
Artificial Empathy in Embodied AI: Pain-Aware Perception as a New Modality for Interaction

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 Prolonged exposure to violent video games has been linked to increased aggression
2 and desensitization, raising concerns about their psychological impact. While ex-
3 isting anti-addiction systems focus on gameplay duration and identity verification,
4 few address players' real-time responses to violent content. We introduce Pain-
5 Mouse, a multimodal embodied AI system that detects desensitization to in-game
6 violence and delivers pain-based haptic feedback to promote awareness and self-
7 regulation. Using a dual-path detection framework, the software module combines
8 and compares visual, audio, and behavioral data from the game and the player.
9 When emotional detachment is detected, a custom-built haptic mouse delivers
10 proportional feedback via electrical stimulation or mechanical impact. Drawing
11 from associative learning theories, our system explores whether repeated pairing of
12 violent actions with physical discomfort can reshape player sensitivity. This work
13 offers an AI-powered, embodied, behavior-sensitive intervention mechanism that
14 complements existing digital well-being frameworks.

15 1 Introduction

16 In recent years, the widespread presence of violent and gory elements in video games has raised
17 growing concerns about their impact on adolescent mental health. Numerous studies have shown
18 that prolonged exposure to such content can lead to increased aggressive cognition and behavior, as
19 well as desensitization toward real-world violence[Anderson & Bushman, 2001, Gentile et al., 2004].
20 Although many countries have implemented anti-addiction policies—such as real-name registration,
21 curfews, and spending caps—these mechanisms primarily focus on regulating gameplay duration
22 and verifying user identity[Xiao & Leon Y, 2020, Yang et al., 2023], while largely overlooking the
23 psychological and behavioral impact of violent game content itself.

24 Recent systems have incorporated artificial intelligence for user-level monitoring, including facial
25 recognition for age verification and behavioral analysis for addiction detection[Ferreri et al., 2018].
26 However, these approaches target surface-level usage patterns and demographic profiling, with
27 limited attention to the violent actions occurring within games, the player's underlying intent, or the
28 emotional responses these actions provoke. Moreover, most existing systems lack real-time feedback
29 mechanisms to interrupt problematic behavior as it occurs. In contrast, pain-based haptic feedback
30 offers a direct sensory channel to intervene in the moment of violent gameplay. By introducing
31 physical discomfort during violent acts, such feedback can disrupt the pleasure-reward cycle and
32 heighten self-awareness. Drawing from Neuro-Associative Conditioning (NAC) theory—which
33 posits that behavior is shaped by associations between pleasure and pain—we aim to establish a
34 negative link between violent actions and pain, helping players develop a long-term awareness of
35 in-game violence.

36 In this paper, we present PainMouse, a novel multimodal embodied AI system designed to detect
37 desensitization to in-game violence and intervene through pain-based haptic feedback. Unlike
38 prior approaches that focus on performance analytics or usage duration, our system targets the
39 player’s real-time psychological response to violent content. The software component of PainMouse
40 continuously analyzes visual and audio content from the game, as well as the player’s synchronized
41 reactions—including facial expressions, voice tone, and gameplay behavior (e.g., aiming or shooting
42 frequency). A dual-path detection architecture is used: one real-time module enables low-latency
43 feedback based on fast multimodal fusion, while a second, context-aware agent module integrates
44 longer temporal windows to identify subtler behavioral patterns. When the system detects a potential
45 mismatch between violent stimuli and emotional response, indicative of desensitization, it triggers
46 physical feedback through a custom-built haptic device. The hardware module, a modified computer
47 mouse, can deliver different forms of discomfort, such as electrical shocks or mechanical slaps,
48 mapped proportionally to the system’s violence assessment. Drawing from neuro-associative learning
49 theories, we explore whether repeated pairing of violent actions with mild but tangible pain can
50 help reshape the player’s long-term sensitivity and awareness toward digital violence. Rather than
51 punishing players for gameplay itself, our goal is to build an embodied intervention system that
52 sensitizes users to their own behavior and complements existing anti-addiction policies with real-time
53 psychological feedback.

54 **2 Related Work**

55 **2.1 Mental Health and Conditioning**

56 Concerns about the negative impact of video games on the mental health of children and adolescents
57 have grown in recent years. Violent video games, in particular, have been found to motivate players to
58 engage in aggressive behaviors by rewarding them for acquiring violent skills and causing destruction
59 [Gentile & Gentile, 2007]. These games have also been linked to increased incidents of school bullying
60 and reduced social satisfaction [Shoshani, Braverman, & Meiorow, 2021], as well as heightened
61 aggression and diminished prosocial behaviors [Greitemeyer, 2022]. Behavioral psychology suggests
62 that positive reinforcement increases the likelihood of a behavior, while negative consequences
63 reduce it [Akpan, 2020]. In line with this, learning has been conceptualized as the acquisition of new
64 behaviors through mechanisms of conditioning [Säljö, 2021]. Moreover, the influence of evaluative
65 conditioning can persist even when individuals consciously attempt to suppress or control its effects
66 [Moran, Nudler, & Bar-Anan, 2023].

67 In response to concerns over excessive gaming, recent studies have begun to explore the use of
68 artificial intelligence (AI) as a means of preventing gaming addiction. AI-powered systems are being
69 developed to monitor players’ gaming patterns—such as play duration and frequency—in order to
70 identify early signs of addictive behaviors and issue timely prompts or alerts to encourage healthier
71 gaming habits [Zhang, Liu, & Chen, 2024]. In therapeutic settings, AI-based predictive models have
72 been employed to forecast the risk of relapse in behavioral addictions, including gaming disorder. By
73 tracking digital usage, these systems can notify both patients and healthcare professionals of potential
74 triggers, thereby facilitating earlier interventions and improved treatment outcomes [Johnson & Patel,
75 2024].

76 These findings collectively underscore the strong relationship between mental health and behavioral
77 cognition. While the application of AI in this domain shows promising potential, current interventions
78 remain largely limited to passive monitoring and warning mechanisms. There is a notable lack of
79 tangible, interactive strategies that actively guide users toward behavioral change. Future research
80 should focus on developing more embodied and experiential interventions that not only detect risks
81 but also meaningfully influence behavior in real time.

82 **2.2 Pain as Modality**

83 Pain is a continuous and multifaceted phenomenon, encompassing both sensory and emotional
84 experiences [Temel, Erden, & Bağcıer, 2025]. Accurately perceiving pain in others requires complex
85 interpretation and inference processes [Sharkey, 2024]. Traditionally, pain detection has relied on
86 unimodal approaches, using singular features such as facial expressions, language, or body posture
87 cues [Werner et al., 2018; Voytovich & Greenberg, 2021; Waller, Julle-Daniere, & Micheletta,
88 2020]. Recent advances have introduced multimodal methods that combine these external signals

89 with neurophysiological data, including electroencephalography (EEG), electromyography (EMG),
90 electrocardiography (ECG), respiration (RSP), and electrodermal activity (EDA) [Cao et al., 2024;
91 Naeini et al., 2020; Thiam & Schwenker, 2017].

92 Machine learning techniques—ranging from decision trees and Bayesian classifiers to neural net-
93 works—have been widely used in these applications [Lötsch & Ultsch, 2018]. Most existing work
94 is situated in medical or therapeutic contexts, such as postoperative pain monitoring or AI-assisted
95 mental healthcare. For example, facial EMG, ECG, and EDA signals have proven useful in clinical
96 settings [Naeini et al., 2020], while natural language processing (NLP) has shown potential in AI-
97 driven cognitive behavioral therapy for reducing anxiety and depression [Cao et al., 2024; Fitzpatrick,
98 Darcy, & Vierhile, 2017].

99 However, these studies are predominantly passive and observational, often relying on static video,
100 voice, or neurophysiological data collected in controlled experimental scenarios. There remains a
101 significant gap in research on pain as an active, embodied modality. That is, rather than treating
102 pain merely as an object to be inferred from external signs, few studies explore how pain itself—its
103 subjective and internal dynamics—can be directly modeled, represented, or communicated. There
104 is also a lack of tangible, interactive systems that engage with pain as a communicable modality in
105 real-world, socially embedded contexts.

106 2.3 AI-Driven Haptic Interaction

107 Recent research at the intersection of artificial intelligence (AI) and haptics has revealed promising
108 directions for enhancing user experience in virtual and augmented reality (VR/AR) environments.
109 Several studies have explored how AI-driven haptic feedback can contribute to immersive interaction,
110 improved decision-making, and enhanced human-AI alignment. For instance, real-time analysis
111 of user behavior by AI algorithms enables the generation of dynamic haptic feedback, heightening
112 the sense of realism and engagement in virtual settings. One emerging direction involves haptic
113 repurposing, where generative AI transforms physical objects into adaptive haptic interfaces for
114 interacting with AI-generated virtual assets, thereby extending the tactile landscape of mixed reality
115 applications [Zhang et al., 2022; Fedoseev et al., 2022].

116 Other studies emphasize haptics as a critical channel for achieving perceptual alignment between
117 humans and AI. Multimodal feedback—especially the integration of visual and haptic stimuli—has
118 been shown to improve user performance, satisfaction, and mutual understanding in collaborative
119 tasks [Wu et al., 2025]. Research has also highlighted the role of haptic feedback in supporting shared
120 embodiment, co-presence, and body ownership in avatar-based virtual environments, demonstrating
121 its importance in fostering social presence and embodied interaction [Venkatraj et al., 2023]. Addi-
122 tionally, muscle-based input systems such as EMG-driven force interfaces offer a more natural and
123 intuitive way for users to physically engage with virtual objects, further underscoring the relevance of
124 internal bodily signals in human-computer interaction.

125 These developments illustrate the growing recognition of haptics not just as an output modality, but as
126 an essential component of embodied, context-aware AI interaction. Haptic feedback enriches sensory
127 realism, strengthens user engagement, and facilitates bidirectional, affective experiences. However,
128 most existing work remains focused on delivering haptic information to the user. There is a notable
129 lack of research on closed-loop systems that interpret haptic feedback from the user as a cognitive
130 signal. That is, while touch is increasingly used as a medium for AI to communicate with humans,
131 few systems leverage tactile input as a meaningful expression of internal states—such as uncertainty,
132 discomfort, or attention. Future research may benefit from reframing haptics not only as a channel for
133 output, but also as a perceptual modality through which user cognition and affect can be dynamically
134 sensed and interpreted.

135 In conclusion, current research has explored mental health, pain detection, and haptic feedback. How-
136 ever, most approaches rely on passive observation or focus on output-only feedback, without engaging
137 pain as an active, communicable signal. Notably, the idea of haptic pain as a new modality—using
138 tactile expression of pain to bridge behavior and cognition—has yet to be fully explored. This presents
139 a valuable opportunity: by leveraging haptic pain as an interactive and embodied channel, future
140 research could support behavioral understanding and emotional communication, offering significant
141 potential for improving children’s mental health.

142 **3 Problem Statement**

143 This research investigates whether controlled, repeated pain-based haptic feedback can modulate
 144 players’ cognitive and bodily responses to violent gameplay. Specifically, we aim to detect and
 145 address potential desensitization—situations where a player fails to exhibit appropriate emotional or
 146 behavioral responses to violent or graphic in-game content. To this end, we propose a multimodal
 147 system that evaluates both the *external stimulus* and the *internal response* of the player in real time.

148 At each time step t , the input to our system is a tuple:

$$\mathbf{x}^t = \{\mathbf{v}^t, \mathbf{a}^t, \mathbf{b}^t\}$$

149 where $\mathbf{v}^t \in \mathbb{R}^{H \times W \times 3}$ is a gameplay video frame, $\mathbf{a}^t \in \mathbb{R}^L$ is the audio segment (player speech or
 150 game sound), and $\mathbf{b}^t \in \mathbb{R}^K$ is a behavioral feature vector (e.g., gaze fixation rate, mouse tapping).

151 We define two scalar functions:

$$\begin{aligned} s_v^t &= w_1 \cdot \text{CLIP}(\mathbf{v}^t) + w_2 \cdot \text{CLAP}(\mathbf{a}^t) \\ s_r^t &= w_3 \cdot \text{AU}(\mathbf{v}^t) + w_4 \cdot \text{SER}(\mathbf{a}^t) + w_5 \cdot \text{Behavior}(\mathbf{b}^t) \end{aligned}$$

152 where $s_v^t \in [0, 4]$ represents the external violence intensity, and $s_r^t \in [0, 4]$ quantifies the player’s
 153 internal response level. The weights $\{w_i\}$ are manually tuned for decision-level fusion.

154 We define the desensitization gap as:

$$\Delta^t = s_v^t - s_r^t$$

155 Large values of Δ^t indicate a lack of emotional engagement in response to intense stimuli, which
 156 may trigger pain-based haptic feedback. Our objective is not to train a traditional supervised model
 157 (i.e., no ground-truth y^t), but to use this gap to study behavioral change over time through repeated
 158 exposure and feedback. The pain feedback is delivered through a hardware device (PainMouse),
 159 which activates when Δ^t exceeds a predefined threshold:

$$\delta(\Delta^t) = \begin{cases} 0, & \text{if } \Delta^t \leq \tau \\ \min(4, \text{round}(\Delta^t)), & \text{otherwise} \end{cases}$$

160 By measuring and modulating the gap Δ^t during gameplay, we aim to evaluate whether negative
 161 reinforcement leads to re-sensitization or altered engagement with violent actions across repeated
 162 sessions.

163 **4 Proposed Approach**

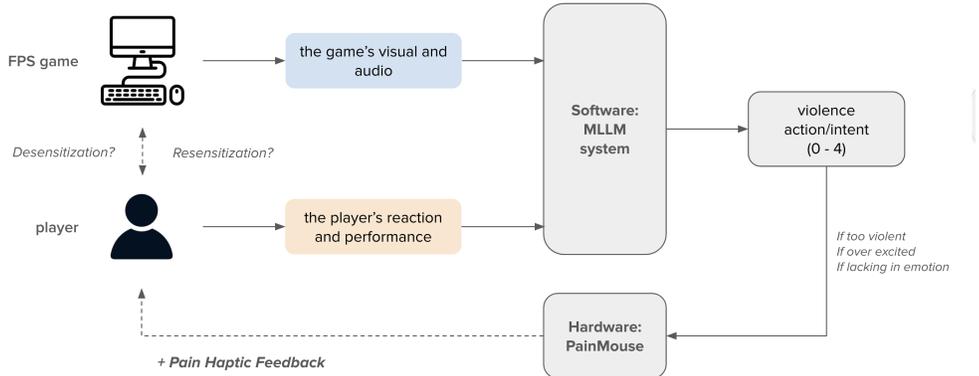


Figure 1: System overview. The PainMouse system consists of two main components: a multimodal violence detection software and a pain-based haptic feedback device.

164 We propose PainMouse, a multimodal embodied AI system designed to evaluate and intervene in
 165 players’ cognitive responses to violent gameplay. As shown in Figure 1, the system consists of two

166 primary components: 1) a software module that performs real-time multimodal analysis of both the
 167 game and the player’s reactions, and 2) a hardware module that delivers pain-based haptic feedback.
 168 Together, they form a feedback loop aimed at reinforcing the player’s sensitivity to violent content
 169 through aversive conditioning.

170 **1) Software: Dual-Path Multimodal Detection System for Violence Detection and User State**
 171 **Evaluation.**

172 The software component processes both the visual and auditory content of the game as well as the
 173 player’s synchronized emotional and behavioral signals. Here we propose two detection mechanisms,
 174 one for real-time monitoring, the other is agent-based. As mentioned in the previous section, at
 175 each time step t , the system receives an input tuple $x(t) = \{v(t), a(t), b(t)\}$, where $v(t)$ is the
 176 current RGB video frame from the game, $a(t)$ is an audio segment, and $b(t)$ is a vector of structured
 177 behavioral features such as gaze fixation rate and shooting frequency. In real-time detection, these
 178 modalities are analyzed using pre-trained foundation models: CLIP for visual violence, CLAP and
 179 SER for audio understanding, and OpenFace for facial expression analysis. The two scalar scores,
 180 $s(v, t)$ and $s(r, t)$, are computed using a manually weighted decision-level fusion strategy at the
 181 moment. The desensitization gap $\Delta_t = s(v, t) - s(r, t)$ captures the degree to which a player may
 182 be underreacting to violent content. A high value of Δ_t suggests potential desensitization and is
 183 used to trigger physical feedback. This detection pathway operates at a high temporal resolution (1s),
 184 allowing for low-latency intervention through the hardware interface.

185 To further enhance accuracy and flexibility, we introduce a Gemini-based fusion mode as an advanced
 186 extension of our agent-based system. In this approach, every 10 seconds, the system synthesizes
 187 three modalities—gameplay visual context, audio understanding, and facial expressions—into textual
 188 descriptions via Gemini-2.0-Flash. These textual summaries, combined with numeric interaction
 189 scores derived from player mouse interactions, form a comprehensive multimodal prompt.

190 Gemini dynamically adjusts modality weights based on contextual importance inferred from real-time
 191 conditions. For example, if a player consistently exhibits minimal facial expression, Gemini assigns
 192 greater emphasis to the visual and audio inputs to accurately estimate violence intensity. The resulting
 193 violence intensity score, produced as JSON-formatted output, ranges from 0.0 (peaceful) to 4.0
 194 (extreme violence), explicitly structured as:

195 `{"violence_score": <float>}`

196 Compared to the purely real-time fusion mode, the Gemini-based approach offers notable advantages:

- 197 • **Dynamic Contextual Weighting:** Automatically and adaptively adjusts modality weights
 198 according to real-time contextual relevance, significantly enhancing prediction accuracy.
- 199 • **Enhanced Multimodal Reasoning:** Capitalizes on Gemini’s advanced multimodal language
 200 understanding capabilities, enabling more nuanced and contextually accurate assessments.
- 201 • **Robustness and Interpretability:** Offers improved robustness and clearer interpretabil-
 202 ity through textual modality summaries, supporting easier diagnostics and validation of
 203 inference logic.

204 Formally, our multimodal fusion at timestep t is defined as:

$$x_t = \{I_{\text{screen}}, I_{\text{face}}, A_{\text{audio}}, s_{\text{interaction}}\} \tag{1}$$

205 Gemini serves as the multimodal fusion function f_θ :

$$y_t = f_\theta(x_t) \in [0, 4] \tag{2}$$

206 Every five timesteps, we compute an aggregated wave-level violence score \bar{y}_w :

$$\bar{y}_w = \frac{1}{5} \sum_{i=1}^5 y_{t_i} \tag{3}$$

207 This Gemini-based fusion mode enhances our system’s ability to flexibly handle ambiguous or noisy
 208 multimodal inputs, providing a robust, context-aware, and interpretable violence intensity evaluation.

209 **2) Painmouse: A Haptic Device for Negative Reinforcement.**

210 To deliver physical feedback, we designed the **PainMouse**, a haptic interface embedded with mul-
 211 tiple pain-inducing mechanisms. It supports both low-latency *electrical stimulation* and delayed
 212 *mechanical impact* (e.g., rubber band lash). The feedback intensity is modulated as a function of Δ^t ,
 213 using:

$$\delta(\Delta^t) = \begin{cases} 0 & \text{if } \Delta^t \leq \tau \\ \min(4, \text{round}(\Delta^t)) & \text{otherwise} \end{cases}$$

214 Pain signals are injected into interaction flows (e.g., mouse clicks and mouse movement), allowing
 215 players to develop embodied associations between specific violent behaviors and discomfort. Pain
 216 feedback can be delivered in several ways: (1) immediately following a detected violent act, (2)
 217 embedded within subsequent user inputs (e.g., next few mouse clicks), or (3) as a patterned sequence
 218 that distributes pain across multiple actions. These variations are intended to create associations
 219 between complex behavioral patterns and pain, moving beyond simple stimulus-response conditioning.
 220 Combining real-time reactivity and longer-term contextual awareness, the dual-path design allows
 221 us to explore different modes of behavioral conditioning and their effects on sensitivity to virtual
 222 violence.

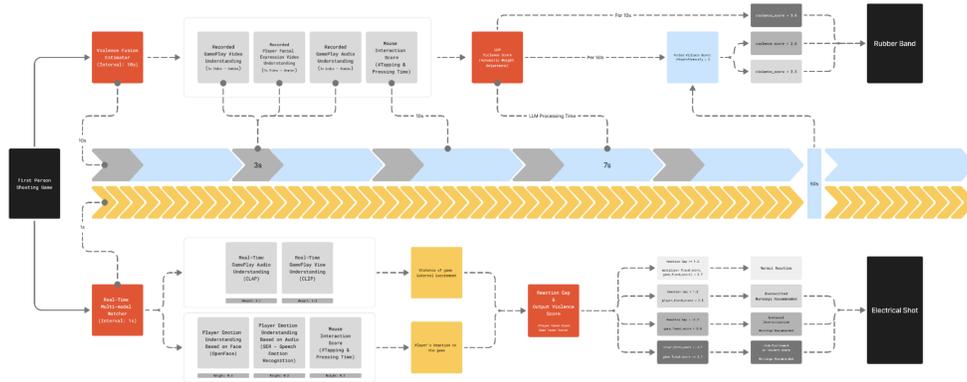


Figure 2: Dual-path detection architecture. The system includes two parallel detection paths: a real-time monitoring module and a delayed agent module. These two paths respectively trigger fast-response (e.g., electrical) and delayed (e.g., mechanical) pain feedback via the PainMouse device.

223 By combining low-latency detection with delayed contextual evaluation, our system supports both
 224 reactive and reflective feedback (see Figure 2). This dual-path structure enables us to investigate the
 225 different psychological effects of fast vs. delayed punishment, and to explore how such feedback
 226 influences long-term sensitivity toward digital violence.

227 **5 Experimental Methodology**

228 Our evaluation leverages real-time generated datasets captured from live gameplay sessions in *Call of*
 229 *Duty Black Ops 2: Zombies*. The dataset comprises four distinct input modalities:

- 230 • **Gameplay Visuals:** Static frame captured every 1 second for real-time analysis; continuous
 231 10-second screen recordings used for agent-based inference, using the `mss` library.
- 232 • **Player Facial Expressions:** Webcam snapshots captured every 1 second for real-time
 233 emotion analysis; sequences aggregated over 10 seconds for longer-term interpretation,
 234 processed via `MediaPipe`.
- 235 • **Audio Inputs:** 1-second voice and environmental audio segments for real-time speech
 236 emotion recognition; 3-second audio windows analyzed in batch for agent-level inference.

- **Player Interaction Metrics:** Mouse click rate recorded per second for real-time behavioral feedback; summarized over 10-second windows to inform delayed context-aware analysis.

239 **Complementarity of Modalities**

240 An essential characteristic of our multimodal dataset is the complementary nature of its modalities.
241 Each modality alone may not sufficiently capture the complexity and nuance of gameplay violence.
242 For instance, some players naturally exhibit fewer emotional expressions, making visual and audio
243 modalities critical. Likewise, ambiguous visual scenes or quiet audio segments necessitate relying on
244 facial expressions and interaction scores for accurate assessment. Thus, integrating diverse modalities
245 ensures comprehensive violence intensity estimation.

246 **Experimental Protocol**

247 To evaluate the potential conditioning effects of pain-based feedback, we will conduct a three-phase
248 user study with repeated gameplay sessions. Participants will be stratified based on their self-reported
249 exposure to violent video games (beginner, intermediate, advanced) to analyze how prior experience
250 influences receptivity to pain-based feedback.

251 Each participant will complete three rounds:

- 252 • **Round 1 (Baseline):** Participants play the game three times using a standard mouse. No
253 pain feedback is provided. Gameplay and physiological signals (e.g., skin conductance
254 level) are recorded.
- 255 • **Round 2 (Intervention):** Participants repeat the three games using the Painmouse (See
256 Figure 1) with active AI detection and pain feedback. The AI model infers \hat{y}^t from visual
257 inputs and triggers feedback via $\delta(\hat{y}^t)$.
- 258 • **Round 3 (Post-test):** Participants return to using a standard mouse, with no pain stimulus.
259 Data is recorded to evaluate any residual behavioral or physiological changes.

260 We will compare violence scores s^t , frequency and intensity of in-game aggressive actions, and
261 skin conductance responses across all three rounds. We hypothesize that participants will exhibit
262 increased awareness and emotional sensitivity to violent content in Round 3 compared to Round
263 1, reflecting conditioning effects from Round 2. We also expect stratification by experience level
264 to reveal differential susceptibility to this conditioning. Subjective feedback will be gathered via
265 post-study surveys.

266 This methodology allows us to investigate the temporal and embodied effects of negative reinforce-
267 ment, using both algorithmic predictions and physiological observations to evaluate our system.



Figure 3: Left: Painmouse device; Right: Experimental setup.

268 **6 Results and Discussion**

269 For each user, our system records the image-based violence score, facial expression score, mouse
270 interaction score, overall game violence level, and player reaction level. We also collected user
271 questionnaire data. In the first session (with a regular mouse), users answered questions about
272 game-play experience, perceived violence, and FPS familiarity. In the second session (with the
273 PainMouse), we added specific questions about physical pain feedback.

274 In Figure 4, we visualize how the system delivers pain-based feedback. Rubber band whipping occurs
 275 roughly every 50 seconds. Based on the intensity of game violence and player reaction, the number of
 276 whips ranges from one to three. Additionally, when both the game violence level and player reaction
 277 level exceed the warning threshold (set at 2.7), the system triggers a single instance of electric shock.
 278 For the player in Figure 4, six electric shocks were triggered. In most cases, we observe a positive
 279 correlation between the game violence level and the player's reaction level, suggesting the system
 280 accurately detects synchronized peaks of aggression and emotional arousal.



Figure 4: Violence Scene vs Player Reaction Over Time

281 When comparing two users with different levels of FPS gaming experience, as shown in Figure 5,
 282 an interesting contrast emerges. Even though both players were exposed to similar game violence
 283 levels, the user with extensive game-play experience displayed a significantly lower reaction level
 284 than the novice. However, contrary to our initial expectations, the game violence level itself was not
 285 significantly different between the two users. One hypothesis is that the game difficulty was relatively
 286 low, preventing the experienced player from expressing more complex or aggressive behaviors
 287 compared to the novice.

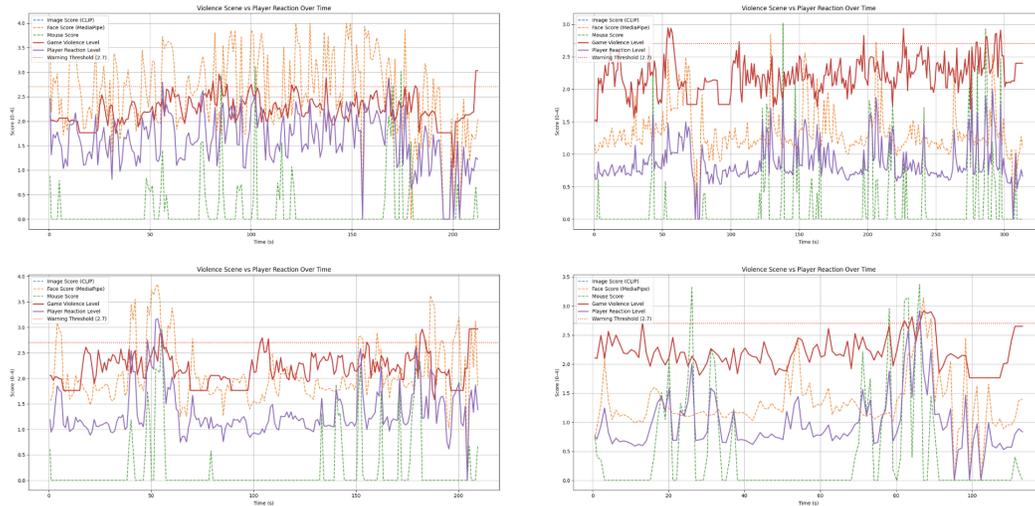


Figure 5: Comparison of Different Game-play Experience users

288 An interesting finding relates to the timing of pain delivery. The electric shock, administered in real
289 time during game-play, was often interpreted as part of the challenge, while the rubber band, applied
290 after the game round, was seen as a punishment. This contrast supports the hypothesis that in-game
291 pain integrates into game-play, whereas post-game pain feels disciplinary. As one player reported,
292 “the electric shock doesn’t feel like a punishment, but the rubber band does.” This suggests that the
293 temporal placement of the feedback may impact players’ emotional framing of pain, influencing how
294 pain is cognitively processed in relation to agency and immersion.

295 Additionally, emotional habituation was observed in some cases, where player reaction levels seemed
296 to decrease following the initial shocks. This raises the possibility that players became desensitized
297 to repeated pain cues, especially the electric feedback. As such, over time, players may become
298 desensitized to pain cues, indicating a need for adaptive feedback intensity to maintain behavioral
299 relevance. A future direction could involve tracking the time to first shock versus the total number of
300 shocks to model this adaptation.

301 We also examined cross-modality alignment. Peaks in facial expression score and mouse interaction
302 intensity often co-occurred with elevated reaction levels, indicating that multi-modal fusion effectively
303 captured players’ internal arousal states. This supports the use of multi-modal input for the reliable
304 detection of emotionally charged moments.

305 Finally, the questionnaire responses revealed an important behavioral insight: the presence of pain
306 influenced players’ decisions. For example, one participant explicitly stated they behaved “less
307 violently, because I feared being shocked.” This implies that pain feedback not only reflects but
308 actively modulates player behavior, functioning as a form of operant conditioning within the game
309 experience. These qualitative reflections further support the notion that pain is not perceived uniformly
310 but is shaped by timing, integration, and user expectation.

311 7 Conclusion and future directions

312 Our project Pain Haptics explores the potential of conditioning users to recognize violent behaviors
313 in video games by associating those moments with physical pain. In our first proof-of-concept, we
314 selected *Call of Duty Black Ops 2: Zombies* due to its testability and the clear, gradable instances
315 of user-initiated violence that can be categorized by AI. While this game allowed us to validate the
316 basic functionality and feasibility of our approach, we identified three key insights that will inform
317 future development. First, we found it challenging to design our system in a way that evokes a
318 sense of meaningful, emotionally grounded supervision. Our goal was to emulate a metaphorical
319 parent figure—an AI presence that intervenes when gameplay becomes excessively violent, much
320 like a concerned guardian stepping in. However, the current implementation feels more like a
321 reactive sensor: users quickly learn to anticipate which actions trigger the pain feedback, leading to a
322 mechanistic and predictable experience. Second, we suspect that this sensor-like response stems not
323 only from system design but also from the nature of the game itself. *Call of Duty* offers limited moral
324 complexity and rarely presents players with meaningful choices. In contrast, open-world games
325 like *Grand Theft Auto (GTA)* offer multiple pathways to achieve a goal, with some options being
326 more violent or destructive than others. We believe future iterations of our system could benefit from
327 games with greater narrative and ethical flexibility, allowing AI to assess context and intent more
328 deeply and deliver pain feedback based on more nuanced behavioral analysis. Third, we initially
329 assumed that pain haptic feedback would be universally perceived as negative. However, some users
330 described the pain as an engaging or even enjoyable part of the experience. This suggests broader
331 potential applications for pain feedback beyond aversive conditioning. In future work, we plan to
332 explore its use in educational training, physical rehabilitation, and the development of human-AI
333 empathy. Looking further ahead, we aim to challenge conventional roles between humans and AI.
334 Rather than designing AI systems that simply obey human commands, we are interested in exploring
335 AI agents that act in the user’s long-term interest—even if that involves short-term discomfort. This
336 opens a philosophical dimension to our work, where AI can exercise its own “judgment” in ways that
337 encourage behavioral growth, reflection, or restraint.

338 References

339 Gentile, D. A., & Gentile, J. R. (2007). Violent Video Games as Exemplary Teachers: A Conceptual Analysis.
340 *Journal of Youth and Adolescence*, 37(2), 127-141. <https://doi.org/10.1007/s10964-007-9206-2>

- 341 Shoshani, A., Braverman, S., & Meirow, G. (2021). Video games and close relations: Attachment and empathy
342 as predictors of children's and adolescents' video game social play and socio-emotional functioning. *Computers*
343 *in Human Behavior*, 114, 106578. <https://doi.org/10.1016/j.chb.2020.106578>
- 344 Greitemeyer, T. (2022). The dark and bright side of video game consumption: Effects of violent and prosocial
345 video games. *Current Opinion in Psychology*, 46, 101326. <https://doi.org/10.1016/j.copsyc.2022.101326>
- 346 Akpan, B. (2020). Classical and operant conditioning—ivan pavlov; burrhus skinner. *Science education in*
347 *theory and practice: an introductory guide to learning theory*, 71-84.
- 348 Säljö, R. (2021). The conceptualization of learning in learning research 1: From introspectionism and conditioned
349 reflexes to meaning-making and performativity in situated practices. In *Learning as social practice* (pp. 146-168).
350 Routledge.
- 351 Moran, T., Nudler, Y., & Bar-Anan, Y. (2023). Evaluative Conditioning: Past, Present, and Future. *Annual*
352 *Review of Psychology*, 74(1), 245-269. <https://doi.org/10.1146/annurev-psych-032420-031815>
- 353 Zhang, Y., Liu, M., & Chen, R. (2024). Using AI to Monitor and Prevent Gaming Addiction: A Behavioral
354 Analysis Approach. *BMC Public Health*, 24(1), Article 20028.
- 355 Johnson, D., & Patel, S. (2024). AI-Based Relapse Prediction in Behavioral Addictions. *Frontiers in Psychiatry*,
356 15, Article 11572328.
- 357 Sharkey, A. (2024). Could a robot feel pain? *AI & Society*. <https://doi.org/10.1007/s00146-024-02110-y>
- 358 Temel, M. H., Erden, Y., & Bağcier, F. (2025). Lumbar radicular pain in the eyes of artificial intelligence: Can
359 you 'imagine' what I 'feel'? *World Neurosurgery*, 193, 309–314. <https://doi.org/10.1016/j.wneu.2024.09.075>
- 360 Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with
361 symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized
362 controlled trial. *JMIR Mental Health*, 4(2), e7785. <https://doi.org/10.2196/mental.7785>
- 363 Werner, P., Al-Hamadi, A., Limbrecht-Ecklundt, K., Walter, S., & Traue, H. C. (2018). Head movements and
364 postures as pain behavior. *PLOS ONE*, 13(2), e0192767. <https://doi.org/10.1371/journal.pone.0192767>
- 365 Voytovich, L., & Greenberg, C. (2021). Natural language processing: Practical applications in medicine and
366 investigation of contextual autocomple. In *Machine Learning in Clinical Neuroscience: Foundations and*
367 *Applications* (pp. 207–214). Springer International Publishing.
- 368 Waller, B. M., Julle-Daniere, E., & Micheletta, J. (2020). Measuring the evolution of fa-
369 cial 'expression' using multi-species FACS. *Neuroscience & Biobehavioral Reviews*, 113, 1–11.
370 <https://doi.org/10.1016/j.neubiorev.2020.02.031>
- 371 Cao, S., Fu, D., Yang, X., Wermter, S., Liu, X., & Wu, H. (2024). Pain recognition and pain empathy from a
372 human-centered AI perspective. *iScience*, 27(8), 110570. <https://doi.org/10.1016/j.isci.2024.110570>
- 373 Naeini, E. K., Jiang, M., Syrjälä, E., Calderon, M. D., Mieronkoski, R., Zheng, K., ... & Rahmani, A. M. (2020).
374 Prospective study evaluating a pain assessment tool in a postoperative environment: Protocol for algorithm
375 testing and enhancement. *JMIR Research Protocols*, 9(7), e17783. <https://doi.org/10.2196/17783>
- 376 Thiam, P., & Schwenker, F. (2017, November). Multi-modal data fusion for pain intensity assessment and
377 classification. In *2017 Seventh International Conference on Image Processing Theory, Tools and Applications*
378 *(IPTA)* (pp. 1–6). IEEE. <https://doi.org/10.1109/IPTA.2017.8310115>
- 379 Lötsch, J., & Ultsch, A. (2018). Machine learning in pain research. *Pain*, 159(4), 623–630.
380 <https://doi.org/10.1097/j.pain.0000000000001118>
- 381 Zhang, Y., Liang, B., Chen, B., et al. (2022). Force-aware interface via electromyography for natural VR/AR
382 interaction. *arXiv*. <https://arxiv.org/abs/2210.01225>
- 383 Fedoseev, A., Baza, A., Gupta, A., Liu, L., Zhu, Y., & Jiang, Y. (2022). DandelionTouch: High fidelity haptic
384 rendering of soft objects in VR by a swarm of drones. *arXiv*. <https://arxiv.org/abs/2209.10503>
- 385 Venkatraj, K. P., Meijer, W., Perusquía-Hernández, M., Huisman, G., & El Ali, A. (2023). Investigating haptic
386 feedback and agency in virtual avatar co-embodiment. In *Proceedings of the 2023 CHI Conference on Human*
387 *Factors in Computing Systems* (Paper No. 3642425). ACM. <https://doi.org/10.1145/3613904.3642425>
- 388 Wu, J., Sun, B., You, H., & Du, J. (2025). Enhancing human-AI perceptual alignment through visual-haptic
389 feedback system for autonomous drones. *SSRN*. <https://doi.org/10.2139/ssrn.5198137>
- 390 Anderson, C. A., & Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cog-
391 nition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review. *Psychological*
392 *Science*, 12(5), 353–359. <https://doi.org/10.1111/1467-9280.00366>

- 393 Gentile, D. A., Lynch, P. J., Linder, J. R., & Walsh, D. A. (2004). The effects of violent video game habits
394 on adolescent hostility, aggressive behaviors, and school performance. *Journal of Adolescence*, 27(1), 5–22.
395 <https://doi.org/10.1016/j.adolescence.2003.10.002>
- 396 Zhao, Y., & Wang, Y. (2020). Exploring anti-addiction systems in Chinese online gaming industry:
397 A policy review. *International Journal of Environmental Research and Public Health*, 17(11), 3946.
398 <https://doi.org/10.3390/ijerph17113946>
- 399 Li, W., Zhao, X., & Zhang, Y. (2021). Youth online gaming and regulation in China: The clash of industry, society,
400 and government. *Media International Australia*, 181(1), 40–53. <https://doi.org/10.1177/1329878X211006446>
- 401 Yang, S., Kim, Y., & Jeong, S. (2022). Artificial intelligence in digital game addiction detection and intervention:
402 A systematic review. *IEEE Access*, 10, 20385–20396. <https://doi.org/10.1109/ACCESS.2022.3151422>
- 403 Robbins, A. (2001). *Awaken the Giant Within: How to Take Immediate Control of Your Mental, Emotional, Physical and*
404 *Financial Destiny!*. New York: Free Press.
- 405 Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*.
406 Oxford University Press.