"I See You!": A Design Framework for Interface Cues about Agent Visual Perception from a Thematic Analysis of Videogames

Matthew Rueben
Matthew R. Horrocks
mrueben@nmsu.edu
mattrh@nmsu.edu
New Mexico State University
Las Cruces, New Mexico, USA

Jennifer Eleanor Martinez
Embry-Riddle Aeronautical University
Daytona Beach, FL, USA
martj326@my.erau.edu

Michelle V. Cormier
mishavc@nmsu.edu
New Mexico State University
Las Cruces, New Mexico, USA

Nicolas LaLone
Department of Information Systems and Quantitative Analysis
University of Nebraska at Omaha
Omaha, NE, USA
nlalone@unomaha.edu

Marlena R. Fraune
Department of Psychology
New Mexico State University
Las Cruces, NM, USA
mfraune@nmsu.edu

Phoebe O. Toups Dugas
Department of Computer Science
New Mexico State University
Las Cruces, NM, USA
phoebe.toups.dugas@acm.org

ABSTRACT

As artificial agents proliferate, there will be more and more situations in which they must communicate their capabilities to humans, including what they can "see." Artificial agents have existed for decades in the form of computer-controlled agents in videogames. We analyze videogames in order to not only inspire the design of better agents, but to stop agent designers from replicating research that has already been theorized, designed, and tested in-depth. We present a qualitative thematic analysis of sight cues in videogames and develop a framework to support human-agent interaction design. The framework identifies the different locations and stimulus types – both visualizations and sonifications – available to designers and the types of information they can convey as sight cues. Insights from several other cue properties are also presented. We close with suggestions for implementing such cues with existing technologies to improve the safety, privacy, and efficiency of human-agent interactions.

CCS CONCEPTS

• **Human-centered computing** → **HCI theory, concepts and models**; *Ubiquitous and mobile computing*; • **General and reference** → *Design*.

KEYWORDS

human-robot interaction, thematic analysis, videogames, sight cues, qualitative methodologies

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3517699

ACM Reference Format:
Matthew Rueben, Matthew R. Horrocks, Jennifer Eleanor Martinez, Michelle V. Cormier, Nicolas LaLone, Marlena R. Fraune, and Phoebe O. Toups Dugas. 2022. "I See You!": A Design Framework for Interface Cues about Agent Visual Perception from a Thematic Analysis of Videogames. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 22 pages. https://doi.org/10. 1145/3491102.3517699

1 INTRODUCTION

As artificial agents (e.g., robots, virtual assistants) proliferate, there will be a growing need to interact with them. A key part of interaction is situation awareness – understanding what others are or will be doing. We envision a need for agents to share information about what they are able to perceive and to what they are attending with collocated humans. To support better design of these types of interactions, we look to a domain that has long designed them – videogames. From videogames, we extract sight cues – the ways that games make agents' visual awareness perceptible to players – and use them to assemble a design framework for human-agent interactions.

In human-agent social interaction, agents can be difficult to understand as they are complex systems of hardware and software. The humans who make artificial agents understand how they tend to work, but, as with most programmed entities, the theory of human activity that an agent represents is difficult to pass on to other humans [77]. Despite this difficulty, there is existing guidance for the processes of humans and programmed entities to get to know one another.

Norman considers how the designer and the user each bring their own *conceptual models* [78] – special *mental models* [42, 88] of how the system works that enable them to reason about what the system will do and how it will act. Mental models are formed by previous experiences; humans, as users, will bring these experiences to every product, thus making it difficult to design a system that has to be learned. It is a paradoxical relationship in that learning to use a

Figure 1: Example of a sight cue. [SMW] Visual Target Behaviors: *Visual Target Acts Differently*. In *Super Mario World* (1990) (and all Mario Games), these ghosts, known as Boos, will not attack the player [LEFT] unless the player has their back to them. When the player has their back to the Boos, their mouths open and their hands move slightly back as they begin floating toward the player [CENTER]. If they strike the player, the player will lose a life or power-ups depending on their status. To stop the Boos' advance, the player can turn around [RIGHT]. The Boos' behavior is a cue about whether Mario (the perceiver) sees the Boos (the visual targets). (Screenshot taken 🅭🅯 the authors.)

product would save time, yet the investment of time to learn that product is too much overhead up front [18].

Taking robots as an example class of agents, these conceptual models can include the user's perception of what the robot can sense [65], what it knows [66], and other abilities and attributes [22]. When working with robotic teammates, human partners need situation awareness [35, 36] to understand what the robots know and intend to do next [56, 94]. Robots deployed in the home or in public places will need to communicate their sensing capabilities to users – e.g., so people can make informed judgments about personal privacy [92]. *Vision* is a popular sensing modality in robotics applications [96] due to the availability of small, inexpensive, high resolution RGB cameras and powerful computer vision techniques [41]. Helping humans form *useful* conceptual models of what robots can "see" is critical for a number of activities with which robots could help.

A promising strategy when designing for human-agent interaction is to pull on prior examples of communications between agents and humans that have succeeded despite these difficulties. If agent and robot designers pull on *designs from videogames*, we expect users to be able to assemble more functional conceptual models. This paper contributes a design framework from a rich historical description of routinized ways that designers communicate to players in videogames, with a focus on sight cues. We expect this framework to seed human-agent interaction with previously existing, successful cues to help humans learn about agents in the wild. While prior work has looked at the use of animation principles for design [44, 70, 84, 86, 100, 102], games are an exciting inspiration in that they are *interactive*, often confronting the player with AI agents. Thus, in this paper, we analyze sight cues that indicate that an agent has visually sighted another agent, including the player's agent.

Our research goal is to understand how to better integrate humans and robots in social contexts. As humans work with robots, there will be a greater need for robots to provide sight cues to humans in the same space, serving as a form of awareness work [49, 105]. To this end, we aim to understand sight cues from a variety of intelligences and media; as one step toward this, we look at the *design of sight cues in videogames* to answer the following research question:

> **Which prior design strategies can be found regarding human-agent interactions such that agents supply cues in support of human situation awareness?**

1.1 Research Approach

What we call *sight cues* have been a core part of feedback [91] in gameplay; these appear in numerous titles that have been carefully designed by expert teams and playtested thoroughly before going to market in order to make fun and interesting games. In this paper, we analyze this design space for the purpose of designing human-agent interactions, which is, necessarily, exploratory and in need of a suitable approach. We develop a thematic analysis [15] of videogames [106] to understand how sight cues among agents have been designed. This approach is qualitative – we aim to build a representative landscape [14] of existing designs to understand what they are and how they are useful for designing for human-agent interaction in general, and human-robot interaction (HRI) specifically. The approach is similar to that used in prior research deriving design insights from videogames (e.g., [85, 101, 106]). We collected sight cues from videogames spanning decades and analyze these.

The games from which we assembled the data items in our data corpus are provided in the Ludography and are referenced in the text in the same way as texts, but with an abbreviation of the game's title, e.g., *Animal Crossing: New Horizons* [AC:NH]. Sight cues were analyzed in-depth through an iterative process that involved identifying cues, re-visiting identified cues, adding new ones, and developing themes. We looked at these themes through a number of lenses to understand how to apply them to human-agent interaction. The resulting framework enables designers to answer questions about what designs are possible and what they communicate.

1.2 Positionality Statement

It is important to understand the researchers, especially in qualitative work that is highly interpretive and wherein the researchers' framings matter. Most of the researchers on this project are *well played* [3, 30] – they have decades of experience playing games, working in games culture, researching game design, and otherwise being immersed in the games space. This is not a call to authority, but, rather, to assure the reader that we start this work with a grounding in the types of games out there and how they are designed. Further, the work relies heavily on experience of gameplay. While we do not aim to be exhaustive in our data corpus, we do expect the data corpus, which seeds from the researchers' knowledge, to be reflective of the space and the general landscape [14] of agent-based cues.

To further introduce the authors: we also bring backgrounds in HRI, human-computer interaction, and information technology for search and rescue. We are situated across the United States of America. Other researchers with other backgrounds may bring different foci, perspectives, and/or concerns into an analysis, arriving at different conclusions.

2 CONTEXT & TERMINOLOGY

In the present analysis, we use several terms intentionally. The human playing the video game is the *player*. The player typically controls a character or an avatar [8], which, for the purposes of this paper, we call a the *player's agent*. Those agents within a game not controlled by the player are *non-player agents*. We chose to use "agents" to describe objects in videogames in this way in order to more directly connect human-agent interaction to video game entities and sidestep arguments about the differences between avatars, characters, loci of control, etc. [8, 19].

We define a *sight cue* as any stimulus perceptible by the player that contains information about *what an agent can or does perceive visually*. Sight cues do not necessarily happen in the visual channel – the "sight" component refers to how an agent perceives another (e.g. ghosts (or Boos) in *Super Mario World* [SMW] seen in Figure 1). This act of "sight" is best described as a mathematical calculation that accounts for the geometry and lighting in the virtual environment and the game's rules for visual perception to simulate one agent seeing another. This event of visual perception we call a *sighting*. Within games, a good example of a *sight cue* is a non-player agent indicating to the player that they saw another agent by turning their head toward them.

In a sight cue, the agent that perceives visually is the *perceiver* and the agent or object being seen the *visual target*. When the information is about what an agent *can* perceive visually, we use the term *field of view (FoV)*. We also use the phrase *line of sight (LoS)* to describe an imaginary ray originating at the agent's eyes (or equivalent) and extending along their gaze direction until it terminates at an occluding obstacle or the edge of their FoV.

A key term used in games and derived from film is the notion of *diegesis* [52, 54, 79]. Elements of a game are diegetic if they exist within and are consistent with the fiction of the game's world (e.g., the agent "perceives" them and the agent responds to them as such). Elements are non-diegetic if they exist for the player, but not the agents (e.g., the game's soundtrack; most head-up display (HUD) information in a third-person perspective).

3 RELATED WORK

We begin our review of prior work by describing how human teams communicate. Next, we describe prior work on our problem domain: mental model formation about the perceptual capabilities of artificial agents. The final two subsections use existing research to highlight the promise of videogames as inspiration for designing sight cues.

3.1 Communication in Human Teams

Team members need to communicate about and coordinate their work (and play); highly effective teams implicitly coordinate [37, 39, 71, 105] – they reduce the need to explicitly communicate by developing shared mental models [17, 31, 37, 46, 72, 94, 99] and situation awareness [35–37, 51]. A key part of implicit coordination is awareness of activities being performed [47, 49, 50]. An understanding of what tasks others are performing, what elements of a workspace they are interacting with, and where they are in the environment enable coordination and situation awareness [35–37, 51]. In many scenarios, human team members render their work highly perceptible to others, often through announcing activities [49]. At the same time, workers develop attention to the sometimes-subtle cues of others [49, 105]. This *awareness work* is a form of implicit communication that supports people in working together effectively [49] and in developing shared mental models [57, 72, 99]. The present research studies how artificial agents can support coordination and increase situation awareness in human-agent interactions by emitting cues about what they can see.

3.2 Human-like Communication with Embodied Conversational Agents

As artificial agents become capable of mimicking humans in more ways, researchers have studied and applied human-human communication patterns to human-agent interactions – embodied conversational agents (ECAs) are a good example [20]. The addition of a body enables an ECA to be more multimodal and expressive than a chatbot [83]. This also, however, makes for a more complex design challenge. All the different types of facial displays and hand gestures are possible in addition to the verbal and nonverbal components of speech, and synchronization between all these modalities is important [20]. Another example modality is gaze, which can communicate several different things and involves not just the eyes but also the head and body orientation [82, 108]. This work leverages the applicability of videogame design to the design of ECAs and other artificial agents: videogames often feature ECAs, and take place in a controlled environment [97] and sometimes even in virtual or augmented reality [4] like some ECAs do.

(1) Familiarization: read, observe, and immerse oneself in the collected data, taking notes.	(4) Reviewing Themes: check themes against the corpus.
(2) Initial Coding: connect codes to observations and collect features relevant to each code.	(5) Defining and Naming Themes: specify each theme and relate them to each other.
(3) Searching for Themes: connect codes, identify higher level structures.	(6) Producing the Report: assemble the themes together with extracts from the corpus.

Table 1: Braun and Clarke's general phases of thematic analysis [15].

This work therefore draws on examples from the vast set of extant videogames to inform human-agent interaction design. We also focus on a relatively unexplored part of human-agent communication – namely, communicating about the agent's perceptual state and perceptual capabilities – as opposed to the content or process of the conversation, or expressing artificial emotions [21].

3.3 Forming Mental Models of Agents' Visual Faculties

Forming mental models of an agent's status and capabilities is part of human-agent teaming [114]. This requires the agent to engage in effective communication with humans. For example, a telepresence robot could use motion to signal that the human using it is about to start speaking [48]. The present research identifies how designers might render perceptible agent's perceptual capabilities – what they can detect, and how well.

A related area of research is explainable AI, which seeks to support mental model formation about agents by humans via automatically generated *explanations* of the AI's recommendations or other behaviors [26, 43]. Several researchers have summarized the abundance of existing literature on how humans create and process explanations in social psychology, cognitive psychology/science, and philosophy [74, 110]. Work on explanations about AI systems has looked at, for example, the soundness or completeness of the explanation [59, 60], how it compares to a human's explanation [61], and the different properties of the AI's performance that influence the human's mental model formation [6]. This work also focuses on supporting mental model formation about artificial agents by humans, but about an agent's perceptual capabilities in particular, and via cues instead of explanations. We leverage videogames as contexts in which the formation of accurate mental models by players about certain aspects of AI agents is often crucial for satisfying play.

Non-teaming scenarios also involve multiple stakeholders. Humans riding in an autonomous car, driving other cars, and walking on the sidewalk all might be stakeholders as autonomous vehicles emerge, creating an urgent need to communicate about their sensing capabilities. Researchers have studied ways for an autonomous vehicle to communicate when it has detected a pedestrian and intends to stop for them [38, 40].

HRI researchers have also begun studying this problem. There have been several studies of how first-time users of a robot determine what it can sense and to what it responds. Three different studies report that participants experimented with the robot by touching it, talking to it, waving, or snapping their fingers [75, 80, 93]. Rueben et al. tracked the development of mental models about a new robot over six weeks [90]. Lee et al. highlighted the importance of good cues about robots' sensing capabilities in their study of privacy concerns about a robot called "Snackbot" [65]. Very few of the people who were interviewed about Snackbot could identify what sorts of sensors it had – they were especially surprised to learn about the omnidirectional camera – or guessed that it was recording audio and video. This work draws insights from videogames about how robots and other artificial agents can emit helpful cues about what they can see to avoid privacy violations and other problems.

Robots have many ways they can signal verbally [7] or nonverbally [23] that might be useful for communicating about their visual perceptual capabilities [1]. Information can even be implicit in dialogue [58] or motion [32]. Agents in videogames are good analogs for robots because they possess many of the same affordances for interacting with humans (touch is an exception). Videogames are also interactive like robots are, setting them apart from animated films, for example. Signaling about an agent's vision contributes to recent research on transparency in both human-robot interaction specifically [13, 67, 111, 112] and human-agent interaction more generally [25].

3.4 Videogames as Inspiration for Human-Agent Interaction

Videogames contain design patterns [12, 89] that could be useful for human-agent interaction design [103]. Some virtual agents are inspired by videogames; some are even animated using game engines. For example, Constable et al. used video game-like interactions between animated agents to represent dominance and discord in a conversation between two humans [28]. As early as 2006 there had been HRI research that took inspiration from videogames [87]. User interfaces for robot operators have been inspired by first-person shooter game interfaces [55], as well as by specific user interface elements such as minimaps and arrows pointing to off-screen agents [81, 95, 109]; similar approaches have been employed when designing for teleoperation [85]. A recent workshop aimed to catalyze further research into how video game design can inspire HRI design [76]. We expect further cross-fertilization as games serve as an important platform for virtual agents, where they might interact with and help players [33, 34, 69]. This work contributes in the specific area of what insights videogames would have for human-agent interaction designers about how to communicate agents' sensing capabilities – specifically their vision – to humans.

3.5 Cues in Videogames

Videogames provide players with a variety of informative cues. Some cues support the player's situation awareness in games [53,

Cue Origination: What agent or other system supplies the sight cue to the player?	Cue Visualization: Is the sight cue visualized? How? Where is the visualization?
Information Type: What does the sight cue indicate about the agent doing the sighting?	Cue Sonification: Is the sight cue sonified? How? From where does the audio originate?
Information about Perceiver: Does the sight cue indicate to the player which agent sighted another? How?	Diegesis: Is the cue diegetic? How?

Table 2: Questions for each sight cue from phase 3 that describe the cue. Each question's title is used as a property in phase 4.

113]; others are derived from players' use of the affordances of the game to communicate with each other [64, 104]. Players invent impromptu cues when the game does not provide them with the cues they need [104]. Players can also use different channels and pay attention to different things to get the information they want [27]. To our knowledge, this is the first study that focuses specifically on cues about what agents in videogames can *see*. We perform a thematic analysis of such cues from a variety of mostly single-player games – our focus was on cues crafted by game designers, with an emphasis on how computer-controlled agents interact, though we also attended to what players improvise [104].

4 METHODOLOGY: ITERATIVE THEMATIC ANALYSIS OF VIDEOGAMES

We co-opt Braun and Clarke's thematic analysis (Table 1) [15], much like Toups Dugas et al. [106], to investigate sight cues in videogames. Unlike quantitative approaches, where statistical analysis lends credibility to the work, qualitative researchers must lay out their processes transparently. As with any qualitative work, other re-searchers may bring different perspectives and *a priori* assumptions into an identical analysis and data corpus. They might expand the data corpus from their own knowledge or searches and arrive at alternative explanations than we do. The goal here is *not* to identify replicable design approaches, but to provide a landscape for future designers and researchers to use. As such, the focus is on creat-ing something useful, but not necessarily an absolutely exhaustive representation of how the world is.

Thematic analysis is a qualitative approach to understanding a data corpus. A *data corpus* is a collection of various data items. A *data item* could be anything at a particular level of analysis, often text; we later discuss the use of sight cues as data items. The objective of thematic analysis is to draw out *themes* – qualitative connections between data items. Thematic analysis can be useful for exploring a new research space and for developing a landscape – identifying areas of the space that are useful for research and design.

4.1 Inclusion, Exclusion, & Stopping Criteria

As part of building a corpus of sight cues, it is necessary to identify what data to analyze. Inclusion and exclusion criteria emerged throughout the research process during repeated iterations of phases 1–5 (Table 1). Stopping criteria – how we know when we have found "enough" data for analysis – emerged through phase 4. What is presented here is a summative and reflective assessment of how inclusion, exclusion, and stopping criteria worked at the point of phase 6.

For a sight cue to be included, there must be a *sighting*: an instance where a perceiver sees a visual target. Often, the player's agent is being seen by a non-player agent, but sometimes the sighting is between two non-player agents. The visual target could be an inanimate object or another agent. For a sighting to occur, an agent must have some means of "visually perceiving" visual targets. We put "visually perceiving" in quotation marks because vision is simulated in digital games (e.g., by collision detection with invisible 2D or 3D view cone objects, or presence of agents at virtual coordinates); we require that the game have some simulation of sight to be included.

When a sighting occurs, there must be an associated stimulus for the player. Stimuli were considered broadly, including agent behavior in addition to user interface components that communicate game state. Sight cues could be designed by game developers or improvised by players, and could be displayed in any way that is accessible to the player: e.g., an agent's movements or speech, sound effects, music, or non-diegetic effects such as those on the HUD. Occurrences that did not meet the above inclusion criteria – e.g., if a cue was about an agent hearing another agent, or something else entirely – would be excluded. No additional exclusion criteria were developed.

As we identified sight cues, we evaluated them for novelty in phase 4 with regards to those previously observed. Our stopping criterion was when we could no longer find cues that were significantly novel compared to those already in our data corpus. Cues that had been seen before were given lower priority, or, potentially not analyzed. Thus, it is worth noting that the data corpus is not, and *is not intended to be*, exhaustive; it is meant to be representative. Some cues that could potentially be included are not because prior cues give us sufficient information about the design space.

4.2 Process

The process is iterative; we regularly returned to earlier steps in the process, then followed the subsequent steps again from there. The initial analysis was carried out January–May 2021 and was revisited in November 2021–January 2022. We had weekly all-team meetings to discuss the direction of the research, the overall state of the corpus, and to discuss inclusion and exclusion criteria. All authors brainstormed cues to potentially include in the data corpus. Focused, paired meetings were carried out at least bi-weekly wherein particular cues were analyzed in-depth by playing the game or poring over videos that showed the cue. All meetings were carried out remotely over video conferencing software due to a combination of the COVID-19 Pandemic and the remote locations of the researchers.

Figure 2: [MK8] Perceiver Behaviors: *Turning Towards :: Eyes Only*; *Turning Towards :: Head or Body.* In *Super Mario Kart 8* players race their go-karts around fantastical tracks against characters from Mario games [LEFT]. When the player approaches other racers, they look toward that character and the character being passed looks back [CENTER]. If they hit each other, the sizes (small, medium, large) of the two characters are compared and if there is a difference, the larger character will make the smaller one stop [RIGHT]. (Screenshot taken 🅭🅯 the authors.)

Phase 1: Familiarization. We worked from an initial research goal of understanding sight cues in games, relating them to human-agent interaction use cases. As the researchers looked for and added new cues, we considered future designs. We reflected upon our experience as well-played scholars to identify games with relevant designs, then identify sight cues associated with those games. We discussed inclusion criteria and identified types of cues that we had seen or heard about in games. Each team member worked from their experience independently, recalling cues, and then these cues sparked further ideas from the rest of the team. Much of the work involved identifying a game and sight cue then collecting data about it, e.g., a gameplay video, manual, descriptive details. In keeping with landscape sampling, we focused on finding the most novel cues and adding them to our data corpus; when cues were too similar to previous cues, they were excluded. Much of the initial thinking in this space included games with stealth-style mechanics, where the player's agent needs to avoid enemy agents. We also reflected upon how tabletop role-playing games use awareness and maps [62, 107].

Phase 2: Initial Coding. Our work in phase 2 led to the development of our definition of what constituted a sight cue. **A sight cue is any stimulus perceptible by the player that contains information about what an agent can or does perceive visually.** What level is atomic for a sight cue – i.e., what constitutes a single cue versus several – was also determined through repeated, close study of the data corpus.

Phase 3: Searching for Themes. As we looked at cues, we identified characteristics shared across them, looking for *themes*. These themes began as notes on each cue, describing it, that then became more codified. The outcome was a series of questions about each cue that we could answer in a free-form way; Table 2 provides a selected subset.

Phase 4: Reviewing Themes. From the questions in phase 3, we worked through the free-form answers for connections between cues across games and were able to identify themes. Our data corpus transitioned from descriptions of cues to a set of *properties* (indicated in Small Caps, as in other qualitative work), each of which could be set to a single *option* (*italicized*). The set of possible options for each property was developed iteratively both by adding new cues to the data corpus and by brainstorming. The set of properties we directly used to build the framework we present in this paper are: Cue Origination, Information Type, Information about Perceiver, Cue Visualization, Cue Sonification, and Diegesis (Table 2). We evaluated sight cues for potential addition to the data corpus based on whether there were already examples in the corpus of cues with the same properties – this was how we evaluated cue novelty for our stopping criterion as mentioned above.

Phase 5: Defining and Naming Themes. We considered ways to organize the properties and identify their relevance to the design of human-agent interaction. The primary analysis, which was intertwined with the others, focuses on how we can best support designing human-agent interactions, which resulted in Table 3.

Phase 6: Producing the Report. Phase 6 involved identifying the interesting components of the framework and applying them to interactions with agents. We also selected which of our several framings of the corpus to present (in Section 7) beyond that of our main framework in Table 3.

5 SUMMATIVE DATA CORPUS CHARACTERISTICS

The data corpus was developed over months of iterative analysis; at intermediate phases, it looked different from what is reported here. This is a characterization of the data corpus as it existed during phase 6. The Ludography provides the list of all games in the corpus. The data corpus contained 60 sight cues across 39 separate games and 2 game series (that featured the same sight cue across all of their entries). We placed just one cue in the corpus from most games and the game series, comprising 28 of the cues. Sometimes we selected multiple cues from the same game: 2 cues from 10 games (20 cues

Figure 3: [AC:NH] Perceiver Behaviors: *Gesture*; *Locomotion :: Approach, Retreat, Leaving Spot or Patrol.* In *Animal Crossing: New Horizons*, villagers living on your island will periodically try to get your attention [LEFT] by yelling for you. When the villager yells at you, they run toward you [CENTER]. Once they reach your location, they will do one of a number of pre-programmed activities such as displaying a reaction (e.g., shocked, bored, happy), offering to sell you an item, or offering to buy an item you have [RIGHT]. (Screenshot taken ⓒⓘ the authors.)

total) and 4 cues from 3 games (12 cues total). The games’ release dates range from 1975–2020.

To better understand the characteristics of the games in the data corpus, we looked up each one of them in MobyGames[1], a widely used database that crowdsources, with rigorous standards and checks, data on every game ever published. We chose MobyGames because the data have very specific definitions that are consistent across the database. We considered GENRE, USER INTERFACE PERSPECTIVE, GAMEPLAY TYPE, and INTERFACE TYPE for the games in the data corpus. Each of these fields might have multiple values for a single game because different parts of the game might play differently (e.g., the player directs agents on a map *and also* controls a single agent), some fields are game configuration settings (e.g., third-person-perspective games that can be played in first-person by changing a setting or clicking a button), and/or games may fit into multiple GENRES. While the categories are largely self-explanatory, additional details can be found on MobyGames’ Genre Definitions page[2].

In our present data corpus, most of the games were in the *Action* GENRE, followed by *Role-Playing (RPG)* and *Strategy / tactics*; a few games fell into other genres. *1st-person* games featured heavily in the data corpus, but many of these also offer a *Behind view* (a special form of *3rd-person*); many also used a *Bird’s-eye view*. The data corpus featured many GAMEPLAY TYPES, but *Shooter* was the most common, followed by *Stealth* and *Puzzle elements* evenly with just one *Racing game* seen in figure 2. Finally, most games had *Direct control* INTERFACE TYPES or INTERFACE TYPE was not an entered field for them.

On reflection, the types of games covered largely make sense given the inclusion and exclusion criteria. *Action*, *Role-Playing (RPG)*, and *Strategy / tactics* games will all rely heavily on a player being aware whether non-player agents can or do see the player’s agent. Similarly, *Shooter* and *Stealth* GAMEPLAY TYPE requires a player to know how hidden their agent is, and when they have been spotted by enemies. We included the *Racing / driving* game [MK8] (Figure 2) to deepen our landscape and better represent games that are competitive against a mixture of agents and humans. These characteristics of our data corpus also, however, influence the types of sight cues we found, a fact that we revisit later in the Limitations (Section 8.3).

6 FRAMEWORK OF CUE PROPERTIES FOR HUMAN-AGENT INTERACTION DESIGNERS

Table 3 shows our framework of three core cue properties and some relationships between them. The table is organized around the stimulus (visualization and sonification) types in the center column and their associated locations/origins in the leftmost column. Each cell in the remaining two columns on the right shows the prospects for creating a cue with a certain property (according to the cell’s column) out of a certain type of stimulus (according to its row) based on our research.

The following subsections use examples from our corpus to first describe all of the stimulus types, then the ways to indicate an agent’s *FoV* or *LoS*. Throughout the present research, we have provided 13 examples of cues including Figure 3 which shows an animal from *Animal Crossing: New Horizons* being triggered by FoV and LoS and running to the player. Figures 1–9 and 12–15 provide images of some especially interesting sight cues in our corpus. We conclude the section by looking at additional observations from our data.

6.1 Visualizations Used as Sight Cues

Here we introduce the STIMULUS TYPES that we identified as we built our corpus. We use example cues that indicate *Detection* or *Attention* since there is at least one example cue in the corpus with that property for every stimulus type. We work from the LOCATION

[1] https://www.mobygames.com
[2] https://www.mobygames.com/glossary/genres/

Figure 4: [LZ:WW] Visual Effects: *Viewport/UI Effects*; *Icon :: Symbol.* In *Legend of Zelda: Wind Waker,* the player assumes the role of Link, a small boy who grew up on an island. One day, a giant bird comes to the island and takes Link's sister. Link leaves to go and rescue his sister and goes on a world-changing adventure. Each Zelda game is independent of each other Zelda game. In *Wind Waker,* Link wanders around areas exploring and solving puzzles [LEFT]. When a monster appears, the viewport narrows and highlights the enemy [CENTER]. After the viewport shift and throughout the rest of the dungeon or area Link is wandering in, instead of a viewport shift, the monster is highlighted when targeted by an arrow which allows Link to circle the monster in order to fight it [RIGHT]. (Screenshot taken 🅭🅯 the authors.)

of Visualization property to organize our discussion of the table and explain specific examples that drive the data on the table.

Perceiver Behaviors. Any *Locomotion* by the perceiver that is oriented around their visual target's location implies detection. Some perceivers run away (e.g., animals [LZ:BW]); others come closer (e.g., enemies [EQ], as in Figure 12). Detection can also cause perceivers to break from their spot or patrol, such as to sound an alarm [W2:TNC]. *Other Perceiver Behavior Changes* might include moving more quickly or altering patrol paths (Figure 7) [MGS2]. *Turning Towards* the visual target (i.e., orienting gaze direction towards them) is another common behavior that indicates detection and attention (Figure 2) [Pik]. The perceiver might also make *Gestures*, such as facial expressions (Figure 3) [CTow, AC:NH] and threatening hand gestures [CTaxi]. *Changing Pose* includes standing up [CTaxi], crouching (Figure 13), and stiffening with surprise [AC:NH]. *Other Actions* by the perceiver include opening doors [CTow], blowing a horn (Figure 5), and attacking the visual target (Figure 12), such as by shooting at them [JB:NF].

Visual Target Behaviors. There are probably many ways that the visual target can react to being sighted, like the diversity of ways the perceiver can react that we just described, but we found very few sight cues featuring the visual target for our corpus. One is in *Super Mario World*: ghost-like enemies ("Boos") change pose, change facial expression, and give chase when the player's agent looks away from them (Figure 1) [SMW].

Visual Effects involve ways in which user interface and user-interface-like elements of the game are employed as cues.

For *Viewport/UI effects* the entire screen does something, like flashing [DQ], letterboxing [LZ:OT], zooming in (Figure 4), or tinting gray. *Captions* (e.g., in Figure 7) are joint with sonifications, rendering as text what is being said in a dialogue (Stimulus Type :: Sonification: *Voice Interaction :: Verbal*). A perceiver can either say outright (e.g., "I can see you!") or imply that they have seen their visual target. *Text* can be rather simple, such as a single word placed in the middle of the viewport: "[HIDDEN]", "[CAUTION]" when the perceiver is alerted, or "[DANGER]" when they have fully detected the player's agent [FO3].

Some *Icons*[3] depict the perceiver or some real object (e.g., the eye of Sauron [LoTR:RotK] as in Figure 6, or pirate ships [SoT]), while others are more abstract symbols such as eyes [ES4] or exclamation points (Figure 8) [MGS2]. *Gizmos*[4] are graphics, often non-diegetic, that are attached to agents or objects. They might not appear except in certain game modes or when an object is selected. Examples include view cones (Figure 9) [Des3] and cursors bracketing an enemy [LZ:OT]. *Light Beams* such as diegetic flashlight (Figure 8) [MN] or searchlight beams [LZ:WW] were also used to show an agent's attention (and FoV). Additionally, when a guardian agent detects the player's agent in *The Legend of Zelda: Breath of the Wild*, it shines a laser on them – this also shows its LoS [LZ:BW].

Characteristics of Visual Effects describe how *visual effects* change responsively. *Colors* are often used to indicate detected versus undetected states (e.g., green for undetected, yellow for alerted, red for detected [Des3], as in Figure 9), as are *Shining, Glowing, or Flashing* of icons or other game elements [DQ].

User Perspective or Affordance Changes. Showing a *First Person Perspective* of an agent in the game gives the player direct access to everything that agent can see. *Camera Movements* in a first-person perspective can indicate attention or detection. E.g., in *The Legend of Zelda: The Wind Waker*, when the player is seeing the agent's view through a spyglass, the camera indicates that the agent has noticed a giant bird by suddenly centering on it and zooming in [LZ:WW]. A cutscene or other *Interruption of Player Control* does not indicate by itself that a sighting has occurred, but rather directs

[3] A discussion of semiotics of icons used for sight cues is beyond the scope of this work.

[4] as defined in the Unity game development platform: https://docs.unity3d.com/Manual/GizmosMenu.html

Figure 5: [LZ:BW] Perceiver Behaviors: *Other Actions*; *Locomotion: Approach, Retreat, Leaving Spot or Patrol*; Visual Effects: *Icon :: Symbol.* In *Legend of Zelda: Breath of the Wild*, the player assumes control of Link, a hero who failed and perished and was placed in a healing chamber for centuries. Upon waking, they enter a hostile world filled with aggressive creatures who wish to end the player [LEFT]. While the creatures do not actively seek out the player, they will react upon the player breaching their field of view. In the [CENTER] picture, the player has breached the lookout's field of view. Upon breaching that FoV, the agent blows a horn, alerting the other agents nearby [RIGHT]. Not seen is the agents running toward the player upon being alerted. (Screenshot taken ⓒⓘ the authors.)

the player's attention to other cues, or emphasizes them [LZ:OT]. A sighting can also lead to *Changes to the Player's Controls or Affordances*, such as when combat options become available when spotted by an enemy [dnd].

6.2 Sonifications Used as Sight Cues

Audio can also function as a sight cue, whether it comes from an agent or from the game. We break down this discussion by the Origin of Sound property.

Agent Sounds. *Voice Interaction* cues can be verbal (i.e., dialogue [Por] as in Figure 7) or nonverbal (e.g., screams, gasps [CTaxi]). Unlike other sounds, agent sounds can be used to indicate *FoV* or *LoS* as an agent might say or shout something in response to sighting an agent. *Action Sounds* are the sounds of actions that function as sight cues, such as the perceiver shooting [JB:NF] or running towards or away from the visual target [LZ:BW], or the sounds of the creature's scissors snapping in *Clock Tower* (Figure 14).

Atmospheric Sounds. *Sound Effects* and *Music Changes* can also indicate detection or attention by a perceiver, such as in *Clock Tower* [CTow] when a startling sound and creepy music coincides with the emergence of a monster from its hiding place (Figure 14).

6.3 Sight Cues with Information about Field of View or Line of Sight

We saw two main ways that information about an agent's *FoV* or *LoS* was available to players. The first way was to directly display the FoV or LoS, either through a non-diegetic *Gizmo* [LZ:MM, Des3] or by a diegetic stand-in such as a flashlight beam [MN] or eye laser beam [LZ:BW]. The second main way to show FoV also includes LoS: showing the agent's viewpoint via a *First Person Perspective*. The player simply sees what the agent sees, akin to viewing a robot's camera video feed.

An agent's FoV or LoS, or perhaps just approximate gaze direction, can also be implied by the orientation of their eyes, head, or body (e.g., the Beamos enemy in *The Legend of Zelda* series [LZ:OT]). The visual target can also imply a perceiver's FoV or LoS by responding when it includes or contacts them, even if they have not been detected (e.g., the Boos in *Super Mario World* [SMW]). FoV or LoS can also be implied via written text or narration [M2:SB], and presumably also in dialogue. Some stimulus types are generally too vague to describe an FoV or LoS, including nonverbal utterances, action sounds, other sound effects, and music changes.

7 OBSERVATIONS ON CUE PROPERTIES & CONSIDERATIONS

Several additional interesting properties of sight cues appeared in the analysis beyond those used in our framework in Table 3. These data points provide further ways that designers might build interactions and provide additional insights into the design of sight cues. Here we describe three: how cues specify perceivers, stimuli whose primary purpose is not to be a sight cue, and non-binary sighting information.

7.1 Whether the Cue Specifies the Perceiver

One of our questions in Table 2 was, does the sight cue indicate to the player which agent sighted another? Many of the ways that cues in our corpus did this were straightforward: the perceiver was the agent performing the action (*Perceiver Behaviors*), *Visual Effects* were pinned to or emanating from the perceiver, or the perceiver was referred to by *Text* or *Verbal* dialogue.

We also observed some more creative ways that sight cues specify the perceiver. For example, cues can refer to the perceiver via a picture or logo, or indirectly by their associates. In the *Lord of the Rings: Return of the King*, Sauron is identified as the perceiver both by an *Icon :: Depiction* of his distinctive eye and by his agents

Location of Visualization	Stimulus Type :: Visualization	Information Contained	
		Detection / Attention	*FoV / LoS*
Perceiver Behaviors	*Locomotion :: Approach, Retreat, Leaving Spot or Patrol*	•	×
	Turning Towards :: Eyes Only	•	•
	Turning Towards :: Head or Body	•	○
	Other Perceiver Behavior Change	•	?
	Gesture	•	○
	Changing Pose	•	?
	Other Action	•	?
Visual Target Behaviors	*Visual Target Acts Differently*	•	•
Visual Effects (any location)	*Viewport/UI Effects*	•	?
	Captions, Other Text	•	•
	Icon :: Depiction	•	?
	Icon :: Symbol	•	?
	Gizmos	•	•
	Light Beams	•	•
Characteristics of Visual Effects	*Colors*	•	?
	Shining, Glowing, Flashing	•	○
User Perspective or Affordance Change (location: HUD or viewport)	*First-Person Perspective*	•	•
	Camera Movement	•	•
	Interrupts Player Control (e.g., Cutscene)	•	•
	Other Change to Player Controls or Affordances	•	•

Origin of Sound	Stimulus Type :: Sonification	Information Contained	
		Detection / Attention	*FoV / LoS*
Agent Sounds (including narration)	*Voice Interaction :: Verbal*	•	•
	Voice Interaction :: Nonverbal	•	×
	Action Sounds	•	×
Atmospheric Sounds	*Sound Effect*	•	×
	Music Changes	•	×

Table 3: Table of sight cue properties to guide sight cue design for human-agent interaction. The table is organized around the leftmost column: the location of visualization or origin of sonification, which are broken down in the next column that identifies a stimulus type. The two rightmost columns show the prospects from our research of creating a cue containing the information indicated at the top of that column using the stimulus type in that row. Green cells with a darkened dot (•) mean we have an example from our corpus of that combination. Orange cells with an open dot (○) mean we believe that combination to be feasible from our experience doing this research and considering the feasibility of hypothetical cues. White cells with a question mark (?) indicate we are not sure if the cue would work (and have no data on it). Cells with an × indicate combinations we identify as infeasible.

the Ringwraiths that appear when he has detected the player's agent [LoTR:RotK].

When communicating an agent's FoV via a first-person view, additional work must be done to identify that agent [SM64,MGS2]. This is because from a first-person perspective, unless there is a mirror, the identity of the perceiver is not easy to see. One solution is for the game's camera to "fly" into the perceiver's head as a transition to the first-person perspective. *Super Mario 64* [SM64] does something similar by first showing a non-player agent (a Lakitu) that is holding a video camera such that it is clear what the camera would be seeing, then emitting a camera shutter sound as the player's perspective switches to the camera's. The player is reminded that they are now seeing the feed from that video camera by an icon on the HUD depicting the non-player agent (Lakitu) that is holding it.

A sight cue can contain partial information about the identity or location of the perceiver. Implying only the direction to the perceiver is common in first-person shooter games: e.g., when the player's agent is attacked, a flashing visual effect such as a red wedge or blood splatter affects the side of the viewport from which the attack came [JB:NF], enabling the player to quickly orient on the direction of the attack.

Some sight cues indicating detection leave the perceiver unidentified, including some that use text [FO3], symbols [ES4], or colors [W2:TNC] in the HUD. Sight cues from sounds effects and music rarely identify the perceiver [CTow, LZ:BW]; we did not find, for

Figure 6: [LoTR:RotK] Visual Effects: *Icon :: Symbol.* In *Lord of the Rings: Return of the King*, the player follows a loose depiction of the events from the movie of the same name. While playing, the player's agent will sometimes encounter enemies that can alert Sauron, the big evil of the Lord of the Rings intellectual property. This is first depicted as a small eye in the upper right of the screen [LEFT]. As the player continues, if they do not kill the agents that are agents of Sauron (in this case, the crows), the eye grows [CENTER]. Once it is at its largest, the eye will continue to grow until it turns Grey [RIGHT]. Once the eye turns grey, a Ring Wraith (a powerful enemy the player most likely cannot kill) will be summoned. (Screenshot taken 🅭🅯 the authors.)

example, any cases wherein the perceiver has a signature sound or musical theme that identified them in a sight cue.

7.2 Stimuli Whose Primary Purpose is Not to Be a Sight Cue

Our data revealed an interesting element of stimulus design – some stimuli designed for purposes other than communicating sight cues can function as a sight cue. It is important for human-agent interaction designers to recognize how their designs might communicate sighting information in addition to its primary purpose. Such designs might be leveraged on purpose, but they could also be accidental side effects of other design choices. For example, an agent attacking another carries information about how these agents are interacting with each other, but the observer can also likely infer that at least one of those agents has sighted the other.

We found examples of every one of our stimulus types except *Viewport/UI Effects*, *Music Changes*, and *Sound Effects* functioning as a sight cue alongside some other primary purpose. These included perceivers approaching [EQ], gazing towards [LZ:OT, Pik], shooting [JB:NF], gesturing towards [Des], and engaging in dialogue with [Por] the visual target. A perceiver's emotional reaction or other responses to a sighting also indicate that the sighting has occurred [CTaxi, CTow, AC:NH]. Showing the player an agent's visual perspective has multiple purposes as well: the player can use what they see in addition to learning about the agent's *LoS*, *FoV*, and *attention* [MGS2, LZ:WW, SM64]. Lights that were primarily used by perceivers to see more clearly [MN, LZ:WW] also communicated their *FoV*, *LoS*, and *attention. Changes to Player Controls or Affordances* can also indicate a sighting in addition to their primary purpose. In our data corpus, this was often a series of cues signaling and enacting a switch to battle mode when the player's agent is sighted by an enemy [DQ].

Layering a sight cue on top of an existing stimulus was the main way to produce *diegetic* sight cues in our corpus. There were very few *diegetic* stimuli whose primary purpose was to be a sight cue; they served another purpose but provided sight-cue information.

Recognizing how sight cues interact with other components of feedback is important because designers should understand, holistically, how their designs are communicating attention information to humans. A robots' physical movements or posture, a virtual agent's response to something unexpected, etc. all provide insights into what the agent is attending to and how it is processing its environment. Ideally, agents should be providing the right insights into that which they are attending.

7.3 Non-Binary Sighting Information

Not all sight cues that indicate *detection* do so as if it were binary – i.e., as if an agent is either detected, or not detected. In our corpus, non-binary detection was usually implemented as an "alert" state wherein perceivers have vaguely seen or heard something suspicious, but have not fully detected anything yet. In *Mark of the Ninja* (for a visual example, see Figure 8), guards alerted by movement say, "what was that?", turn and walk towards it, and have a question mark icon above their heads [MN]. In *Fallout 3*, the word "[HIDDEN]" in the HUD turns to "[CAUTION]" when an enemy becomes alerted while the player's agent is in sneak mode [FO3]. Players are often given a time limit to stop being so visible to the perceiver or else they will be detected. The timer can be visualized on the question mark icon [MN, LZ:BW] or on a view cone [Des3] (seen in Figure 9).

Simulating in-between states gives the visual target a chance to avoid being seen (or unseen) if desired, or explore the boundaries of the perceiver's *FoV* without being seen (or unseen). This offers a lower-stakes way to build a mental model of the perceiver's visual faculties.

Figure 7: [MGS2] Perceiver Behaviors: *Locomotion: Approach, Retreat, Leaving Spot or Patrol*; *Other Perceiver Behavior Change*; Visual Effects: *Captions, Other Text*; *Icon :: Symbol*; *Gizmos*; Characteristics of Visual Effects: *Colors*; Agent Sounds: *Voice Interaction :: Verbal.* In *Metal Gear Solid 2: Sons of Liberty*, the player takes on the role of a spy infiltrating a terrorist organizations in an effort to uncover a plot to attack New York City. The game relies on cues in the form of guards on patrol who will seek out disturbances and potentially call for help. On the [LEFT], we see a guard in the bottom left hear the player in the center of the picture knocking on a piece of metal. The exclamation point and line of sight in the upper right turns yellow to indicate caution. Those visual cues are followed by additional sound cues: the soldiers say something to the effect of, "what was that?!" In the [CENTER], guards will run off once they discover the player to radio for help. This will result in the radar (upper right) being shut off with the player being forced to fight for their lives and eventually hide. On the [RIGHT], the player has shot a guard with a tranquilizer dart. While this results in the guard's attention (indicated by the exclamation point and the yellow line of sight on the radar), the guard will fall unconscious which is visualized by a number of Z's floating around the guard's head. The number of Z's decreases over time or if the player begins to move the body. Another guard could also wake the unconscious guard up resulting in both guards calling for help. (Screenshot taken 🅭🅯 the authors.)

Figure 8: [MN] Perceiver Behaviors: *Locomotion :: Approach, Retreat, Leaving Spot or Patrol*; Visual Effects: *Icon :: Symbol*; *Light Beams*; Characteristics of Visual Effects: *Colors.* In *Mark of the Ninja*, the player assumes the role of a Ninja. Immediately after starting the game, the player is introduced to agents who are out to kill Ninjas where players can see where they are looking [LEFT]. As the agents wander the map, if you as a Ninja make a sound (symbolized by the large circle in the lower right of [CENTER], the agents come to investigate that sound. Should the player be seen by the agent, the ? becomes a red ! and the agent begins to attack [RIGHT]. While these cues are seen in a number of games of different types, their use within each game in unique and contextualized to those specific games. (Screenshot taken 🅭🅯 the authors.)

8 DISCUSSION

We discuss how interaction designers can use insights from the present research to build sight cues, followed by expected use cases. We further address the limitations of the study and pointers to future work derived from challenges discovered through the analysis.

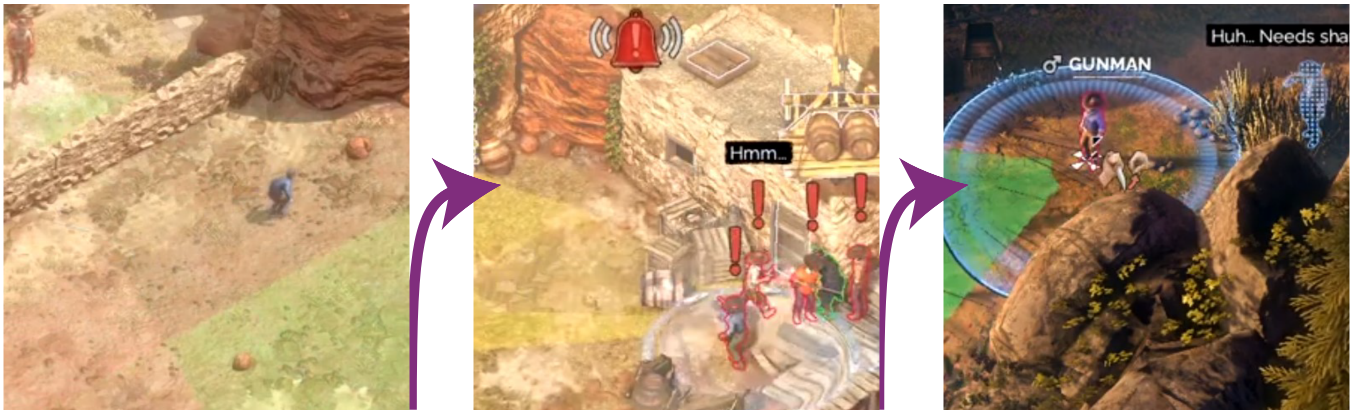

Figure 9: [Des3] Visual Effects: *Icon :: Symbol*; *Gizmos*; Characteristics of Visual Effects: *Colors*. In *Desperados III*, the player assumes the role of a cowboy who has to sneak and kill their way through a bandit group responsible for killing his father. Since *Desperados III* is one of the newest games in our corpus, the view cones and other cues are more integrated with the game experience. On the [LEFT], you can see a small spot of green on the ground to the upper left. That view cone extends after a small space that is created by the fence just in front of the agent. The cone continues further down the path (lower right) giving the player a small area to sneak through. Since this game is a tactical experience more than a realtime one, when the player is discovered by the bandits (noted in the [CENTER]), the player will most likely perish. Going back to the fact that this is one of the newest games on this corpus, many of the cues are more complex since players of video games will have been exposed to them. Seen on the [RIGHT], the player is hidden in a bush (noted on the upper right and by the blue dashes). In the center of the image, the agent has a blue circle around them noting that if the player makes a sound anywhere within that circle, the player will be detected. The goal and use of these cues is to allow the player to see and understand how to sneak up on the agent without them noticing in an effort to incapacitate them. (Screenshot taken 🅭🅯 the authors.)

8.1 Implementing Sight Cues in Human-Agent Interaction

8.1.1 How To Use the Framework. The purpose of the designed framework is to enable designers to answer questions about how best to represent agents' attention to a human. Through drawing on examples from game design, we have identified previously successful ways of doing this in games.

The framework, while necessarily ordered on the page, is not intended to be used in a particular order. That is, a designer should "enter" the framework based on their design constraints, preferences, or vision, and explore the table from there, using it to answer questions about the potential design. Example questions that can be answered with framework include:

- What information could I convey with a *Sound Effect*?
- What visuals or sounds can be used to indicate an agent's *FoV* or *LoS*?
- Given that our device is capable of a particular set of effectors, what information can we communicate?

We expect the framework to be a resource and a call to action. That is, it is increasingly important for agents to communicate where their attention is placed and whether they aware aware of human(s) in the environment – so we hope that designers keep these concerns in mind as they build new systems, and use the framework to design those things well. We also hope that designers keep the framework in mind as they work – using it to raise and answer questions about what they *should* be doing when making new systems. All this said, the framework is not a panacea and forms the basis of future work that might draw on inspiration from many places.

8.1.2 Technology Considerations. Adapting sight cues in games for virtual agents or robots is nontrivial, but all identified methods could be employed in the real world in some form. We briefly discuss how technologies might be used to reify the VISUALIZATIONS and SONIFICATIONS shown in Table 3; we first visit virtual agents, then consider robots.

Considerations for Virtual Agents. Virtual agents already exist through various display technologies, such as monitors, speakers, smartphones, and projected images. Adding sight cues to them is relatively straightforward – the agent's avatar can respond with surprise or *gestures*. With sufficient sensing capabilities, an agent's eyes or head could orient to track a visual target. *Non-diegetic* visualizations can be added to the display to indicate status (e.g., icons appearing alongside the agent on a display).

The range of SONIFICATIONS can also be directly applied, as most devices with virtual agents will already employ sound to communicate. Voice assistants already do this to indicate when they are available to respond to a request, as recommended in prior work [9].

In many ways, sight cues from games serve as ready-made [68] examples that could be directly imported to agent designs. The primary challenges in employing them are in how to sense relevant context from humans and/or the environment, and in selecting cue types that will be effective in that context.

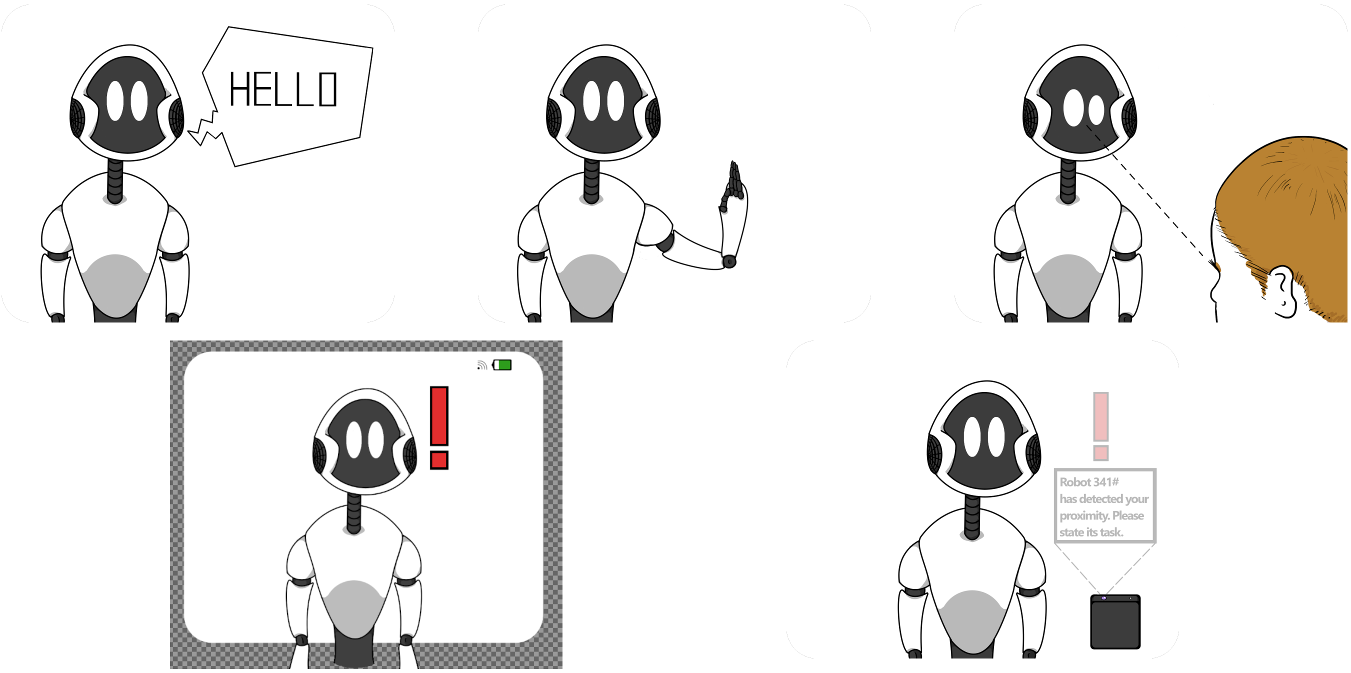

Figure 10: Mockups of interactions with a robot employing sight cues to indicate that a human has been acknowledged in the environment. On the top, the robot uses *Voice Interaction :: Verbal* [TOP LEFT], a combination of *Gesture* and *Changing Pose* [TOP CENTER], or *Turning Towards :: Eyes Only* [TOP RIGHT]. On the bottom, augmentations are used so that the view through a human’s smart phone or glasses shows a Visual Effect – an *Icon :: Symbol* – when the robot is attentive [BOTTOM LEFT]; on the [BOTTOM RIGHT] a projector in the environment provides the Visual Effects: an *Icon :: Symbol* as well as *Captions, Other Text* shown on the wall of the environment. (Original artwork 🅭🅯 by author Cormier.)

Considerations for Robots. Sight cues for robots are more challenging because of their physical embodiments; Figure 10 illustrates potential examples. A robot body, itself, can be used, such as for *locomotion* or *turning towards* the visual target (Figure 10, TOP RIGHT). A robot could also use *gestures* (Figure 10, TOP CENTER), such as pointing in its gaze direction or tracing its *FoV* with an appendage. For overlaying visualizations on the physical environment, mixed reality [5, 11, 73] is needed, such as through a headset like the Microsoft Hololens[5] (Figure 10, BOTTOM LEFT; Figure 11, LEFT). Projectors can also display imagery on the physical environment (Figure 10, BOTTOM RIGHT); further examples include work on using drones to project a user interface on the ground [16] or projecting *onto* a drone [29].

Projectors could visualize the robot’s *FoV*, perhaps via a multi-featured view cone like the one in *Desperados III* [Des3], as in Figure 11, RIGHT. Detected objects and people could be projected upon or pointed to by projections on other surfaces. HUD elements, expressive avatars, and other visualizations bearing sight cues could be projected on the floor or other nearby surfaces.

In trusted environments, personal mobile devices (e.g., smartphones) can serve as proxy user interfaces, providing sight cues from local robots (Figure 10, BOTTOM LEFT; Figure 11, LEFT). Such environments must be trusted because providing such information through personal devices would involve data exchange between the personal device and whatever systems are providing computation for the robot and/or its local environment. Large displays could serve similar purposes.

Displays, projectors, and other indicators (e.g., lights) can be mounted on robots, and have the advantage over headsets and handheld devices of being visible to everyone in the area (Figure 10, BOTTOM RIGHT; Figure 11, RIGHT). Projectors could also provide functional lighting for robots and other agents in ill-lit locations, which doubles as a cue about *FoV* and *attention*. Other flashlights and searchlights mounted to a robot or other apparatus in the environment could also serve as cues, but moreso if the agent can move it, and projectors have the added advantage of controlling the shape of the beam – e.g., a robot could project multiple spotlights on multiple detected people.

The SONIFICATIONS in Table 3 could be played to human users via headphones, head-mounted displays, or handheld devices, or to anyone within hearing distance via speakers mounted on a robot or elsewhere in the environment.

Public and Private Considerations. It is worth noting that virtual agents are more likely to be personal, while robots, generally, more public (at least, in the current technological landscape). Because of this, it is worthwhile to consider how humans might personalize their interactions with devices, especially personal virtual assistants. Designers should offer a range of ways to personalize the

[5] https://www.microsoft.com/hololens/

Figure 11: Mockups of interactions with a robot employing sight cues to indicate the robot's FoV with different implementations of Visual Effects: *Light Beams.* Using augmented reality through a smartphone, a translucent solid cylinder is rendered over the robot, showing the robot's FoV around it on demand [LEFT]. Using a projector mounted on the robot, the robot projects an arc shape indicating where its FoV is for humans to either come into to interact with it, or avoid [RIGHT]. (Original artwork 🅭🅯 by author Cormier.)

visualization and sonification of sight cues and enable humans to select a subset.

In public places especially, designers should consider multiple modalities as the default. These environments likely prevent personalization and need to be widely accessible. Using a single channel (e.g., only visualizations or only sonifications) is discouraged, as it is less accessible to a range of people. On the other hand, agents in public contexts should avoid polluting shared space with visuals and sounds that could be annoying, distracting, or overwhelming to people who are not using them.

Communicating Capabilities. Agents' sensors and detectors might only be able to "pick up" certain things about humans in an environment – for example, it might be possible to detect the presence of a person, but not their face or other features. Designers likely need to consider how much sight cues might need to be faked when interacting with humans in the environment and/or how much of their sensing capabilities and ranges to reveal. For example, is it better to offer a conservative sense of the FoV of a robot, or a correct one? The answer is likely driven by context. If a search and rescue robot is attempting to locate victims in a disaster, the design might communicate the narrowest space in which the sensors can 100% spot a person to ensure clear communication with victims, but it might still act on any and all data about detected people, even on the edge of the detection area, to maximally prevent loss of life. At the same time, a floor cleaning robot in a grocery store might show a much larger area than it can reasonably detect to ensure that shoppers keep away from it.

Seamful design offers useful insights when considering the degree to which designers might communicate about sensing capabilities. This work comes from mixed realities – such as those we are discussing with robots – and considers how the technological breaks in our mixed reality environments might be revealed, leveraged, ignored, and/or hidden to/from humans [10, 11, 24]. This points to places in which ambiguities in sensing might, or might not, be valuable.

8.1.3 The Role of Mental Models. Interpreting some sight cues requires the player to have an accurate mental model of how the game, or an agent in the game, behaves under different conditions. For example, guards in *Metal Gear Solid 2* [MGS2] only break from their rigidly defined patrol paths if they spot an intruder. Similarly, if an enemy in *Everquest* [EQ] spots the player's agent through a window or upper ledge, it might abruptly begin running in a seemingly random direction, but experienced players understand that it is heading for the nearest door or staircase that leads to the player's agent. An example of this can be seen in Figure 12 with the player running their aggressor toward a different agent meant to help players in areas for new players. This requirement to have a certain mental model to understand a sight cue makes the cue less accessible to novice users. On the other hand, some sight cues might be instances of patterns in video game design that are familiar to anyone with sufficient experience playing videogames (e.g., the use of question marks and exclamation points to indicate possible detection and detection, respectively). Interaction designers could avoid the paradox of the active user [18] by adapting such cues to leverage this existing user knowledge.

Observers of sight cues based on agent behaviors need to know how an agent behaves when it detects someone, or even how it reacts differently to different types of people (e.g., children versus adults) or to different individuals. For example, a real life agent could have a different reaction – including emotional expressions, sounds, and music – depending on its relationship with the human who just entered a space.

8.2 Use Cases

The present research was motivated by a number of use cases for humans interacting with artificial agents. We point the reader to situations in which sight cues are especially valuable and we expect to see agents proliferate.

Search and rescue is a domain wherein robots might help find people and remove them from dangerous situations [2, 63]. Such robots need to be approachable and be designed to avoid scaring humans or breaching ethical norms. Sighting is key in search and rescue scenarios. For one thing, a missing person will want to know if a rescue robot is aware of them or not. In the case that the robot is not, sight cues should help the person know how to grab the

Figure 12: [EQ] Perceiver Behaviors: *Locomotion :: Approach, Retreat, Leaving Spot or Patrol*; *Other Action.* In *Everquest*, players assume the role of a character that is completely customized to represent what the player wants. Upon starting the game, players are introduced to a brutal and harsh world that is complicated, vast in that it takes hours to walk across, and filled with content to explore. As the player explores, they embark on a journey – sometimes called the "immigrant's dream" – wherein they arrive with nothing and seek to amass everything [98]. The cues in this game predate many of the cues seen in later games but much of the same type of content is present. For example, on the [LEFT], the player is able to wander freely and come upon all manner of agents. Some agents, for example in the [CENTER], this Gloomingdeep Warrior, will detect and begin attacking the player if their character moves across its field of view (which it has begun to do here). In this case, the Gloomingdeep Warrior is far stronger than the player, and the player will die if they try to fight it. In this case, since the character "Mattreuben" is in a starting area, there are guards who will attack enemies that enter their field of view. As a result, Guard Hobart is a hero to Mattreuben and dispatches the Gloomingdeep Warrior [RIGHT]. (Screenshot taken 🅭🅯 the authors.)

robot's attention, such as knowing where to move to so that the robot can perceive them.

Other places wherein humans might come into contact with robots in somewhat unstructured environments include public places (e.g., shopping malls, airports, sidewalks), farms, warehouses, and the home [45]. It can be important for co-located humans to know when a robot *is aware* of them to prevent privacy violations, and to know when the robot *is not aware* of them to avoid injury.

As these use cases proliferate, we expect a growing need for humans to understand the visual perception capabilities of robots. While there are many potential hostile use cases where a robot designer might *not* desire such interactions, we focus on places where humans and robots act in concert. We hope that as designers build new agents they consider how to share with humans what their agents are aware of in ways that are understandable and helpful.

8.3 Limitations

While we expect our findings to be of great use for interaction designers, they are not a panacea, and should be taken as one part of a larger design context. Future work could expand the framework using insights from other domains to address the limitations here, especially through design workshops and user studies.

8.3.1 Characteristics of the Corpus and Cooperation. We hope for human-agent interactions to be collaborative rather than oppositional. Because many games with sight cues fall under the *Action* genre with *Shooter* and *Stealth* gameplay types, the perceiver is often an enemy. Thus, a high proportion of our cues have an urgent and/or negative connotation – an enemy has spotted the player's agent, so the player has either lost the game or needs to act quickly to win a battle or escape. For a non-combat example, see Figure 13 which has the player sneaking up on an agent that is meant to be the main character's wife for a kiss. There are surely more cooperative sight cues to consider, but we do not expect to find them in the current landscape of games. As an aside, this might also point to a deficit in game design to which the industry might attend.

8.3.2 Focus on Vision. We only looked at cues about visual perception; we did not consider cues about agents perceiving each other via hearing, feeling, smelling, or tasting. This aligns well with agents, which often use a form of visual perception to interact with the world, but fails to capture the full range of human experience. We saw hints that these different senses might lend themselves to different types of cues from the ones we included in our corpus – e.g., ears perking up for hearing, or body movements in response to feeling.

We collected data through online videos of gameplay via Youtube and other video services like Twitch in addition to playing games ourselves. Data collected through video did not reproduce any stereo or 3D sound from the original games, so we were not able to observe any directionality to the sound. Directional sound could help make *Sound Effects* and *Action Sounds* less ambiguous by specifying who is making them. We were also limited by the fact that we did not play all of the games represented in the corpus. For some games, those of us who had played it could interpret the cues in

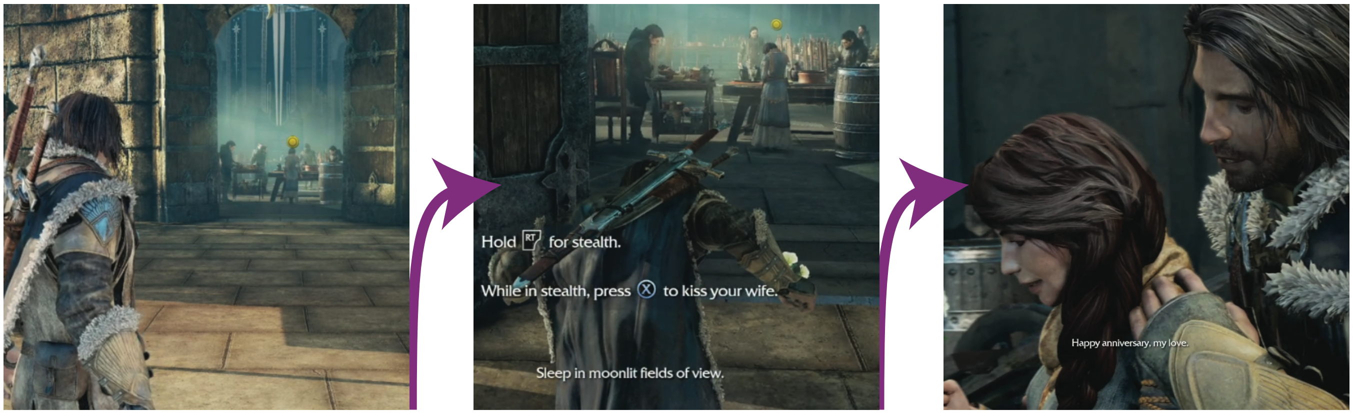

Figure 13: [MESOM] Perceiver Behaviors: *Changing Pose*; *Other Action*; Visual Effects: *Captions, Other Text*; *Gizmos*. In *Middle-Earth: Shadow of Mordor*, the player controls Talion, a ranger of Gondor, as they seek revenge against Sauron for the death of their spouse and child. This game relies on stealth and surprise. In order to teach these behaviors, the player's first task of stealth is to sneak up to their spouse and kiss them. The player begins sneaking up on the agent after being placed on a walkway [LEFT]. Shortly after beginning, text prompts the character to hold a button to sneak [CENTER]. Once sneaking, the agent is highlighted just as they are able to execute their attack. In this case, they kiss their spouse [RIGHT]. (Screenshot taken ©ⓘ the authors.)

the videos we found, but in certain games that we had not played as much our understanding of the cues was uncertain.

8.3.3 Game Design. Finally, games, even those that are highly rated, might be poorly designed, which should be kept in mind when using these sight cues as inspiration for human-agent interaction design. For example, in *Everquest*, there is no way to know where the edges of an enemy's FoV are except by being abruptly detected, attacked, and chased around the entire zone [EQ] (seen in Figure 12).

8.4 Combining Multiple Cues: Future Work

Identifying the atomic level of a cue is nontrivial and, for this work, was based heavily on an understanding of the data corpus and the wider context of games. While we considered sight cues as individual data items, in many cases a bundle of multiple cues was emitted in the period of several seconds or so following a sighting. An example sighting we studied in *Clock Tower* [CTow] includes 11 sight cues in ten seconds using a variety of stimulus types (the below is seen visually in Figure 14):

> The player's agent enters a room. Upon entering a room, a creature crashes into the center of it with a large pair of scissors impaled through a character's torso. Upon landing, they turn toward the player's agent and begins to move toward them. The player's facial expression contorts in fear followed by a scream for the character who is killed. The player runs their agent away. In the next room, the scissor-holding creature can be heard running down the hall toward the player's agent. Again, the player's agent's picture contorts in fear. The player moves their agent into a room, a bathroom, and locks the door. The music swells and the scissor-holding creature walks into the room and begins to stab the door with its scissors. The player's agent is killed by being stabbed by scissors resulting in that playthrough being over. [Sight cue description for *Clock Tower* [CTow]]

Future work should study how cues of different stimulus types, including different channels such as video and audio, can be combined to improve accessibility, salience, readability, the emotional tone or connotations of the sighting, and the amount of information communicated. Additionally, *Clock Tower* [CTow] provides a number of cues as clues for solving future puzzles and is unique in how it uses cues (seen in Figure 15). Considering how influential *Clock Tower* [CTow] is to videogame design, it seems as though videogames could once again offer insights from designers thoughtfully combining sight cues.

9 CONCLUSION

Since artificial agents, including robots, are becoming more pervasive in a variety of application areas, we have presented a framework to supply human-agent interaction designers with a resource for design: a list of known visual and audio stimulus types for sight cues and whether they can indicate what all an agent can see in addition to whether the agent has currently detected something. We also presented the iterative process by which this framework was formed: a qualitative, thematic analysis of a corpus of sight cues – by which a player understands what an agent can see – in videogames.

Videogames proved to be a rich resource for studying sight cues, not only by yielding the diversity of stimulus types in our framework but also by showing us several additional, complicating nuances of sight cues. For example, there are sight cues that indicate that a sighting occurred without specifying who sighted what. Other sight cues can indicate states in between visual detection and lack of visual detection, such as when an agent is uncertain

Figure 14: [CTow] Combining multiple sight cues: an example. Perceiver Behaviors: *Locomotion :: Approach, Retreat, Leaving Spot or Patrol*; Characteristics of Visual Effects: *Colors*; Agent Sounds: *Action Sounds*; Atmospheric Sounds: *Music Changes*. In *Clock Tower*, the player assumes control of an agent named Jennifer Simpson. She is an orphan who has been invited to a large house to live. Upon entering the house, the social worker goes missing and Jennifer has to investigate. Upon entering the main room, a creature with a large pair of scissors (aptly named Scissorman) lets the player know that this house is dangerous and the player must survive to discover its secrets [LEFT]. There are additional cues to consider here. The agent's portrait, seen [LEFT] has different colored backgrounds. Depending on the background, when Scissorman appears and Jennifer begins to run, movement can be difficult if Jennifer is low in health. In this case, with a blue background (meaning full health) Jennifer enters a hallway and the music indicates to the player that they are being chased by the creature; they can also hear the creature's scissors snapping [CENTER]. Upon heading into a room, they lock the door and hide in the bathroom. The music swells, the creature enters the room, breaks down the door, and kills Jennifer, thus ending the game [RIGHT]. Cues like this for survival horror games not only indicate to the player which actions they need to take, but additionally provide motivation to move forward. (Screenshot taken 🅭🅯 the authors.)

Figure 15: [CTow] Perceiver Behaviors: *Gestures*; *Locomotion: Approach, Retreat, Leaving Spot or Patrol*. *Clock Tower* is a game that inspired a number of cues in other games. Some cues from agents are indicative of other actions a player or user can do. For example, on the [LEFT] the player is provided a clue in the form of a parrot flying around the room. This creature will later provide an escape from the Scissorman creature that constantly chases Jennifer. While these clues are useful, in the [CENTER] the parrot that was released in the room constantly screams, "I'll kill you!" and begins to hurt Jennifer, so not every cue provides a beneficial result for the player or user. Jennifer escapes the room with moderate damage (note that on the [RIGHT] her picture's background is red) and walking down the hall ominous music begins to play. She walks into a steamy bathroom and sees that a shower is still on and that the player can examine it. After examining it, she will find the body of one of her friends as Scissorman emerges from the bathtub and begins to chase her. In this case, because of the cue to run from Scissorman Jennifer can run back to the parrot and hopefully escape Scissorman in order to solve the mystery of the mansion. (Screenshot taken 🅭🅯 the authors.)

whether it saw something interesting or not. Sight cues can even be present in agent behaviors whose primary purpose is something besides providing information about a sighting – i.e., one behavior or feature can serve multiple purposes. Lastly, there is a whole new dimension of possibilities when multiple cues are used for a single sighting: cues can be combined to improve clarity or communicate additional nuances such as urgency, or cues with multiple stimulus types can provide redundant information to improve accessibility.

Future work should continue to study all of these factors of sight cues in videogames, but our vision is chiefly that the findings would be applied in other areas. Designing, implementing, and deploying sight cues in key contexts that involve mobile robotics in public spaces, for example, will likely reveal a host of technical challenges related to the hardware being used to render the cues and sensing difficulties in the environment. Furthermore, evaluation of interactions with users will help us understand not just what types of cues are possible, but what makes them effective in these different contexts. Such work would contribute to a future in which humans interact with artificial agents effectively, without any misconceptions about what they can sense.

ACKNOWLEDGMENTS

We also acknowledge Andrew Perez for making the Ludography and helping to analyze cue combinations, as well as Jacob Rueben and Armando Nelson for early vision and cue ideas.

This material is based upon work supported by the National Science Foundation under Grant Nos. IIS-2106380, IIS-2106402, IIS-1651532, 1950121, and 1849591; the US Department of Education under Grant No. P200A180005; the New Mexico Space Grant Consortium (NMSGC) under NASA grant No. 80NSSC20M0034; and the New Mexico State University Office of the Vice President of Research and Graduate School (NMSU VPRGS).

REFERENCES

[1] Henny Admoni, Bradley Hayes, David Feil-Seifer, Daniel Ullman, and Brian Scassellati. 2013. Are you looking at me? Perception of robot attention is mediated by gaze type and group size. In *8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 389–395.

[2] Sultan A. Alharthi, Olaa Alsaedi, Phoebe O. Toups Dugas, Theresa Jean Tanenbaum, and Jessica Hammer. 2018. Playing to Wait: A Taxonomy of Idle Games. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 621:1–621:15.

[3] Sultan A. Alharthi, Phoebe O. Toups Dugas, Olaa Alsaedi, Theresa Jean Tanenbaum, and Jessica Hammer. 2018. The Pleasure of Playing Less: A Study of Incremental Games through the Lens of Kittens. *Well Played Singles* (2018). DOI: http://dx.doi.org/https://doi.org/10.1184/R1/6686957.v1

[4] Mahoro Anabuki, Hiroyuki Kakuta, Hiroyuki Yamamoto, and Hideyuki Tamura. 2000. Welbo: an embodied conversational agent living in mixed reality space. In *Extended Abstracts on Human Factors in Computing Systems (CHI '00)*. ACM Press, The Hague, The Netherlands, 10. DOI: http://dx.doi.org/10.1145/633292.633299

[5] Ronald T. Azuma. 1997. A Survey of Augmented Reality. In *Presence: Teleoperators and Virtual Environments*, Vol. 6. 355–385.

[6] Gagan Bansal, Besmira Nushi, Ece Kamar, Walter S Lasecki, Daniel S Weld, and Eric Horvitz. 2019. Beyond accuracy: The role of mental models in human-AI team performance. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 7. 2–11. Issue: 1.

[7] Christoph Bartneck, Tony Belpaeme, Friederike Eyssel, Takayuki Kanda, Merel Keijsers, and Selma Šabanović. 2020. *Human-robot interaction: An introduction.* Cambridge University Press.

[8] Peter Bayliss. 2007. Beings in the Game-World: Characters, Avatars, and Players. In *Proceedings of the 4th Australasian Conference on Interactive Entertainment (IE '07)*. RMIT University, Melbourne, AUS.

[9] Victoria Bellotti, Maribeth Back, W. Keith Edwards, Rebecca E. Grinter, Austin Henderson, and Cristina Lopes. 2002. Making Sense of Sensing Systems: Five Questions for Designers and Researchers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. Association for Computing Machinery, New York, NY, USA, 415–422. DOI: http://dx.doi.org/10.1145/503376.503450

[10] Steve Benford, Andy Crabtree, Martin Flintham, Adam Drozd, Rob Anastasi, Mark Paxton, Nick Tandavanitj, Matt Adams, and Ju Row-Farr. 2006. Can You See Me Now? *ACM Trans. Comput.-Hum. Interact.* 13, 1 (March 2006), 100–133. DOI: http://dx.doi.org/10.1145/1143518.1143522

[11] Steve Benford and Gabriella Giannachi. 2011. *Performing Mixed Reality.* MIT Press.

[12] Staffan Björk and Jussi Holopainen. 2006. Games and Design Patterns. In *The Game Design Reader: A Rules of Play Anthology*, Katie Salen and Eric Zimmerman (Eds.). 410–437.

[13] Gabriele Bolano, Christian Juelg, Arne Roennau, and Ruediger Dillmann. 2019. Transparent Robot Behavior Using Augmented Reality in Close Human-Robot Interaction. In *28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1–7. DOI: http://dx.doi.org/10.1109/RO-MAN46459.2019.8956296 ISSN: 1944-9437.

[14] Nathan Bos, Ann Zimmerman, Judith Olson, Jude Yew, Jason Yerkie, Erik Dahl, and Gary Olson. 2007. From Shared Databases to Communities of Practice: A Taxonomy of Collaboratories. *Journal of Computer-Mediated Communication* 12, 2 (2007), 652–672.

[15] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. DOI: http://dx.doi.org/10.1191/1478088706qp063oa

[16] Anke M. Brock, Julia Chatain, Michelle Park, Tommy Fang, Martin Hachet, James A. Landay, and Jessica R. Cauchard. 2018. FlyMap: Interacting with Maps Projected from a Drone. In *Proceedings of the 7th ACM International Symposium on Pervasive Displays (PerDis '18)*. Association for Computing Machinery, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3205873.3205877

[17] Janis A Cannon-Bowers, Eduardo Salas, and Sharolyn Converse. 1993. Shared Mental Models in Expert Team Decision Making. In *Individual and Group Decision Making: Current Issues*, N John Castellan Jr (Ed.). Earlbaum Associates, Hillsdale, NJ, USA, 221–246.

[18] John M Carroll and Mary Beth Rosson. 1987. Paradox of the active user. In *Interfacing thought: Cognitive aspects of human-computer interaction.* 80–111.

[19] Marcus Carter, Martin Gibbs, and Michael Arnold. 2012. Avatars, Characters, Players and Users: Multiple Identities at/in Play. In *Proceedings of the 24th Australian Computer-Human Interaction Conference (OzCHI '12)*. Association for Computing Machinery, New York, NY, USA, 68–71. DOI: http://dx.doi.org/10.1145/2414536.2414547

[20] Justine Cassell. 2000. Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. In *Embodied conversational agents*, Justine Cassell, Joseph Sullivan, Elizabeth Churchill, and Scott Prevost (Eds.). Vol. 1. MIT Press.

[21] Justine Cassell and Kristinn R. Thorisson. 1999. The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence* 13, 4-5 (May 1999), 519–538. DOI: http://dx.doi.org/10.1080/088395199117360

[22] Elizabeth Cha, Anca D. Dragan, and Siddhartha S. Srinivasa. 2015. Perceived robot capability. In *24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 541–548.

[23] Elizabeth Cha, Yunkyung Kim, Terrence Fong, and Maja J. Mataric. 2018. A survey of nonverbal signaling methods for non-humanoid robots. *Foundations and Trends® in Robotics* 6, 4 (2018), 211–323. Publisher: Now Publishers, Inc.

[24] Matthew Chalmers and Areti Galani. 2004. Seamful interweaving: Heterogeneity in the theory and design of interactive systems. In *DIS '04: Proceedings of the 2004 Conference on Designing Interactive Systems.* ACM Press, 243–252. DOI: http://dx.doi.org/10.1145/1013115.1013149

[25] Jessie Y.C. Chen and Michael J. Barnes. 2015. Agent Transparency for Human-Agent Teaming Effectiveness. In *IEEE International Conference on Systems, Man, and Cybernetics.* 1381–1385. DOI: http://dx.doi.org/10.1109/SMC.2015.245

[26] Hao-Fei Cheng, Ruotong Wang, Zheng Zhang, Fiona O'Connell, Terrance Gray, F. Maxwell Harper, and Haiyi Zhu. 2019. Explaining Decision-Making Algorithms through UI: Strategies to Help Non-Expert Stakeholders. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* ACM, Glasgow, Scotland, UK, 1–12. DOI: http://dx.doi.org/10.1145/3290605.3300789

[27] Victor Cheung, Y.-L. Betty Chang, and Stacey D. Scott. 2012. Communication channels and awareness cues in collocated collaborative time-critical gaming. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work.* 569–578.

[28] Martin Constable, Justin Dauwels, Shoko Dauwels, Rasheed Umer, Mengyu Zhou, and Yasir Tahir. 2016. Modelling Conversation. In *Context Aware Human-Robot and Human-Agent Interaction*, Nadia Magnenat-Thalmann, Junsong Yuan, Daniel

Thalmann, and Bum-Jae You (Eds.). Springer International Publishing, Cham, 81–111. DOI: http://dx.doi.org/10.1007/978-3-319-19947-4_4

[29] Rajkumar Darbar, Joan Sol Roo, Thibault Lainé, and Martin Hachet. 2019. DroneSAR: Extending Physical Spaces in Spatial Augmented Reality Using Projection on a Drone. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia (MUM '19)*. Association for Computing Machinery, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3365610.3365631

[30] Drew Davidson. 2008. Well Played: Interpreting *Prince of Persia: The Sands of Time*. *Games and Culture* 3, 3-4 (July 2008), 356–386.

[31] Leslie A DeChurch and Jessica R. Mesmer-Magnus. 2010. Measuring shared team mental models: A meta-analysis. *Group Dynamics: Theory, Research, and Practice* 14, 1 (2010), 1–14.

[32] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. 2013. Legibility and predictability of robot motion. In *8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 301–308.

[33] Katharina Emmerich and Maic Masuch. 2016. The Influence of Virtual Agents on Player Experience and Performance. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '16)*. Association for Computing Machinery, New York, NY, USA, 10–21. DOI: http://dx.doi.org/10.1145/2967934.2968092

[34] Katharina Emmerich, Patrizia Ring, and Maic Masuch. 2018. I'm Glad You Are on My Side: How to Design Compelling Game Companions. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '18)*. Association for Computing Machinery, New York, NY, USA, 141–152. DOI: http://dx.doi.org/10.1145/3242671.3242709

[35] Mica R. Endsley. 1995. Toward a theory of situation awareness in dynamic systems. *Human Factors* 37, 1 (1995), 32–64.

[36] Mica R. Endsley. 2000. Theoretical Underpinnings of Situation Awareness: A Critical Review. In *Situation Awareness Analysis and Measurement*, Mica R. Endsley and D J Garland (Eds.). Lawrence Erlbaum Associates, Mahwah, NJ, USA, 3–6.

[37] Elliot E. Entin and Daniel Serfaty. 1999. Adaptive Team Coordination. *Human Factors* 41, 2 (June 1999), 312–325.

[38] Michael R. Epke, Lars Kooijman, and Joost C. F. de Winter. 2021. I See Your Gesture: A VR-Based Study of Bidirectional Communication between Pedestrians and Automated Vehicles. *Journal of Advanced Transportation* 2021 (April 2021), e5573560. DOI: http://dx.doi.org/10.1155/2021/5573560 Publisher: Hindawi.

[39] J. Alberto Espinosa, F. Javier Lerch, and Robert E. Kraut. 2004. Explicit versus implicit coordination mechanisms and task dependencies: One size does not fit all. In *Team Cognition: Understanding the factors that Drive Process and Performance*. American Psychological Association, 107–130.

[40] Evelyn Florentine, Mark Adam Ang, Scott Drew Pendleton, Hans Andersen, and Marcelo H Ang Jr. 2016. Pedestrian notification methods in autonomous vehicles for multi-class mobility-on-demand service. In *Proceedings of the Fourth International Conference on Human Agent Interaction*. 387–392.

[41] David Forsyth and Jean Ponce. 2011. *Computer vision: A modern approach*. Prentice hall.

[42] D Gentner and A L Stevens. 1983. *Mental Models*. Lawrence Earlbaum Associates, Hillsdale, NJ, USA.

[43] Katy Ilonka Gero, Zahra Ashktorab, Casey Dugan, Qian Pan, James Johnson, Werner Geyer, Maria Ruiz, Sarah Miller, David R. Millen, Murray Campbell, Sadhana Kumaravel, and Wei Zhang. 2020. Mental Models of AI Agents in a Cooperative Game Setting. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–12. DOI: http://dx.doi.org/10.1145/3313831.3376316

[44] Michael J. Gielniak and Andrea L. Thomaz. 2012. Enhancing Interaction through Exaggerated Motion Synthesis. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '12)*. Association for Computing Machinery, New York, NY, USA, 375–382. DOI: http://dx.doi.org/10.1145/2157689.2157813

[45] Michael A Goodrich and Alan C Schultz. 2008. *Human-robot interaction: a survey*. Now Publishers Inc.

[46] Andrea Gurtner, Franziska Tschan, Norbert K Semmer, and Christof Nägele. 2007. Getting groups to develop good strategies: Effects of reflexivity interventions on team process, team performance, and shared mental models. *Organizational Behavior and Human Decision Process* 102 (2007), 127–142.

[47] Carl Gutwin and Saul Greenberg. 2004. The Importance of Awareness for Team Cognition in Distributed Collaboration. In *Team Cognition: Understanding the Factors that Drive Process and Performance* (1st ed.), Eduardo Salas and Stephen M. Fiore (Eds.). American Psychological Association, Washington, DC, USA, 177–201.

[48] Komei Hasegawa and Yasushi Nakauchi. Preliminary Evaluation of a Telepresence Robot Conveying Pre-motions for Avoiding Speech Collisions. (????), 4.

[49] Christian Heath and Paul Luff. 2000. Team work: Collaboration and control in London Underground line control rooms. In *Technology in Action*, Roy Pea, John Seely Brown, and Christian Heath (Eds.). Cambridge University Press, Cambridge, UK, 88–124.

[50] Eva Hornecker, Paul Marshall, Nick Sheep Dalton, and Yvonne Rogers. 2008. Collaboration and Interference: Awareness with Mice or Touch Input. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. Association for Computing Machinery, New York, NY, USA, 167–176. DOI: http://dx.doi.org/10.1145/1460563.1460589

[51] Edwin Hutchins. 1988. *The Technology of Team Navigation*. ICS Report. Technology and Cooperative Work.

[52] Ioanna Iacovides, Anna Cox, Richard Kennedy, Paul Cairns, and Charlene Jennett. 2015. Removing the HUD: The Impact of Non-Diegetic Game Elements and Expertise on Player Involvement. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '15)*. Association for Computing Machinery, New York, NY, USA, 13–22. DOI: http://dx.doi.org/10.1145/2793107.2793120

[53] Colby Johanson and Regan L. Mandryk. 2016. *Scaffolding Player Location Awareness through Audio Cues in First-Person Shooters*. Association for Computing Machinery, New York, NY, USA, 3450–3461. https://doi.org/10.1145/2858036.2858172

[54] Kristine Jørgensen. 2013. *Gameworld Interfaces*. MIT Press, Cambridge, MA, USA.

[55] M. Waleed Kadous, Raymond Ka-Man Sheh, and Claude Sammut. 2006. Effective user interface design for rescue robotics. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 250–257.

[56] Sara Kiesler. 2005. Fostering common ground in human-robot interaction. In *IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 729–734.

[57] Richard Klimoski and Susan Mohammed. 1994. Team Mental Model: Construct or Metaphor? *Journal of Management* 20, 2 (1994), 403–437.

[58] Ross A Knepper, Christoforos I Mavrogiannis, Julia Proft, and Claire Liang. 2017. Implicit communication in a joint action. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 283–292.

[59] Todd Kulesza, Simone Stumpf, Margaret Burnett, and Irwin Kwan. 2012. Tell me more?: the effects of mental model soundness on personalizing an intelligent agent. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin Texas USA, 1–10. DOI: http://dx.doi.org/10.1145/2207676.2207678

[60] Todd Kulesza, Simone Stumpf, Margaret Burnett, Sherry Yang, Irwin Kwan, and Weng-Keen Wong. 2013. Too much, too little, or just right? Ways explanations impact end users' mental models. In *2013 IEEE Symposium on Visual Languages and Human Centric Computing*. IEEE, 3–10.

[61] Johannes Kunkel, Tim Donkers, Lisa Michael, Catalin-Mihai Barbu, and Jürgen Ziegler. 2019. Let Me Explain: Impact of Personal and Impersonal Explanations on Trust in Recommender Systems. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow, Scotland, UK, 1–12. DOI: http://dx.doi.org/10.1145/3290605.3300717

[62] Nicolas LaLone. 2019. A Tale of Dungeons & Dragons and the Origins of the Game Platform. *Analog Game Studies* 3, 6 (2019).

[63] Nicolas LaLone, Sultan A. Alharthi, and Phoebe O. Toups Dugas. 2019. A Vision of Augmented Reality for Urban Search and Rescue. In *Proceedings of the Halfway to the Future Symposium 2019*. 1–4.

[64] Alex Leavitt, Brian C. Keegan, and Joshua Clark. 2016. *Ping to Win? Non-Verbal Communication and Team Performance in Competitive Online Multiplayer Games*. Association for Computing Machinery, New York, NY, USA, 4337–4350. https: //doi.org/10.1145/2858036.2858132

[65] Min Kyung Lee, Karen P. Tang, Jodi Forlizzi, and Sara Kiesler. 2011. Understanding users' perception of privacy in human-robot interaction. In *Proceedings of the 6th International Conference on Human-Robot Interaction (HRI)*. ACM, 181–182.

[66] Sau-lai Lee, Ivy Yee-man Lau, Sara Kiesler, and Chi-Yue Chiu. 2005. Human mental models of humanoid robots. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2767–2772.

[67] Michael Lewis, Huao Li, and Katia Sycara. 2021. Chapter 14 - Deep learning, transparency, and trust in human robot teamwork. In *Trust in Human-Robot Interaction*, Chang S. Nam and Joseph B. Lyons (Eds.). Academic Press, 321–352. DOI: http://dx.doi.org/10.1016/B978-0-12-819472-0.00014-9

[68] Lucy R. Lippard. 1971. Marcel Duchamp. In *Dadas on Art: Tzara, Arp, Duchamp, and Others*, Lucy R. Lippard (Ed.). Dover Publications, Inc., 139–154.

[69] Stefan Liszio, Katharina Emmerich, and Maic Masuch. 2017. The Influence of Social Entities in Virtual Reality Games on Player Experience and Immersion. In *Proceedings of the 12th International Conference on the Foundations of Digital Games (FDG '17)*. Association for Computing Machinery, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3102071.3102086

[70] Michal Luria. 2018. Designing Robot Personality Based on Fictional Sidekick Characters. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18)*. Association for Computing Machinery, New York, NY, USA, 307–308. DOI: http://dx.doi.org/10.1145/3173386.3176912

[71] Jean MacMillan, Elliot E. Entin, and Daniel Serfaty. 2004. Communication Overhead: The Hidden Cost of Team Cognition. In *Team Cognition: Understanding the Factors that Drive Process and Performance*. American Psychological Association, 61–82.

[72] John E. Mathieu, Gerald F. Goodwin, Tonia S. Heffner, Eduardo Salas, and Janis A. Cannon-Bowers. 2000. The Influence of Shared Mental Models on Team Process and Performance. *Journal of Applied Psychology* 85, 2 (2000), 273–283.

[73] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE Trans. Information Systems* E77-D, 12 (1994), 1321–1329.
[74] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence* 267 (Feb. 2019), 1–38. DOI: http://dx.doi.org/10.1016/j.artint.2018.07.007
[75] Brian Ka-Jun Mok, Stephen Yang, David Sirkin, and Wendy Ju. 2015. A place for every tool and every tool in its place: Performing collaborative tasks with interactive robotic drawers. In *24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 700–706.
[76] John Edison Muñoz Cardona and Kerstin Dautenhahn. 2021. Robo Ludens: Game Design Techniques Applied in HRI Experiments. In *Companion of the ACM/IEEE International Conference on Human-Robot Interaction*. 691–693.
[77] Peter Naur. 1985. Programming as theory building. *Microprocessing and microprogramming* 15, 5 (1985), 253–261.
[78] Don Norman. 2013. *The Design of Everyday Things: Revised and Expanded Edition* (2013 ed.). Basic Books, New York City, USA.
[79] Oxford English Dictionary. 2017. "diegesis, n.". Dictionary Definition. (June 2017). Retrieved September 15, 2017 from http://www.oed.com/view/Entry/52402.
[80] Steffi Paepcke and Leila Takayama. 2010. Judging a bot by its cover: an experiment on expectation setting for personal robots. In *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE Press, 45–52.
[81] Ramviyas Parasuraman, Sergio Caccamo, Fredrik Baaberg, Petter Ogren, and Mark Neerincx. 2017. A New UGV Teleoperation Interface for Improved Awareness of Network Connectivity and Physical Surroundings. *Journal of Human-Robot Interaction* 6, 3 (2017), 48–70.
[82] Tomislav Pejsa, Sean Andrist, Michael Gleicher, and Bilge Mutlu. 2015. Gaze and Attention Management for Embodied Conversational Agents. *ACM Transactions on Interactive Intelligent Systems* 5, 1 (March 2015), 1–34. DOI: http://dx.doi.org/10.1145/2724731
[83] Catherine Pelachaud. 2005. Multimodal expressive embodied conversational agents. In *Proceedings of the 13th annual ACM international conference on Multimedia - MULTIMEDIA '05*. ACM Press, Hilton, Singapore, 683. DOI: http://dx.doi.org/10.1145/1101149.1101301
[84] Muhammet Ramoğlu, Idil Bostan, Mohammad Obaid, Tilbe Göksun, and Oğuzhan Özcan. 2016. Interaction between Abstract Agents: Increasing the Readability of Causal Events with Animation Principles. In *Proceedings of the 20th International Academic Mindtrek Conference (AcademicMindtrek '16)*. Association for Computing Machinery, New York, NY, USA, 272–280. DOI: http://dx.doi.org/10.1145/2994310.2994345
[85] Daniel J. Rea. 2020. *Now You're Teleoperating with Power: Learning from video games to improve teleoperation interfaces*. Ph.D. Dissertation. University of Manitoba.
[86] Tiago Ribeiro and Ana Paiva. 2012. The Illusion of Robotic Life: Principles and Practices of Animation for Robots. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '12)*. Association for Computing Machinery, New York, NY, USA, 383–390. DOI: http://dx.doi.org/10.1145/2157689.2157814
[87] Justin Richer and Jill L. Drury. 2006. A video game-based framework for analyzing human-robot interaction: characterizing interface design in real-time interactive multimedia applications. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on human-robot interaction*. 266–273.
[88] Anna L. Rowe and Nancy J. Cooke. 1995. Measuring Mental Models: Choosing the Right Tools for the Job. *Human Resource Development Quarterly* 6, 3 (1995), 243–255. http://lib-ezproxy.tamu.edu:2048/login?search.ebscohost.com/login.aspx?direct=true&db=bth&AN=14858776&site=ehost-live
[89] Adam W. Ruch. 2011. Videogame interface: Artefacts and tropes. In *Videogames Studies: Concepts, Cultures, and Communication*. Brill, 1–9.
[90] Matthew Rueben, Jeffrey Klow, Madelyn Duer, Eric Zimmerman, Jennifer Piacentini, Madison Browning, Frank J Bernieri, Cindy M Grimm, and William D Smart. 2021. Mental Models of a Mobile Shoe Rack: Exploratory Findings from a Long-term In-the-Wild Study. *ACM Transactions on Human-Robot Interaction (THRI)* 10, 2 (2021), 1–36.
[91] Katie Salen and Eric Zimmerman. 2004. *Rules of Play: Game Design Fundamentals*. MIT Press, Cambridge, MA, USA.
[92] Burkhard Schafer and Lilian Edwards. 2017. "I spy, with my little sensor": fair data handling practices for robots between privacy, copyright and security. *Connection Science* 29, 3 (2017), 200–209.
[93] Mark Scheeff, John Pinto, Kris Rahardja, Scott Snibbe, and Robert Tow. 2002. Experiences with Sparky, a social robot. In *Socially Intelligent Agents*. Springer, 173–180.
[94] Matthias Scheutz, Scott A. DeLoach, and Julie A. Adams. 2017. A framework for developing and using shared mental models in human-agent teams. *Journal of Cognitive Engineering and Decision Making* 11, 3 (2017), 203–224.
[95] Stela H. Seo, James E. Young, and Pourang Irani. 2017. Where are the robots? In-feed embedded techniques for visualizing robot team member locations. In *26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 522–527.
[96] Bruno Siciliano, Oussama Khatib, and Torsten Kröger. 2008. *Springer handbook of robotics*. Vol. 200. Springer.
[97] Jan Smeddinck, Kamila Wajda, Adeel Naveed, Leen Touma, Yuting Chen, Muhammad Abu Hasan, Muhammad Waqas Latif, and Robert Porzel. 2010. QuickWoZ: a multi-purpose wizard-of-oz framework for experiments with embodied conversational agents. In *Proceedings of the 15th international conference on Intelligent user interfaces - IUI '10*. ACM Press, Hong Kong, China, 427. DOI: http://dx.doi.org/10.1145/1719970.1720055
[98] Lizzie Stark. 2012. We Hold these Rules to be Self-Evident: Larp as Metaphor for American Identity. *States of Play: Nordic Larp around the World* (2012), 184–189.
[99] Renee J. Stout, Janis A. Cannon-Bowers, Eduardo Salas, and Dana M. Milanovich. 1999. Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors* 41, 1 (March 1999), 61–71.
[100] Leila Takayama, Doug Dooley, and Wendy Ju. 2011. Expressing Thought: Improving Robot Readability with Animation Principles. In *Proceedings of the 6th International Conference on Human-Robot Interaction (HRI '11)*. Association for Computing Machinery, New York, NY, USA, 69–76. DOI: http://dx.doi.org/10.1145/1957656.1957674
[101] Theresa Jean Tanenbaum, Nazely Hartoonian, and Jeffrey Bryan. 2020. "How Do I Make This Thing Smile?": An Inventory of Expressive Nonverbal Communication in Commercial Social Virtual Reality Platforms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. DOI: http: //dx.doi.org/10.1145/3313831.3376606
[102] Frank Thomas and Ollie Johnston. 1995. *The Illusion of Life: Disney Animation*. Disney Editions.
[103] Güliz Tokadlı, Kaitlyn Ouverson, Chase Meusel, Austin Garcia, Stephen B Gilbert, and Michael C Dorneich. 2018. An Analysis of Video Games Using the Dimensions of Human-Agent Interaction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62. SAGE Publications Sage CA: Los Angeles, CA, 716–720.
[104] Phoebe O. Toups Dugas, Jessica Hammer, William A. Hamilton, Ahmad Jarrah, William Graves, and Oliver Garretson. 2014. A Framework for Cooperative Communication Game Mechanics from Grounded Theory. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-human Interaction in Play (CHI PLAY '14)*. ACM, New York, NY, USA, 257–266. DOI: http://dx.doi.org/10.1145/2658537.2658681
[105] Phoebe O. Toups Dugas and Andruid Kerne. 2007. Implicit Coordination in Firefighting Practice: Design Implications for Teaching Fire Emergency Responders. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. Association for Computing Machinery, New York, NY, USA, 707–716. DOI: http://dx.doi.org/10.1145/1240624.1240734
[106] Phoebe O. Toups Dugas, Nicolas Lalone, Sultan A. Alharthi, Hitesh Nidhi Sharma, and Andrew M. Webb. 2019. Making Maps Available for Play: Analyzing the Design of Game Cartography Interfaces. *ACM Trans. Comput.-Hum. Interact.* 26, 5 (July 2019). DOI: http://dx.doi.org/10.1145/3336144
[107] Phoebe O. Toups Dugas, Nicolas LaLone, Katta Spiel, and Bill Hamilton. 2020. Paper to Pixels: A Chronicle of Map Interfaces in Games. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference (DIS '20)*. Association for Computing Machinery, New York, NY, USA, 1433–1451. DOI: http://dx.doi.org/10.1145/3357236.3395502
[108] Roel Vertegaal, Robert Slagter, Gerrit van der Veer, and Anton Nijholt. 2001. Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '01*. ACM Press, Seattle, Washington, United States, 301–308. DOI: http://dx.doi.org/10.1145/365024.365119
[109] Michael Walker, Hooman Hedayati, Jennifer Lee, and Daniel Szafir. 2018. Communicating robot motion intent with augmented reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 316–324.
[110] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y. Lim. 2019. Designing Theory-Driven User-Centric Explainable AI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–15. DOI: http://dx.doi.org/10.1145/3290605.3300831
[111] Robert H. Wortham and Andreas Theodorou. 2017. Robot transparency, trust and utility. *Connection Science* 29, 3 (July 2017), 242–248. DOI: http: //dx.doi.org/10.1080/09540091.2017.1313816 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/09540091.2017.1313816.
[112] Julia L. Wright, Jessie Y. C. Chen, and Shan G. Lakhmani. 2020. Agent Transparency and Reliability in Human–Robot Interaction: The Influence on User Con-fidence and Perceived Reliability. *IEEE Transactions on Human-Machine Systems* 50, 3 (June 2020), 254–263. DOI: http://dx.doi.org/10.1109/THMS.2019.2925717 Conference Name: IEEE Transactions on Human-Machine Systems.
[113] Jason Wuertz, Sultan A. Alharthi, William A. Hamilton, Scott Bateman, Carl Gutwin, Anthony Tang, Phoebe O. Toups Dugas, and Jessica Hammer. 2018. A Design Framework for Awareness Cues in Distributed Multiplayer Games. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–14. DOI: http://dx.doi.org/10.1145/3173574.3173817
[114] Greg Zacharias. 2019. *Autonomous horizons: the way forward*. Air University Press ; Curtis E. LeMay Center for Doctrine Development and Education, Maxwell Air Force Base, Alabama.

LUDOGRAPHY

AC:NH Nintendo EPD. *Animal Crossing: New Horizons*. Game [Nintendo Switch], March 2020.

Aka Richard Garriott. *Akalabeth: World of Doom*. Game [Apple II, DOS], 1979.

AoE2 Ensemble Studios. *Age of Empires II: Age of Kings*. Game [Windows], September 1999.

BW2 Lionhead Studios. *Black and White 2*. Game [Android], October 2007.

CTaxi Hitmaker. *Crazy Taxi*. Game [Windows], February 1999.

CTow Human Entertainment. *Clock Tower*. Game [Super Nintendo], September 1995.

DKC2 Rare. *Donkey Kong Country 2: Diddy's Kong Quest*. Game [Super Nintendo], November 1995.

DQ MachineGames. *Dragon Quest*. Game [Nintendo Enterntainment System], October 2017.

DR Irem. *Disaster Report*. Game [Wii Switch], November 2018.

Des Bungie. *Destiny*. Game [Playstation 4], September 2017.

Des3 Mimimi Games. *Desperados III*. Game [Windows], June 2020.

EFT Battlestate Games. *Escape from Tarkov*. Game [Windows], July 2017. Battlestate Games, London, UK.

EQ Sony Online Entertainment. *EverQuest*. Game [Windows], March 1999. Sony Online Entertainment, San Diego, California, USA.

ES4 Bethesda Game Studios. *The Elder Scrolls IV: Oblivion*. Game [Windows], March 2006.

FO3 Bethesda Game Studios. *Fallout 3*. Game [Xbox 360], October 2008. Bethesda Softworks, Rockville, Maryland, USA.

H:ZD Guerrilla Games. *Horizon: Zero Dawn*. Game [PS4], February 2017. Sony Interactive Entertainment, San Mateo, California, USA.

Ins New World Interactive. *Insurgency*. Game [Windows], December 2018.

JB:NF EuroCom. *James Bond 007: Nightfire*. Game [GameCube], November 2002.

LM Nintendo EAD. *Luigi's Mansion*. Game [GameCube], November 2001.

LZ:BW Nintendo EPD. *The Legend of Zelda: Breath of the Wild*. Game [Switch], March 2017. Nintendo, Kyoto, Japan.

LZ:MM Nintendo EAD. *The Legend of Zelda: Majora's Mask*. Game [Nintendo 64], April 2000.

LZ:OT Nintendo EAD. *Ocarina of Time*. Game [Nin64], November 1998.

LZ:WW Nintendo EAD. *The Legend of Zelda: The Wind Waker*. Game [GameCube], December 2002.

LoTR:RotK EA Redwood Shores. *Lord of the Rings: The Return of the King*. Game [Game Boy Advance], December 2003.

M2:SB Bungie. *Myth 2: Soulblighter*. Game [PC], December 1998.

MESOM Monolith Productions, Inc. *Middle-Earth: Shadow of Mordor*. Game [Windows], September 2014.

MG Bluepoint Games. *Metal Gear*. Game [Nintendo Enterntainment System], July 1987.

MGS2 Konami. *Metal Gear Solid 2 HD*. Game [Nintendo Enterntainment System], July 1990.

MK8 Nintendo EPD. *Mario Kart 8*. Game [Nintendo Switch], April 2017.

MN Klei Entertainment. *Mark of the Ninja*. Game [Nintendo Switch], September 2012.

Mad21 EA Tiburon. *Madden 21*. Game [Xbox 360], August 2020.

Mor Triternion. *Mordhau*. Game [Windows], April 2019.

Pac NAMCO. *Pac-Man*. Game [Atari], June 1980.

Pik Nintendo EAD. *Pikmin*. Game [GameCub], October 2001.

Por Valve. *Portal*. Game [Windows], October 2007.

RCT Chris Sawyer. *RollerCoaster Tycoon*. Game [Windows], October 2002.

SH Superhot Team. *Superhot*. Game [Windows], February 2016.

SM64 Nintendo EAD. *Super Mario 64*. Game [Nintendo 64], October 2017.

SMW Nintendo EAD. *Super Mario World*. Game [Super Nintendo], November 1990.

SS Croteam. *Serious Sam HD*. Game [Windows], March 2001.

SoT MachineGames. *Sea of Thieves*. Game [Xbox One]], March 2018.

T1 Looking Glass Studios. *Thief*. Game [PS4], February 2014. Eidos Interactive, London, UK.

T2 Looking Glass Studios. *Thief II*. Game [Windows], March 2000. Eidos Interactive, London, UK.

TWAU Telltale Games. *The Wolf Among Us*. Game [Windows], October 2013.

VC Sega. *Valkyria Chronicles*. Game [Playstation 3], November 2008.

W2:TNC MachineGames. *Wolfenstein II: The New Colossus*. Game [Windows], October 2017.

dnd Ray Wood and Gary Whisenhunt. *dnd: The Game of Dungeons*. Game [PLATO], 1975.