Evaluation of virtual agents' hostility in video games

Remi Poivet, Alexandra de Lagarde, Catherine Pelachaud, & Malika Auvray

Abstract—Non-Playable Characters (NPCs) are a subtype of virtual agents that populate video games by endorsing social roles in the narrative. To infer NPCs' roles, players evaluate NPCs' appearance and behaviors, usually by ascribing human traits to NPCs, such as intelligence, likability and morality. In particular, hostile NPCs in video games are essential to build the games' inherent challenges. The three experiments reported here investigated the extent to which the perception of hostility in a military shooter game (including both threat of appearance and aggressiveness in behaviors) is influenced by the appearance and the behaviors of NPCs thanks to perceived intelligence, likability and morality-related questionnaires. Our results first show that hostility is efficiently conveyed through NPCs' behaviors, but not significantly by their appearance. Second, our study allows identifying the main predictors of hostility perception, namely unfriendliness, knowledge and harmfulness.

Index Terms—Affective issues in enhancing machine/robotic intelligence, Cognition, Entertainment, Games
Keywords—Video games, Hostility, Likability, Perceived intelligence, Non-playable characters

1 INTRODUCTION

In video games, the virtual agent embodied by the human player (referred to as the main character) interacts with virtual agents that populate the game: the non-playable characters (NPCs). These NPCs, which can be considered as a subtype of highly interactive virtual agents, are defined by their role in the game's narrative. For instance, some NPCs can be designed as friendly artificial agents aiming to help the main character in pursuing the game. This can be done by offering positive outputs for the player. On the other hand, some NPCs can be hostile toward the main character, and thus have negative interactions with them. The NPC's role shapes the type of interaction it will have with the main character. For designers, this implies manipulating specific factors to convey the intended gaming experience.

In most video games, hostility is crucial, as enemies endorse the challenge of the game. Hostility in games is conveyed through an antagonistic stance, which implies threatening appearances and aggressive actions towards the main character. Game designers manipulate the appearance and behaviors of hostile NPCs to create diverse enemy types, enhancing the immersive gaming experience.

To express NPCs' hostile intentions, game designers rely on visual cues associated with hostile stereotypes and aggressive behaviors. Thus, creating NPCs involves manipulating their design factors during the production of the game. However, how players experience and evaluate the hostility of the enemies they encounter remains under-investigated so far. In particular, the respective impact of visual and behavioral designs on players' evaluation of NPCs' hostility is yet to be unveiled. Beyond the video game context, investigating how humans evaluate virtual agents' hostility through their appearance and actions would deepen our understanding of affective interactions with artificial agents. For example, simulating hostility in police academy training through virtual reality can aid trainees in conflict management [1]. Consequently, designers of virtual agents can leverage our findings to better control the factors influencing humans' evaluations of their interactions.

Manipulating the factors conveying NPCs' roles and hostility has an influence on player's expectations of NPCs' personality traits (see Figure 1). For instance, in the game's narrative, morality holds substantial importance and NPCs designed in line with their roles, such as friendly versus hostile NPCs, should be perceived differently in terms of their moral trait. Similarly, traits such as likability and perceived intelligence can vary depending on NPCs' appearance or behaviors. These traits contribute to an immersive experience and foster a positive appreciation of interactions with NPCs. Consequently, creating engaging enemies requires careful consideration of those factors that convey hostility as well as other traits, such as their likability, morality, and perceived intelligence, which actively shape players' overall gaming experience. To investigate the factors influencing players' perception of NPCs, in particular their hostility, three experiments were conducted in the context of a military shooter game. In the next section, we define hostility and list its characteristics related to appearance and behavior. Section 3 introduces the three personality traits our study focuses on, namely morality, likability, and perceived intelligence. From Section 4 onward, we present the three conducted studies. Section 5 describes the methodology. Sections

6 to 8 detail each of the experiments and section 9 discusses their results. Finally, the paper ends by presenting some limitations and grounds for future works.

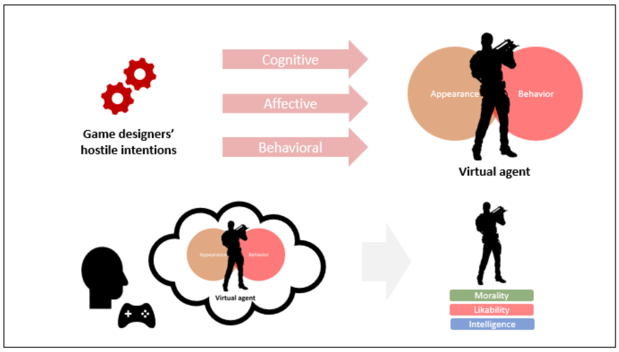

Fig. 1. The game designers' intention to create a hostile NPC influences their choice of design factors that affect players' attribution of personality traits.

2 Background

Antagonists are designed to play a unique role inside a video game's narrative and they need to look and act like the "bad guys". The worst of them, the "final boss", is reached after interactions with less hurtful antagonists. The hierarchy of challenges during the narrative allows keeping the game entertaining. To better understand how antagonists are conceived by game designers, the definition of hostility for virtual agents needs to be clarified. To do so, the parameters that allow virtual agents, including NPCs, to convey hostility will be detailed below.

2.1 Definition of hostility and its characteristics.

Most studies on human-agent interaction try to develop artificial agents that adopt a positive stance towards humans (e.g., chatbots designed for elderly care, see [2]). Nonetheless, the creation of virtual 'villains' could have certain advantages, in particular when designing hostile NPCs.

Hostility has been the focus of studies in clinical psychology where it is usually defined as an antagonistic attitude towards people including cognitive, affective, and behavioral components [3], [4]. Cognitive aspects of hostility refer to cynical beliefs and mistrust of others (e.g., believing that an agent has negative intentions). The affective dimension covers the negative emotions induced by the antagonistic attitude (e.g., anger, disgust). The behavioral component involves verbal or physical aggressiveness which can be defined as impulsive or goal-oriented behaviors towards others motivated by a desire to harm them. Hostility involves two core characteristics: threat and aggressiveness. Threat refers to the perceived potential for harm or danger posed by an agent or a situation [5]. It is often associated with the anticipation of negative outcomes or the activation of self-protective mechanisms. On the other hand, aggressiveness primarily focuses on verbal or physical actions directed toward others with harmful intentions [6].

In social psychology research, hostility perception in others is characterized by the detection of physical and verbal aggressiveness [7], [4]. However, it is important to note that aggressiveness is a narrower construct than hostility, as it primarily encompasses observable behaviors. In contrast, hostility involves a broader range of negative attitudes, emotions, and antagonistic actions directed towards others. When interacting with a hostile person, individuals often infer the inner states of this person based on visual cues that communicate potential aggressiveness, which is perceived as a threat (e.g., facial expressions associated with negative emotions). These cues allow individuals to adapt their attitudes and behaviors accordingly. Then, if the person's subsequent behavior is aggressive, this confirms the initial inference.

Perceiving hostility in virtual agents relies on the same mechanism of decoding a combination of visual cues related to threatening appearances and aggressive behaviors. Blankendaal et al. created aggressive virtual agents to measure their effects on human perception and to assess their potential benefits for clinical applications [8]. The authors compared the level of stress induced in participants facing an aggressive human versus an aggressive virtual agent. In this study, both the virtual agent and the human adopted an aggressive stance towards the participant (e.g., screaming at them). Both interactions generated physiological stress in participants, although the induced stress was lower when interacting with a virtual agent than when interacting with a human. Bosse et al. accounted for this result by the fact that virtual 'villains' are non-consequential, whereas human-to-human interactions can have real consequences for humans [9]. In the field of video games, designing hostile NPCs is highly relevant to increasing the challenging and engaging aspects of the game. Hostile NPCs can inflict damage on the main character, which results in the loss of their "health points". Interacting with a hostile NPC can have multiple consequences, such as the need to find a friendly NPC to heal the main character's wounds or even its death. For instance, in the shooter game Ready or Not, the main character is heading a SWAT team in multiple crisis scenarios (e.g., hostage rescue or bomb defusing). Here, it is crucial for the player to rapidly decide whether the NPCs they encounter are hostile or not.

2.2 Threatful appearance

The decision process is essential when facing potential threats. In video games, choosing whether or not to interact with NPCs is often influenced by visual cues from their appearance. In Ready or Not, the main character may enter a room with hostile NPCs ready to attack them, which can lead to the end of the game session if the player is too slow. Detecting NPCs' social roles has been studied by Rogers et al.across different game types, revealing that players rely on stereotypical visual cues to understand NPCs' roles [10]. For example, NPCs having the appearance of mentors wear a beard, which conveys older age and higher competence [11]. These visual cues aim to activate stereotypes in the player's mind [12]. In Super Mario Bros., the aggressive and angry nature of the Goombas is conveyed through their slanting eyebrows. Furthermore, Ferstl et al. examined the effect of manipulating facial features on humans' attribution of personality traits to virtual agents [13]. The authors found that wider faces influence the perception of trustworthiness, aggressiveness, and dominance. The narrative context also plays a role in determining the appearances that are perceived as a threat. In a military shooting setting like Call of Duty 4: Modern Warfare, human soldiers are considered as enemies, whereas in science fiction shooter games like Halo, they can be perceived as allies. Apart from stereotypes, the mere presence of a weapon in an NPC's appearance is often sufficient to communicate a threat in players' minds. Detecting the presence of a weapon provides information about potential behavioral patterns, attack range, frequency, and any special ability the NPC may possess [14]. Consequently, when the main character encounters NPCs for the first time, players can make predictions about their abilities, intentions, and emotional states, which significantly influence their subsequent interactions.

2.3 Aggressive behaviors

For virtual agents, behaviors mainly refer to the nonverbal abilities that turn their static appearance into a dynamic stance. Nonverbal behaviors include facial expression [15], gaze direction [16], and different types of gestures [17]. As nonverbal communication adds supplementary information to the verbal communication, it enhances the believability when interacting with the agent [18]. The implementation of such nonverbal cues on virtual agents influences human perception, as shown in Kim et al.'s study [19]. In their experiment, users evaluated their interactions with virtual agents more positively when they perceived them as more reactive and socially present. Hence, in the eyes of the users, the interaction with virtual agents gains relevance when they are designed with behaviors such as facial expressions or gestures.

NPCs can be seen as interactive virtual agents, which can be designed with complex behaviors. In that sense, NPCs' behaviors refer to the sets of predefined sequences of actions leading to interactions with their surroundings. In shooter games, hostile NPCs are characterized by their aggressiveness toward the main character. Game designers manipulate the aggressiveness of enemies by utilizing a classification system relying both on their behaviors and their weapons [14]. For example, enemies armed with shotguns follow a sequence of actions that involve shortening the distance to target, due to the weapon's limited range. On the contrary, enemies equipped with sniper rifles behave differently and tend to maintain their distance from the main character. By considering these parameters, game designers allow players to anticipate the behaviors of enemies just by identifying their weapons.

In addition, game designers have the flexibility to manipulate other parameters, such as enemies' shooting precision (i.e., their accuracy in hitting targets), the damage they can inflict, and even their decision-making processes as they approach the main character. Decision-making particularly affects how enemies manage cover sequences, as game environments often include cover walls that create safe zones for both the main character and NPCs. For instance, enemies using cover tactics generate less pressure on the main character compared to those who recklessly charge without seeking cover. In conclusion, game designers shape the behaviors of hostile NPCs to convey the game's level of challenge by carefully manipulating a combination of parameters and action sequences. These design choices directly influence the gaming difficulty, as a more aggressive NPC presents a greater threat to the main character and increases the risk of defeat.

3 PERSONALITY TRAITS

First impressions play a crucial role in evaluating others, as they activate our knowledge and concepts to predict interactions and influence decision-making [20]. This extends to artificial agents, as humans tend to perceive computers and robots as social actors [21]. NPCs in video games can also make first impressions on players, who infer both basic (good or bad) and complex (personality traits) information about them. Game designers aim to convey the three components of hostility (i.e., affective, cognitive, and behavioral hostility) through the manipulation of NPCs' appearance and behaviors. When the main character encounters an NPC for the first time, players' preconceived notions of artificial agents come into play, shaping their expectations for future interactions and influencing their decision to engage. For example, in the comparison between games like Call of Duty 4: Modern Warfare and Halo, players might be more inclined to approach visually similar soldiers in Halo due to their knowledge of the game's context, where enemies are non-humanoid. On the other hand, encountering a humanoid soldier in Call of Duty might represent a threat in players' minds. Thus, visual cues from NPCs' appearance allow players to ascribe personality traits, such as morality, and predict their role in the game [22]. Visual cues and the game context can lead players to infer that a character is friendly and increase the likelihood of their interaction, while NPCs perceived as unfriendly may be approached with caution. These assumptions rely on stereotypes associated with likability and morality, which are evaluated based on NPCs' appearance [23]. During gameplay, players update their mental representations of NPCs by comparing their expectations with the behaviors they observe. In particular, players update their evaluation of intelligence, likability, and morality. Consequently, the ascription of personality traits to NPCs, encompassing the three aforementioned traits, is a crucial aspect of the gaming experience, increasing the engaging aspect of the interaction [24]. Beyond video games, understanding how appearance influences morality ascription is useful for serious games aiming at decreasing stereotypes [1]. To investigate further the contributions of morality, likability, and perceived intelligence, our study analyzes scales' items that focus on these dimensions.

3.1. Morality.

As most video games are narrative-oriented, the perception of NPCs' roles is important for players' mental representations. The role of a character in fictional media can be described thanks to their morality traits within the narrative. According to Long and Sedley, morality can be defined as "the differentiation of intentions, decisions, and actions between those that are distinguished as proper and those that are improper" [25]. In the field of video games, morality has been investigated through the antagonistic characters' physical appearances [22]. The concept of antagonist is bonded to the notion of morality ascription as they are introduced as the villain of the story ([26], [27], [28]). The morality of antagonists is based on the evaluation of their actions and the nature of their intentions. To assess fictional characters' perceived morality, Grizzard et al. [29] introduced a questionnaire based on the Moral foundation theory [30], relying on five-dimension items: Harm/Care, Fairness/Reciprocity, Ingroup/Loyalty, Authority/Respect, Purity/Sanctity. Using these scales, Pradantyo et al. identified visual cues from the appearance that trigger mental representations associated with morality [22]. The authors highlighted that players are able to ascribe moral intention to characters they never saw before using only their physical appearance. Their results show that certain stereotypical items extracted from NPCs' appearance are associated with a perception of immorality (i.e., covered faces, dark colors clothes, skin problems, shape dysmorphia). Therefore, as part of the ascribed human traits, morality perception influences players' mental representations of NPCs and subsequently empowers their decision to interact with them and the way they do so.

3.2. Likability.

One dimension which has been shown to influence human's perception of agents is their likability. In that sense, Haidt and Joseph [31] developed a likability evaluation scale made of five items: Like/Dislike, Kind/Unkind, Pleasant/Unpleasant, Friendly/Unfriendly, Nice/Awful. This scale has been used to assess the first impression that humans have of robots' likability [32] and it can be used for NPCs in the context of video games. In the case of the Goombas in Super Mario Bros. introduced earlier, the choice of certain visual cues (i.e., the slanting eyebrows) aims to be automatically perceived as unlikeable by players. Having this negative first impression solely based on the Goombas' appearance allows players to rapidly detect the Goombas' intentions and to expect subsequent unlikable behaviors. Therefore, very much like what happens when humans meet other humans, NPCs such as Goombas can be perceived as more or less likable. This dimension can be manipulated by game designers to convey NPCs' role within the game.

3.3. Perceived intelligence.

Humans attribute intelligence to artificial agents based on their appearance and behavior. People's evaluation of intelligence relies on two main dimensions: understandability and performance. On the one hand, Bartneck et al. [33] underlined the difficulty of understanding the underlying intention of an artificial agent from its behaviors. For instance, in video games, NPCs have limited behaviors (i.e., sets of available interactions) while players are unpredictable. The combination of limited abilities and unpredictability can lead NPCs to have inaccurate behaviors, which are seen as mistakes by players. When players cannot explain NPCs' behaviors, this negatively influences their perceived intelligence and thus their overall evaluation.

On the other hand, Koda and Maes [34] explained that virtual agents' performance in completing their tasks is crucial to users' evaluation. For instance, hostile NPCs are designed to create difficulty by endorsing the role of opponents. Players have expectations about what opponents should be able to do in each context. Then, if hostile NPCs are not performing well in this role (e.g., they cannot defeat the main character), they are perceived as lacking intelligence. In 2002, the video game company Bungie conducted an experiment while creating their game Halo [35]. They manipulated game sessions' difficulty by creating different hostile characters who could be defeated more or less easily and who could inflict different levels of damage. The experimenters asked players to judge the presence or absence of intelligence in NPCs on a binary scale and found that NPCs that are more challenging for players are perceived as more intelligent.

Bartneck et al. [33] proposed a scale to assess users' evaluation of artificial agents' perceived intelligence that consists of five semantic differential items: Incompetent/Competent, Ignorant/Knowledgeable, Irresponsible/Responsible, Unintelligent/Intelligent, Foolish/Sensible. Using the scores of each of the items, the scale has the advantage of covering the evaluation of both understandability (Responsible, Intelligent, and Sensible scores) and performance (Competent and Knowledgeable scores).

4 EXPERIMENTS ON HOSTILE NPCS

To summarize, players rely on their knowledge and stereotypes to infer NPCs' role inside the game by attributing human traits to them and infer potential hostile intentions based on their appearance and behavior. The perception of threat, inferred from NPCs' appearance, help the player make the best decision, namely, to decide whether to interact or not with them. Then, during interactions, players update their judgment about NPCs. For instance, interacting with hostile NPCs might result in negative consequences for the main character, such as a loss of health points or a defeat. Such experience induces the player to categorize NPCs as aggressive. The present research aims at investigating how hostility is conveyed by NPCs. We hypothesize that hostility evaluation in virtual agents is a combination of their appearance, activating threatful concepts, and their aggressive behaviors. Three experiments were conducted in the context of a Ubisoft military shooter game to investigate the weight of appearance and behaviors involved in hostility evaluation. Hostility evaluation was investigated through three dimensions: Likability, Perceived Intelligence, and Morality. More precisely, the first experiment focuses on the evaluation of threats conveyed by NPCs' appearances. The second experiment assessed the evaluation of aggressiveness in NPCs' behaviors. The last experiment used the results from the previous two experiments; the appearance and behavior of NPCs were manipulated to evaluate their impact on hostility evaluation. In addition, we looked at which scales' items predict the perception of hostility.

5 GENERAL METHODS

Participants

For each study, participants were volunteer players registered at the Ubisoft User Research Laboratory platform (i.e., a mailing platform where players can deliberately enter their information to participate in Ubisoft's research). All the participants were contacted by email and were provided with information about the content of the research. The email informed participants of the free nature of their participation and that they could stop at any time they wanted. Each experiment took approximately 15 minutes to complete and was conducted in accordance with the Declaration of Helsinki.

Procedure and materials

The three experiments were conducted online. The participants were first welcomed and introduced to the topic of the research. Then, they were presented with the stimuli and asked to answer questionnaires.

To create the stimuli, NPCs from the military video game *Ghost Recon Breakpoint* were used. This is a realistic military game in which the main character fights against hostile soldiers and interacts positively or in a neutral way with civilians. In the game, there are more soldiers than civilians, as in military shooter games soldiers are at the core of the action. In our study, this distribution was respected, and hence, six different archetypes of NPCs were selected: five soldiers and one civilian. Archetypes are represented through an appearance that involves specific visual cues and behaviors that aim to express NPCs' role inside the game (see Figure 2 for an illustration and Table 1 for the details). For each archetype, three variations were chosen to display different samples of appearance (i.e., 3 different civilians that differ by the color of their outfit).

Fig. 2. NPCs' design from left to right: Civilian (one of the three displayed civilians, friendly role, will get on their knees if the main character targets them with their weapon), Drone carrier (enemy role, spawn drones to attack as their main behavioral pattern but can use a rifle at close ranges), Rifleman (enemy role, use a rifle), Rocket launcher (enemy role, use a rocket as their main weapon but can use a rifle at close ranges), Rusher (enemy role, use a shotgun and will run toward a main character located at a close range), and Sniper (enemy role, use a sniper but can use a rifle at close ranges).

6 EXPERIMENT 1: THREAT OF NPCS' APPEARANCES

Experiment 1 aims at evaluating players' perception of the threat conveyed by NPCs based on their appearance (i.e., NPCs' shapes and the threatening cues from their outfit). As the aim of our study is to investigate players' evaluation of NPCs from their appearance, all soldiers' were displayed without their weapon to limit participants' inference of their abilities and focus on NPCs' shape and outfit. Following their role in the game, our hypothesis is that the appearance of civilians should be perceived as less threatful than soldiers' ones. This difference would reflect the type of interaction expected in the game, as civilians are neutral or friendly characters while soldiers are enemies. In addition, soldiers should be perceived at a distinct level of threat based on the visual cues (e.g., military gears) from their appearance.

Participants, stimuli, and procedure

Thirty-nine French participants completed Experiment 1 (thirty-six men, one woman and two others, mean age = 31.4 years old, SD = 9.2). The gender repartition can be explained by the significant part of men registered on the Ubisoft User Research mailing list and that the survey was randomly sent to players and was about a military game (i.e., potentially interested

TABLE 1
DESCRIPTION OF THE APPEARANCES OF NPCS SELECTED FOR EXPERIMENT 1

Archetype of the NPC	Role in the game	Head	Body	Accessories
Civilian	Not opponent	Uncovered face	Unmuscular shape. Recreational light-colored clothes (jacket or t-shirt with graphics, classic pants, and baskets shoes)	None
Drone Carrier	Opponent	Covered face (Balaclava)	Athletic shape. Black t-shirt, blue jeans, black boots, and gloves	Ammunition harness (dark colors), backpack, empty holster, and small screen on the right wrist
Rifleman	Opponent	Military cap (dark colors)	Athletic shape. Dark t-shirt, blue jeans, dark boots, and gloves	Light harness and empty holster. Headset and sunglasses
Rocket Launcher	Opponent	Partly covered face (dark colors). Red military cap	Athletic shape. Grey t-shirt, blue jeans, and black boots	Backpack and empty holster. Headset
Rusher	Opponent	Partly covered face (dark scarf on the lower face)	Athletic shape. Black t-shirt, blue jeans, black boots	Ammunition harness (light colors), empty holster and ski goggles
Sniper	Opponent	Uncovered face	Athletic shape. Dark t-shirt, blue jeans, dark boot.	Ammunition harness (dark colors), a tight protection and empty holster

gamers were mostly men). In Experiment 1, the materials consisted of static images of the six standing NPCs (see Figure 2 and Table 1). Each NPC is shown with a neutral posture, no visible weapon (i.e., soldiers wore an empty gun holster and different ammunition belts), no visible background, and controlled skin color (see Table 1 for a detailed description of NPCs' appearance). NPCs' appearances were randomly presented three times each. As said earlier, there were more stimuli of soldiers than of civilians to respect the repartition in the original game. For each presentation, participants rated the amount of threat they perceived on a continuous scale (from '0 - not threatful at all' to '100 - totally threatful') and responded to the Likability and Perceived Intelligence scales (Godspeed, [33]) and to the morality questionnaire (CMFQ-S, [28]).

Results and discussion

Scores were obtained from participants' ratings of NPCs' threat in appearance, and the scales' items of Likability, Perceived intelligence and Morality. In order to maintain the representativeness of the dataset, no explicit outlier treatment was applied, hence all the collected data was included in the statistical analysis. The threat's scores did not follow a normal distribution. Therefore, a one-way ANOVA on ranks (Kruskal-Wallis H test) was conducted in order to investigate the significant difference of threat between NPCs' appearances ($df = 5$, $p < .001$, $\eta^2 = 0.429$). Civilian's appearance was perceived as significantly less threatful than soldiers' ones. Moreover, Drone carriers' appearance was perceived as the most threatful among all the soldiers (see Figure 3). A significant correlation was found between threat and the items from the Likability scale "Friendly" ($r = -0.364$), "Kind" ($r = -0.235$), "Pleasant" ($r = -0.17$), "Nice" ($r = -0.351$), "Intelligent" ($r = 0.249$), "Competent" ($r = 0.43$), as well as with the Morality scales regarding "Ingroup/Loyalty" ($r = 0.261$), "Harm/Care" ($r = 0.64$), "Authority/Respect" ($r = 0.67$) and "Fairness/Reciprocity" ($r = 0.515$). A multiple linear regression was conducted to identify the predictors of the score of threat perception. All the items explained 56.2% of the variability of the threat score ($R^2 = 0.562$). However, the items "Competent" ($\beta = 0.143$, $p = .034$), "Harm/Care" ($\beta = 0.255$, $p < .001$) and "Authority/Respect" ($\beta = 0.357$, $p < .001$) were significant predictors of threat perception. A multiple linear regression with those items showed that they predict 53.5% of the variability of the threatening score ($R^2 = 0.535$).

Experiment 1's results show a significant difference in threat perception between the NPCs (see Figure 3). The civilian's appearance was perceived as less threatful compared to all soldiers' appearances. This difference is explained by very distinct characteristics in appearance, such as casual versus military clothes and differences in body shape. The overall shape of soldiers

TABLE 2
DESCRIPTIVE STATISTICS FOR THREAT SCORES

	Civilian	Drone Carrier	Rifleman	Rocket Launcher	Rusher	Sniper
N	78	78	78	78	78	78
Mean	1.417	51.2	29.75	30.883	42.017	29.317
SD	2.872	20.757	18.636	19.297	18.813	19.827
Min	0	14	0	0	8	0
Max	11	94	77	68	84	91

is more squared with a prominence of muscles. These results are in line with [36], according to whom square shapes in fictional characters refer to strength and stability while round shapes are linked to safety and friendliness. In addition, soldiers' appearances include military accessories which are associated with concepts of war and violence. The linear regression showed a granularity of the threat attribution within soldiers' appearances. NPCs with the most threatful appearances were perceived as more competent, dangerous, and antipathetic. However, none of the items from the Likability scale significantly predicted the threat score, suggesting the low importance of this dimension to evaluate enemies' threat. Furthermore, one of the differences between soldiers' appearance was the visibility of their military gear. Soldiers perceived as the most threatful had also their face hidden (see the general discussion in Section 9, for further discussion of the results).

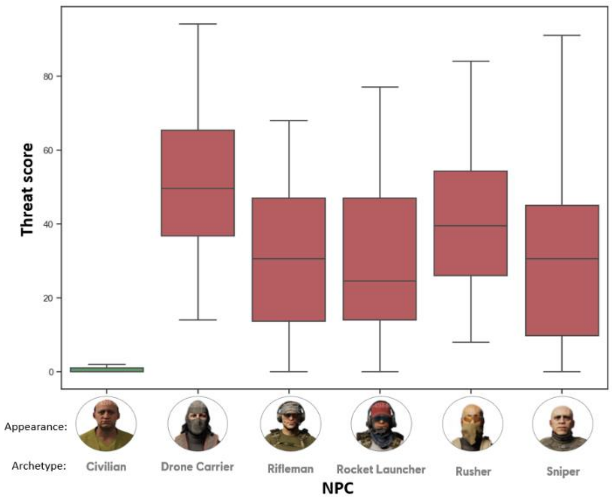

Fig. 3. Boxplots of NPCs' threat scores: boxplots' colors represent NPCs' role inside the narrative (Civilian in green and Soldiers in red). The boxplots show the distribution of the threat score by NPCs' appearances, their median and interquartile range. Black error bars represent the standard error of the mean.

7 Experiment 2: Aggressiveness of NPCs' Behaviors

Experiment 2 investigates how players evaluate the aggressiveness of several NPCs' behaviors. Based on their roles and the intended challenge within the narrative, NPCs behave differently to create different types of interactions with the main character. Civilians are friendly NPCs in military games, they should not be perceived as aggressive. Soldiers are the opponents in the narrative and their goal is to try and kill the main character. Therefore, unlike civilians, soldiers' behaviors should be perceived as aggressive. In addition, as hostile NPCs have different behaviors to express their stance toward the main character, participants should rate their aggressiveness differently.

Fig. 4-a. (left). In-game screen capture of NPCs' appearance used for the video. Note that all NPCs had the same appearance (same face, body shape, clothes, and weapon). Fig. 4-b. (right). Illustration of the task: the main character moves toward the NPC inside a corridor. Here the main character is in the background while the NPC is in the foreground. Note that for each image, watermarks are automatically generated and required by Ubisoft.

Participants, stimuli, and procedure

Forty-three French participants completed Experiment 2 (thirty-eight men, four women and one other, mean age = 31.9, SD = 8.2). In Experiment 2, we controlled the appearance of the NPCs; the six NPCs' appearances were identical (see Figure 4-a), which means that all NPCs' appearances consisted of the same face, body shape, clothes and weapon (a rifle, which is the most common weapon in military shooter games). The only difference between them was their behaviors, that is, their sequence of actions (see Table 2). For instance,

TABLE 3
DESCRIPTION AND CONSEQUENCES OF THE BEHAVIOR OF NPCS SELECTED FOR EXPERIMENT 2

Archetype of the NPC	Role	Shooting distance	Length of shooting sequences (seconds)	Cover sequences	Main character's state at the end of the interaction
Civilian	Not opponent	Not applicable	0	Get on their knees	Alive
Drone Carrier	Opponent	Long distance	6	Hide behind a wall	Alive
Rifleman	Opponent	Long distance	3	Hide behind a wall	Dead
Rocket Launcher	Opponent	Long distance	10	Hide behind a wall	Alive
Rusher	Opponent	Short distance	3	Rush toward the main character	Dead
Sniper	Opponent	Long distance	5	Hide behind a wall	Dead

TABLE 4
DESCRIPTIVE STATISTICS FOR AGGRESSIVENESS SCORES

	Civilian	Drone Carrier	Rifleman	Rocket Launcher	Rusher	Sniper
N	43	43	43	43	43	43
Mean	1.154	43.154	51.654	39.135	84.442	47.942
SD	4.399	26.216	27.189	25.463	18.156	22.282
Min	0	1	0	2	2	10
Max	31	100	100	100	100	100

civilians in military games get on their knees while soldiers attack the main character. For soldiers, the behaviors differ as a function of two parameters: shooting (number of shooting sequences, precision and distance) and cover sequences (hiding behind a wall or rushing toward the main character). As a result, the manipulated parameters lead to different consequences for the main character's state (i.e., dead or alive). NPCs' behaviors were shown to the participants via videos of pre-recorded interactions between the main character (controlled by the experimenter) and each of the NPCs. In each sequence, the main character moves toward the NPC inside a corridor. The video camera was placed in a specific angle to help participants focus on NPCs' behaviors (see Figure 4-b). For each video, the participants completed the same scales as those used in Experiment 1, except the threat score which was replaced by aggressiveness (from '0 - not aggressive at all' to '100 - totally aggressive').

Results and discussion

Scores were obtained from participants' ratings of aggressiveness of NPCs' behaviors and their ratings in the scales' items of Likability, Perceived intelligence and Morality. In order to maintain the representativeness of the dataset, no explicit outlier treatment was applied, hence all the collected data was included in the statistical analysis. The perceived aggressiveness's scores did not follow a normal distribution. Therefore, a one-way ANOVA was conducted on ranks (Kruskal-Wallis H test) to assess the significant difference of aggressiveness between NPCs (df=5, $p < .001$, $\eta^2 = 0.555$). The less aggressive behavior was of the civilian, which is characterized by getting on his knees when the main character comes too close. The most aggressive behavior was the Rusher's, which is characterized by a reduction of the distance between him and the main character. A significant correlation was found with aggressiveness for the items: "Friendly" (r= -0.657), "Kind" (r = -0.415), "Nice" (r = -0.492), "Responsible" (r = 0.311), "Sensible" (r = 0.384), Knowledgeable (r = 0.503), "Intelligent" (r = 0.494), "Competent" (r = 0.558), "Harm/Care" (r = 0.739), "Purity/Sanctity" (r = 0.505), "Authority/Respect" (r = 0.688) and "Fairness/Reciprocity" (r = 0.632). A multiple linear regression was conducted to identify the predictors of the score of aggressiveness. All the items explained 71.1% of the variability of the aggressiveness score ($R^2 = 0.711$). However, among them, only the items "Friendly" ($\beta = -0.279$, $p < .001$), "Knowledgeable" ($\beta = 0.120$, $p = .043$, "Harm/Care" ($\beta = 0.214$, $p < .001$), "Authority/Respect" ($\beta = 0.167$, $p = .005$), and "Fairness/Reciprocity" ($\beta = 0.148$, $p = .004$) were significant predictors of aggressiveness perception. Indeed, a multiple linear regression with those items showed a prediction of the variability of 69.5% of the aggressiveness score ($R^2 = 0.695$).

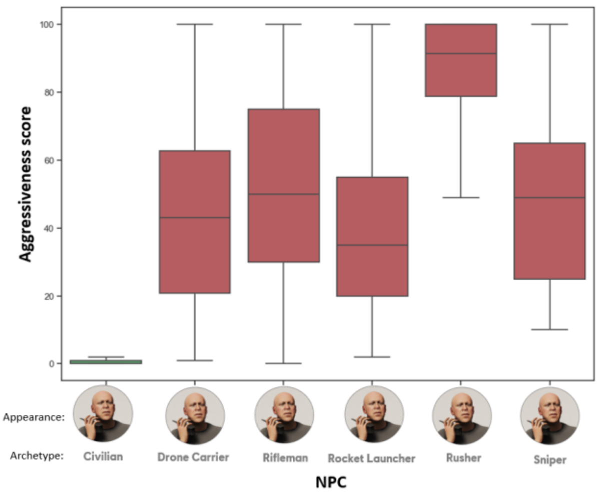

Fig. 5. Boxplots of NPCs' aggressiveness scores: boxplots' colors represent NPCs' role inside the narrative (the green color for Civilian and red for Soldiers). The boxplots show the distribution of the aggressiveness score by NPC behavior, their median and interquartile range. Black error bars denote the standard error of the mean.

The results from Experiment 2 indicate a significant difference of aggressiveness evaluation between all the NPCs. The difference of aggressiveness between civilian and soldiers' behaviors is explained by the lack of damage resulting from the civilian's activity. Indeed, the combination of the items "Friendly", "Knowledgeable", "Harm/Care", "Authority/Respect", and "Fairness/Reciprocity" refers to the perceived efficiency to harm the main character's health. The behavior perceived as the most aggressive one has the particularity of distance reduction between the two agents (i.e., rushers' behavior). This behavior induces in participants' the perception of an intrusion into their main character's safety zone. Other soldiers' behaviors were defined as less mobile (i.e., hiding behind a wall), meaning that the NPC is less intrusive into the main character's safety zone.

8 Experiment 3: Contribution of Appearance and Behavior to the Perception of Hostility

Experiment 3 investigates the influence of NPCs' appearance and behaviors on hostility perception. More precisely, the aim is to investigate within the predictors of threat and aggressiveness, those influencing players' evaluation of hostility. During this experiment, the participants completed the individual items that were found to be predictors of threat ("Competent", "Harm/Care" and "Authority/Respect") and aggressiveness ("Friendly", "Knowledgeable", "Harm/Care", "Authority/Respect" and "Fairness/Reciprocity") in Experiments 1 and 2.

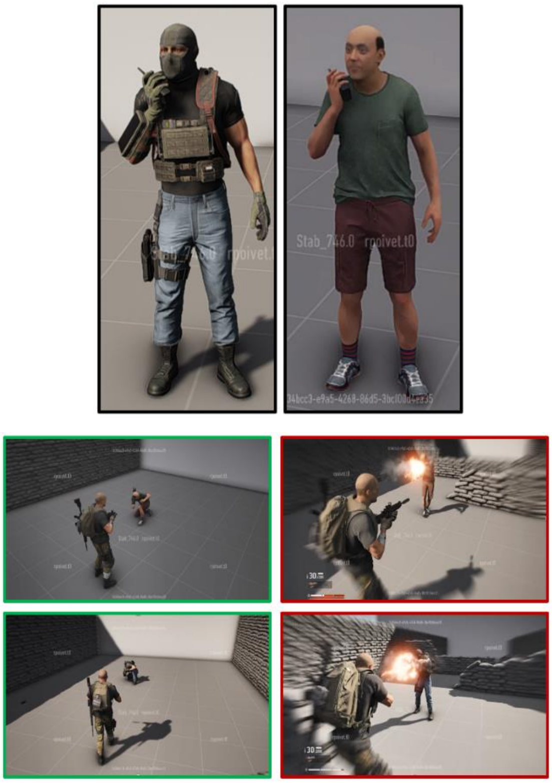

Fig. 6-a. (top) Illustration of the NPCs' appearance (most threatful appearance on the left: Drone Carrier and least on the right: Civilian), made of a screenshot of the NPCs encountered in the video. Fig. 6-b. (bottom) Screenshots from the videos illustrating the interactions between the main character and the NPCs. The two pictures on the top depict an NPC with a civilian appearance behaving non-aggressively (left) and aggressively (right) toward the main character. On the bottom, an NPC with a soldier appearance follows the same dichotomy. Same as E2, watermarks are automatically generated and required by Ubisoft when screens are captured in the game's editor.

Participants, stimuli, and procedure

Fifty-one French participants completed Experiment 3 (forty-six men, four women and one other, mean age = 30.03, SD = 6.4). Four NPCs were created based on the results from the two previous experiments. The appearances and behaviors with the highest and lowest threat and aggressiveness scores from Experiments 1 and 2 were used. From these, two congruent NPCs (most and least threatful appearance and aggressive behavior combined) and two incongruent NPCs (most threatful appearance and least aggressive behavior, and vice versa) were created. For each NPC, an interaction with the main character was recorded. The position of the camera was placed behind the main character, as it is the case in most video games. The videos stop after the interaction (i.e., either the main character is killed by the NPC or the main character reaches the crouched NPC). Before watching each video, the participants first had to rate the threat of the appearance (from '0 - not threatful at all' to '100 - totally threatful'), based on an in-game picture of the NPC (see Figure 6-a). Then, the participants watched the videos and evaluated NPCs' aggressiveness in behavior (from '0 - not aggressive at all' to '100 - totally aggressive') and hostility in NPCs (from '0 - not hostile at all' to '100 - totally hostile'). After that, the participants were asked to complete the following individual items: "Friendly", "Competent", "Knowledgeable", "Harm/Care", "Authority/Respect" and "Fairness/Reciprocity".

Results and discussion

No explicit outlier treatment was applied; hence all the collected data was included in the statistical analysis. The hostility scores did not follow a normal distribution. Therefore, a two-way ANOVA on ranks (Kruskal-Wallis H test) was conducted to assess the significant difference in hostility perception between NPCs ($df = 3$, $p < .001$). No significant difference between appearances was found, while there was a significant difference between behaviors ($p < .001$) (see Figure 7). There was no significant interaction between appearance and behavior. A multiple linear regression was conducted to identify the predictors of the score of hostility. All the items taken together explained 47.9% of the variability of the hostility score ($R^2 = 0.479$). However, only the items "Friendly" ($\beta = -0.253$, $p = .004$) "Knowledgeable" ($\beta = 0.214$, $p = .051$) and "Harm/Care" ($\beta = 0.39$, $p < .001$) were significant predictors of hostility perception. A multiple linear regression with these predictors showed a prediction of 47% of the variability of the hostility score ($R^2 = 0.47$). A two-way ANOVA was conducted to assess the significant difference between appearance and behavior for the items "Friendly", "Knowledgeable"

TABLE 5
DESCRIPTIVE STATISTICS FOR HOSTILITY SCORES

	Not aggressive soldier	Aggressive soldier	Not aggressive civilian	Aggressive civilian
N	51	51	51	51
Mean	44.740	87.960	38.200	87.700
SD	33.245	15.585	36.006	25.463
Min	0	46	0	21
Max	100	100	100	100

and “Harm/Care”. No significant effect of appearance nor interaction between appearance and behavior were observed for the item Friendly, while a significant effect was found for the behavior ($p < .001$). No significant effect of appearance and behavior was found for the item “Knowledgeable”. Significant effects of appearance ($p = .015$), behavior ($p < .001$) and an interaction between Appearance and Behavior ($p = .02$) were found for the item “Harm/Care”. Finally, a linear regression with the scores of threat and aggressiveness was conducted to assess their impact on the hostility score. The two scores explained 48.7% of the variability of the final score ($R^2 = 0.487$). Only the score of aggressiveness was a significant predictor ($\beta = 0.690$, $p < .001$).
The results from Experiment 3 show that hostility perception is different for the four NPCs. This difference is due to the variability in the aggressiveness of the behaviors. Moreover, friendliness, knowledge and harm’s perception were the main predictors of hostility attribution. Those results follow findings from Experiment 2 and confirm the importance of behaviors for hostility attribution.

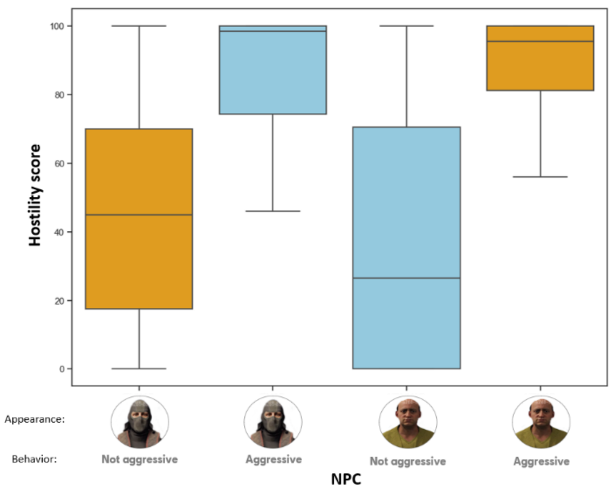

Fig. 7. Boxplots of NPCs’ hostility scores: boxplots’ colors represent NPCs’ congruence during the experiment (the blue color for congruent and orange for incongruent). The boxplots show the distribution of the hostility score by NPCs’ manipulated parameters, their median and interquartile range. Black error bars denote the standard error of the mean.

9 GENERAL DISCUSSION

Our study aimed at investigating, through three experiments, how video game players evaluate NPCs’ hostility using soldiers and civilians in a military game. To tackle the perception of hostility, two dimensions were considered here, namely appearance and behavior, while measuring personality traits attributed to NPCs. The first experiment dug into the perception of threat conveyed by NPCs’ appearance. The second experiment focused on the perception of aggressiveness in NPCs’ behaviors. The third experiment looked at the perception of hostility both in appearance and behavior.

9.1. Perception of hostility in appearance, behavior, and when the two are combined

The first result (Experiment 1) showed that NPCs’ appearance influences the perception of threat. This occurs at three levels: (i) there was a significant difference in threat evaluation between NPCs looking like civilians versus soldiers, (ii) there were significant differences in threat perception among soldiers (differing in gears or covered faces), and (iii) the level of threat conveyed by the different NPCs rely on the morality and competence ascribed to them. More precisely, differences in threat evaluation were observed among NPCs depending on their archetype. Stereotypes were inferred from NPCs’ body shapes, clothes, and gears. First, the plain and neutral looks of civilian NPCs must have been implicitly associated with the harmless environment of everyday life, while the gears worn by the soldiers must have been associated with danger and violence, thus with threat. Within soldiers, those with covered faces and dark clothes were perceived as more threatening than the other ones. These results are in line with Pradantyo et al.'s study [22], which outlined the visual cues influencing the perception of immorality within antagonists in video games, and among them covered faces and darker colors.
The second result (Experiment 2) showed differences in the perception of aggressiveness in behavior as a function of sequences of action, even when all NPCs have the same appearance. First, NPCs displaying a submissive stance (crouching, stereotypical of civilians) are evaluated as less aggressive than NPCs attacking the main character (shooting, stereotypical of soldiers). Second, the perception of aggressiveness increases as a function of the NPCs’ ability to defeat the main character. For instance, some NPCs are not

accurate with their weapon, their shooting resulting in a small loss in the main character's health points. Third, the NPC whose behavior is characterized by a reduced distance from the main character is evaluated as the most aggressive. Social spaces can be defined by the distance between agents [37]. NPCs and the main character possess their own personal space, that is, the distance in which they each feel comfortable. When NPCs rush toward the main character and intrude on the latter's personal space, their behavior is perceived as aggressive.

The third result (Experiment 3) revealed that when the perception of hostility is assessed both through appearance and behaviors, hostility is conveyed through aggressive behaviors rather than threatful appearances. Hence, when there are no interactions (as in Experiment 1), NPC's appearance effectively communicates threat. However, when there is an interaction, NPCs' hostility is primarily communicated through their behaviors.

9.2. The predictors of hostility perception

Our study also aimed at highlighting the main predictors characterizing hostility perception through the threat of appearance and the aggressiveness in behavior. The first experiment focused on appearance. Only three items turned out to be significant: "Competent" (from the Likability scale), "Harm/Care" and "Authority/Respect" (from the Morality scale). Hence, NPCs that are perceived as the most threatening are also perceived as the most competent, the most harmful, and the most inclined to cause chaos and disorder, thus as the most immoral. Civilians are expected to be less competent in a military game and thus less threatening than soldiers who evolve in familiar territory.

The second experiment focused on the aggressiveness in behavior. Five items turned out to be significant: "Knowledgeable" (from the Perceived Intelligence scale), "Friendly" (from the Likability scale), and "Harm/Care", "Authority/Respect", "Fairness/Reciprocity" (from the Morality scale). "Knowledgeable" refers to the perceived efficiency of NPCs to achieve their task. In our game context, this corresponds to the abilities and skills required for the military (e.g., weapon usage, personality, and emotional control). Regarding friendliness, the more aggressive NPCs are evaluated, the less friendliness they are ascribed to. The last three predictors concern the ascription of interpersonal traits related to violence and authority. This can be explained by players' mental representations of the military world in which soldiers are opponents of the main character, as they endorse an antagonistic stance through the improper nature of their acts (e.g., killing civilians) which triggers the main character's moral compass (e.g., soldiers oppressing civilians). To summarize, NPCs' behaviors perceived as knowledgeable, unfriendly, harmful, authoritarian, and unfair are also perceived as the most aggressive.

In the third experiment, when evaluating the perception of both the threat of appearance and aggressiveness in behavior, there was no significant predictor of the former (see Figure 7). While the threat of appearance conveyed information about the role and intention of the characters, only behaviors conveyed hostility. The items "Friendly", "Harm/Care", and "Knowledgeable" were significant predictors of the latter, in line with Experiment 2. In other words, NPCs that are perceived as the most hostile are perceived as the most unfriendly, harmful, and knowledgeable. In the absence of interaction with a virtual agent, it seems that its appearance effectively communicates threat through the anticipation of hostility. However, in the context of interaction, the virtual agent's hostility is primarily communicated through the nature of its behavior. This finding is in line with the assumption made by Bosse et al. [9] and confirms that behaviors that have consequences (such as damages inflicted on the main character) are the main parameters that influence humans' evaluation of virtual agents' hostility.

9.3. Recommendations to designers

Creating hostile virtual agents is complex and is a crucial aspect when designing most video games. Enemies are part of the inherent challenges of these games and might represent most of the players' interactions. It is thus essential to understand how appearance and behaviors affect players' evaluation of hostility and personality traits ascription, such as likability, intelligence, and morality. The results from our study allow drawing recommendations to game designers. First, NPCs' appearances that are evaluated as more competent and immoral are also perceived as more threatening and thus, these NPCs would be expected to be potential enemies in the narrative. These personality traits can be used to predict NPCs' abilities and intentions. Second, game design would benefit from implementing players' evaluation of enemies' aggressiveness to convey an increasing challenge in the game experience. For instance, at the start of the game, the main character may encounter soldiers who are less accurate shooters and tend to stay in cover. Then, as the main character progresses toward the end of its mission, it may come across soldiers with greater shooting accuracy and enhanced mobility, demonstrated by a reduction in the distance between them. Third, designing incongruent NPCs can have interesting outcomes as it can create surprises in players' minds. For instance, designers could manipulate the appearance and behaviors of NPCs inside a mission to create ambiguity and challenge, without affecting players' evaluation of hostility. Hence, by considering the interplay between visual design, behavioral design, and players' evaluation, game designers can create enemies that evoke the desired emotional responses and enhance the overall gameplay experience.

10 FUTURE RESEARCH AND CONCLUSIONS

10.1. Limitations

The main limitations of this research, which open the road to future studies, touch on three areas: (i) participants, (ii) material, and (iii) measures. First, the participants completing the studies belong to the Ubisoft user research database and are known to be highly engaged players. To investigate whether the predictors of hostility differ between highly trained players and more naive ones, it would be interesting to conduct the experiment with groups of players with different levels of gaming experience. Moreover, as most of the participants were men, it would be important for a full generalization of the result to investigate the influence of gender on the evaluation of hostility in video games. Understanding the difference between genders in hostility evaluation could improve the design of more nuanced enemies. Second, our three experiments were conducted in the context of a military video game. It would be interesting to investigate the extent to which the obtained results generalize to other contexts in video games and beyond. For instance, enemies in a more cartoonish context might trigger different hostility perceptions, such as in Super Mario Odyssey, where enemies are non-humanoid agents. Additionally, this study focused on non-verbal behaviors, it will be important in future research to explore how verbal behaviors (such as verbal aggressiveness during conversation with NPCs) influence hostility evaluation, as this modality of communication has been shown to be influenced by players' expectations [38]. Third, in our studies, participants were watching videos of interactions with NPCs. Future research should investigate whether players observing versus directly engaging with NPCs perceive hostility differently. Moreover, this study used explicit scales of measure, which require participants to consciously rationalize their often unconscious perception. It would be interesting in future experiments to add indirect measures such as electrophysiological ones (e.g., heart rate, skin-conductance response) to investigate implicit evaluations of players' state in video games. This would allow comparison of hostility as explicitly perceived (through questionnaires) and implicitly felt (through electrophysiological recordings), a direction that is currently being explored in our team.

10.2. Future directions and conclusions

Our study paves the way for future research in contexts beyond military videogames. In particular, our findings reveal that players' evaluation of NPCs' hostility is more influenced by their behaviors than by their appearances. This result raises the fundamental theoretical question of the differences in the perceptual content formed in front of a virtual agent versus a real human. This is in particular the case regarding the influence of first impressions. Indeed, in a real environment, initial impressions heavily influence human-to-human evaluations and engagements. Individuals construct mental models of their surroundings, encompassing all perceived elements through various sensory channels. These mental models serve as specific cognitive structures facilitating the processing, organization, evaluation, and interaction with environmental stimuli. According to the spreading activation theory, these mental models are spatially organized, such that exposure to one concept activates closely associated, albeit unconsciously related concepts [39]. This mechanism likely underlies the formation of first impressions, where the mere presentation of someone's appearance triggers a cascade of associated ideas based on prior experiences, knowledge, or assumptions (see [40], for the influence of faces). This cognitive framework appears to extend to the video game environment proposed here: portraying an NPC as an armed soldier might suffice to evoke the perceived violence associated with the military world as humans conceptualize it. It remains to explore whether this phenomenon would extend to other types of virtual environments, other agents (including those that are non-realistic), and other modalities (e.g., different types of voice).

Going further, the organization of mental models is not strictly logical; rather, first impressions often arise from attentional biases and implicit stereotypes, leading to random and irrational connections between concepts. Research by Griffin and Langlois [41] demonstrated how individuals, both adults and children, exposed to unattractive female faces tend to arbitrarily associate them with negative traits. Overcoming these initial biases proves challenging, if not impossible, during human-to-human evaluations, as individuals tend to seek confirmation of their initial impressions even in the face of contradictory behavior. It is intriguing to note that in our study, participants were able to quickly discard stereotyped first impressions formed about NPCs, instead relying more on their observed behaviors when evaluating them. This finding is all the more surprising as virtual agents remain essentially the same throughout interactions, compared to the ever-changing nature of humans. One would expect initial impressions to be more firmly established regarding a pre-programmed virtual setting. Hence, it appears that observing NPCs' behavior, which serves as feedback on participants' initial impressions, is sufficient to alter biases and stereotypes. This finding confirms Bosse's assumption [9], that humans' appraisal of virtual agents' hostility focuses more on the outcomes of the interaction.

The adjustment of human mental models based on virtual agents' design factors opens opportunities for investigation in other contexts and other forms of interaction. Here, participants were only observing pre-recorded interactions between the main character and the NPCs. Which assumptions could we thus have on participants interacting in real-time with virtual agents? Dumont et al. [42] showed that proposing to interact with an intentionally stereotyped character modifies the judgments and biases usually observed in these studies. Following this idea, future research could involve participants directly interacting with these agents, thus examining how taking an active role influences their evaluation of the agents. Furthermore, Interactions could be studied through the lens of other communication modalities, in particular the verbal one. Conversational agents are flourishing, and their communication style can convey various intentions, including aggressiveness [38].

The methodology used in our study provides tools for future research in broader contexts. Beyond hostility perception, future studies could explore numerous other dimensions influenced by the first impressions humans form, such as warmth vs. competence [43], happiness, or trustworthiness [44]. These other dimensions might have different effects on users' affective experience as a function of the context. For instance, trustworthiness might be crucial in a medical context to ease the interactions between a user and a virtual therapist [45], [46], [47]. Future studies might also investigate other types of subsequent interactions humans engage in with virtual agents (such as cooperative vs. competitive). Finally, future research could investigate agents designed with other personality traits than those proposed by Bartneck (e.g., competence, knowledge, intelligence, responsibility, sensibility), such as compassion or care. To summarize, our methodology allows measuring the impact of virtual agents' design factors on users' evaluations. It provides resources for future investigations of a broad range of game contexts, virtual agents' personality traits, and types of interactions between users and virtual agents.

.

Acknowledgements

Remi Poivet is funded by the French ANRT and the Company Ubisoft.

References

[1] T. Bosse, C. Gerritsen, and J. de Man, "An intelligent system for aggression de-escalation training," in *ECAI'16*. IOS Press, Pasadena, 2016, pp. 1805-1811, doi: 10.3233/978-1-61499-672-9-1805.

[2] W.-D. Liu, K.-Y. Chuang, and K.-Y. Chen, "The Design and Implementation of a Chatbot's Character for Elderly Care," in *2018 Int. Conf. System Science. and Engineering (ICSSE),* New Taipei, Jun. 2018, pp. 1–5, doi: 10.1109/ICSSE.2018.8520008.

[3] J. C. Barefoot, "Keeping Conflicting Findings in Perspective: The Case of Hostility and Health," *Mayo Clinic Proc*., vol. 68, no. 2, pp. 192–193, Feb. 1993, doi: 10.1016/S0025-6196(12)60170-0.

[4] J. C. Barefoot and I. M. Lipkus, "The assessment of anger and hostility," in *A. W. Siegman & T. W. Smith (Eds.), Anger, hostility, and the heart*, Lawrence Erlbaum Associates, Inc, 1994, pp. 43–66, doi: 10.1037/0278-6133.15.3.200.

[5] W. G. Stephan and C. W. Stephan, "Intergroup Anxiety," *J. of Social Issues*, vol 43, no. 3, pp. 157-175, 1985, doi: 10.1177/1088868314530518.

[6] B. J. Bushman and C. A. Anderson, "Is it time to pull the plug on the hostile versus instrumental aggression dichotomy?," *Psychol. Rev*., vol 108, no. 1, pp. 273–279, 2001, doi: 10.1073/pnas.1713611115.

[7] J. C. Barefoot, "Developments in the measurement of hostility," in *Hostility, coping, & health*., H. S. Friedman, Ed. Washington: American Psychological Association, 1992, pp. 13–31, doi: 10.1037/10105-001.

[8] R. Blankendaal, T. Bosse, C. Gerritsen, T. de Jong, and J. de Man, "Are Aggressive Agents as Scary as Aggressive Humans?," in *Proc. 2015 Int. Conf. Autonomous Agents and Multiagent Systems*, 2015, pp. 553–561, doi: 10.5555/3237383.3237885.

[9] T. Bosse, T. Hartmann, R. A. M. Blankendaal, N. Dokter, M. Otte, et L. Goedschalk, "Virtually Bad: A Study on Virtual Agents that Physically Threaten Human Beings," in *Proc. 17th Int. Conf. Autonomous Agents and MultiAgent Syst*., Richland, SC, Jul. 2018, p. 1258-1266, doi: 10.5555/3237383.3237885.

[10] K. Rogers, M. Aufheimer, M. Weber, and L. Nacke, "Towards the Visual Design of Non-Player Characters for Narrative Roles," in *Graphics Interface 2018*, 2018, pp. 154–161, doi: 10.20380/GI2018.21.

[11] J. A. Reed and E. M. Blunk, "The influence of facial hair on impression formation," *Social Behav. and Personality: an Int. J.*, vol. 18, no. 1, pp. 169–175, Jan. 1990, doi: 10.2224/sbp.1990.18.1.169.

[12] M. Haake and A. Gulz, "Visual Stereotypes and Virtual Pedagogical Agents," *Educational Technology and Society*, vol. 11, no. 4, pp. 1–15, 2008, doi: 10.5555/1517338.1517341.

[13] Y. Ferstl, M. McKay, and R. McDonnell, "Facial Feature Manipulation for Trait Portrayal in Realistic and Cartoon-Rendered Characters," *ACM Trans. Appl. Percept*., vol. 18, no. 4, pp. 1–8, Oct. 2021, doi: 10.1145/3486579.

[14] R. Giusti, K. Hullett, and J. Whitehead, "Weapon design patterns in shooter games," in *Proc. 1st Workshop on Des. Patterns in Games*, 2012, pp. 1-7, doi: 10.1145/2427116.2427119.

[15] C. Becker, S. Kopp, and I. Wachsmuth, "Simulating the Emotion Dynamics of a Multimodal Conversational Agent," in *Tut. and Res. Workshop on Affect. Dialogue Syst.*, in Affect. Dialogue Syst., vol. 3068, E. André, L. Dybkjær, W. Minker, and P. Heisterkamp, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 154–165, doi: 10.1007/978-3-540-24842-2_15.

[16] K. Ruhland et al., "A Review of Eye Gaze in Virtual Agents, Social Robotics and HCI: Behaviour Generation, User Interaction and Perception: A Review of Eye Gaze," *Comput. Graphics Forum*, vol. 34, no. 6, pp. 299–326, Sep. 2015, doi: 10.1111/cgf.12603.

[17] N. Sadoughi and C. Busso, "Speech-driven animation with meaningful behaviors," *Speech Commun.*, vol. 110, pp. 90–100, Jul. 2019, doi: 10.1016/j.specom.2019.04.005.

[18] A. B. Loyall, "Believable Agents: Building Interactive Personalities," Ph.D. dissertation, Carnegie Mellon University, 1997.

[19] K. Kim, L. Boelling, S. Haesler, J. Bailenson, G. Bruder, and G. F. Welch, "Does a Digital Assistant Need a Body? The Influence of Visual Embodiment and Social Behavior on the

Perception of Intelligent Virtual Agents in AR," in *2018 IEEE Int. Symp. Mixed and Augmented Reality (ISMAR),* Munich, Germany, Oct. 2018, pp. 105–114, doi: 10.1109/ISMAR.2018.00039.

[20] T. L. Robbins and A. S. DeNisi, "A closer look at interpersonal affect as a distinct influence on cognitive processing in performance evaluations," *J. of Applied Psychol.*, vol. 79, no. 3, pp. 341–353, Jun. 1994, doi: 10.1037/0021-9010.79.3.341.

[21] B. Reeves and C. Nass, *The media equation: how people treat computers, television, and new media like real people and places Choice Reviews*. Cambridge, UK, 1996.

[22] R. Pradantyo, M. V. Birk, and S. Bateman, "How the Visual Design of Video Game Antagonists Affects Perception of Morality," *Frontiers in Computer Science*, vol. 3, 2021, doi: 10.3389/fcomp.2021.531713.

[23] H. Warpefelt. "Cues and insinuations: Indicating affordances of non-player character using visual indicators," in *Proc. DiGRA 2015: Diversity of play: Games - Cultures – Identities*, May. 2015. [Online]. Available: http://www.digra.org/digital-library/publications/cues-and-insinuations-indicating-affordances-of-non-player-character-using-visual-indicators/.

[24] M. Lee and C. Heeter, "What do you mean by believable characters?: The effect of character rating and hostility on the perception of character believability*," J. of Gaming and Virtual Worlds*, vol 4, no. 1, pp. 81-97, 2012, doi: 10.1386/jgvw.4.1.81_1.

[25] A. A. Long and N. Sedley, *The Hellenistic Philosophers: Volume 2, Greek and Latin Texts with Notes and Bibliography*. Cambridge University Press, 1987.

[26] A. Eden, M. B. Oliver, R. Tamborini, A. Limperos, and J. Woolley, "Perceptions of Moral Violations and Personality Traits Among Heroes and Villains," *Mass Commun. and Soc.*, vol. 18, no. 2, pp. 186–208, Mar. 2015, doi: 10.1080/15205436.2014.923462.

[27] T. Manninen and T. Kujanpää, "The value of virtual assets: The role of game characters in MMOGs," *Int. J. of Bus. Sci & Appl. Manage*., vol. 2, no. 1, pp. 21–33, 2007.

[28] P. Vorderer, T. Hartmann, and C. Klimmt, "Explaining the enjoyment of playing video games: the role of competition," in *Proc. 2nd Int. Conf. Entertainment computing*, 2003, pp. 1–9, doi: 10.1177/0093650209356394.

[29] M. Grizzard et al.,"Validating the extended character morality questionnaire," *Media Psychology*, vol. 23, no. 1, pp. 107–130, Jan. 2020, doi: 10.1080/15213269.2019.1572523.

[30] J. Haidt and C. Joseph, "The Moral Mind: How Five Sets of Innate Intuitions Guide the Development of Many Culture-Specific Virtues, and Perhaps Even Modules" in *Evolution and Cognition. The Innate Mind, Volume 3*, 1st ed., P. Carruthers and S. Laurence, Eds. Oxford University Press New York, 2008, pp. 367–392, doi: 10.1093/acprof:oso/9780195332834.003.0019.

[31] J. L. Monahan, "I Don't Know It But I Like You: The Influence of Non-conscious Affect on Person Perception," *Human Commun. Res*., vol. 24, no. 4, pp. 480–500, Jun. 1998, doi: 10.1111/j.1468-2958.1998.tb00428.x.

[32] C. Bartneck, T. Kanda, H. Ishiguro, and N. Hagita, "My robotic doppelgänger - a critical look at the Uncanny Valley," in *RO-MAN 2009 – The 18th IEEE Int. Symp. Robot and Human Commun.*, 2009, p. 269–276, doi: 10.1109/ROMAN.2009.5326351.

[33] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots," *Int. J. of Social Robot*., vol. 1, no. 1, pp. 71–81, 2009, doi: 10.1007/s12369-008-0001-3.

[34] T. Koda and P. Maes, "Agents with faces: the effect of personification," in *RO-MAN'96 – The 5th IEEE Int. Workshop Robot and Human Commun*., 1996, pp. 189–194, doi: 10.1109/ROMAN.1996.568812.

[35] C. Butcher and J. Griesemer, "The illusion of intelligence: The integration of AI and level design in Halo," in *Game Developers Conference*, San Jose, CA, March , 2002.

[36] T. Bancroft, *Creating characters with personality*. New York: Watson-Guptill, 2006.

[37] R. Gifford, "The experience of personal space: Perception of interpersonal distance," *J. Nonverbal Behav*., vol. 7, no. 3, pp. 170–178, Mar. 1983, doi: 10.1007/BF00986947.

[38] R. Poivet, M. L. Lopez-Malet, C. Pelachaud, and M. Auvray, "The influence of conversational agents' role and communication style on user experience," *Frontiers in Psychol.*, vol 14, 2023, doi: 10.3389/fpsyg.2023.1266186.

[39] A. M. Collins and E. F. Loftus, "A spreading-activation theory of semantic processing," *Psychol. Rev*., vol. 82, no. 6, pp. 407-428, 1975, doi: 10.1037/0033-295X.82.6.407.

[40] C. Y. Olivola, F. Funk and A. Todorov, "Social attributions from faces bias human choices," *Trends in Cogn. Sci.*, vol. 18, no. 11, pp. 566-570, 2014, doi: 10.1016/j.tics.2014.09.007.

[41] A. M. Griffin and J. H. Langlois, "Stereotype directionality and attractiveness stereotyping: is beauty good or is ugly bad?," *Soc. Cogn.*, vol. 24, no. 2, pp. 187-206, 2006, doi: 10.1521/soco.2006.24.2.187.

[42] M. Dumont et al., "Suppression and hypothesis testing: does suppressing stereotypes during interactions help to avoid confirmation biases?," *Eur. J. of Social Psychol.*, vol. 33, no. 5, pp. 659-677, 2003, doi: 10.1002/ejsp.177.

[43] S. Fiske, "Stereotype content: warmth and competence endure," *Current Directions in Psychol. Sci.*, vol. 27, no. 2, pp. 67-73, 2018, doi: 10.1177/0963721417738825.

[44] N. Oosterhof and A. Todorov, "The functional basis of face evaluation," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 32, 2008, doi: 10.1073/pnas.0805664105.

[45] M. Auriacombe et al., "Development and validation of a virtual agent to screen tobacco and alcohol use disorders," *Drug and alcohol dependence*, vol. 193, pp. 1-6, 2018, doi: 0.1016/j.drugalcdep.2018.08.025.

[46] M. de Gennaro, E. G. Krumhuber and G. Lucas, "Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood," *Frontiers in Psychol.*, vol. 10, 2020, doi: 10.3389/fpsyg.2019.03061.

[47] D. S. Choi, J. Park, M. Loeser and K. Seo, "Improving counseling effectiveness with virtual counselors through nonverbal compassion involving eye contact, facial mimicry, and head-nodding," *Scientific Rep.*, vol. 14, no. 506, 2024, doi: 10.1038/s41598-023-51115-y.

Remi Poivet is a PhD candidate at ISIR, Sorbonne University. His research interests include human-autonomous agent interaction. In addition to his academic work, he is a cognitive science researcher at Ubisoft Paris, where he is focused on developing innovative game design that consider players' perception.

Alexandra de Lagarde is a research engineer at ISIR, Sorbonne University. She received her MD in Cognitive Science from Cogmaster (ENS-PSL, Paris, France), in 2022. Her research interests include multisensory perception and interactions, as well as the links between social and spatial perspective taking.

Catherine Pelachaud is CNRS Director of Research at ISIR, Sorbonne University. She received her Ph.D. in CS from University of Pennsylvania in 1991. Her research interest includes socially interactive agent, nonverbal communication (face, gaze, gesture and touch), and social interaction. With her research team, she has been developing an interactive virtual agent platform, Greta, that can display emotional and communicative behaviors. She is co-editor of the ACM handbook on socially interactive agents.

Malika Auvray is CNRS Director of Research at ISIR, Sorbonne University. She received her Ph.D. in Cognitive Sciences from EHESS, Paris in 2004. Her research interests include spatial and social cognition, multisensory perception and augmented cognition. She published more than 50 articles and book chapters and gave more than 300 communications on the topic. With her research team, she currently investigates the links between spatial and social perspective taking and ways to augment cognition by means of multisensory conversion.