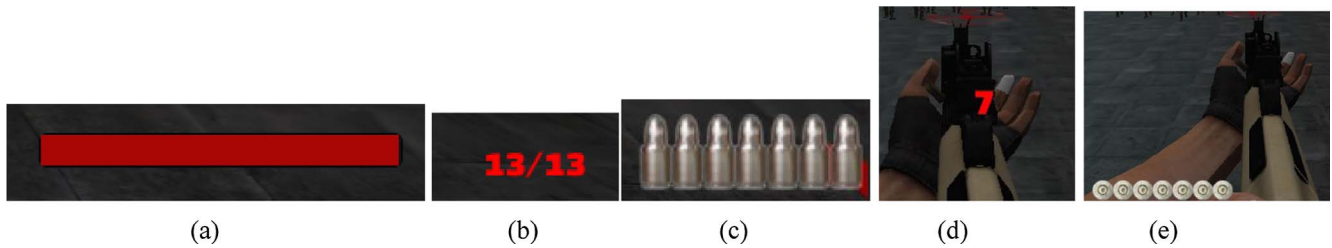


Fig. 6. The five ammo display conditions for Experiment 1. (a) Bar-on-HUD (BH), (b) Number-on-HUD (NH), (c) Icons-on-HUD (IH), (d) Number-in-game (NG), (e) Icons-in-game (IG).

5.4. Design

The study employed a 5 × 15 within-subjects design. The independent variables and levels were as follows:

Ammunition Display: BH, NH, IH, NG, IG
Trial: 1, 2, 3, ... 15

The ammunition displays are depicted in Fig. 6 and described in Section 5.2. Ammunition display ordering was counterbalanced according to a Latin square.

The dependent variables were the number of shots before reload (count) and the time before reload (seconds). Shots before reload was the average number of shots fired from the time the participant ran out of ammunition until the reload button was pushed. Time before reload was the time between running out of ammunition and pushing the reload button.

In total, participants completed 20 × 5 × 15 = 1500 trials.

5.5. Results

5.5.1. Shots before reload

Achieving a low score for shots before reload required the participant to notice they had no ammunition left. A high score thus indicates decreased awareness of ammunition levels. Average shots before reload is summarized for each ammunition display in Fig. 7.

The main effect of ammunition display on shots before reload was statistically significant ($F_{4,19} = 9.22, p < .0001$). A Tukey-Kramer post hoc analysis revealed that the difference between number-in-game (NG) and all other ammunition displays was statistically significant ($p < .05$). The rest of the ammunition displays were not significantly different from each other. The main effect for trial on shots before reload was not significant ($F_{14,19} = 0.94, ns$), nor was the interaction effect between ammunition display and trial ($F_{56,19} = 1.03, p > .05$).

The worst performing ammunition displays were the HUD options (BH, NH, IH). All three had comparable scores (slightly over 1 each) and were not significantly different from one another. Although icons-in-game (IG) performed slightly better, the difference was not significant. The best performing option was number-in-game (NG), which

had 0.68 shots fired before reload. Number-in-game resulted in about 35% fewer shots before reload than the worst performer, icons-on-HUD.

Participants noted that number-in-game (NG) was easy to see, as the ammunition count was almost directly where they were looking while aiming. The HUD-based displays were in the bottom right corner, requiring more glancing. These ammunition displays performed very similarly, suggesting a relationship between performance and display location. We speculate that positioning the HUD in a different location (e.g., another corner of the screen) is unlikely to yield a substantial performance difference, unless they are placed closer to the screen centre. However, it appears that “counting” the icons takes enough time that, in the icons-in-game (IG), their central location was not sufficient to yield a statistically significant performance boost. This suggests that visualization may have a greater influence than position.

5.5.2. Time before reload

Like shots before reload, higher scores were worse: The greater time before reload was, the lower the awareness of the ammunition level. Lower scores indicate a more immediate awareness of low ammo levels. Average time before reload for each ammunition display is depicted in Fig. 8.

There was a significant main effect of ammunition display on time before reload ($F_{4,19} = 4.26, p < .005$). A Tukey-Kramer analysis indicated a significant difference between number-in-game (NG) and all other ammunition displays. The main effect for trial was not significant ($F_{14,19} = 1.61, p > .05$), nor was the interaction between ammunition display and trial ($F_{56,19} = 0.92, ns$).

As with shots before reload, the icons-on-HUD (IH) ammunition display performed worst, and the number-in-game (NG) ammunition display performed best. NG offered the lowest time before reload, with an average of 1.0 s, approximately 20% lower than the next best performing ammunition display, number-on-HUD (NH). The most substantial difference was between icons-on-HUD (IH) and number-in-game (NG) ammunition displays. NG was about 26% faster than IH.

The results are rather consistent for both dependent variables. It appears the central location of the number-in-game ammunition display allows for better performance than the other displays. This is most likely because it reduces the amount of gaze shifting or glancing required. Again, IG seems to require enough additional mental processing

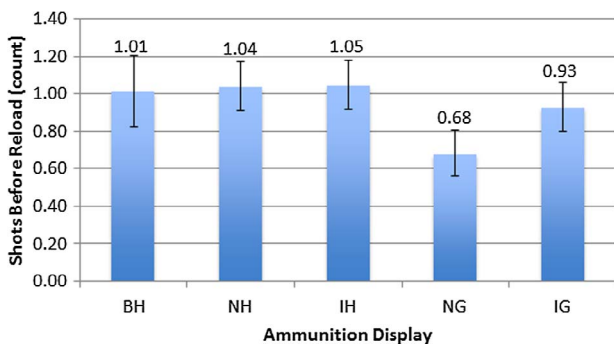


Fig. 7. Shots before reload by ammunition display. Lower scores are better. Error bars show ± 1 SD.

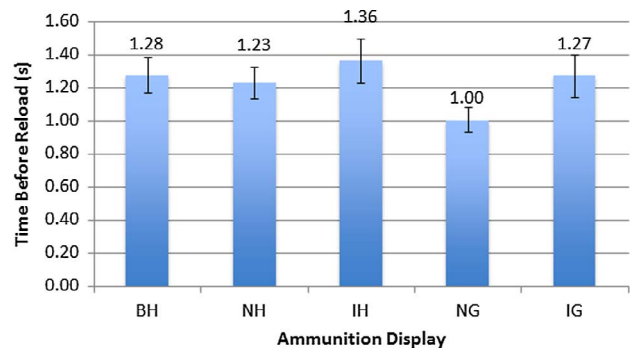


Fig. 8. Time before reload for each ammunition display. Lower scores are better. Error bars show ± 1 SD.

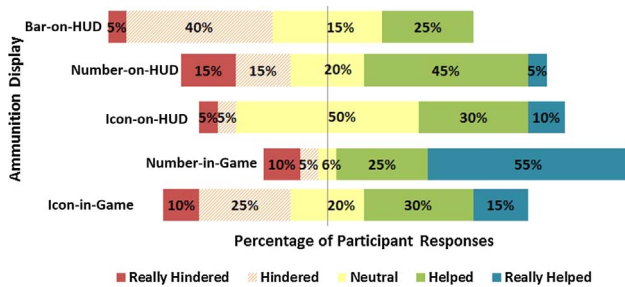


Fig. 9. Participant perceived effectiveness of each ammunition display.

to outweigh this positional advantage, however.

5.5.3. Questionnaire

Participants completed a questionnaire soliciting their feedback on the ammunition displays studied. They were asked to rate their preference towards each ammunition display on a 5-point Likert scale. Specifically, they were asked “Did each of the ammunition displays help or hinder your gameplay?” with response options ranging from “Really hindered” to “Really helped”. Fig. 9 depicts the percentage of participants for each response level.

Overall, the number-in-game (NG) ammunition display was considered the most effective, with 80% of participants reporting they found it helpful or really helpful. Opinions toward icons-in-game (IG), icons-on-HUD (IH), and number-on-HUD (NH) ammunition displays were mixed. The bar-on-HUD (BH) was thought to hinder gameplay by 45% of participants. A Friedman non-parametric test deemed the differences statistically significant ($\chi^2 = 11.56, p < .05, df = 4$). A post hoc analysis revealed significant differences between number-in-game (NG) and bar-on-HUD (BH), number-in-game (NG) and number-on-HUD (NH), and number-in-game (NG) and icons-in-game (IG). Participants tended to correctly identify the most effective display, exhibiting higher subjective preference towards it.

5.6. Summary

Overall, the diegetic number-in-game (NG) display offered significantly better performance than all other displays studied. Participants were also aware of the performance difference, as they ranked the display significantly better in terms of subjective perceived effectiveness. This may be due to the central placement of the display (near the screen centre), made possible by the fact that it was displayed spatially rather than as a HUD. Previous work [7] revealed differences in player performance due to HUD position, hence we included HUD position in the next experiment to investigate further.

6. Experiment 2: Health displays

The purpose of this experiment was to evaluate which of several HUD and diegetic/spatial options presented health information most effectively to the player. The task involved monitoring health levels while being shot at by enemies. Participants had to notice when their health level fell below a certain threshold.

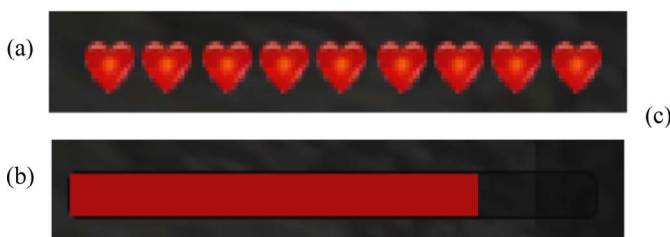


Fig. 10. Health display visualizations including (a) icons (b) bar and (c) numbers.

6.1. Participants

Twenty-four paid participants (21 male, 3 female) took part in the study. Ages ranged from 19 to 53 years (*mean* 25.4, *SD* 7.2). All were regular gamers, with 58% playing video games for more than 10 h per week, and 86% playing first-person shooter games every week. None of the participants had participated in the previous (ammunition) experiment.

6.2. Apparatus

The game was set within a turret with five enemies surrounding the participant’s character in a semi-circle. Movement was disabled. The right trigger button shot and the left bumper button “escaped” when health was low (upon reaching 20% or lower). The player initially had 100 health points. Enemies shot at the player, causing 5 points of damage with each hit. The player would “die” (ending the trial unsuccessfully) upon reaching 0% health.

The participant’s health level was displayed using one of 12 health displays. Each display presented the same information, but visualized it differently. Nine were HUD-based, derived from all combinations of three visualizations (icons, bar, and number) and three display positions (top, bottom, left). The bar-based options presented health as a meter decreasing towards the left as the player took damage. The icons showed hearts, similar to games like *The Legend of Zelda*. The number options presented health as a percentage of their total. Fig. 10 depicts the three different HUD visualizations and Fig. 11 depicts the three display positions. The nine HUD-based options were coded thus:

- bar-on-HUD-bottom (BHB)
- bar-on-HUD-left (BHL)
- bar-on-HUD-top (BHT)
- icons-on-HUD-bottom (IHB)
- icons-on-HUD-left (IHL)
- icons-on-HUD-top (IHT)
- number-on-HUD-bottom (NHB)
- number-on-HUD-left (NHL)
- number-on-HUD-top (NHT)

The bar and number visualizations were selected since they commonly appear in FPS games, while icons are common in games of other genres, but may have potential in FPS games as well.

Three HUD-alternative options were also studied. Two were spatial/diegetic variants of the HUDs described above, icons-in-game (IG), and bar-in-game (BG). They operated the same as the HUD-based variants, but were displayed on the 3D model of the character rather than as a HUD, similar to the number-in-game option in the ammo experiment. The third was blood splatter (S), a commonly used meta-perception which increasingly occludes the participant’s view with simulated blood as they take more damage. These are depicted in Fig. 12.

6.3. Procedure

Participants were instructed to shoot enemies while the enemies shot at them. Each time they hit an enemy, their score increased. The

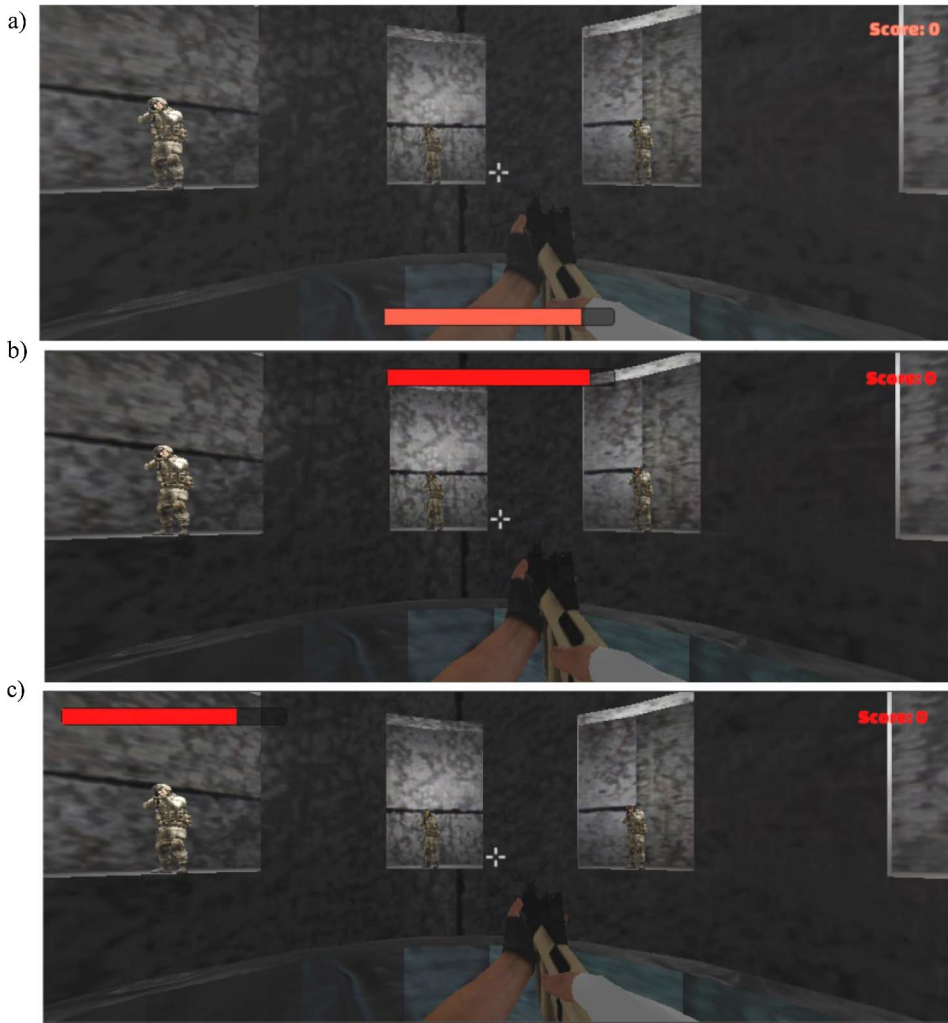


Fig. 11. Health display positions used with each visualization. (a) bottom (b) top (c) left. Note: only bar visualizations are shown here. Number and icons appear in the same locations, but use the visualizations shown in Fig. 10.

point value of hitting an enemy increased with lower health. This encouraged participants to stay as long as possible without reaching 0 health (i.e., dying) and to get as high a score as possible. They were informed that once their health reached 20% or lower, they could “escape” by pressing the escape button, which transported them to a safe location and successfully completed the trial. The primary goal was to escape, while the secondary goal was to get a high score. The overall high score only updated if they escaped before running out of health. The data recorded included participant escape before 0 health, health remaining (zero if they did not escape), and score. Each shot was recorded along with hit or miss and, if it hit an enemy, how many points were earned.

Participants completed 20 trials for each of the 12 health displays, 240 trials in total. After each trial participants could take a break. Each trial took approximately 5 s, or about 25 min for the entire experiment.

6.4. Design

The study employed a 12 × 20 within-subjects design. The independent variables and levels were as follows:

Health Display: BHB, BHL, BHT, IHB, IHL, IHT, NHB, NHL, NHT, BG, IG, S

Trial: 1, 2, 3, ..., 20

The ordering of health display was randomized in order to offset potential learning effects. The dependent variables were escape percentage (%) and health when escaped (%). Escape percentage was calculated as the percentage of trials where participants were able to escape (i.e., without dying). Health when escaped was the health remaining (as a percentage of their initial 100 health points) when participants successfully escaped, avoiding dying in a given trial.



Fig. 12. Alternative displays: (a) bar-in-game, (b) icons-in-game, and (c) splatter (at reaching 0% health).

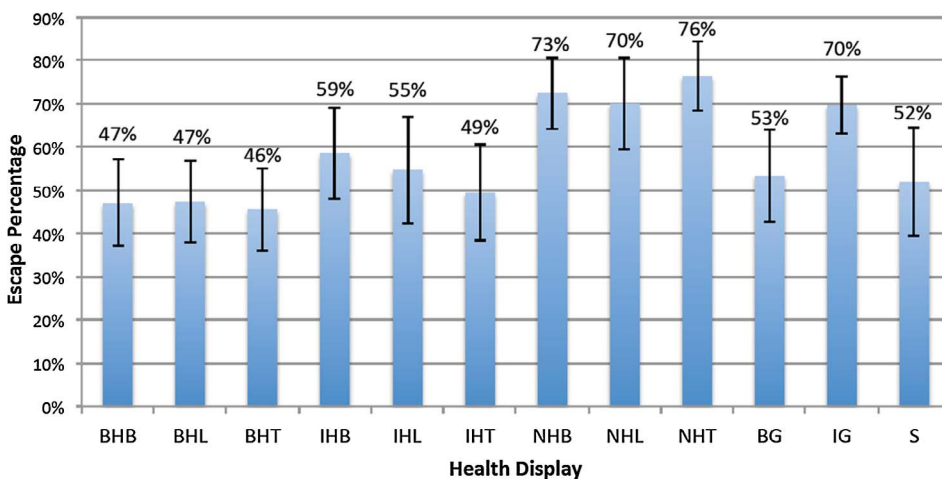


Fig. 13. Average escape percentage by health display. Higher scores are better. Error bars show ± 1 SD.

Participants completed $24 \times 12 \times 20 = 5760$ individual trials.

6.5. Results

6.5.1. Escape percentage

A high escape percentage indicates better performance, as participants could see and understand the health display well, escaping from the trial before dying with greater reliability. The best performing health display was the HUD-based number-on-HUD-top (NHT), with 76%. The worst performer was the bar-on-HUD-top (BHT), with 46%, followed closely by bar-on-HUD-bottom (BHB) and bar-on-HUD-left (BHL), both at 47%. See Fig. 13.

A repeated measures ANOVA revealed a significant main effect for health display on escape percentage ($F_{11,23} = 12.91, p < .0001$). A Fisher LSD pair-wise analysis revealed that the number-on-HUD displays (NHB, NHL, and NHT) all performed significantly better than all other health displays except for icons-in-game (IG). Icons-on-HUD-bottom (IHB) performed significantly better than all bar-on-HUD variants.

In general, the number-based options did well, likely because it was easier to determine when health was between 0% and 20%. The bar and icon options were more ambiguous, although icons were better than bars. Position appeared to matter less than visualization; as seen in Fig. 13, results “cluster” by visualization rather than location.

6.5.2. Health when escaped

Scores for health when escaped are always between 0% and 20%, since participants could only escape upon reaching 20% or lower health (and died upon hitting 0% health). Lower scores are better, since they were tasked with getting the highest score possible while still escaping. Participants were rewarded for staying longer as they were awarded higher scores when their health was low.

Icons-on-HUD-top (IHT) had the best result, with an average health of 6.54% upon escape. Since each enemy shot cost 5 health points, this means participants tended to escape with slightly more than one shot left of health until they would have died in a given trial. This was followed closely by bar-on-HUD-top (BHT), with an average health when escaped of 6.65%. Icons-in-game (IG) had the highest health when escaped, with an average of 12.88%. This was approximately 49% higher than with the icons-on-HUD-top (IHT) health display. All number options were close to 10% (i.e., around 2 enemy shots from dying). See Fig. 14.

There was a significant main effect for health display on health when escaped ($F_{11,23} = 16.01, p < .0001$). In summary, participants escaped with significantly lower health remaining with the bar and icon displays on the HUD. All other health displays had significantly lower health when escaped than the icons-in-game (IG).

6.5.3. Subjective results

Participants were queried the effectiveness of health display on a 5-point Likert scale, ranging from “Really hindered” performance, to “Really helped” performance. Results are summarized in Fig. 15, and suggest strong preference towards the HUD-based options, particularly the Bar-on-HUD variants. The results were found to be statistically significant using a Friedman non-parametric test ($\chi^2 = 15.134, p < .01, df = 5$). A post hoc analysis revealed significantly different preference levels between bar-on-HUD (BH) and bar-in-game (BG), icons-in-game (IG), and spatter (S), as well as between number-on-HUD (NH) and bar-in-game (BG), icons-in-game (IG), and spatter (S). This result shows that the bar and number on HUD were thought to be much more helpful than the alternative displays (BG, IG, and S).

6.6. Summary

Overall, the HUD-based variants offered the best performance, regardless of placement on the screen. In terms of trial success rate, numeric options were best. Only the diegetic icon-in-game option came close. Subjectively, participants preferred the bar-on-HUD display, despite its relatively worse performance.

7. Experiment 3: Weapon displays

The purpose of this experiment was to compare different options for presenting the currently selected weapon. We note that all modern (and most pre-modern) FPS games show the current weapon diegetically in front of the player. However, many include redundant icons, text, or various other HUD representations. Players were tasked with trying to match their weapon to specific enemies (a colour-matching task).

7.1. Participants

The same 24 participants that completed the health display experiment completed this experiment, either right before, or right after. The ordering of the experiments was alternated to eliminate learning/practice effects. Half of the participants completed the health experiment first, followed by this experiment. The other half completed the two experiments in the reverse order.

7.2. Apparatus

The experiment was set in the same warehouse scenario as in the ammunition experiment. Eight groups of three enemies appeared in succession, each group appearing after the last was destroyed by the participant. Enemy groups were coloured red, green, blue, or yellow, and all enemies in a single group were the same colour. See Fig. 16.

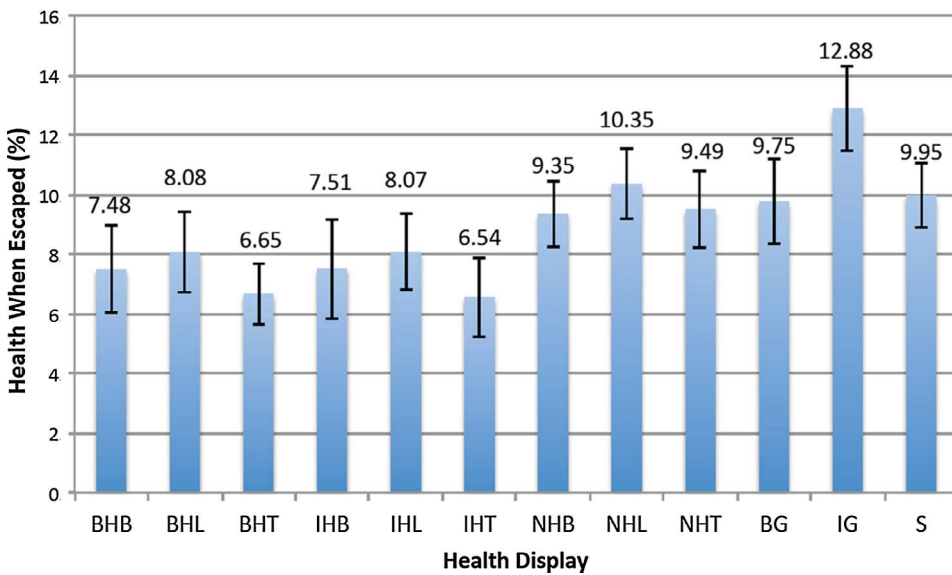


Fig. 14. Health when escaped by health display. Lower scores are better. Error bars show ± 1 SD.

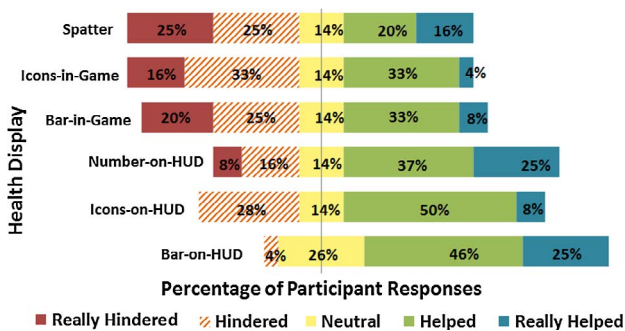


Fig. 15. Participant perceived effectiveness for each health display, showing percentage of participants giving each response.

Player movement was disabled, but the participant could look around and aim using the right analog stick, and shoot using the right trigger button. Each direction pad button mapped to switching to a different weapon colour (i.e., up \rightarrow red, right \rightarrow green, down \rightarrow blue, left \rightarrow yellow; see Fig. 4b). This direct mapping was used instead of pressing

buttons to cycle through weapons, as the latter introduced additional delay in contrast with the immediacy of pressing a direction button.

The UI displayed the participant’s current weapon (colour) using one of six weapon displays: name-on-HUD-right (NHR) name-on-HUD-left (NHL), icon-on-HUD-right (IHR), icon-on-HUD-left (IHL), in-front (IF) and name-on-gun (NG). The “left” and “right” options indicated on which side of the bottom of the screen a HUD-based option was displayed. As with the previous experiments, each weapon display presented the same information, but visualized it differently, either as HUDs via text or an icon, or spatially (i.e., rendered directly on the weapon in the game itself). The displays are seen in Fig. 17. Note that the spatial “in-front” option (which positions the weapon display in front of the player) is consistently used in all modern FPS games. The rest of the weapon displays and positions are found in recent FPS games.

7.3. Procedure

Participants were instructed to match the weapon colour to the enemy colour and shoot as quickly as possible to get a high score.



Fig. 16. Weapon display experiment, showing overview of the task and scene.

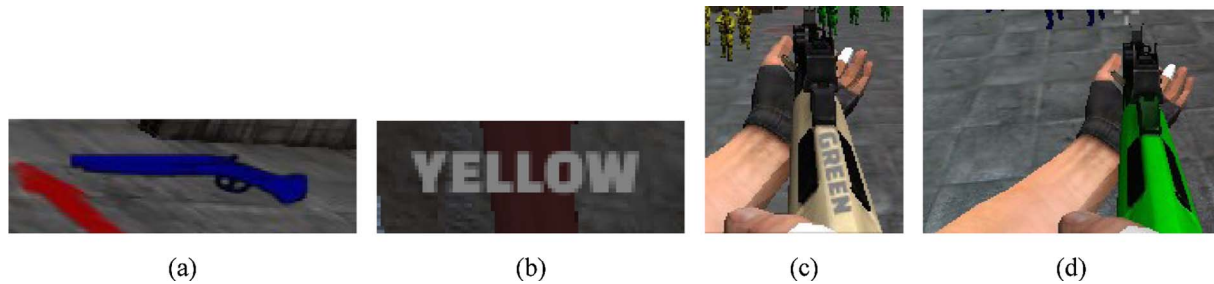


Fig. 17. Weapon displays used in experiment. (a) Icon-on-HUD, (b) Name-on-HUD, (c) Name-on-gun (NG), (d) In-front (IF). The Icon-on HUD and Name-on HUD weapon displays were tested in two display positions, at the bottom right (IHR, NHR) and the bottom left (IHL, NHL).

Shooting an enemy with a gun of the wrong colour had no effect. We used colour matching, rather than different weapons, since matching specific weapons to enemy types would require additional training. Colour matching is more straightforward. Upon killing an enemy group, a new randomly coloured group appeared. A participant could get a “kill streak”, gaining extra points, by killing additional enemies within 750 ms of the previous kill. The high-score goal encouraged speed and, in turn, necessitated quick recognition of the current weapon. A trial finished when all enemy groups were destroyed. Data recorded included number of shots with the wrong colour weapon and time to switch to the correct weapon.

Participants completed 8 trials for each of the six weapon displays, for 48 trials in total. After each trial participants could take a break before continuing. Each trial took about 25 s, or about 25 min for the entire experiment.

7.4. Design

The study employed a 6×8 within-subjects design. The independent variables and levels were as follows:

Weapon Display: NHR, NHL, IHR, IHL, IF, NG
Trial: 1, 2, 3, ..., 8

The ordering of weapon display was counterbalanced with a Latin square.

The dependent variables were shots with wrong weapon (%), and weapon switch time (seconds). Shots with wrong weapon is the percentage of shots fired with a non-matching weapon. Weapon switch time is the time to switch from an incorrect to correct weapon colour.

Overall, participants completed $6 \times 8 \times 24 = 1728$ trials.

7.5. Results

7.5.1. Shots with wrong weapon

Fewer shots with wrong weapon scores are better, as this indicates participants could more quickly identify when using a non-matching weapon on an enemy. Icons-on-HUD-right (IHR) performed best, with 2.12% of shots with wrong weapon. This was closely followed by the weapon in-front (IF) display, with 2.20%. The name-on-gun (NG) display performed worst, with 2.93% of shots with wrong weapon. See Fig. 18. A repeated measures ANOVA revealed that the main effect for trial was significant ($F_{7,23} = 2.44, p < .05$), while the main effect for weapon display was not ($F_{5,23} = 1.74, p > .05$). The interaction effect was not significant ($F_{35,23} = 0.82, ns$).

7.5.2. Weapon switch time

Weapon switch time indicates how quickly participants recognized they were using the wrong weapon and changed. Lower times are better. The icon-on-HUD-right (IHR) again performed best, at an average of 1.01 s to switch to the correct weapon. This was closely followed by name-on-HUD-left (NHL), in-front (IF), then name-on-

HUD-right (NHR). The name-on-gun (NG) weapon display performed worst, with an average of 1.16 s to switch to the correct weapon. This was approximately 11% worse than with the icons-on-HUD-right (IHR) display. See Fig. 19.

The main effect for trial on weapon switch time was significant ($F_{7,23} = 17.33, p < .0001$), while the main effect for weapon display on weapon switch time was not significant ($F_{5,23} = 2.03, p > .05$). The interaction effect between weapon display and trial was significant ($F_{35,23} = 1.73, p < .01$). A Fisher LSD pair-wise analysis revealed a significant difference between NHR trial 1, IHR trial 1, and NG trial 1, and the remaining trials.

7.5.3. Subjective results

Participants ranked their preference towards each weapon display on a 5-point Likert scale, ranging from “Really hindered” (1) to “Really helped” (5). Results are summarized in Fig. 20. Participants found IF and IH most helpful. The results were statistically significant using a Friedman non-parametric test ($\chi^2 = 40.854, p < .0005, df = 3$). Post-hoc analysis revealed the significant pairwise differences between all condition pairs except for name-on-HUD and name-on-gun. Participants performed well with most preferred weapon displays – icon-on-HUD and in-front. This suggests that both of these weapon displays are good choices, and also could be good to use in combination.

7.6. Summary

Results of this experiment revealed little difference between the displays studied. Overall, participants got better over the 8 trials, but none of the displays were significantly different in terms of performance. Participants strongly preferred the “in-front” display, although icon-on-HUD was also positively assessed.

8. Experiment 4: Navigation aid displays

The final task we studied was navigation through the environment. Players were tasked with moving through a maze, with several different navigation aids to assist them, ranging from diegetic compasses, mini-map HUDs, or a “navigation line” modeled after the diegetic way-finding aid found in *Dead Space*, which showed the player the exact path through the maze. While we expected this latter option to perform best, we included it primarily as a “ground truth” or control condition.

8.1. Participants

Twenty-one of the participants who completed the preceding experiments also took part in this experiment. Any participants who completed multiple experiments did so on different days, such that an equal number did each possible ordering of experiments. In total, this experiment included 24 paid participants (22 male, 2 female). Ages ranged from 19 to 53 years (*mean* 25.3, *SD* 7.3). All participants were regular gamers, with 46% reporting they play video games more than 10 h per week.

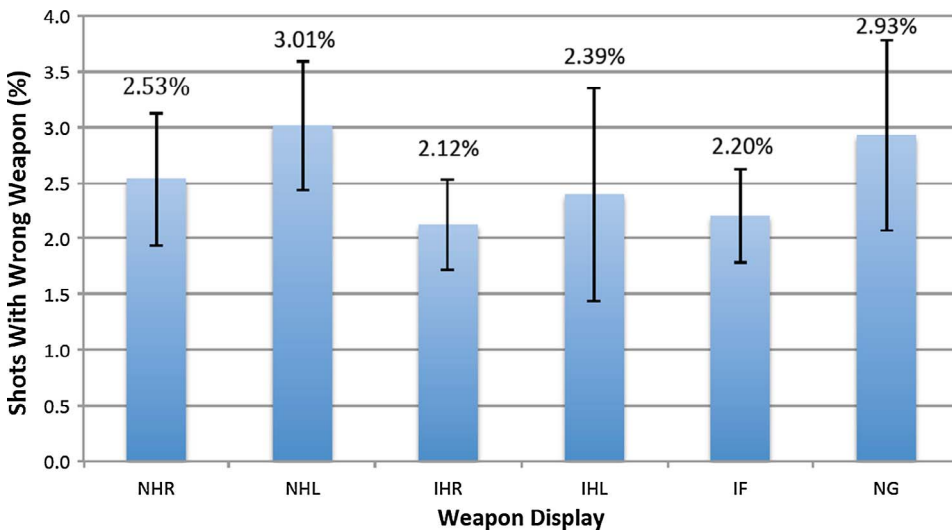


Fig. 18. Shots with wrong weapon by weapon display. Lower scores are better. Error bars show ± 1 SD.

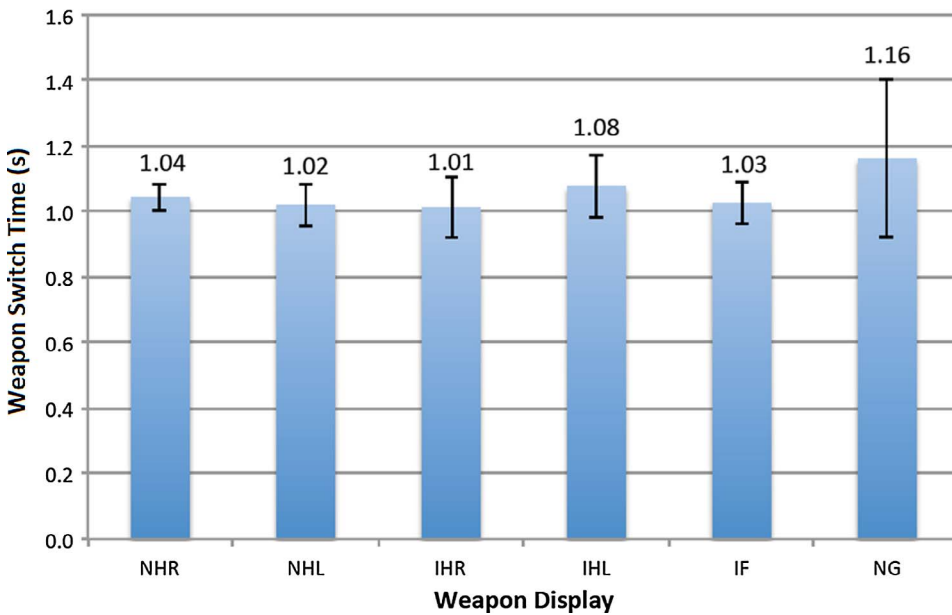


Fig. 19. Weapon switch time by weapon display. Lower scores are better. Error bars show ± 1 SD.

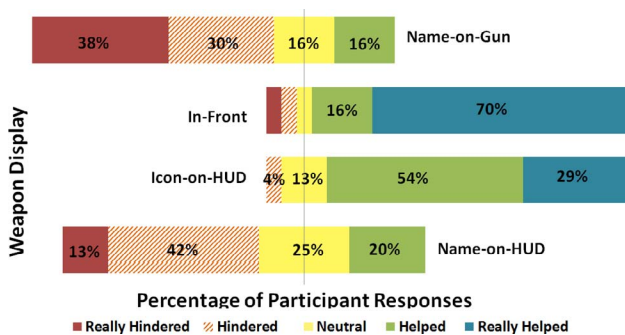


Fig. 20. Subjective preferences for each weapon display, showing percentage of participants giving each response.

8.2. Apparatus

The left analog stick controlled movement. See Fig. 4b. All other controls were disabled; the task *only* required navigating the maze. Navigating the maze required reaching four waypoints (represented by large red spheres floating slightly above the floor) located in the maze,

all separated by equidistant paths. The four waypoints were positioned along the “optimal” path through the maze. Participants needed to reach all four waypoints in order to complete a trial, i.e., the last waypoint was positioned at the very end of the maze. The next waypoint sphere would appear only after the previous waypoint was reached. Participants simply had to walk through a waypoint to reach it. The maze had a grid of “tiles” – 1×1 m squares that made up the map. Distance was judged based on the number of tiles crossed by the participant – hence the number of tiles between each waypoint was the same. The software recorded the number of tiles the participant crossed in reaching each waypoint and the time to get to each waypoint from either the start or the previous waypoint.

The experiment included five navigation displays: mini-map still (MMS), mini-map rotate (MMR), wayfinding arrow (WA), light pillar (LP), compass (C), and navigation line (NL). Each presented the same information, but visualized it differently. See Fig. 21, which also depicts the environment used in the experimental task.

Mini-maps and the wayfinding arrow are HUD elements commonly used in recent FPS games (see Table 1). The mini-maps showed a top-down view of the participant position, the maze, and the position of the next waypoint when it came into the mini-map range. With MMS, north was always up on the mini-map; with MMR, the movement direction

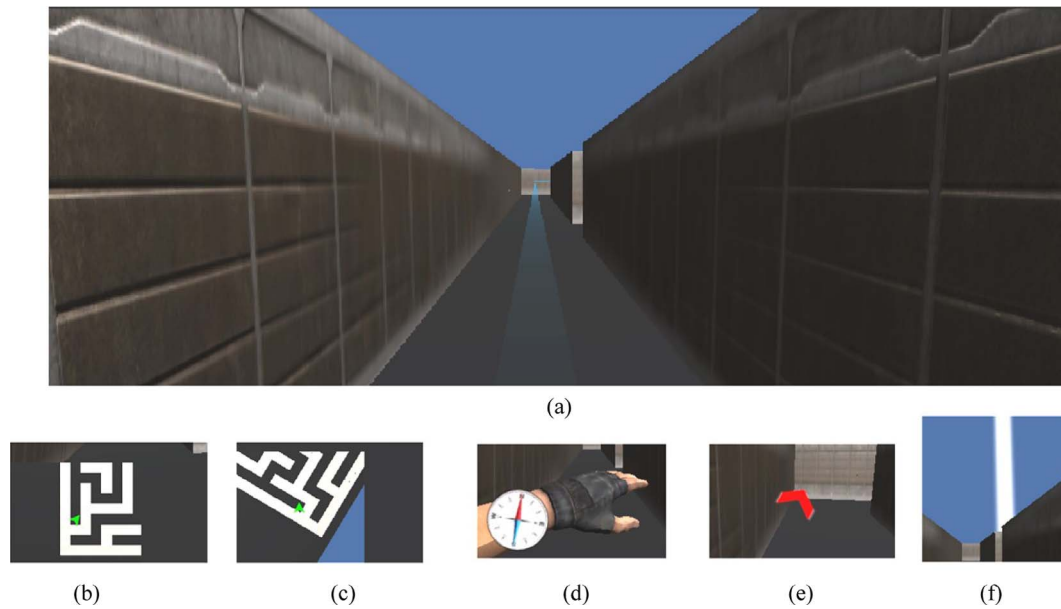


Fig. 21. Navigation displays. (a) Task overview, showing navigation line display (b) mini-map still (c) mini-map rotate (d) compass (e) wayfinding arrow (f) light pillar.

was up. The mini-map rotated such that the player's movement direction (the green arrow in Fig. 21c) always pointed up. The wayfinding arrow (HUD) and compass (diegetic) always pointed to the next waypoint. Such displays are sometimes used in FPS games. The navigation line (Fig. 21a) traced a path (visualized as a semi-transparent blue line) along the floor of the maze directly from the player's position to the next waypoint, turning corners as necessary. Since it was visualized in the game world, it qualifies as a spatial element (as per Fig. 2). The navigation line was based on the navigation aid in the popular third-person shooter game *Dead Space*. Since this display shows the player exactly how to navigate the maze to all waypoints, we included it as a control condition, as it was expected to offer optimal performance. Finally, the light pillar is another spatial option based on navigation aids used in some adventure titles, such as the *Legend of Zelda*. It emitted light straight up from the position of the next waypoint, allowing navigation by a physical landmark, similar to navigating a real city using physical features such as tall buildings. We are unaware of any precedent for such a display in existing FPS games, but included it to assess its potential.

8.3. Procedure

The environment used an outdoor maze (see Fig. 21a). Participants navigated to the end of the maze as quickly and accurately as possible. Accuracy meant minimizing wrong turns and avoiding wandering into unnecessary parts of the maze – in other words, minimizing the number of tiles they crossed while completing the maze.

The same maze was used with all navigation displays, and was completed twice with each. For the first trial, participants navigated the maze in one direction, and in the second trial, they proceeded in the reverse direction. They were not informed that both trials took place in the same maze. This was intended to avoid learning effects that could occur by using the same maze in the same direction in both trials. However, it also ensured that the complexity and distance were equal in both trials, unlike using a different maze. To prevent participants getting lost and wandering endlessly, a 6-min time limit was used for each trial. Trials were marked as “incomplete” if the participant did not reach the maze's end within the time limit.

Participants completed two trials for each of the six navigation displays, completing 12 trials in total. After each trial, participants could take a break before continuing. Each trial took approximately 3

min, for a total experiment time of approximately 36 min.

8.4. Design

The study employed a 6×2 within-subjects design. The independent variables and levels were as follows:

Navigation Display: MMS, MMR, WA, LP, C, NL
Trial: 1, 2

Navigation display order was counterbalanced with a Latin square. The dependent variables were step error rate (path efficiency ratio), waypoint time (in seconds), and incomplete trials (%). Step error rate was calculated by dividing the actual number of tiles the participant crossed in navigating to the next waypoint by the actual distance between waypoints, 60 tiles. Waypoint time is the time to find each waypoint, measured from the start of the maze (for the first waypoint) or the previous waypoint (for subsequent waypoints). Incomplete trials indicated the percentage of trials where the participant ran out of time before finding all waypoints.

Overall, the experiment included 12 participants \times 6 navigation displays \times 2 trials per navigation display = 144 trials in total.

8.5. Results

8.5.1. Waypoint time

Waypoint time indicated how long it took (on average, in seconds), to find the next waypoint. Lower scores are better. Note that this dependent variable *excludes* incomplete trials – if the participant ran out of time, their waypoint times for that trial were not included in the average. Unsurprisingly, the navigation line (NL) performed best, with an average time of 21.6 s to reach each waypoint. The compass (C) performed worst, with an average time of 56.4 s to reach each waypoint. This was about 62% worse than with the navigation line (NL). See Fig. 22.

There was a significant main effect of navigation display on time ($F_{5,23} = 26.77, p < .0001$). A Fisher LSD pair-wise analysis revealed that the mini-map displays (MMR and MMS) and the navigation line (NL) were significantly faster than when using the wayfinding arrow (WA), light pillar (LP), and compass (C).

The main effect for trial was significant ($F_{1,23} = 47.06, p < .0001$),

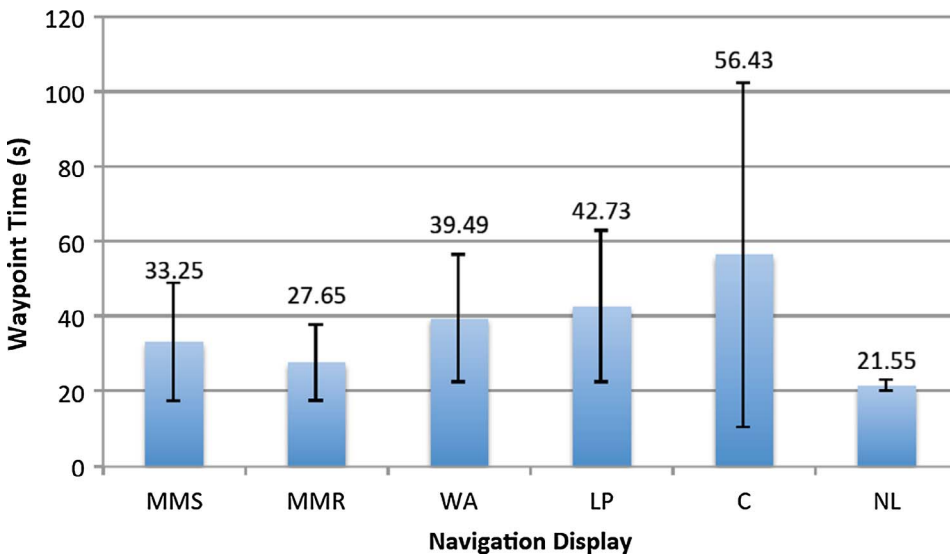


Fig. 22. Time by navigation display. Lower scores are better. Error bars show ± 1 SD.

as was the interaction between navigation display and trial ($F_{5,23} = 9.90, p < .0001$). Participants generally completed the first trial faster than the second trial. Despite being the same map, participants often found completing the maze in reverse direction more difficult than the first-trial maze, taking on average about 9 s longer to complete the maze in the second trial (reverse order). This is perhaps because the end of the maze (with respect to the first trial) was more complex than the start, hence the reverse order began with a more complicated start.

8.5.2. Step error rate

Step error rate yields a ratio of path efficiency, where the best result is 1.0 (i.e., a participant who did not deviate from the optimal path). Scores higher than 1.0 indicate worse performance. See Fig. 23.

The results for step error rate and waypoint time were similar. The navigation line (NL) display (again) performed best, with an average step error rate of 1.03 to reach each waypoint. This is not unexpected, as the display shows participants exactly which way to go. The compass (C) again performed worst, with an average step error rate of 2.43 to reach each waypoint – in other words, participants traversed on average $\approx 2.5 \times$ more distance in the maze as required, clearly becoming lost more often than with other displays. The variability was also high, suggesting the compass was highly ineffective for many

participants.

There was a significant main effect for navigation display on step error rate ($F_{5,23} = 36.32, p < .0001$). A Fisher LSD post hoc test revealed that the navigation line (NL) had significantly lower step error rate than all other navigation displays except for the mini-map rotate (MMR). This is a promising result; the MMR display offered comparable performance to the somewhat unrealistic and arguably unfair navigation line. The compass (C) had significantly higher step error rate than all other navigation displays except for the wayfinding arrow (WA).

The main effect for trial was significant ($F_{1,23} = 132.24, p < .0001$), as was the interaction between navigation display and trial ($F_{5,23} = 5.27, p < .0005$). Like time, participants performed worse in the second trial, with an average step error rate (over all navigation displays) of 1.91 vs. 1.48 in the first trial.

8.5.3. Incomplete trials

Trials were marked “incomplete” if a participant took longer than 6 min to find all four waypoints, reaching the end of the maze. When this occurred it was typically because a participant was hopelessly lost and could not find their way through the maze. The frequency of incomplete trials is expressed as a percentage of all trials in Fig. 24. A lower percentage indicates a better performance. The navigation line (NL)

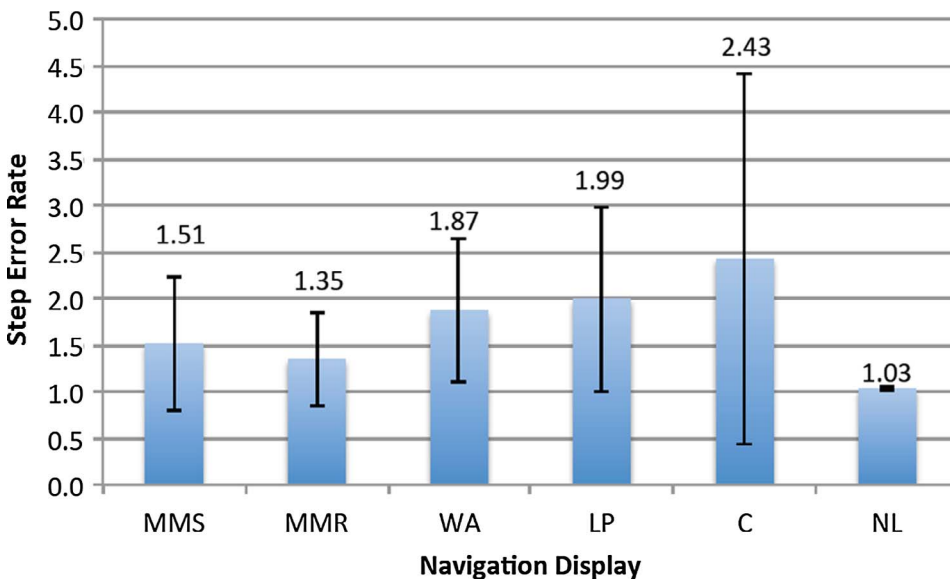


Fig. 23. Step error rate by navigation display. Lower scores are better. Error bars show ± 1 SD.

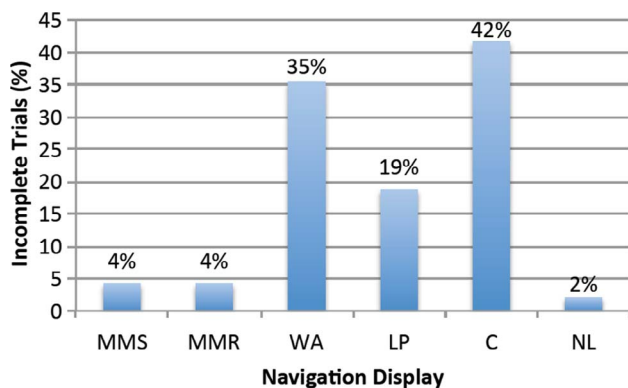


Fig. 24. Incomplete trials by navigation display. Lower scores are better.

performed best – only 2% of trials with this navigation display were incomplete. The compass performed worst, with 42% of trials incomplete.

There was a significant main effect for navigation display on incomplete trials ($F_{5,23} = 18.29, p < .0001$). A Fisher LSD pair-wise test indicated that both mini-maps and the navigation line (NL) all yielded significantly fewer incomplete trials than the wayfinding arrow (WA), light pillar (LP), and compass (C). The wayfinding arrow (WA) and compass (C) both yielded significantly more incomplete trials than the other navigation displays. It is interesting to note again that display effectiveness had little to do with whether the display was rendered spatially (e.g., within the game world) or not. Both spatial and traditional HUDs are represented in the best and worst performers in this experiment.

8.5.4. Subjective results

Participants ranked their perceived effectiveness of each navigation display on a 5-point Likert scale with responses ranging from “Really hindered” (1) to “Really helped” (5). As seen in Fig. 25, participants found NL most effective, followed by MMR and MMS. A Friedman non-parametric test found the results to be statistically significant ($\chi^2 = 71.94, p < .001, df = 5$). A post hoc analysis revealed significance between the following:

- MMS and WA, C, LP, and NL
- MMR and WA, C, and NL
- WA and MMS, MMR, LP, and NL
- C and MMS, MMR, LP, and NL
- LP and MMS, WA, C, and NL
- NL and MMS, MMR, WA, C, and LP

These results indicate that the navigation line received a significantly more positive result than any of the other navigation displays. They also indicate that the mini-maps (MMS and MMR) received

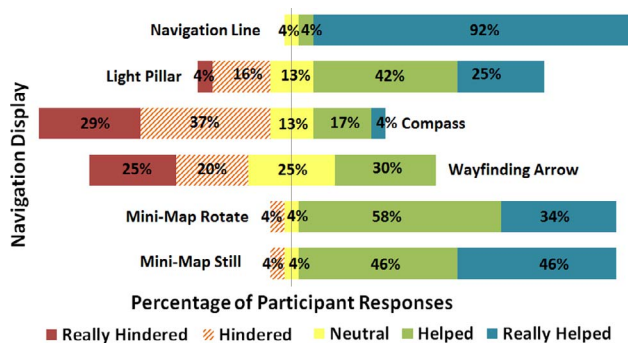


Fig. 25. Subjective preferences for each navigation display, showing percentage of participants giving each response.

statistically similar results, and that mini-map rotate (MMR) and light pillar (LP) also received statistically similar results.

8.6. Summary

Results of this study indicate that, unsurprisingly, the navigation line, functionally equivalent to the diegetic navigation aid in *Dead Space*, offered the best performance in terms of completion time and path efficiency. However, the HUD-based mini-map options also did well, coming close to the navigation line, and significantly better than other HUD, diegetic, or spatial options. Notably, rotating the mini-map (i.e., forward is always up) worked better, despite participant preference for the non-rotating variant.

9. Discussion

Globally, our results prevent us from being categorical as to whether HUDs or alternatives offer better performance. In some cases, alternative displays were best, while in other cases, HUDs were best.

Results of the ammunition experiment (#1) indicate that the number-in-game ammunition display was best in terms of how long it took participants to recognize they were out of ammo. We expected this is due in part to the placement of the display. Since it was co-located with the player’s gun, no additional glancing to HUD elements was required. We note that without eye tracking data, we cannot be certain; this is merely our suspicion based on the primary differences between these displays. In particular, participants were able to effectively track their ammunition while otherwise playing the game normally. This suspicion motivated us to include display position as a condition in the second experiment on health displays.

The health display experiment (#2) revealed that number-based options reliably allowed participants to detect when their health was low, successfully escaping the scenario before reaching 0 health. This result is similar to the ammo experiment, where, as noted above, the “number-in-game” diegetic option performed best. In contrast to the ammo experiment, alternative display options placed near the centre of the screen offered mediocre performance compared to HUDs. In the health experiment, display position seemed less important than display visualization. While there were large differences between types of displays, positioning HUDs at the top, bottom, or side of the screen mattered less. We were surprised by this as we expected a central position (i.e., as presented by the diegetic options, and similar to the ammo study) to reduce glancing to HUDs located on the periphery of the screen. Unlike the ammo experiment, HUDs performed better than alternative displays. Participants also indicated that they found these more effective, likely due to the comparative precision. In particular, numeric display options were well suited to identifying when health was in a certain percentage range. Alternative methods – including less precise iconic and bar HUD options – simply did not offer the same degree of precision as numbers.

Results of the weapon display experiment (#3) are more mixed. The diegetic in-front display performed well, and was strongly preferred by participants. This was expected, as it not only centralizes the information, but utilizes an arguably immersive display for the weapon (i.e., showing the weapon directly in front of the player as though they are holding it). In practice, all modern FPS games display the current weapon this way – but they also frequently show a redundant icon or the name of the weapon. To our surprise, and despite its redundant nature, such an icon display (icon-on-HUD) did very well, offering the best performance of all displays in the experiment. This was unexpected – the icon presented the same information as the weapon in front, yet necessitated extra glancing to effectively use it. In contrast, the weapon-in-front display was visible at all times. This may have improved performance because the redundant icon display provided participants a second place to look and assess their current weapon, thus potentially saving time while glancing around the screen for enemies. Our findings

also suggest that the bottom-right corner is the best location for displaying HUD-based weapon information. This might be due to participant familiarity with this option, as this location is commonly used for such icons in commercial games, as revealed in Section 3. However, we did not exhaustively evaluate the effect of display position in this experiment. This result lends merit to the idea of combining diegetic and HUD-based options in other situations. Testing other combinations of HUDs and diegetic displays is an opportunity for future work, potentially in all kinds of displays (health, ammo, etc.). The HUD-alternative display option, “name-on-gun”, performed worse with participants noting that they did not like it. This could be due to the viewing angle of the text printed on the gun, or to the unusual nature of such a display.

Results of the navigation aid experiment (#4) revealed that the spatial navigation display option offered the best performance for this task, which is unsurprising, given the relationship between the information presented and the task itself. The spatial (HUD-alternative) navigation line (NL) offered high performance due to the detailed path information it provided. Like the mini-map options, the navigation line provided information about direction and distance, making it easy to turn through the maze. In contrast, the worst performing display – the compass (C) – the only diegetic display type included in this test, offered less rich path information (only direction). This suggests the relationship (at least for navigation aids) between alternative display types is limited. Instead, performance is determined by how well an option suits the task. In this study, the navigation line (NL) was the best display for a maze. This may not be the case in other environments (e.g., an open environment). This is something we may wish to explore in future work.

9.1. Limitations and future work

The primary limitation of these experiments is that each display type was studied in isolation. This is appropriate from an experimental control point of view, and thus enhances the internal validity of the results. However, it decreases the generality of the results. Most games show multiple displays at once (e.g., see Fig. 1, usually combinations of health, ammunition, weapon, and navigation displays. Sometimes, these even include combinations of diegetic displays and HUDs. Studying a single display in isolation is not fully representative of this more complex task of monitoring multiple displays simultaneously. However, we expect that, with multiple displays present, those that individually demonstrated better performance are likely to offer better performance together. Hence, we believe studying multiple display types in isolation is worthwhile to maintain high internal validity and to “chip away” at the more complex problem of monitoring multiple displays at once. Future work will focus on this goal. For example, a competitive “death match” style experiment using different combinations of displays could provide great insight into how competitive play is influenced by diegetic displays, HUDs, spatial representations, and meta-perceptions.

As noted earlier, without eye tracking data, many explanations of our results are somewhat speculative. Eye tracking could enable more insightful explanations – for example, that it was indeed the fact that the fixation point of the user’s eyes influenced performance with one display or another. This is a consideration for a future study.

Finally, we note again that our experiments exclusively used “expert” FPS gamers. There are two limitations here. First, expertise was self-assessed by participants; it is difficult to objectively quantify expertise, but on average, participant performance seemed fairly consistent. Second, these results likely do not apply to novice gamers. As described in the introduction, from an experimental control point of view, this was likely the right choice, as it increases the likelihood of detecting significant differences *due to inherent differences in the displays studied*. That said, it decreases the generality (external validity) of our results.

10. Conclusions

Our results suggest that neither traditional HUDs nor alternative display types (e.g., diegetic, spatial, etc.) are best suited to FPS games, but rather specific properties of a given display (i.e., word, icon, number, bar, etc.) and, to a lesser extent, its position (in-game, bottom middle, top left, etc.) *relative to the task at hand* have an effect on player performance. In other words, a proper design methodology would first identify the most important information for player success in completing game tasks, then choose the best method to display that information effectively, and finally understand the best place to put the information display. This of course assumes that the game designer is trying to optimize player efficiency; if the designer is instead trying to optimize “difficulty” (as apparently the designers of *Dark Souls* appear to strive for [9]), then of course one could also use our results in reverse. For example, to make a game harder (e.g., via difficulty modes), a developer could use less effective information displays, or vice versa. This could also be adjusted to the skill level of the player, similar to a suggestion by Iacovides et al. [16].

Our results do not support existing recommendations to use alternative displays whenever possible [10]. Instead, our results indicate that HUD alternatives are not inherently the best option when considering player performance and preference. The game world and tasks should be the key factors in selecting an information display method, and not immersion or following current trends. For example, while it is commonly used, *splatter* as an indicator of health should be avoided, as it hampers players much more than aesthetic considerations can justify. Our results also suggest a strong relationship between performance and perceived performance (preference) – participants fairly consistently identified the better display options presented to them. In general, participants were well aware of which displays let them to perform best, and preferred these over less effective display types.

Finally, we speculate that the over-emphasis on diegetic and alternative displays in the literature might be due to a certain *onlooker effect*, where theories of game design are arrived at by “looking at” a game (as a reviewer would, but also as spectators), rather than through the lens of the player’s experience. For someone who has the leisure to look around the game world at their own pace, of course HUD alternatives will be much more satisfying; our results show that from the players’ perspective, this may not matter at all, and in some cases, their ability to perform well in the game might even be hampered by such a design. Of course, there are game genres, such as e-sports, where one could argue that spectators are as important as players – and in these games, aesthetic considerations may weigh in as importantly as player effectiveness. Our results suggest that the game designer should pay close attention to their intended audience, and tune the game UI to optimize the audience’s experience.

Acknowledgements

This work was supported by the Canada Foundation for Innovation, and the Natural Sciences and Engineering Research Council of Canada.

References

- [1] E. Adams, *Fundamentals of Game Design*, 2nd ed., California, New Riders, 2010.
- [2] J. Babu, Video game HUDs: information presentation and spatial immersion. Rochester, NY, Golisano College of Computing and Information Sciences, Rochester Institute of Technology, Master of Science, 2012.
- [3] D. Bordwell, K. Thompson, *Film Art*, McGraw Hill, New York, 1993.
- [4] B. Bowman, N. Elmqvist, T. Jankun-Kelly, Toward visualization for games: theory, design space, and patterns, *IEEE Trans. Visual Comput. Graph.* 18 (11) (2012) 1956–1968.
- [5] E. Brown, P. Cairns, A grounded investigation of game immersion, Extended Abstracts of the 2004 Conference on Human factors in Computing Systems, Austria, ACM, 2004.
- [6] L. Caroux, K. Isbister, Influence of head-up displays’ characteristics on user experience in video games, *Int. J. Hum Comput Stud.* 87 (2016) 65–79.
- [7] L. Caroux, L. Le Bigot, N. Vibert, Maximizing players’ anticipation by applying the

- proximity-compatibility principle to the design of video games, *Human Factors* 53 (2) (2011) 103–117.
- [8] D. Conroy, P. Wyeth, D. Johnson, Understand player threat responses in FPS games, in: *Proceedings of the 9th Australasian Conference on Interactive Entertainment Matters of Life and Death*, Melbourne, ACM, 2013.
- [9] A. Donaldson, *Dark Souls 3: Miyazaki explains the difference between “difficult” and “unreasonable”*, Designer Interview, 2016, URL: < <https://www.vg247.com/2016/03/02/dark-souls-3-miyazaki-bloodborne-interview/> > (Accessed November 26, 2017).
- [10] E. Fagerholt, M. Lorentzon, Beyond the HUD – user interfaces for increased player immersion in FPS games, Dept. of Telecommunications, Indiana University, M.S., 2002.
- [11] M.A. Federoff, Heuristics and usability guidelines for the creation and evaluation of fun in video games, Dept. of Telecommunications, Indiana University, M.S., 2002.
- [12] S. Fragoso, Interface design strategies and disruptions of gameplay: notes from a qualitative study with first-person gamers, *Human-Computer Interaction. Applications and Services*. M. Kurosu, Springer International Publishing, 2014, vol. 8512, pp. 593–603.
- [13] A.R. Galloway, *Gaming: Essays on Algorithmic Culture*, University of Minnesota Press, Minneapolis, 2006.
- [14] D.J. Hill, Call of Duty Video Games Reaches \$1 Billion In Sales In 16 Days, Faster Than Cameron’s Avatar, Retrieved December 9, 2014, from < <http://singularityhub.com/2012/01/19/call-of-duty-video-game-reaches-1-billion-in-sales-in-16-days-faster-than-camerons-avatar/> > .
- [15] T. Hynninen, First-person shooter controls on touchscreen devices: a heuristic evaluation of three games on the iPod touch. Department of Computer Sciences. Tampere, Finland, University of Tampere. M.Sc. Thesis, 2012, 64 pages.
- [16] I. Iacovides, A. Cox, R. Kennedy, P. Cairns, C. Jennett, Removing the HUD: the impact of non-diegetic game elements and expertise on player involvement, in: *Proceedings of the ACM symposium on computer-human interaction in play – CHI Play 2015*, ACM, London, United Kingdom, New York, 2015, p. 13–22.
- [17] D. Ignacio, Crafting destruction: The evolution of the Dead Space user interface, *Game Developer’s Conference*, 2013. URL: < <https://www.youtube.com/watch?v=pXGWJRV1Zoc> > (Accessed November 26, 2017).
- [18] P. Isokoski, B. Martin, Performance of input devices in FPS target acquisition, in: *Proceedings of the international conference on advances in computer entertainment technology*, Austria, ACM, 2007.
- [19] C. Klochek, I.S. MacKenzie, Performance measures of game controllers in a three-dimensional environment, in: *Proceedings of Graphics Interface 2006*, Canadian Information Processing Society, Toronto, 2006.
- [20] S.C. Llanos, K. Jørgensen, Do players prefer integrated user interfaces? A qualitative study of game UI design issues, in: *Proceedings of the 2011 DiGRA international conference: think design play*, Hilversum, Netherlands, DiGRA, 2011.
- [21] J. Looser, A. Cockburn, J. Savage, On the validity of using First-Person Shooters for Fitts’ law studies, *Proceedings of the British HCI Conference*, Springer, 2005.
- [22] I.S. MacKenzie, *Human-Computer Interaction: An Empirical Research Perspective*, Morgan Kaufman, 2012.
- [23] V. McArthur, S.J. Castellucci, I.S. MacKenzie, An empirical comparison of “Wiimote” gun attachments for pointing tasks, *Proceedings of the ACM Symposium on Engineering Interactive Computing Systems – EICS 2009*, ACM Press, 2009, pp. 203–208.
- [24] E. McDonald, The global games market will reach \$108.9 billion in 2017 with mobile taking 42%. Newzoo Global Games Market Report, April 2017. URL: < <https://newzoo.com/insights/articles/the-global-games-market-will-reach-108-9-billion-in-2017-with-mobile-taking-42/> > (Accessed November 26, 2017).
- [25] R.P. McMahan, E.D. Ragan, A. Leal, R.J. Beaton, D.A. Bowman, Considerations for the use of commercial video games in controlled experiments, *Entertain. Comput.* 2 (1) (2011) 3–9.
- [26] D. Natapov, I.S. MacKenzie, Gameplay evaluation of the trackball controller, in: *Proceedings of the international academic conference on the future of game design and technology – FuturePlay 2010*, ACM, New York, 2010.
- [27] D.A. Norman, *The Design of Everyday Things*, Basic Books, New York, 2002.
- [28] R.J. Pagulayan, K. Keeker, D. Wixon, R.L. Romero, T. Fuller, User-centered design in games, in: J. Jacko, A. Sears (Eds.), *Human-Computer Interaction in Interactive Systems*, Lawrence Erlbaum Associates Inc, 2002, pp. 883–906.
- [29] M. Peacocke, R.J. Teather, J. Carette, I.S. MacKenzie, Evaluating the effectiveness of HUDs and diegetic ammo displays in first-person shooter games, in: *Proceedings of the IEEE Consumer Electronics Society Games, Entertainment, and Media Conference – GEM 2015*, IEEE, Toronto, 2015, p. 138–145.
- [30] N. Schaffer, Heuristics for usability in games white paper, Retrieved April 3, 2015, from < https://gamesqa.files.wordpress.com/2008/03/heuristics_noahschafferwhitepaper.pdf > .
- [31] R.J. Teather, I.S. MacKenzie, Comparing order of control for tilt and touch games, in: *Proceedings of the 2014 conference on interactive entertainment*, ACM, Newcastle, NSW, Australia, 2014, p. 1–10.
- [32] R. Vicencio-Moreira, R.L. Mandryk, C. Gutwin, S. Bateman, 2014. The effectiveness (or lack thereof) of aim-assist techniques in first-person shooter games, in: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, ACM, New York, NY.
- [33] R. Weber, NPD: 2015 video game sales flat compared to 2014, Retrieved March 9, 2016, from < <http://www.gamesindustry.biz/articles/2016-01-14-mpd> > .
- [34] G. Wilson, Off with their HUDs!: Rethinking the Heads-up display in console game design, 2006 (Retrieved November 12, 2014) < http://www.gamasutra.com/view/feature/130948/off_with_their_huds_rethinking_php > .
- [35] V. Zammito, Visualization techniques in video games, in: *Proceedings of Electronic Information, the Visual Arts and Beyond*, London, UK, BCS, 2008.
- [36] A. Zaraneek, B. Ramoul, H.F. Yu, Y. Yao, R.J. Teather, Performance of modern gaming input devices in first-person shooter target acquisition, *CHI’14 Extended Abstracts on Human Factors in Computing Systems*, ACM, 2014.