A Multilayer Systems Approach to Modeling User and Player Experience in Interactive Digital Environments

Dr. Tibor Guzsvinecz

14 April 2026

Contents

1 Introduction **1**
1.1 Perceptual-cognitive foundations 2
1.2 Constraints and inclusivity in digital environments 3
1.3 Structural-affective dynamics and modeling 3
1.4 The proposed model of user and player experience 4
1.5 Structure of this booklet 4
1.6 Statistical reporting 5

2 Thesis group 1: Perceptual-cognitive layer **6**
2.1 The influence of immersion and human characteristics on distance estimation in virtual environments 6
2.2 Affecting sound localization performance in immersive environments 11

3 Thesis group 2: Constraint and inclusivity layer **15**
3.1 Cybersickness as a recurring constraint in commercial VR games 15
3.2 Accessibility as structural participation constraint 19

4 Thesis group 3: Structural-affective layer **22**
4.1 Top-level genre as affective macro-architecture 22
4.2 Player engagement modes as meso-structural affective operators . 24
4.3 Design paradigms and subgenres as micro-structural affective operators 27

5 Discussion and conclusions **31**

1 Introduction

The research presented in this habilitation thesis booklet conceptualizes digital user experience (UX) and player experience (PX) as a complex human-computer

interaction system embedded within a multilayer model. Instead of viewing digital environments and games as cultural or entertainment artifacts, they are examined as interactive information systems. In this model, the player becomes an integral part of the system rather than a passive observer. The game engine provides stimuli, the player perceives and responds, and these responses are processed by the program logic. From an informatics perspective, this interaction can be described as a feedback-oriented system connecting human perception and software architecture.

UX and PX can therefore be treated as an outcome of an integrated interaction system. Rather than separating immersive perception, accessibility, game mechanics and emotional expression into independent research domains, they are examined within a joint model. The approach combines formal modeling with data analysis in order to better understand how these elements interact and influence one another. The proposed model is organized into three distinct, yet dynamically connected layers:

1. **Perceptual-cognitive layer:** In this layer, attention is directed to how immersive technologies influence human perception and cognitive processing. It is examined how spatial awareness, multisensory experience, and response behavior are shaped when interaction takes place in digital environments.

2. **Constraint and inclusivity layer**: The limitations and enabling conditions of participation are considered in this layer. Physiological factors such as cybersickness, as well as accessibility-related design decisions are discussed as elements that shape who can meaningfully participate in a given digital environment and under what conditions.

3. **Structural-affective layer**: Lastly, in this layer, the focus is placed on how gameplay organization and rule structures relate to patterns of emotional expression. Emotional responses observed in large-scale player discourse are interpreted in connection with interaction design, genre characteristics, and game design tendencies.

Together, these layers provide a structured perspective on digital UX and PX. At the most immediate level, perception shapes how interaction is sensed and interpreted. At an intermediate level, limitations and design decisions influence participation. Lastly, at a broader level, gameplay organization and genre characteristics contribute to recurring emotional and evaluative patterns.

1.1 Perceptual-cognitive foundations

At the most immediate level of interaction, immersive digital environments shape how sensory information is presented and experienced. As such, immersive virtual environments (IVEs) influence spatial cues, bodily orientation, and the coordination of visual and auditory signals. As a result, perception in these IVEs differs in important ways from traditional screen-based interaction.

The results presented in this habilitation thesis booklet show that immersion combined with human characteristics affects egocentric and exocentric distance estimation as well as auditory localization [**GT1**–**GT15**]. These differences cannot be explained only by display quality or graphical detail. Instead, they indicate that perception itself is influenced by the conditions under which sensory information is delivered and processed.

From an informatics perspective, immersive systems therefore modify the conditions of interaction within the human-computer relationship. The way information is generated and presented by the software affects how it is interpreted and acted upon by the user. Therefore, UX and PX are shaped at the level of perception, even before game mechanics or narrative elements begin to influence emotional responses.

1.2 Constraints and inclusivity in digital environments

In interactive systems, constraints shape how people are able to engage with them. In digital environments, these limits are shaped both by human characteristics and by design decisions embedded in the software.

One important example is cybersickness in IVEs. The analysis of more than one million user-generated reviews shows that motion-related discomfort is not a rare or incidental complaint, but a recurring and recognizable aspect of virtual reality (VR) experiences [**GT16**]. Such effects influence how long and how comfortably users are able to engage with a system, and therefore have practical implications for the design of IVEs.

Beyond physiological factors, accessibility features also play a central role in shaping participation. Accessibility features are not implemented uniformly across games, nor do their numbers increase over time. Instead, their presence and depth are closely connected to genre characteristics and interaction styles. This is especially true in multiplayer settings. This suggests that accessibility is closely tied to the overall design logic of a game rather than being an optional or secondary addition [**GT17**].

The results presented in this habilitation thesis booklet indicate that inclusive participation emerges from the interaction between human characteristics, possible disabilities, and design choices. Participation is therefore not a fixed property of a system, but a dynamic outcome shaped by how individual capabilities meet structural design elements.

1.3 Structural-affective dynamics and modeling

UX and PX become visible not only in controlled experimental settings but also in large-scale discourse data. The analysis of tens of millions of Steam reviews shows that emotional expression in digital environments follows recognizable patterns and cannot be explained by individual player differences alone. Instead, these patterns are closely connected to structural properties of game design. Here, three distinguishable, yet interconnected structural layers can be identified.

First, top-level genres (TLGs) function as macro-architectures. They have different combinations of gameplay mechanics, pacing structures, interaction densities, and narrative framing. Across millions of textual reviews, TLGs differ from each other in emotional valence, review length, playtime at review, and emotion composition [**GT18**]. Second, player engagement modes (PEMs) such as singleplayer, cooperative, multiplayer formats operate at a meso-structural level. While TLGs describe content-oriented classifications, PEMs define interactional architectures. They structure how players relate to the system and to each other, thereby influencing patterns of emotional articulation, narrative framing, and review timing independently of which TLG they belong to [**GT19**]. Third, specific design paradigms and subgenres operate at a micro-structural level. High-risk mastery-oriented configurations (e.g., Soulslike mechanics), immersive VR games, spatial-ambiguity designs (liminal space games), and games with competitive interaction (esports) influence difficulty, embodiment, social friction, and reward structures. These paradigms generate distinctive, yet stable affective signatures across titles and over time [**GT20**–**GT29**].

These results suggest that emotional discourse in digital environments is shaped by multiple structural conditions. Emotional expression can therefore be understood as the outcome of how rule systems, interaction structures, and evolving design practices shape UX and PX. This perspective connects lower-level perceptual and participation constraints with higher-level patterns observed in collective discourse. Based on this idea, the following subsection introduces a multilayer model.

1.4 The proposed model of user and player experience

The layers described above form a structured system in which each level influences the next. At the most immediate level, perception shapes how interaction is sensed and interpreted. Physiological factors and accessibility considerations influence who can participate and under what conditions. At a broader level, gameplay organization through TLGs, PEMs, and common design approaches contributes to recurring emotional and evaluative patterns in player discourse.

The central proposition of this habilitation thesis booklet is that UX and PX arise from the interaction of these layered conditions. Digital environments should therefore not be seen merely as technical artifacts, but as human-system contexts in which perception, participation, and design are closely interconnected. While the analyses reveal clear and stable tendencies, the system is not deterministic: individual differences and situational factors still introduce variation. This multilayer perspective provides the conceptual foundation for the theses presented in the following sections.

1.5 Structure of this booklet

The following sections present the habilitation theses derived from this integrated perspective. Each of the aforementioned three layers forms a thesis group as well as a section in this booklet. Therefore, Section 2 presents thesis group 1,

Section 3 details thesis group 2, and Section 4 shows thesis group 3. Lastly, a short discussion and conclusions are presented in Section 5. It should be noted that in each thesis group, each subsection corresponds to a thesis as well.

1.6 Statistical reporting

All statistical variables and values are reported following the conventions outlined in the Publication Manual of the American Psychological Association (7th ed.) [1]. Furthermore, the following statistical abbreviations are used throughout this booklet:

- Mean (M),
- Standard deviation (SD),
- Median (Mdn),
- Standard error (SE),
- Confidence interval (CI),
- Sample size (N for participants, R for reviews),
- Probability (P),
- Significance level (p),
- Correlation coefficient (r),
- Spearman rank correlation coefficient (r_s),
- Point-biserial correlation coefficient (r_{pb}),
- Unstandardized regression coefficient (b),
- t statistic (t),
- F statistic (F),
- Mann-Whitney U statistic (U),
- Wilcoxon test statistic (W),
- Cramér's V (V),
- Likelihood-ratio statistic (G),
- Chi-square statistic (χ^2),
- Analysis of variance (ANOVA).

2 Thesis group 1: Perceptual-cognitive layer

2.1 The influence of immersion and human characteristics on distance estimation in virtual environments

Background and related work

Distance estimation is a core component of visuospatial cognition in real and virtual environments (VEs), influencing how spatial relationships are perceived and acted upon, as it directly affects navigation, interaction with objects, and task execution [2]. A widely used conceptual distinction separates egocentric distance estimation (observer-to-target distance) from exocentric distance estimation (distance between two targets independent of the observer). Previous studies have shown that distance estimation in VEs can deviate from real-world perception, and that immersive display technologies may alter the reliability and structure of available depth cues [3–6]. However, few controlled studies have jointly investigated egocentric and exocentric distance estimation in identical environments while modeling immersion as a structural modifier of the underlying distance estimation process [7]. The research related to the first thesis addresses this gap by examining how different levels of immersion influence accuracy and response behavior, while also considering human characteristics as moderating factors.

Materials and methods

The measurements were conducted in a VE developed in Unity. The virtual room measured 12 m $\times$ 12 m and contained minimal environmental cues to reduce anchoring effects. Two versions of the VE were implemented: one for desktop (controlled with the keyboard and mouse) and an immersive version for the Gear VR head-mounted display (HMD). Participants were positioned at the center of the virtual room and were not allowed to locomote; only the rotation of the head/camera was permitted. The height of the virtual camera was calibrated to match each participant's real-world height. Object types were randomized between cubes, cylinders and spheres to reduce shape-based estimation bias.

The sample size consisted of $N = 229$ participants, out of which 157 ($M_{age} = 19.80, SD_{age} = 2.09$) used the desktop display and 72 ($M_{age} = 22.51, SD_{age} = 6.63$) the Gear VR HMD. Out of males, 128 used the desktop display, while 49 used the Gear VR HMD. In the case of females, 29 and 23 used these two display devices, respectively. During the measurements, immersion level was the primary factor under investigation. In the case of egocentric distance estimation, target objects were placed at distances between 25 cm and 160 cm in 15 cm increments. Each distance was presented twice, once with a virtual ground scale and once without. During the exocentric distance estimation, two objects were placed in front of the participant, and the distance between them varied between 60 cm and 150 cm in 10 cm increments. Participants reported their distance estimates in centimeters, and response time was recorded in milliseconds, but was converted to seconds later.

Let i denote the participant index and j the trial index. Let d_j denote the true distance presented in trial j, and let $\hat{d}_{ij}$ denote the estimate reported by participant i. The relative estimation error is defined as shown in Equation (1).

$$\epsilon_{ij} = \frac{\hat{d}_{ij} - d_j}{d_j}. \tag{1}$$

An estimate is classified as accurate if the absolute value of the relative estimation error does not exceed 10%. The accuracy indicator $Y_{ij} \in \{0, 1\}$, where 1 represents an accurate estimate, is defined as shown in Equation (2).

$$Y_{ij} = \begin{cases} 1 & \text{if } |\epsilon_{ij}| \leq 0.10, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The probability of accurate estimation was modeled using logistic regression [8]. The logit formula is defined as shown in Equation (3).

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right). \tag{3}$$

The probability of accurate estimation was modeled using logistic regression as shown in Equation (4).

$$\text{logit}(P(Y_{ij} = 1)) = \beta_0 + \beta_1 I_c + \beta_2 G_i + \beta_3 D_i + \beta_4 H_i + \beta_5 W_i + \beta_6 V_i + \beta_7 d_j, \tag{4}$$

where

- Y_{ij} is the accuracy indicator for participant i in trial j,
- $I_c \in \{0, 1\}$ denotes the immersion level associated with display device c,
- G_i denotes participant gender,
- D_i denotes dominant hand,
- H_i denotes participant height,
- W_i denotes weekly gaming hours,
- V_i denotes prior VR experience,
- d_j denotes the true distance in trial j,
- $\beta_0, \ldots, \beta_7$ are regression coefficients.

Although Equation (4) is written in additive form for baseline modeling, extended models including interaction terms were also estimated to examine moderation effects. This allows immersion to be interpreted not only as an additive predictor, but as a factor whose effects depend on its interaction with other variables.

Lastly, distance estimation time T_{ij} was modeled using linear regression, as shown in Equation (5).

$$T_{ij} = \alpha_0 + \alpha_1 I_c + \alpha_2 G_i + \alpha_3 D_i + \alpha_4 H_i + \alpha_5 W_i + \alpha_6 V_i + \alpha_7 d_j + \varepsilon_{ij}, \quad (5)$$

where T_{ij} denotes the response time of participant i in trial j, $\alpha_0, \ldots, \alpha_7$ are regression coefficients, and ε_{ij} denotes the residual term. Other variables are defined above.

Emerged egocentric and exocentric patterns

The results show that immersion levels influenced the accuracy of distance estimates. In the case of egocentric distances, participants with the Gear VR HMD underestimated distances between approximately 40 cm and 130 cm, whereas desktop display users produced more accurate or slightly overestimated estimates at larger distances (130-160 cm). Estimation times were significantly faster when using the HMD, with reductions between 35% and 57% depending on distance category. However, in the case of exocentric distances, the use of the Gear VR HMD reduced the probability of accurate estimation compared to the desktop display. Depending on the target distance, the decrease in accuracy probability ranged between 11.86% and 17.08%.

Importantly, the effect of immersion was not the same between the participants. Gender and gaming experience significantly moderated accuracy. At shorter distances (60 cm and 70 cm), male participants were 80.35% and 90.65% more likely to provide accurate estimates, respectively. At larger distances (110 cm and 120 cm), participants who played 3-4 hours per week were 83.49% and 92.00% more likely to estimate distances accurately. In contrast, some gaming groups showed lower accuracy. Compared to non-players, participants who played 11-19 hours per week, were 65.94% less likely to estimate accurately at 40 cm. Those playing 1-2 hours per week were also less accurate, being 64.97% and 55.38% less likely to provide accurate estimates at 40 cm and 145 cm, respectively. Prior VR experience had a positive effect at shorter distances, with such participants being 68.55% more likely to provide accurate estimates at 40 cm.

When interaction terms were examined, the magnitude of these effects increased significantly. For example, participants playing 1-2 hours per week were 88.88% less likely to estimate accurately at 55 cm, while participants playing 11-19 hours per week without prior VR experience were 84.85% less likely to provide accurate estimates at 100 cm. In some combined factor configurations, the decrease in accuracy probability exceeded 90%.

These results indicate that immersion does not affect estimation accuracy uniformly. Instead, its effects depend on factors such as the distance range, the type of distance estimation, and individual user characteristics.

The relationship between speed and accuracy

Based on the results, it can be stated that the average response time was lower under immersive VR ($I_c = 1$) than with the desktop display ($I_c = 0$). This can be written formally as in Equation (6).

$$M(T_{ij} \mid I_c = 1) < M(T_{ij} \mid I_c = 0). \tag{6}$$

However, the probability of accurate estimates decreased under immersion ($I_c = 1$) compared to the non-immersive version ($I_c = 0$), as shown in Equation (7).

$$P(Y_{ij} = 1 \mid I_c = 1) < P(Y_{ij} = 1 \mid I_c = 0). \tag{7}$$

According to the results of Spearman's correlation analysis as shown in Table 1, no meaningful association was observed between response time and the overall number of accurate egocentric/exocentric distance estimates.

Table 1: The correlation between response time and the number of accurate distance estimates.

	Egocentric estimation	Exocentric estimation
Full dataset	$r_s(4578) = -.067$, $p < .001$	$r_s(4578) = .080$, $p < .001$
Desktop display	$r_s(3138) = -.075$, $p < .001$	$r_s(3138) = .070$, $p < .001$
Gear VR HMD	$r_s(1438) = .000$, $p = .989$	$r_s(1438) = .010$, $p = .520$

Interpretation of the results

Instead of introducing a constant bias, immersion level changes how people perceive and interpret spatial information. This effect appears in both egocentric and exocentric distance estimation tasks. Importantly, these effects are not identical across participants. The magnitude and direction of immersion-related changes depend on individual characteristics such as gender, gaming experience, and prior VR experience. Although response times were generally shorter in the IVE, accuracy did not follow the same trend. The absence of a meaningful correlation between response time and accuracy suggests that immersion does not produce a classical speed-accuracy trade-off. Instead, immersion may influence temporal and spatial processing independently.

Let $\mathbf{H}_i = (G_i, D_i, H_i, W_i, V_i)$ denote the vector of human characteristics for participant i. Let Q_{I_c} represent an immersion-dependent transformation operator that alters how distance estimates are distributed under immersion level I_c. The effect of immersion level on distance estimates is formalized in Equation (8).

$$P(\hat{d}_{ij} \mid d_j, I_c, \mathbf{H}_i) = \mathsf{Q}_{I_c}\left(P(\hat{d}_{ij} \mid d_j, \mathbf{H}_i)\right). \tag{8}$$

Therefore, immersion level, human characteristics, and distance together shape the distribution of estimation errors. Consequently, VEs influence not

only how stimuli are presented, but also how they are interpreted. Based on these observations, the first thesis (T1) can be formulated.

T1. Immersion level influences how users estimate egocentric and exocentric distances and how quickly these estimates are made. These changes are not uniform, as they can be further affected by individual characteristics. Perceptual differences in virtual environments therefore result from the interaction between immersion level and the user.

Publications supporting T1

T1 is supported by 10 publications which are the following [**GT1–GT10**]:

[GT1] **T. Guzsvinecz**, J. Szűcs, and E. Perge, "Examining the influence of egocentric distance on exocentric distance estimation in virtual environments," *Acta Polytechnica Hungarica*, vol. 22, no. 7, pp. 163–179, 2025, doi: 10.12700/APH.22.7.2025.7.9. **(Q_{2024}: Q2 IF_{2024}: 1.8)**

[GT2] **T. Guzsvinecz**, J. Szűcs, and E. Perge, "The effects of immersion level, human characteristics, and a virtual scale on exocentric distance perception," *Infocommunications Journal*, vol. 17, Special Issue on Cognitive Infocommunications, pp. 24–31, 2025, doi: 10.36244/ICJ.2025.6.4. **(Q_{2024}: Q3 IF_{2024}: 1.3)**

[GT3] **T. Guzsvinecz**, J. Szűcs, and E. Perge, "How egocentric distance estimation changes in virtual environments by using a desktop display or the Gear VR," *Acta Cybernetica*, vol. 27, no. 1, pp. 15–41, 2025, doi: 10.14232/actacyb.307740. **(Q_{2024}: Q3 IF_{2024}: –)**

[GT4] **T. Guzsvinecz**, J. Szűcs, and E. Perge, "Examination of accurate exocentric distance estimates in a virtual environment using a desktop display and the Gear VR," *Electronics*, vol. 13, no. 8, Art. no. 1536, 2024, doi: 10.3390/electronics13081536. **(Q_{2024}: Q2 IF_{2024}: 2.6)**

[GT5] **T. Guzsvinecz**, J. Szűcs, and E. Perge, "Towards understanding exocentric distance estimation skills of university students in virtual reality," in *Proc. 2024 IEEE 22nd World Symp. Applied Machine Intelligence and Informatics (SAMI)*, 2024, pp. 481–486, doi: 10.1109/SAMI60510.2024.10432814.

[GT6] **T. Guzsvinecz** and J. Szűcs, "Az egocentrikus és exocentrikus távolságok kapcsolata az e-sporttal," in *II. PE ZEK E-Sport Szakmai Napok: Összefoglaló kiadvány*, Zalai Egyetemi Sportegyesület, 2024, p. 8.

[GT7] **T. Guzsvinecz**, E. Perge, and J. Szűcs, "Examining the results of virtual reality-based egocentric distance estimation tests based on immersion level," *Sensors*, vol. 23, no. 6, Art. no. 3138, 2023, doi: 10.3390/s23063138. **(Q_{2023}: Q1 IF_{2023}: 3.4)**

[GT8] **T. Guzsvinecz**, E. Perge, and J. Szűcs, "Analyzing accurate egocentric distance estimates of university students in virtual environments with a desktop

display and Gear VR display," *Electronics*, vol. 12, no. 10, Art. no. 2253, 2023, doi: 10.3390/electronics12102253. **(Q_{2023}: Q2 IF_{2023}: 2.6)**

[GT9] **T. Guzsvinecz**, J. Szűcs, and E. Perge, "Egocentrikus és exocentrikus távolságok vizsgálata virtuális környezetben," in *Orvosi informatika: A XXXVI. Neumann Kollokvium konferencia-kiadványa*, 2023, pp. 40–46, doi: 10.5281/zenodo.10432880.

[GT10] **T. Guzsvinecz**, J. Szűcs, and E. Perge, "An overview of display parameters and devices that affect egocentric distance estimation in virtual environments," in *9th Winter School for PhD Students in Informatics and Mathematics*, Association of Hungarian PhD and DLA Students, 2023, p. 17.

2.2 Affecting sound localization performance in immersive environments

Background and related work

Sound localization plays an important role in spatial orientation in both physical and virtual environments. Although auditory perception is generally less dominant than vision, humans are capable of identifying sound sources with considerable accuracy under ordinary conditions, mainly within the horizontal plane [9, 10]. However, in VEs, spatial hearing acquires additional importance since users depend on digitally generated cues to interpret and navigate simulated spaces [11].

Research has shown that localization in VR is influenced by multiple interacting factors [12, 13]. These include well-known auditory mechanisms such as interaural time and level differences as well as technological components like head-related transfer functions (HRTFs) and display characteristics. At the same time, human characteristics such as gender, prior experience with VR, or sensory conditions may also influence performance.

Previous studies have examined desktop and immersive displays as well as human characteristics and acoustic manipulation [14, 15], but these factors have rarely been investigated together within the same controlled experimental setting. Therefore, the results presented here address this gap by examining how immersion and human characteristics jointly influence horizontal sound localization performance in a dynamically scaled VE.

Materials and methods

A closed, rectangular virtual room was implemented in Unity for a desktop display (with mouse interaction) and the Gear VR HMD (head-tracked orientation). Eight identical spherical sound sources were positioned on a horizontal circle at equal angular intervals, each located 8.3 virtual meters from the position of the participant. During each trial, one randomly selected sphere emitted a spatialized sound. No visual cue indicated the correct source. The same Sony PlayStation 4 Platinum Wireless headset (with virtual 7.1 mode) was used in the case of both display devices.

To introduce controlled acoustic variation, the room size was randomly scaled between 100% and 250% of its base dimensions after each trial. This manipulation altered reverberation timing and reflection structure while preserving spatial layout. A total of $N = 110$ participants completed 10 trials per display device. Their characteristics were as follows:

- 86 male, 24 female,
- 103 right-handed, 7 left-handed,
- 59 wore glasses, 51 did not,
- 4 reported hearing issues, 106 reported none,
- 18 had prior VR experience, 92 had none,
- mean age: 24.92 years ($SD = 7.42$).

Localization accuracy was analyzed at two complementary levels. First, trial-level correctness was examined using logistic regression [8]. As before, let i denote the participant index and j the trial index. Let $Y_{ij} \in \{0, 1\}$ denote whether the response of participant i in trial j was correct (0 = incorrect, 1 = correct). Let $c \in \{0, 1\}$ denote the display device (0 = desktop display, 1 = HMD), and let I_c denote the associated immersion level. Let T_{ij} denote response time and let R_{ij} denote the room scaling factor for participant i in trial j. The probability of a correct localization is modeled using logistic regression, as shown in Equation (9).

$$\text{logit}(P(Y_{ij} = 1)) = \beta_0 + \beta_1 I_c + \beta_2 T_{ij} + \beta_3 R_{ij}, \tag{9}$$

where $\beta_0, \ldots, \beta_3$ are regression coefficients.

Second, participant-level accuracy was examined using linear regression on the number of correct responses [8]. Let K_{ic} denote the number of correct localizations provided by participant i using display device c. Human characteristics were incorporated as predictors as can be seen in Equation (10).

$$K_{ic} = \beta_0 + \beta_1 I_c + \beta_2 G_i + \beta_3 D_i + \beta_4 Z_i + \beta_5 U_i + \beta_6 A_i + \beta_7 V_i + \varepsilon_{ic}, \tag{10}$$

where I_c denotes immersion condition, G_i gender, D_i dominant hand, Z_i the use of glasses, U_i hearing problems, A_i age, V_i prior VR experience, and $\beta_0, \ldots, \beta_7$ are regression coefficients.

Third, χ^2 tests [16] were used to assess the proportions of inaccurate sound localizations.

Response time was analyzed using linear regression [8]. Let T_{ic} denote the total completion time of participant i using display device c. The regression model is defined as in Equation (11).

$$T_{ic} = \alpha_0 + \alpha_1 I_c + \alpha_2 G_i + \alpha_3 D_i + \alpha_4 Z_i + \alpha_5 U_i + \alpha_6 A_i + \alpha_7 V_i + \varepsilon_{ic}, \tag{11}$$

where $\alpha_0, \ldots, \alpha_7$ are regression coefficients and ε_{ic} denotes the residual term.

Lastly, the relationship between response time and localization accuracy was examined using logistic regression, as shown in Equation (12).

$$\text{logit}(P(Y_{ij} = 1)) = \gamma_0 + \gamma_1 T_{ij}, \tag{12}$$

where γ_0 and γ_1 are regression coefficients.

Accurate sound localization

Overall, the descriptive statistics of the results were similar across immersion levels ($M_{PC} = 7.94$, $SD_{PC} = 2.51; M_{HMD} = 7.81, SD_{HMD} = 2.43$). The regression analysis revealed that several human characteristics significantly influenced localization accuracy. Male participants achieved higher accuracy than female participants, with approximately 87% higher odds of producing a correct response. On the desktop display, males significantly outperformed females ($W = 1562.5$, $p < .001$). Using the HMD, this gap disappeared due to a 25% increase in the accuracy of females. This increase was significant between the two groups of females, $W = 33$, $p = .002$.

Prior VR experience had a similarly strong positive effect, more than doubling the odds of a correct answer (approximately +111%). In contrast, wearing glasses reduced accuracy, corresponding to an approximately 32% decrease in the odds of a correct response. The numerical difference was significant between these two groups as well ($W = 1110.5, p = .015$). Age also showed a gradual negative association with performance, with the odds of a correct answer decreasing by roughly 2.8% per additional year.

In contrast, immersion level itself did not significantly affect correctness probability. Similarly, dominant hand and self-reported hearing problems did not produce statistically reliable effects on localization accuracy. Room size scaling between 100% and 250% also did not significantly influence correctness probability.

These results indicate that sound localization accuracy in the present task was structured primarily by human characteristics rather than by the immersion level or spatial scaling manipulation.

Structure of inaccurate sound localizations

Beyond binary correctness, the proportions of incorrect responses were analyzed. Using the desktop display, most errors occurred in adjacent frontal positions, $\chi^2(6, N = 227) = 243.91$, $p < .001$. The "in front, to the right" and the "in front, to the left" directions were the two most frequently chosen inaccurate directions. With the Gear VR HMD, this error distribution shifted slightly, as "in front, to the left" became the most frequent error direction, $\chi^2(6, N = 241) = 400.18$, $p < .001$. A cross-condition Pearson chi-squared test confirmed that the proportions differed significantly between immersion levels, $\chi^2(6, N = 468) = 20.848$, $p = .001$.

These results show that immersion level does not merely scale performance but modifies the spatial structure of localization errors. Thus, immersion appears to influence the internal spatial reference frame rather than simply improving or degrading overall accuracy.

Results regarding response time

The regression results revealed a pattern in how human characteristics influenced localization accuracy. Male participants completed the localization task significantly faster than female participants, requiring on average about 30% less total time, corresponding to a reduction of roughly 33 seconds. Previous VR experience showed a similar effect. Participants familiar with immersive systems completed the task approximately 26% faster than inexperienced users. Age, in contrast, was positively associated with response duration: each additional year increased total completion time by about 2.9 seconds. Participants reporting hearing problems required significantly more time to complete the task (approximately 52% longer total completion time). However, this result should be interpreted cautiously due to the small subgroup size.

In contrast, immersion level itself did not have a statistically significant effect on response time. Although participants were on average about 5% slower when using the HMD, this difference was small and not statistically significant.

Finally, the regression coefficient linking response time to correctness was not statistically different from zero ($\gamma_1 \approx 0$). This indicates that longer response times did not significantly affect the probability of a correct response. In other words, a classical speed-accuracy trade-off was not observed in this task.

Interpretation of the results

The results show that immersion does not automatically improve or worsen sound localization accuracy. Instead, performance differences are primarily explained by human characteristics. Gender, prior VR experience, age, and whether a participant wore glasses significantly influence localization accuracy and response efficiency. Immersion level itself does not independently determine overall performance, but it modifies the spatial structure of localization errors. At the same time, response time differences are mainly associated with user characteristics rather than with immersion level.

Let E_{ij} denote the categorical error in trial j of participant i, representing the deviation between the reported and the true direction. Let $\mathbf{H}_i = (G_i, D_i, Z_i, U_i, A_i, V_i)$ denote the vector of human characteristics for participant i. Let Q_{I_c} denote the immersion-dependent perceptual transformation introduced in the previous subsection. The effect of immersion on the distribution of directional errors is formalized in Equation (13).

$$P(E_{ij} \mid I_c, \mathbf{H}_i) = \mathsf{Q}_{I_c}\left(P(E_{ij} \mid \mathbf{H}_i)\right). \tag{13}$$

As shown in Equation (13), the model parallels Equation (8) in T1 by demonstrating that immersion changes the pattern of sound localization accuracy

rather than simply improving or reducing overall performance. Based on these results, the second thesis (T2) can be formulated.

T2. In a dynamically scaled virtual environment, sound localization is influenced primarily by the characteristics of the user. Immersion level does not independently determine overall accuracy, but acts primarily as a condition-dependent transformation factor to directional errors and performance differences. Variations in response speed reflect perceptual adaptation processes rather than inherent limitations of the display device.

Publications supporting T2

T2 is supported by five publications which are the following **[GT11–GT15]**:

[GT11] **T. Guzsvinecz**, "The effects of display devices on locating sound sources in virtual environments," *Acta Polytechnica Hungarica*, vol. 22, no. 6, pp. 153–170, 2025, doi: 10.12700/APH.22.6.2025.6.10. **(Q_{2024}: Q2 IF_{2024}: 1.8)**

[GT12] **T. Guzsvinecz**, "How sound localization is affected in virtual environments," *Infocommunications Journal*, **under review. (Q_{2024}: Q3 IF_{2024}: 1.3)**

[GT13] A. Szelinger and **T. Guzsvinecz**, "Hanghatás alapú akadálymentesítési vizsgálat virtuális környezetben," in *Orvosi informatika: A XXXVI. Neumann Kollokvium konferencia-kiadványa*, 2023, pp. 27–33, doi: 10.5281/zenodo.10461807.

[GT14] **T. Guzsvinecz**, M. Szabó, B. Halmosi, and C. Sik-Lányi, "The virtual reality sound-based spatial orientation project," in *Proc. 12th IEEE Int. Conf. Cognitive Infocommunications (CogInfoCom 2021)*, 2021, pp. 11–16.

[GT15] **T. Guzsvinecz**, M. Szabó, and C. Sik-Lányi, "Háromdimenziós virtuális környezet fejlesztése hanghatás alapú HCI és akadálymentesítési vizsgálatra," in *Orvosi informatika 2020 – A XXXIII. Neumann Kollokvium konferencia kiadványa*, 2020, pp. 22–27.

3 Thesis group 2: Constraint and inclusivity layer

3.1 Cybersickness as a recurring constraint in commercial VR games

Background and research gap

In thesis group 1, immersion was described as a factor that changes how users perceive space. The results showed measurable differences in spatial accuracy and response time depending on immersion level. Cybersickness (also referred to as visually induced motion sickness, VIMS) can be understood as a related

consequence of the same process. While T1 and T2 identified perceptual changes that generally remained within tolerable limits, the present thesis suggests that under certain combinations of design parameters and individual susceptibility these effects may exceed physiological tolerance. In such cases, the effect is no longer limited to changes in spatial judgement or response time, but can appear as physical discomfort. From this perspective, cybersickness is not treated as an isolated anomaly, but as a consequence of immersion that appears when the interaction between technological conditions and human characteristics goes beyond what users can comfortably adapt to.

Cybersickness is commonly explained by sensory conflict and postural instability theories [17]. Laboratory research has shown that HMDs produce higher symptom intensity than desktop displays [18]. This is primarily due to intensified optic flow and discrepancies between visual and vestibular input [19, 20]. Still, most studies are conducted in controlled laboratory environments with small samples. While such environments are well suited to isolating causal mechanisms, they offer only a partial understanding of how cybersickness unfolds in commercial gaming contexts. There, exposure is prolonged, interaction patterns are diverse, and individual susceptibility varies.

Consequently, the role of cybersickness outside laboratory environments has received limited attention. To fill this research gap, the present thesis relies on a corpus of more than one million user-generated video game reviews of first-person (FPS) VR titles. This shift from laboratory environments allows cybersickness to be understood not only as a perceptual problem, but also as a factor that can limit participation in immersive human-computer systems. As a result, it can influence how users engage with, evaluate, and adapt to such systems.

Materials and methods

The dataset consists of R = 1,375,125 English-language product reviews of FPS VR games. The reviews were scraped from the Steam digital video game distribution platform using the Steam API and automated scraping methods implemented in the R environment. For each identified FPS VR title, up to 1,000 of the most recent English-language reviews were collected to ensure recency and consistency across titles. Duplicate and empty entries were removed, resulting in the final corpus. Topic modeling (BERTopic) combined with keyword-based filtering was applied to identify cybersickness-related content. Reviews assigned to topics with at least 20% motion-sickness keyword alignment were retained, resulting in a focused subset of 1,898 reviews. Each review could easily be grouped by polarity (positive or negative) as it was available in the metadata and set by the reviewer.

In this section, the following statistical methods are used: χ^2-tests were used to assess the proportions of the reviews [16]; significant differences between groups were investigated with the Mann-Whitney U-test [21]; and point-biserial correlation [22] was used to assess the relationship between review polarity and playtime at review as well as between review polarity and review length (number of words). The Valence Aware Dictionary and sEntiment Reasoner (VADER)

compound sentiment scores were also assessed [23].

Let r denote the review index and let $i(r)$ denote the author of review r. Let $Y_r \in \{0, 1\}$ denote review polarity, where 0 indicates a negative review and 1 a positive review. Let $CS_r \in \{0, 1\}$ denote whether cybersickness is mentioned in review r (0 = no mention, 1 = mentioned). Let τ_r denote playtime at the time of review and let L_r denote the length of review r measured in words.

Analyzing the reviews of cybersickness

In the focused subset ($R = 1{,}898$), cybersickness was mentioned in 28.40% of negative reviews and 15.62% of positive reviews. The difference in proportions was statistically significant ($\chi^2(1, R = 1{,}898) = 34.95$, $p < .001$, $V = .135$), indicating a systematic but small association. Thus, cybersickness increases the probability of negative evaluation, yet it does not deterministically imply dissatisfaction. The relationship between cybersickness mention and negative evaluation is formalized in Equation (14).

$$P(CS_r = 1 \mid Y_r = 0) > P(CS_r = 1 \mid Y_r = 1). \tag{14}$$

Negative reviews containing cybersickness-related content were written earlier ($Mdn = 49.00$, $M = 909.60$, $SD = 6{,}024.32$ minutes) than positive reviews that also contain such content ($Mdn = 322.00$, $M = 1{,}740.00$, $SD = 7{,}592.63$ minutes). The difference was statistically significant, $U = 445{,}167$, $p < .001$, $r = .31$. The observed playtime asymmetry based on sample averages can be expressed as shown in Equation (15).

$$M(\tau_r \mid CS_r = 1, Y_r = 0) < M(\tau_r \mid CS_r = 1, Y_r = 1). \tag{15}$$

Although the point-biserial correlation between playtime and cybersickness mention was statistically significant ($r_{pb}(1{,}891) = -.074$, $p = .001$, 95% CI $[-.119, -.030]$), the effect size was negligible. Nevertheless, the negative direction suggests that players mentioning cybersickness tended to have lower playtime at the time of writing, which is consistent with the pattern of earlier disengagement observed among negative reviewers.

Negative reviews mentioning cybersickness were longer ($Mdn = 37.50$, $M = 72.26$, $SD = 115.74$) than positive reviews mentioning it ($Mdn = 21.00$, $M = 51.61$, $SD = 86.45$). The difference was statistically significant, $U = 252{,}627$, $p < .001$, $r = .14$. The observed length asymmetry based on sample averages is formalized in Equation (16).

$$M(L_r \mid CS_r = 1, Y_r = 0) > M(L_r \mid CS_r = 1, Y_r = 1). \tag{16}$$

The point-biserial correlation between review polarity and review length within the cybersickness-related subset was positive and statistically significant, $r_{pb}(1{,}896) = .122$, $p < .001$, 95% CI $[.078, .167]$. Although the effect size was small, the results suggest that negative experiences tend to produce longer and more detailed reviews.

Let V_r denote the VADER compound sentiment score of review r. The mean VADER compound sentiment score in the cybersickness-related subset – which includes both positive and negative reviews – was .360 which indicated that discomfort may coexist with positive appraisal. This average-based relationship is defined in Equation (17).

$$M(V_r \mid CS_r = 1) > 0. \tag{17}$$

Interpretation of the results

Building on the perceptual model introduced in T1 and T2, immersion can be represented as the operator $\mathsf{Q}_{I_{i(r)}}$ which influences how sensory information is processed. Cybersickness occurs when these changes exceed the user's physiological tolerance.

In the corpus related to this thesis, all analyzed titles were immersive VR games. Therefore, the immersion level was fixed at 1 for all textual reviews. Variation in cybersickness occurrence can thus be attributed to the interaction between design parameters and individual susceptibility.

Let $g(r)$ denote the game reviewed in r. Let $D_{g(r)}$ denote the design parameter configuration of game $g(r)$ (e.g., locomotion mode, optic flow intensity, comfort settings), and let $S_{g(r)}$ denote the gameplay stimulus profile that generates sensory motion cues. Let $\mathsf{T}_{C_{i(r)}}$ denote the physiological tolerance operator parameterized by the susceptibility of the author of review r. The probability of cybersickness-related content can then be expressed as shown in Equation (18).

$$P\big(CS_r = 1 \mid D_{g(r)},\ \mathsf{T}_{C_{i(r)}}\big(\mathsf{Q}_{I_{i(r)}}(S_{g(r)})\big)\big) = g\big(D_{g(r)},\ \mathsf{T}_{C_{i(r)}},\ \mathsf{Q}_{I_{i(r)}}\big)\,, \tag{18}$$

where $g\big(D_{g(r)},\ \mathsf{T}_{C_{i(r)}},\ \mathsf{Q}_{I_{i(r)}}\big)$ summarizes how game design, individual tolerance, and immersion level jointly influence the probability that cybersickness is mentioned in a review. Therefore, in the multilayer model, cybersickness represents the point at which perceptual change becomes a constraint on participation, thereby activating the constraint and inclusivity layer.

T3. Cybersickness is a recurring limitation in commercial VR games. It arises from the interaction between immersive presentation and individual susceptibility, increases the likelihood of negative reviews and earlier disengagement, and often leads to more detailed descriptions of user experience. However, it does not necessarily prevent positive emotional responses. This suggests that cybersickness acts as a participation barrier that affects users differently in immersive virtual environments.

Publication supporting T3

T3 is supported by one publication **[GT16]**:

[GT16] **T. Guzsvinecz**, K. Németh, and J. Szűcs, "Cybersickness in Commercial First-Person Virtual Reality Video Games: Analysis of Player Experience in

Steam Reviews," *Virtual Reality*, vol. 30, art. 83, 2026, doi: 10.1007/s10055-026-01353-y. **(Q_{2024}: Q1 IF_{2024}: 5)**

3.2 Accessibility as structural participation constraint

Research background and gap

Accessibility in video games has usually been examined from guideline-based or assistive-technology perspectives with a strong focus on sensory, motor, and cognitive disability categories [24]. Although several design methodologies including the guidelines of the International Game Developers Association Game Accessibility Special Interest Group (IGDA GA SIG) exist [25], commercial video games are not universally accessible [26]. This is partly because mainstream games contain interaction structures and design priorities that differ from controlled laboratory or serious-game contexts [27].

Most previous studies either investigated serious games or analyzed a limited number of mainstream titles [28]. As a result, there is little large-scale evidence on how accessibility features are distributed across TLGs, release periods, and engagement modes in commercial games as well as digital environments. This creates a gap between accessibility recommendations and how these features are actually implemented across TLGs in practice. This thesis addresses this gap.

Materials and methods

A dataset was constructed consisting of $N = 260$ commercial PC video games (25 per TLG, except Experimental games with $n = 10$), categorized into 11 TLGs defined by the Steam platform [29]. Release years ranged from 2002 to 2023. Accessibility features were coded manually and grouped into three disability domains as seen in the literature: sensory, motor, and cognitive accessibility features. Coding was based on in-game inspection where possible, supplemented by official documentation and store descriptions when necessary.

Let $F_g^{(k)} \in \mathbb{N}_0$ denote the number of accessibility features of domain k in game g, where $k \in \{1, 2, 3\}$ corresponds to sensory, motor, and cognitive features, respectively. The total accessibility load of a game is defined as shown in Equation (19).

$$F_g^{\text{total}} = \sum_{k=1}^{3} F_g^{(k)}. \tag{19}$$

Games were grouped by TLG, release year, and mode of engagement (singleplayer only or having a multiplayer mode). Differences in proportions were examined using log-likelihood ratio tests (G-tests) [30], with effect sizes reported using Cramér's V [31]. Regression analyses and one-way ANOVA were used to identify predictors of accessibility feature counts, while a two-way ANOVA examined the interaction between TLG and accessibility feature type.

The proportion of accessibility features

Significant differences were observed in the proportions of accessibility features between TLGs. This is true in the case of the full dataset ($G(160, N = 1{,}257) = 371.99, p < .001, V = .170$), sensory accessibility features ($G(80, N = 666) = 222.59, p < .001, V = .195$), and cognitive accessibility features ($G(50, N = 157) = 110.51, p < .001, V = .369$). On the contrary, the proportion of motor accessibility features did not differ significantly between the TLGs, $G(10, N = 434) = 4.79, p = .904, V = .150$.

Release year did not significantly influence the proportion of accessibility features (all $p > .90$ in the G-tests). However, modes of engagement showed significant differences between the features of various accessibility types: full dataset ($G(16, N = 936) = 1{,}119.8, p < .001$), sensory features ($G(8, N = 487) = 164.2, p < .001$), motor features ($G(1, N = 333) = 448.02$, $p < .001$), and cognitive features ($G(5, N = 116) = 30.334, p < .001$).

The effects of TLGs, release dates, and modes of engagement on accessibility features

Regression analyses showed that TLG had a significant effect on the number of accessibility features across all three types. Compared to action games (reference category), most other TLGs contained significantly fewer sensory accessibility features. One-way ANOVA confirmed significant TLG-level differences in sensory accessibility features, $F(10,\ 249) = 6.928, p < .001$. A similar pattern emerged for motor accessibility features, $F(10,\ 249) = 2.296, p = .013$, and cognitive accessibility features, $F(10,\ 249) = 2.011, p = .038$.

A two-way ANOVA examining the interaction between TLG and accessibility feature type showed significant main effects for TLG, $F(10,\ 812) = 17.300$, $p < .001$, and accessibility feature type, $F(2,\ 812) = 148.200, p < .001$. These results indicate that accessibility features vary across TLGs and feature type rather than being randomly distributed.

Release year did not have a significant effect on the number of accessibility features, $b = -0.003$, $SE = 0.015$, $t = -0.256$, $p = .798$. Although descriptive trends suggest a gradual increase over time, the statistical tests did not confirm a systematic time-related effect in this dataset. Therefore, the growth in accessibility features cannot be explained simply by the passage of time.

Regarding modes of engagement, regression results showed that compared to singleplayer-only games, games supporting multiplayer contained significantly more motor accessibility features, $b = 0.37$, $SE = 0.18$, $t = 2.10$, $p = .036$. No significant mode-related effects were found for the other accessibility feature types.

Interpretation of the results

Within the multilayer model presented in this booklet, accessibility features can be understood as design elements that broaden participation. Since immersion influences perceptual processing (T1-T2), and cybersickness occurs when these

changes exceed physiological tolerance (T3), accessibility features play a complementary role in the system. They provide alternative ways of interacting with digital environments or games, helping to reduce barriers and enabling a wider range of players to remain engaged.

Let TLG_g, MOE_g, and RY_g denote the top-level genre, mode-of-engagement class, and release year of game g, respectively. The structural determinants of accessibility implementation are expressed as can be observed in Equation (20).

$$F_g^{\text{total}} = h(TLG_g, MOE_g, RY_g), \tag{20}$$

where h captures TLG- and mode-related design tendencies. The effect of accessibility features in the multilayer model is represented by the operator A_g, defined in Equation (21).

$$\mathsf{A}_g = \mathsf{A}\big(F_g^{\text{total}}\big) = \mathsf{A}(h(TLG_g, MOE_g, RY_g))\,. \tag{21}$$

Accessibility should not be understood merely as a checklist of features. Rather, it influences how easily players with different abilities and sensitivities can remain engaged. Although participation was not directly measured in this dataset, the multilayer model suggests the monotonic relationship shown in Equation (22). Let $X_i \in \{0, 1\}$ denote continued participation of player i, where 0 indicates discontinuation and 1 indicates continued play, let $\mathbf{H}_i$ denote the vector of player characteristics, and let $g(i)$ denote the game played by player i.

$$P(X_i = 1 \mid F_{g(i)}^{\text{total}}, \mathbf{H}_i) = f(F_{g(i)}^{\text{total}}, \mathbf{H}_i), \tag{22}$$

where f is expected to increase with $F_{g(i)}^{\text{total}}$, meaning that with more accessibility features – holding other factors constant – a wider range of players can remain engaged. In this sense, accessibility acts as a design layer that helps reduce participation barriers caused by perceptual changes and physiological limits. Based on these results, the fourth thesis (T4) can be formulated.

T4. Accessibility implementation in commercial video games follows patterns related to top-level genre and modes of engagement rather than appearing randomly. Within the multilayer model of user and player experience, accessibility features broaden participation by reducing barriers related to perceptual differences and physiological limits.

Publication supporting T4

T4 is supported by one publication **[GT17]**:

[GT17] **T. Guzsvinecz**, "Video game accessibility in the top-level genres," *Universal Access in the Information Society*, vol. 24, pp. 1479–1494, 2025, doi: 10.1007/s10209-024-01154-7. **(Q_{2024}: Q2 IF_{2024}: 2.7)**

4 Thesis group 3: Structural-affective layer

4.1 Top-level genre as affective macro-architecture

Research background and gap

The previous thesis groups presented a multilayer structure of UX and PX by describing how these experiences are shaped by perceptual processes and constraints of the system. However, these layers primarily address the internal organization of experience. They do not fully explain how such structured experiences become visible in player communication. Player expression therefore represents an additional analytical layer: the discursive-affective layer. This layer becomes observable through large-scale textual review corpora, where players articulate their experiences, evaluations, and emotions.

TLGs are more than simple classificatory labels. They reflect recurring combinations of game mechanics, pacing, interaction patterns, and narrative structures. In doing so, they shape the experiential conditions under which perception, participation, and tolerance limits operate. Despite their central role in both game design and player expectations, large-scale comparative analyses of emotional expression across TLGs remain relatively rare. Existing studies have typically focused on individual game mechanics, specific franchises, few TLGs and subgenres, or only on small samples of titles [32–34]. Less attention has been devoted to whether emotional valence, review length, and sentiment composition differ between the TLGs in video games. The present habilitation thesis addresses this gap by analyzing more than 30 million Steam reviews across 11 TLGs, thereby conceptualizing TLG as an affective macro-architecture that shapes players' emotional discourse.

Materials and methods

Using the Steam API, the most recent 1,000 English-language reviews per game were scraped between August and September 2022. Reviews were grouped according to the 11 TLGs defined in the Steamworks documentation [29]: action, adventure, casual, experimental, puzzle, racing, role-playing games (RPGs), simulation, sports, strategy, and tabletop.

Sentiment analysis was conducted using the NRC Emotion Lexicon via the syuzhet package in R [35, 36]. For each review r, emotional valence and eight discrete emotion categories (anger, fear, anticipation, trust, surprise, sadness, joy, disgust) were extracted. Review length (word count), review polarity (positive or negative), and playtime at review were also gathered. Depending on the type of the statistical analysis, Wilcoxon rank-sum tests [21], Wilcoxon signed rank [37], and ANOVA [38] were used to compare the reviews between the TLGs. Emotional trajectories across narrative time were estimated by segmenting each review into 25 equal-length chunks and computing mean valence per segment.

Genre-level differences in affective expression

Across the full corpus of over 30 million reviews, significant differences were observed between TLGs in emotional valence, review length, playtime at review, and discrete emotion composition.

Review length differed significantly between TLGs, $F(10, 31{,}653{,}105) = 1{,}604$, $p < .001$. Mean word counts ranged from $M = 58.9$ (racing) to $M = 71.4$ (RPG), with a global median of 22 words. When grouped by polarity, negative reviews were consistently longer than positive ones across all TLGs, $815{,}690{,}850 \leq W \leq 4{,}286{,}245{,}480{,}401$, all $p < .001$. The median word count was 19 for positive reviews and 40 for negative reviews. Mean values were $M_{pos} \in [52.7, 64.5]$ for positive and $M_{neg} \in [83.4, 100.3]$ for negative reviews, indicating increases between 33.28% and 39.45% depending on the TLG.

Playtime at review also differed significantly between TLGs, $F(10, 31{,}566{,}213) = 1{,}024$, $p < .001$. Mean playtime ranged from 708.5 minutes (experimental) to 2,285.3 minutes (simulation). The overall median playtime was 344 minutes. When grouped by polarity, the comparison showed that positive reviews were written significantly later than negative ones in every TLG, $416{,}352{,}427 \leq W \leq 2{,}286{,}496{,}892{,}581$, all $p < .001$. The median playtime was 406 minutes for positive reviews and 124 minutes for negative reviews.

Emotional valence decreased over narrative time across all TLGs. The largest decrease between the beginning and end of reviews occurred in experimental games (40.66%), followed by puzzle (30.30%) and adventure (23.93%). The smallest decrease was observed in sports games (0.40%). When grouped by review polarity, the valence differences between positive and negative reviews were significant in all TLGs, $300 \leq V \leq 325$, all $p < .001$.

Emotion percentages differed between the TLGs. For example, anger was highest in action games (10.62%); fear peaked in action games (12.24%); joy (18.43%), sadness (13.32%), and surprise (9.20%) were the highest in experimental games; trust was the highest in tabletop (19.054%) and simulation (19.053%); disgust was most prominent in racing (6.29%) and sports games (6.21%). Most pairwise TLG comparisons were strongly significant (all $p < .001$), with only a limited number of non-significant contrasts. These results show that different TLGs have different review length, temporal engagement patterns, emotional valence trajectories, and various emotion distributions.

Interpretation of the results

Within the multilayer model presented in this habilitation thesis booklet, TLGs can be understood as macro-level structures that shape how experiences of users and players are expressed publicly. The results show that differences between TLGs are not only statistically significant but also have different patterns. Dissatisfaction tends to lead to longer and more detailed reviews, whereas positive ones are usually expressed more concisely. Positive reviews appear after longer playtime, while negative ones tend to form earlier. Reviews also often begin with a relatively positive tone, with sentiment becoming more negative as

the text progresses. These patterns suggest that TLGs do not directly determine individual emotions. Instead, they influence how emotional experiences are expressed and structured in the reviews of players. These results indicate that large-scale emotional patterns in reviews reflect underlying gameplay structures rather than random fluctuations in the review data.

Let r denote the review index, $i(r)$ the player who authored review r, $g(r)$ the game evaluated in review r, and $\ell(r)$ the TLG of game $g(r)$. Let $Y_r \in \{0, 1\}$ denote review polarity, where 0 indicates a negative and 1 a positive review, L_r review length, τ_r playtime at the time of review, and $\mathcal{E}_r$ the extracted emotional profile of review r. The operators $\mathsf{Q}_{I_{i(r)}}$, $\mathsf{T}_{C_{i(r)}}$, and $\mathsf{A}_{g(r)}$ keep their earlier meanings, representing the effects of immersion on perception, physiological tolerance differences between users, and the role of accessibility features. Let $\mathsf{G}_{\ell(r)}$ represent the structural influence of TLGs. Under this formulation, the discursive-affective output of review r can be expressed as the layered transformation shown in Equation (23).

$$\mathbf{D}_r = \mathsf{G}_{\ell(r)}\Big(\mathsf{A}_{g(r)}\big(\mathsf{T}_{C_{i(r)}}\big(\mathsf{Q}_{I_{i(r)}}(S_{g(r)})\big)\big)\Big) = (Y_r, L_r, \tau_r, \mathcal{E}_r), \tag{23}$$

where $S_{g(r)}$ denotes the gameplay structure of game $g(r)$ that gives rise to the player's experience.

T5. Top-level video game genres function as affective macro-structures within digital gaming environments. They influence the length, polarity, playtime at review, and emotional composition of player reviews by shaping the conditions in which players experience and evaluate games.

Publication supporting T5

T5 is supported by one publication [**GT18**]:

[GT18] **T. Guzsvinecz** and J. Szűcs, "Length and sentiment analysis of reviews about top-level video game genres on the Steam platform," *Computers in Human Behavior*, vol. 149, p. 107955, 2023, doi: 10.1016/j.chb.2023.107955. **(Q_{2023}: D1 IF_{2023}: 9)**

4.2 Player engagement modes as meso-structural affective operators

Background and research gap

While PX research has examined the emotional and psychological effects of video games [39], previous studies have mainly focused on genre-level differences, specific mechanics, or individual design elements [32, 40, 41]. However, little attention has been devoted to the structural role of PEMs, such as singleplayer, cooperative, or other, multiplayer formats. While TLG describe content-oriented macro-categories, PEMs define interactional architectures. They structure how players relate to each other, to the system, and to progression dynamics. These

interactional constraints can reorganize emotional articulation independently of whether a game belongs to a TLG.

Previous literature has shown that review length, polarity, and emotional tone vary across some TLGs and subgenres [32]. However, few studies have examined how different engagement modes influence discursive-emotional expression [42–44]. Most existing work considers only a limited set of engagement modes rather than the full range. This gap is important because contemporary digital environments increasingly blur genre boundaries while still maintaining distinct interaction formats. To address this issue, this habilitation thesis examines how the 10 PEMs relate to review length, playtime at review, emotional valence trajectories, and discrete emotion distributions across more than 15 million Steam reviews.

Materials and methods

A total of $R = 15{,}023{,}567$ textual reviews were scraped using the Steam API, restricted to the 1,000 most recent reviews per game. Reviews were grouped into 10 PEM categories based on platform metadata tags [29]: 4 player local, asynchronous multiplayer, co-op, co-op campaign, local co-op, local multiplayer, massively multiplayer, multiplayer, online co-op, and singleplayer.

Text preprocessing included English language filtering, lowercasing, punctuation removal, stopword removal, and token normalization. Emotional analysis was conducted using the NRC Emotion Lexicon via the syuzhet package in R [35, 36]. As with T5, the number of eight different emotions (anger, anticipation, disgust, fear, joy, sadness, surprise, trust) and sentiments (positive/negative) was extracted.

Statistical procedures included: one-way ANOVA for PEM-level comparisons as well as two-way ANOVA for including review polarity [38], Tukey's Honestly Significant Difference (HSD) [45] and Fisher's Least Significant Difference (LSD) [46] post hoc tests, Spearman rank correlation [47] for correlation between review length and playtime at review, as well as narrative valence trajectory modeling using 25 equally partitioned text segments [36].

Results regarding PEMs

Review length differed significantly between the PEMs, $F(9,\ 15{,}032{,}557) = 1{,}418, p < .001$. Review type also showed significant effects, $F(1,\ 15{,}032{,}547) = 205{,}558.53, p < .001$. It also had a significant interaction, $F(9,\ 15{,}032{,}547) = 82.77, p < .001$. Negative reviews were consistently longer between the PEMs. Median word counts ranged between $14 - 22$ words for positive reviews and $30 - 51$ words for negative reviews. Asynchronous multiplayer showed the highest mean review length ($M = 67.26$), whereas local multiplayer showed the lowest ($M = 50.17$).

Playtime at review also had significant differences between the PEMs, $F(9,\ 15{,}000{,}879) = 2{,}567, p < .001$. Two-way ANOVA confirmed significant main effects of review polarity as well, $F(1,\ 15{,}000{,}869) = 5{,}632.20, p < .001$. It also had significant interactions, $F(9,\ 15{,}000{,}869) = 43.03, p < .001$. Median playtime

before review ranged between 392-565 minutes for positive reviews and between 126-196 minutes for negative reviews. Massively multiplayer games had the highest mean playtime at review ($M = 5{,}984.80$ minutes).

Spearman correlation coefficients indicated no meaningful association between review length and playtime at review, despite statistical significance due to large sample size, $r_s \in [-.078, .012], p \leq .041$.

Emotional valence declined across narrative time segments. The largest decrease was observed in asynchronous multiplayer reviews (49.88%). In contrast, small increases occurred in co-op campaign (+0.32%), massively multiplayer (+6.71%), and online co-op (+6.91%).

All eight NRC emotions differed significantly across the PEMs. Anger was around 10% across PEMs, while anticipation ranged between approximately $18 - 19\%$. Trust reached its highest level in the co-op PEM (19.00%), and joy was most prominent in local multiplayer (18.22%). Fear was highest in co-op campaign (12.08%), whereas disgust reached its highest value in massively multiplayer (6.50%). Finally, asynchronous multiplayer had the highest mean number of positive sentiment words ($M = 3.33$).

Interpretation of the results

Let r denote the review index, $i(r)$ the author of review r, $g(r)$ the game evaluated in review r, $\ell(r)$ the top-level genre of game $g(r)$, and $k(r)$ its player engagement mode. Let $\mathbf{D}_r = (Y_r, L_r, \tau_r, \mathcal{E}_r)$ denote the discursive-affective output associated with review r. Furthermore, let $\mathsf{M}_{k(r)}$ represent the structural influence of PEM. The operators $\mathsf{Q}_{I_{i(r)}}$, $\mathsf{T}_{C_{i(r)}}$, $\mathsf{A}_{g(r)}$, and $\mathsf{G}_{\ell(r)}$ retain their earlier meanings, referring to immersion-related perceptual changes, differences in physiological tolerance, accessibility features, and TLG-level structure. Under this formulation, the discursive-affective output of review r can be expressed as shown in Equation (24).

$$\mathbf{D}_r = \mathsf{M}_{k(r)}\Big(\mathsf{G}_{\ell(r)}\big(\mathsf{A}_{g(r)}\big(\mathsf{T}_{C_{i(r)}}\big(\mathsf{Q}_{I_{i(r)}}(S_{g(r)})\big)\big)\big)\Big) \tag{24}$$

where $\mathbf{D}_r$ refers to the review-level discursive-affective output introduced earlier, including review polarity $Y_r \in \{0, 1\}$, review length L_r, playtime at review τ_r, and emotional profile $\mathcal{E}_r$. Lastly, $S_{g(r)}$ denotes the gameplay structure of game $g(r)$ from which the player's experience emerges. The PEM-level operator does not replace the TLG-level structure but further shapes how player experiences are expressed in public reviews. Across engagement modes, three regularities appear: negative reviews tend to be longer, they are written earlier in playtime, and emotional profiles differ between PEMs. Based on these results, the sixth thesis (T6) can be formulated.

T6. Player engagement modes form an intermediate structural layer in digital gaming environments. In addition to top-level genre, they influence review length, when they are written, and their emotional expression. As a result, they shape how player experiences are communicated in public reviews.

Publication supporting T6

T6 is supported by one publication [**GT19**]:

[GT19] **T. Guzsvinecz** and J. Szűcs, "Emotions in Play: A Length and Sentiment Analysis of Steam Reviews by Player Engagement Mode," *Heliyon*, **under review**. (**Q_{2024}: Q1 IF_{2024}: 3.6**)

4.3 Design paradigms and subgenres as micro-structural affective operators

Background and research gap

While T5 conceptualized TLG as a macro-level affective architecture and T6 showed that PEMs operate as meso-structural interactional regulators, neither level fully explains how design paradigms shape the ways players articulate their experiences. Game design paradigms such as Soulslikes, liminal-space games, immersive VR titles, or competitive esports are more than simple subgenres. They represent various design philosophies. Each establishes its own logic of difficulty, space, narrative clarity or ambiguity, embodiment, social dynamics, and risk-reward balance. In doing so, they subtly but consistently structure how players think, feel, and speak about their experience.

A series of earlier studies addressed multiple dimensions of design paradigms in digital games from complementary perspectives. Research in this area examined the affective characteristics of liminal-space games [**GT20**]. Related work explored thematic configurations in Soulslike games through topic modeling [**GT21**] and conceptualized the process of "Soulsification" at the industry level [**GT22**]. Further studies analyzed emotional structures in VR-related player discourse, including the distribution and thematic organization of emotions in VR reviews [**GT23**] and the relationship between design-mechanic combinations and positive reception from a game-theoretic perspective [**GT24**]. The distribution and thematic structure of emotions in VR reviews was also examined in a complementary analysis [**GT25**], together with emotional and textual asymmetries in VEs [**GT26**]. Additional work investigated the structure of competitive player discourse through a preliminary analysis of esports reviews [**GT27**] and analyzed sentiment structures and emotional patterns in Soulslike games [**GT28**]. Finally, the relationship between specific design-mechanic combinations and positive reception was examined from a design perspective [**GT29**].

Although these studies examined different design paradigms and methodological perspectives, they were conducted independently. What is still missing is a unified explanation of how such paradigms shape emotional experiences in digital environments. The present habilitation thesis addresses this gap by identifying regular patterns across these paradigms and placing them within the multilayer operator model.

Materials and methods

In this section, the results derived from multiple large-scale textual corpora are synthesized. The analysis draws on 993,932 Soulslike reviews examined through sentiment analysis, 418,483 English-language Soulslike reviews processed via topic modeling, and 1,635,919 VR reviews analyzed in terms of sentiment, lexical frequency, and correlational structure. The dataset further includes 66,858 liminal-space game reviews and 8,425 e-sport reviews covering Counter-Strike 2, Apex Legends, PUBG: Battlegrounds, and Dota 2.

As previously, all corpora were collected from the Steam platform. Emotional valence and eight different NRC emotions were calculated using the syuzhet package in R [35, 36]. Topic structures were identified with Latent Dirichlet Allocation (LDA) [48], while correlation [47] and regression analyses [8] were used to examine the relationships between playtime, review polarity, and sentiment composition.

The goal of this seventh habilitation thesis is not to restate individual results, but to identify structural regularities across these design paradigms. Although each paradigm shows its own affective and discursive configuration, their comparison reveals how differently structured design logics generate measurably different emotional and textual patterns.

Results regarding various design paradigms

Soulslike games showed strong emotional polarity separation. Negative reviews showed below-zero valence, while positive reviews had elevated joy, trust, and anticipation. Topic modeling revealed that positive reviews concentrate on atmosphere, aesthetics, and challenge mastery, whereas negative ones focused on game mechanics and technical frustration. Moreover, slight-to-moderate correlations were found between increased playtime and the probability of a positive review which indicated that prolonged exposure may transform frustration into mastery. Similarly, different game mechanics yielded distinct PX patterns and emotional profiles in the reviews. From a game-theoretic perspective, Soulslike mechanics instantiate repeated strategic adaptation cycles. Frequent failure followed by retry creates iterative emotional recalibration. This design structure intensifies both frustration and eventual reward. Repetition built into the mechanics therefore shapes the emotional arc of the experience.

VR reviews showed similar length asymmetry as negative reviews were longer and written earlier than positive ones. Emotion distributions were dominated by anticipation, trust, and joy. Unlike Soulslike games, VR dissatisfaction often centers on physical discomfort and technical instability rather than narrative or mechanical opacity. Here, physiological tolerance thresholds (T3) became dominant modulators.

Liminal-space games evoked anticipation, trust, and joy, but also fear and surprise. Emotional intensity emerged from spatial ambiguity and ontological uncertainty rather than explicit threat. Narrative opacity and architectural repetition reorganized cognitive mapping processes, which consequently produced

a distinctive affective topology.

Competitive esports games showed a different temporal pattern in how emotions were expressed. As in the other examined subgenres, negative reviews were significantly longer than positive ones. However, unlike in VR and Soulslike games, negative and positive reviews were written after roughly the same amount of playtime. This suggests that in competitive games, dissatisfaction does not necessarily lead to early abandonment. Instead, players often continue to play despite negative experiences, meaning that frustration and long-term engagement can coexist. Negative reviews also frequently emphasized cheating, hacking, bans, server instability, lag, and competitive imbalance. These patterns indicate that emotional articulation in competitive games is linked to persistent adversarial interaction and fairness perception rather than to embodied discomfort or narrative ambiguity. Emotional intensity is generated through social-friction mechanisms embedded in multiplayer architectures.

Interpretation of the results

Based on the results, design paradigms such as Soulslike games, VR games, liminal-space aesthetics, or competitive esports games should not be seen merely as stylistic variations or subgenre labels. Instead, they are built on consistent rule structures that shape aspects such as difficulty, embodiment, spatial coherence, narrative structure, social interaction, and risk-reward balance. Through these structural differences, they influence how emotional intensity develops, how positive and negative reviews change over time, and how players describe their experiences in reviews.

Let r denote the review index, $i(r)$ the author of review r, and $g(r)$ the reviewed game. Let $\ell(r)$ denote the top-level genre of game $g(r)$, $k(r)$ its player engagement mode, and $m(r)$ the dominant design paradigm associated with the game. Let $\mathsf{R}_{m(r)}$ denote the operator corresponding to this design paradigm. The discursive-affective output of review r can then be expressed as shown in Equation (25). The operators $\mathsf{Q}_{I_{i(r)}}$, $\mathsf{T}_{C_{i(r)}}$, $\mathsf{A}_{g(r)}$, $\mathsf{G}_{\ell(r)}$, and $\mathsf{M}_{k(r)}$ retain their earlier meanings, referring to immersion-related perceptual changes, differences in physiological tolerance, accessibility features, TLG structure, and PEM structure.

$$\mathbf{D}_r = \mathsf{R}_{m(r)}\Big(\mathsf{M}_{k(r)}\big(\mathsf{G}_{\ell(r)}\big(\mathsf{A}_{g(r)}\big(\mathsf{T}_{C_{i(r)}}\big(\mathsf{Q}_{I_{i(r)}}(S_{g(r)})\big)\big)\big)\big)\Big). \tag{25}$$

where $\mathbf{D}_r = (Y_r, L_r, \tau_r, \mathcal{E}_r)$ denotes the review-level discursive-affective output introduced earlier, consisting of review polarity $Y_r \in \{0, 1\}$, review length L_r, playtime at review τ_r, and emotional profile $\mathcal{E}_r$. Lastly, $S_{g(r)}$ denotes the gameplay structure of game $g(r)$ from which the player's experience emerges.

The paradigm-level operator does not replace TLG and PEM structures; instead, it further specifies how gameplay is experienced within them. Different design paradigms influence how players encounter challenges, interpret failure, perceive space, and experience social tension during play. In mastery-oriented designs built around repetition and gradual improvement, emotional experience often alternates between frustration and achievement. In IVEs, embodiment

and physiological limits become central to how affect is expressed. Spatially ambiguous games can alter cognitive orientation and narrative closure, giving rise to particular patterns of anticipation, unease, and curiosity. Competitive game designs may sustain long-term engagement even when players feel dissatisfied, as emotional intensity becomes linked to perceptions of fairness and ongoing competition.

Across these paradigms, emotional discourse is not random. Instead, it reflects recurring patterns of PX rooted in rule systems and interaction structures, even if these patterns are subsequently modulated by individual differences. Paradigm-level operators therefore represent the lowest design-structural layer considered in this habilitation thesis before individual variation enters the model. In this sense, emotional intensity, valence differences, temporal development, and the level of narrative detail can be understood as structured outcomes of underlying design paradigms. On this basis, the seventh thesis (T7) can be formulated.

T7. Various game design paradigms and subgenres function as micro-level affective layers within digital game environments. By shaping how difficulty is experienced, how embodiment and space are organized, and how competition or ambiguity appear during play, they influence the intensity, tone, timing, and expression of emotions beyond top-level genres and player engagement modes.

Publications supporting T7

T7 is supported by 10 publications [**GT20–GT29**]:

[GT20] **T. Guzsvinecz**, "Reactions to video games that evoke a feeling of liminality," *Entertainment Computing*, p. 101048, 2025, doi: 10.1016/j.entcom.2025.101048. **(Q_{2024}: Q2 IF_{2024}: 2.4)**

[GT21] **T. Guzsvinecz**, "Topic modeling of positive and negative reviews of soulslike video games," *Computers*, vol. 14, no. 8, p. 339, 2025, doi: 10.3390/computers14080339. **(Q_{2024}: Q2 IF_{2024}: 4.2)**

[GT22] **T. Guzsvinecz**, "The soulsification of video games," *Multimedia Tools and Applications*, vol. 84, no. 15, pp. 14827–14853, 2025, doi: 10.1007/s11042-024-19628-4. **(Q_{2024}: Q1 IF_{2024}: - IF_{2023}: 3)**

[GT23] **T. Guzsvinecz**, K. Németh, and J. Szűcs, "Topic modeling of virtual reality game reviews using latent Dirichlet allocation," in *Proceedings of the 2025 IEEE 29th International Conference on Intelligent Engineering Systems*, 2025, pp. 255–260, doi: 10.1109/INES67149.2025.11078210.

[GT24] **T. Guzsvinecz** and J. Szűcs, "Game theory and strategic decision-making in the Dark Souls," in *Game Theory – Computational Aspects and Applications*, p. 79, 2025, doi: 10.5772/intechopen.1008456.

[GT25] **T. Guzsvinecz** and J. Szűcs, "Investigation of sentiments in virtual reality game reviews," in *Proceedings of the 2024 IEEE 15th International Conference on Cognitive Infocommunications*, 2024, pp. 11–16, doi: 10.1109/CogInfoCom63007.2024.10894725.

[GT26] **T. Guzsvinecz** and J. Szűcs, "Textual analysis of virtual reality game reviews," *Infocommunications Journal*, Joint Special Issue on Cognitive Infocommunications and Cognitive Aspects of Virtual Reality, pp. 84–91, 2024, doi: 10.36244/ICJ.2024.5.10. **(Q_{2024}: Q3 IF_{2024}: 1.3)**

[GT27] J. Szűcs and **T. Guzsvinecz**, "A preliminary analysis of e-sports game reviews," in *Proceedings of the 2024 IEEE 18th International Symposium on Applied Computational Intelligence and Informatics*, 2024, pp. 377–382, doi: 10.1109/SACI60582.2024.10619731.

[GT28] **T. Guzsvinecz**, "Sentiment analysis of souls-like role-playing video game reviews," *Acta Polytechnica Hungarica*, vol. 20, no. 5, pp. 149–168, 2023, doi: 10.12700/APH.20.5.2023.5.10. **(Q_{2023}: Q2 IF_{2023}: 1.4)**

[GT29] **T. Guzsvinecz**, "The correlation between positive reviews, playtime, design and game mechanics in souls-like role-playing video games," *Multimedia Tools and Applications*, vol. 82, no. 3, pp. 4641–4670, 2023, doi: 10.1007/s11042-022-12308-1. **(Q_{2023}: Q1 IF_{2023}: 3)**

5 Discussion and conclusions

The results presented in this habilitation thesis booklet support the hypothesis that UX and PX in digital games are best understood as a layered interaction system rather than the outcome of isolated design variables. Evaluative behavior and emotional expression emerge from the interaction of several structural factors, including perceptual transformations under immersion, physiological limits, accessibility-related participation, gameplay structure, TLG and PEM architectures, and game design paradigms that evolve over time.

At the perceptual-cognitive layer, immersive technologies influence how users process spatial and auditory information. Distance estimation and sound localization are not neutral cognitive operations in VEs, as they are affected by immersion levels and human characteristics. In this sense, immersion changes how users perceive and process information rather than simply improving performance.

The second layer concerns participation constraints and inclusivity. Here, embodied limits appear as boundaries to participation. Cybersickness emerges as a recurring and measurable constraint in commercial VR discourse, reflecting cases where the effects of immersion exceed individual tolerance. Accessibility features provide a complementary design response by broadening participation and offering alternative ways for different players to remain engaged. Importantly, their distribution across games is not random, nor is it explained solely by release year; instead, accessibility implementation aligns with TLGs and engagement modes.

At the structural-affective layer, TLGs differ not only in review polarity and length but also in how reviews evolve over time and how discrete emotions are expressed. PEMs further shape these patterns by reorganizing the social and interactional contexts in which experiences are narrated, without replacing the broader TLG structure. Finally, design paradigms and subgenres influence PX

at the most concrete design level considered in this booklet. Differences in game mechanics affect the intensity, tone, timing, and form of emotional expression, even within the same TLG and PEM.

Overall, these results motivate a unified model in which experiential input is transformed through nested structural layers before becoming observable in review-level discourse. Formally, this multilayer transformation can be summarized as shown in Equation (26).

$$\mathbf{D}_r = \mathsf{R}_{m(r)} \circ \mathsf{M}_{k(r)} \circ \mathsf{G}_{\ell(r)} \circ \mathsf{A}_{g(r)} \circ \mathsf{T}_{C_{i(r)}} \circ \mathsf{Q}_{I_{i(r)}}\big(S_{g(r)}\big), \tag{26}$$

where $\mathsf{Q}_{I_{i(r)}}$ represents the effect of immersion on perception, $\mathsf{T}_{C_{i(r)}}$ physiological tolerance, $\mathsf{A}_{g(r)}$ the role of accessibility features, $\mathsf{G}_{\ell(r)}$ the structural influence of TLG, $\mathsf{M}_{k(r)}$ the structural influence of PEM, $\mathsf{R}_{m(r)}$ the influence of the design paradigm, and $S_{g(r)}$ denotes the gameplay structure of game $g(r)$ from which the player's experience emerges. Furthermore, $\mathbf{D}_r = (Y_r, L_r, \tau_r, \mathcal{E}_r)$ also denotes the review-level discursive-affective output introduced earlier, consisting of review polarity $Y_r \in \{0, 1\}$, review length L_r, playtime at review τ_r, and emotional profile $\mathcal{E}_r$.

Equation (26) summarizes the central theoretical contribution of this habilitation thesis booklet. It describes UX and PX in digital games as the outcome of several interacting layers. Perceptual processes determine how players sense and interpret the VE. Physiological limits and accessibility features influence how long and under what conditions players can remain engaged. Gameplay structures, TLGs, PEMs, and design paradigms then shape how these experiences are interpreted and expressed in player discourse. These layers should not be seen as independent components, but as interconnected conditions that jointly structure player experience. Together, they explain how UX and PX emerge in interactive digital environments.

Implications for research and practice

The results of this habilitation thesis booklet have implications for scientific research on UX and PX. The multilayer model provides a unified analytical model that connects perceptual processes, physiological constraints, accessibility-related participation, TLGs, PEMs, and design paradigms within a single conceptual system. For the scientific community, this model helps bridge previously separate research traditions in human-computer interaction, game studies, and computational discourse analysis. In this sense, the model also outlines the foundations of a coherent research program for studying UX and PX in interactive digital environments.

Methodologically, the results show the value of combining perceptual experiments, large-scale discourse analysis, and structural modeling. Such integration allows UX and PX to be studied not only as individual psychological states but also as emergent properties of complex interactive systems. The multilayer model therefore supports more cumulative and comparable research programs in the scientific study of digital experience.

Beyond academia, the results also have practical implications for the design and deployment of digital games and immersive systems. Immersive technology should be developed with awareness that immersion changes how users perceive and process information. Increasing immersion does not automatically improve experience, since users differ in their perceptual and physiological responses. Accessibility planning is most effective when it is integrated into the structural design of games. TLGs and PEMs influence which accessibility features are useful in practice. Finally, specific design choices such as punishment structures, progression systems, comfort settings, or competitive balance affect how players evaluate and describe their experiences. These elements can therefore be treated as intentional design parameters rather than incidental side effects.

Future research directions

Several research directions follow from the multilayer model proposed in this booklet. An important next step is to examine how the different layers interact with each other, especially in immersive environments where perceptual changes may combine with game mechanics or design paradigms in complex ways. Such research could further clarify how structural properties of games shape player experiences across different technological contexts.

Longitudinal research could also help to understand how design paradigms change across the industry and how players' emotional expectations develop over time. Combining physiological measurements with large-scale analysis of player reviews could further clarify the relationship between players' physical experiences and how they describe them in reviews and community discussions.

Finally, the multilayer approach may be extended beyond entertainment games to other interactive digital environments. Serious games, training simulations, and educational platforms similarly involve layered constraints and structures that affect UX. Applying the proposed model in these contexts may help reveal common principles governing UX and PX across different forms of interactive systems.

General conclusions

Digital UX and PX are not single, isolated attributes but emerge from the interaction of multiple system layers. The emotional patterns observable in player discourse are therefore not random fluctuations or purely subjective accounts. Instead, they reflect structured effects shaped by perceptual conditions, forms of participation, and underlying design logic. By bringing these layers together within a common model, this booklet offers an informatics perspective on UX and PX grounded in empirical analysis. The results show that experiential phenomena in digital games can be studied as structured effects of technological, physiological, and design-related factors. In this sense, UX and PX are best understood not as independent effects, but as emergent outcomes of layered interactive systems.

In conclusion, these results show that UX and PX in digital environments can be explained by the interaction of perceptual, participatory, and design-related structures within a unified multilayer model.

References

[1] American Psychological Association, *Publication manual of the American Psychological Association*, 7th ed. American Psychological Association, 2020. doi: 10.1037/0000165-000 .

[2] C. L. Miller and G. R. Bertoline, "Spatial visualization research and theories: Their importance in the development of an engineering and technical design graphics curriculum model." *Engineering Design Graphics Journal*, vol. 55, no. 3, pp. 5–14, 1991.

[3] S. H. Creem-Regehr, J. K. Stefanucci, and B. Bodenheimer, "Perceiving distance in virtual reality: theoretical insights from contemporary technologies," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 378, no. 1869, 2023. doi: 10.1098/rstb.2021.0456 .

[4] R. V. Kenyon, M. Phenany, D. Sandin, and T. Defanti, "Accommodation and size-constancy of virtual objects," *Annals of biomedical engineering*, vol. 36, no. 2, pp. 342–348, 2008. doi: 10.1007/s10439-007-9414-7 .

[5] F. El Jamiy and R. Marsh, "Distance estimation in virtual reality and augmented reality: A survey," in *2019 IEEE international conference on electro information technology (EIT)*. IEEE, 2019, pp. 063–068.

[6] P. Willemsen and A. A. Gooch, "Perceived egocentric distances in real, image-based, and traditional virtual environments," in *Proceedings IEEE Virtual Reality 2002*. IEEE, 2002, pp. 275–276. doi: 10.1109/VR.2002.996536 .

[7] J. W. Kelly, "Distance perception in virtual reality: A meta-analysis of the effect of head-mounted display characteristics," *IEEE transactions on visualization and computer graphics*, vol. 29, no. 12, pp. 4978–4989, 2022. doi: 10.1109/TVCG.2022.3196606 .

[8] R. J. Freund and W. J. Wilson, *Statistical methods*. Elsevier, 2003.

[9] J. Blauert, *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.

[10] J. C. Middlebrooks, "Sound localization," *Handbook of clinical neurology*, vol. 129, pp. 99–116, 2015. doi: 10.1016/B978-0-444-62630-1.00006-8 .

[11] D. Hug, "How do you sound design? an exploratory investigation of sound design process visualizations," in *Proceedings of the 15th International Audio Mostly Conference*, 2020, pp. 114–121. doi: 10.1145/3411109.3411144 .

[12] M. Berger and R. Bill, "Combining vr visualization and sonification for immersive exploration of urban noise standards," *Multimodal technologies and interaction*, vol. 3, no. 2, p. 34, 2019. doi: 10.3390/mti3020034 .

[13] W. Wang, N. Xu, S. Dang, X. Hong, and J. Qu, "Induction mechanism of auditory-assisted vision for target search localization in mixed reality (mr) environments," *Aerospace*, vol. 9, no. 7, p. 340, 2022. doi: 10.3390/aerospace9070340 .

[14] C. Rajguru, G. Brianza, and G. Memoli, "Sound localization in web-based 3d environments," *Scientific Reports*, vol. 12, no. 1, p. 12107, 2022. doi: 10.1038/s41598-022-15931-y .

[15] A. Ahrens, K. D. Lund, M. Marschall, and T. Dau, "Sound source localization with varying amount of visual information in virtual reality," *PloS one*, vol. 14, no. 3, p. e0214603, 2019. doi: 10.1371/journal.pone.0214603 .

[16] W. G. Cochran, "The $\chi 2$ test of goodness of fit," *The Annals of mathematical statistics*, pp. 315–345, 1952.

[17] C. Curry, R. Li, N. Peterson, and T. A. Stoffregen, "Cybersickness in virtual reality head-mounted displays: examining the influence of sex differences and vehicle control," *International Journal of Human–Computer Interaction*, vol. 36, no. 12, pp. 1161–1167, 2020. doi: 10.1080/10447318.2020.1726108 .

[18] C. Yildirim, "Cybersickness during vr gaming undermines game enjoyment: A mediation model," *Displays*, vol. 59, pp. 35–43, 2019. doi: 10.1016/j.displa.2019.07.002 .

[19] B. Keshavarz and H. Hecht, "Axis rotation and visually induced motion sickness: the role of combined roll, pitch, and yaw motion," *Aviation, space, and environmental medicine*, vol. 82, no. 11, pp. 1023–1029, 2011. doi: 10.3357/asem.3078.2011 .

[20] E. Chang, H. T. Kim, and B. Yoo, "Virtual reality sickness: a review of causes and measurements," *International Journal of Human–Computer Interaction*, vol. 36, no. 17, pp. 1658–1682, 2020. doi: 10.1080/10447318.2020.1778351 .

[21] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The annals of mathematical statistics*, pp. 50–60, 1947. doi: 10.1214/aoms/1177730491 .

[22] R. C. MacCallum, S. Zhang, K. J. Preacher, and D. D. Rucker, "On the practice of dichotomization of quantitative variables." *Psychological methods*, vol. 7, no. 1, p. 19, 2002.

[23] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the international AAAI conference on web and social media*, vol. 8, no. 1, 2014, pp. 216–225. doi: 10.1609/icwsm.v8i1.14550 .

[24] B. Yuan, E. Folmer, and F. C. Harris Jr, "Game accessibility: a survey," *Universal Access in the information Society*, vol. 10, no. 1, pp. 81–100, 2011. doi: 10.1007/s10209-010-0189-5 .

[25] K. Bierre, M. Hinn, T. Martin, M. McIntosh, T. Snider, K. Stone, and T. Westin, "Accessibility in games: Motivations and approaches," *White paper, International Game Developers Association (IGDA)*, 2004.

[26] M. C. Barlet and S. D. Spohn, "Includification: A practical guide to game accessibility," *Charles Town: The Ablegamers Foundation*, 2012.

[27] K. Miesenberger, R. Ossmann, D. Archambault, G. Searle, and A. Holzinger, "More than just a game: accessibility in computer games," in *Symposium of the Austrian HCI and usability engineering group*. Springer, 2008, pp. 247–260. doi: 10.1007/978-3-540-89350-9_18 .

[28] Y. Francillette, E. Boucher, B. Bouchard, K. Bouchard, and S. Gaboury, "Serious games for people with mental disorders: State of the art of practices to maintain engagement and accessibility," *Entertainment Computing*, vol. 37, p. 100396, 2021. doi: 10.1016/j.entcom.2020.100396 .

[29] Valve Corporation. (2026) Steam tags. Steamworks Documentation. Accessed: March 1, 2026. [Online]. Available: https://partner.steamgames.com/doc/store/tags

[30] J. H. McDonald, *Handbook of biological statistics*. Sparky House Publishing, Baltimore, Maryland. USA, 2014.

[31] H. Cramér, *Mathematical methods of statistics*. Princeton university press, 1999, vol. 9.

[32] D. Lin, C.-P. Bezemer, Y. Zou, and A. E. Hassan, "An empirical study of game reviews on the steam platform," *Empirical Software Engineering*, vol. 24, no. 1, pp. 170–207, 2019. doi: 10.1007/s10664-018-9627-4 .

[33] S. Petrosino, E. Loria, A. Kainz, and J. Pirker, "The panorama of steam multiplayer games (2018-2020): A player reviews analysis," in *Proceedings of the 17th International Conference on the Foundations of Digital Games*, 2022, pp. 1–7. doi: 10.1145/3555858.3555922 .

[34] Y. Yu, B.-H. Nguyen, F. Yu, and V.-N. Huynh, "Esports game updates and player perception: Data analysis of pubg steam reviews," in *2021 13th international conference on knowledge and systems engineering (KSE)*. IEEE, 2021, pp. 1–6. doi: 10.1109/KSE53942.2021.9648670 .

[35] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word–emotion association lexicon," *Computational intelligence*, vol. 29, no. 3, pp. 436–465, 2013. doi: 10.1111/j.1467-8640.2012.00460.x .

[36] M. Jockers, "Package 'syuzhet'," *URL: https://cran. r-project. org/web/packages/syuzhet*, 2017.

[37] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics bulletin*, vol. 1, no. 6, pp. 80–83, 1945. doi: 10.2307/3001968 .

[38] H. Scheffe, *The analysis of variance*. John Wiley & Sons, 1999.

[39] I. Granic, A. Lobel, and R. C. Engels, "The benefits of playing video games." *American psychologist*, vol. 69, no. 1, p. 66, 2014. doi: 10.1037/a0034857 .

[40] P. Jagoda, "On difficulty in video games: Mechanics, interpretation, affect," *Critical inquiry*, vol. 45, no. 1, pp. 199–233, 2018. doi: 10.1086/699585 .

[41] S.-C. Liao, F.-G. Wu, C.-H. Feng, and C.-Y. Shuai, "A research on how to enhance user experience by improving arcade joystick in side-scrolling shooter games," in *International conference on human-computer interaction*. Springer, 2020, pp. 431–437. doi: 10.1007/978-3-030-60700-5_55 .

[42] M. Mora-Cantallops and M.-Á. Sicilia, "Exploring player experience in ranked league of legends," *Behaviour & Information Technology*, vol. 37, no. 12, pp. 1224–1236, 2018. doi: 10.1080/0144929X.2018.1492631 .

[43] J. Fox, M. Gilbert, and W. Y. Tang, "Player experiences in a massively multiplayer online game: A diary study of performance, motivation, and social interaction," *New Media & Society*, vol. 20, no. 11, pp. 4056–4073, 2018. doi: 10.1080/0144929X.2018.1492631 .

[44] J. Zhu and S. Ontañón, "Experience management in multi-player games," in *2019 IEEE Conference on Games (CoG)*. IEEE, 2019, pp. 1–6. doi: 10.1109/CIG.2019.8848030 .

[45] J. W. Tukey, "Comparing individual means in the analysis of variance," *Biometrics*, pp. 99–114, 1949. doi: 10.2307/3001913 .

[46] R. A. Fisher, *The design of experiments*. Oliver & Boyd, Edinburgh & London, 1937.

[47] C. Spearman, "The proof and measurement of association between two things," in *Studies in Individual Differences: The Search for Intelligence*, J. J. Jenkins and D. G. Paterson, Eds. Appleton-Century-Crofts, 1961, pp. 45–58. doi: 10.1037/11491-005 .

[48] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003.