RESEARCH GROUP SOCIAL COMPUTING

TECHNICAL UNIVERSITY OF MUNICH

Master of Science in Informatics: Games Engineering

Predicting BFI-2, UPPS and BIS-11 Psychometric Scores Through Gameplay: Virtual Reality and Machine Learning for Assessing Impulsivity, Risk-Taking, Cue Reactivity and Cognitive Profiles

B.Sc. Imène (Imen) Hellali

RESEARCH GROUP SOCIAL COMPUTING

TECHNICAL UNIVERSITY OF MUNICH

Master of Science in Informatics: Games Engineering

Predicting BFI-2, UPPS and BIS-11 Psychometric Scores Through Gameplay: Virtual Reality and Machine Learning for Assessing Impulsivity, Risk-Taking, Cue Reactivity and Cognitive Profiles

BFI-2-, UPPS- und BIS-11-Psychometrische Scores durch Gameplay: Virtuelle Realität und Maschinelles Lernen zur Einschätzung von Impulsivität, Risikobereitschaft, Hinweisreiz-Reaktivität bei Suchtverhalten und Kognitiven Profilen

Author:	B.Sc. Imène (Imen) Hellali
Advisor:	Prof. Dr. med. Michael Soyka, Prof. Dr. Georg Groh
Submission Date:	17 March 2025

I confirm that this master of science in informatics: games engineering is my own work and I have documented all sources and material used.

Munich, 17 March 2025 B.Sc. Imène (Imen) Hellali

Acknowledgments

I extend my deepest gratitude to Prof. Dr. Georg Groh and Prof. Dr. med. Michael Soyka for their invaluable support throughout this project. Their guidance, trust, and encouragement have been without a doubt the most important aspect in bringing this work to life. From providing the necessary hardware to offering constructive feedback, and before anything for trusting me and giving me the total freedom, their support has gone far beyond expectations. They not only allowed me to develop my ideas independently but also continuously motivated me with their expertise and guidance to push forward. This project would not have been possible without their patience, and encouragement. For that, I am forever grateful.

Abstract

Traditional impulsivity assessment tools although reliable [1] they still heavily rely on self-reported questionnaires, which may be subjected to personal biases [2]. Additionally, many standard clinical tasks for evaluating decision-making and cognitive flexibility lack real-time behavioral tracking [3] and are criticized to be invalid outside the controlled clinical setting (ecological validity) [4]. Therefore, with this project we introduce a virtual reality (VR) game-based assessment that integrates eye-tracking, gaze fixation, reaction latency, and decision patterns to measure impulsivity and cognitive abilities. Each in-game task is designed to extract specific cognitive and behavioral measures based on validated research. Then a machine learning-based prediction model maps these extracted values to validated psychometric scores, specifically to the German versions of BIS-11, UPPS, and BFI-2. This report will outline mainly the conceptualization of the game, argue the validity of each task and the extracted measures, then present the prediction model. We omitted chapters regarding testing with a control group (CG) and evaluation of our prediction model to include them in further reports. However to test the output of the prediction model given our measured in-game extracted values, we compared some real world UPPS, BIS-11, BFI-2 scores with those predicted by the model from fellow students.

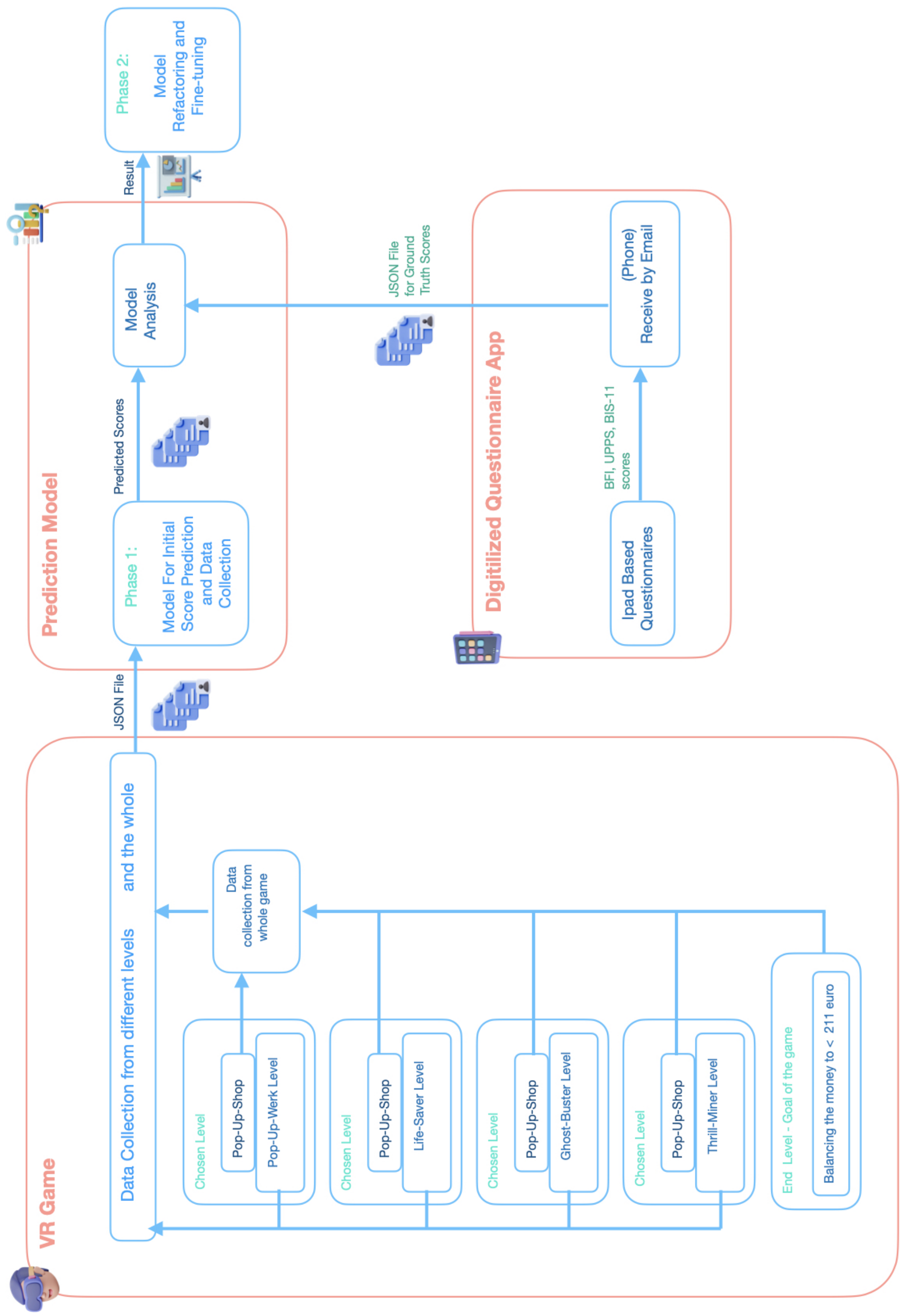


Figure 1: Overview of the VR and ML based solution workflow and the different components constituting the project.

Contents

Abbreviations & Definitions

- VR - Virtual Reality
- AUD - Alcohol Use Disorder
- CG - Control Group: Individuals with no history of alcohol addiction
- RTT - Risk-Taking Task
- IAT - Implicit Association Test
- WoF - Wheel of Fortune Task
- METH-VR - Methamphetamine Craving Virtual Environment
- EEG - Electroencephalography (Measures electrical activity in the brain)
- fMRI - Functional Magnetic Resonance Imaging (Measures brain activity by detecting changes in blood flow)
- CBT - Cognitive Behavioral Therapy (Psychotherapeutic approach for addiction and impulse control)
- Nyctophobia - Fear of darkness
- Claustrophobia - Fear of enclosed spaces
- Teratophobia - Fear of deformed or monstrous figures
- Thanatophobia - Fear of death or dead things
- Glabella - The smooth prominence between the eyebrows
- Salient visual cues - Stimuli that attract attention due to their high perceptual significance (e.g., reward-related cues in addiction studies)
- UPPS - Urgency, Premeditation (lack of), Perseverance (lack of), Sensation Seeking (Personality scale measuring impulsivity dimensions)
- BFI - Big Five Inventory (Personality trait assessment measuring Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism)
- BIS-11 30 item - Barratt Impulsiveness Scale (30-item version used for clinical and research assessment of impulsivity)

- SUD - Substance Use Disorder (Chronic, relapsing disorder characterized by compulsive addictive element use despite adverse consequences)
- CGT - Computerized Gambling Task (Assesses decision-making under risk, linked to ventromedial prefrontal cortex and reward processing)
- Saccades - Rapid eye movements between fixation points, often used as a marker of cognitive control in addiction studies
- Oculomotor control - Regulation of voluntary and involuntary eye movements, including saccades and smooth pursuit, often studied in addiction research
- Neurophysiological responses - Brain and autonomic nervous system reactions (e.g., SCR, fMRI, EEG) associated with addiction and impulse control
- Attentional bias - A cognitive tendency to focus on addiction-related cues over neutral stimuli, often studied using eye-tracking in SUD individuals
- VM-PFC - Ventromedial Prefrontal Cortex (Involved in decision-making, reward evaluation, and impulse control; often impaired in SUD individuals)
- rIFG - Right Inferior Frontal Gyrus (Critical for response inhibition; reduced activation linked to impaired impulse control in SUD individuals)
- Anticipatory saccades - Predictive eye movements occurring before a stimulus appears, indicating cognitive and motor planning deficits in addiction

1 Introduction

This chapter introduces the fundamental knowledge behind the chosen metrics for our Virtual Reality (VR) game and technologies used to derive our prediction model. Please note that argumentation regarding the correctness of the chosen metrics will be discussed in chapters 2 and 3.

1.1 Correlation of Eye-Tracked Data

Eye-tracked measures (e.g. gaze fixation duration and pupil dilation) are quantifiable measures reflecting cognitive processes, attentional biases, and implicit decision-making [5, 6, 7, 8]. When individuals with history of addiction are exposed to reward cues, their gaze subconsciously shift toward these stimuli[6, 8, 7]. The subconscious shift in gaze reflects attentional bias and impaired disengagement from salient visual cues [9, 8, 10, 5]. The time required to disengage from these cues correlate with the individual's recovery stage [11]. Prolonged disengagement indicates a higher likelihood of relapse[11, 12].

1.2 Correlation of UPPS, BFI-2 & BIS-11

The UPPS, BFI-2, and BIS-11 German versions are consistently used scales to assess impulsivity, personality traits, and decision-making. UPPS offers the most insights regarding impulsivity in addiction-related research. It assesses five distinct impulsivity types:

· Negative Urgency (acting impulsively under distress)

· Positive Urgency (acting impulsively under positive emotions)

· Lack of Premeditation (difficulty considering consequences before actions)

· Lack of Perseverance (difficulty maintaining focus on tasks)

· Sensation Seeking (preference for novel and stimulating experiences)

Big Five Inventory (BFI-2) provides a broad personality profile to capture long-term behavioral tendencies. It assesses five dimensions:

· Extraversion (Extraversion)
· Verträglichkeit (Agreeableness)
· Gewissenhaftigkeit (Conscientiousness)
· Negative Emotionalität (Neuroticism)
· Offenheit (Openness)

Additionally, BFI-2 includes fifteen facets:

· Geselligkeit (Sociability)
· Durchsetzungsfähigkeit (Assertiveness)
· Aktivität (Energy Level)
· Mitgefühl (Compassion)
· Höflichkeit (Politeness)
· Zwischenmenschliches Vertrauen (Interpersonal Trust)
· Ordnungsliebe (Orderliness)
· Fleiß (Diligence)
· Verlässlichkeit (Reliability)
· Ängstlichkeit (Fearfulness)
· Niedergeschlagenheit (Dispiritedness or Downheartedness)
· Unbeständigkeit der Gefühle (Emotional Volatility)
· Ästhetisches Empfinden (Aesthetic Sensitivity)
· Intellektuelle Neugierde (Intellectual Curiosity)
· Kreativer Einfallsreichtum (Creative Imagination)

BIS-11 (30-item version) remains widely used for clinical screening of impulsivity. It assesses six distinct factors:

· Aufmerksamkeit (Attention)
· Kognitive Komplexität (Cognitive Complexity)

· Motorische Impulsivität (Motor Impulsivity)

· Selbst Kontrolle (Self-Control)

· Beharrlichkeit (Perseverance)

· Kognitive Instabilität (Cognitive Instability)

Research has has shown, that certain scores from UPPS, BFI-2, and BIS-11 are correlated [13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27]. Taking the example of, negative Urgency strongly correlates with high Neuroticism in BFI-2 and elevated BIS-11 scores. Sensation Seeking is linked to high Extroversion in BFI-2. Conscientiousness negatively correlates with BIS-11 Motor Impulsivity and UPPS Lack of Premeditation. Motor Impulsivity in BIS-11 is directly associated with UPPS Lack of Perseverance. Non-Planning Impulsivity in BIS-11 negatively correlates with Conscientiousness in BFI-2.

Therefore we deduce that if certain measured variables can reliably predict a score from one questionnaire, they can be used to infer an estimate of a correlated score.

1.3 Multimodal Machine Learning

Multimodal Machine Learning (MML) integrates information from multiple data modalities. MML was a necessary choice to combine heterogeneous behavioral metrics (e.g., eye-tracking, reaction times, decision-making strategies) with semantic textual data to accurately predict participants' psychometric scores.

Transformer-Based Embeddings

Transformer architectures are used for generating embeddings by modeling dependencies between input elements through self-attention [28]. SciBERT is a transformer model pretrained to produce contextually rich embeddings from literature and questionnaires [29]. The self-attention mechanism computes attention as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1.1}$$

where Q, K, and V denote query, key, and value matrices, respectively, and d_k is the dimensionality of keys and queries.

Self-Supervised Learning (SSL)

Self-supervised learning is commonly used with unlabeled data. SSL reduces dependency on labeled datasets making it a best fit for newly introduced data (e.g. the in-game extracted measures). It defines auxiliary tasks, delivering inherent structural information, enabling the Neural Network (NN) to learn meaningful feature representations [30]. Common self-supervised tasks are: Context prediction; Masked input reconstruction and contrastive learning.

Contrastive Learning

Contrastive learning trains NNs by minimizing distance (maximizing agreement) between semantically similar (positive) pairs and maximizing distance (minimizing agreement) between semantically dissimilar (negative) pairs [31]. Contrastive learning effectively captures nuanced behavioral patterns (e.g. the in-game extracted measures in context of addiction and psychometric scores).

$$L_{\text{InfoNCE}} = -\log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_{k=1}^{N} \exp(\text{sim}(z_i, z_k)/\tau)} \tag{1.2}$$

where z_i, z_j, and z_k are learned embeddings, $\text{sim}(\cdot)$ is the cosine similarity, and τ is a temperature parameter controlling contrastiveness. Contrastive learning was not used during the implementation of the prediction model, but it is a direction which should be followed in the future.

1.3.1 Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation (RAG) retrieves relevant context from external data sources to augment the inputs provided to the generative model [32]. RAG reduces hallucination of Large Language Models (LLMs) by enhancing factual accuracy [32]. In our prediction model, embeddings generated from extracted in-game measures are the retrieval context. The similarity between a query embedding e_q and dataset embeddings e_{c_i} is computed using cosine similarity:

$$\text{sim}(e_q, e_{c_i}) = \frac{e_q \cdot e_{c_i}}{\|e_q\| \|e_{c_i}\|} \tag{1.3}$$

Contexts meeting a specified similarity threshold (θ) are selected as:

$$C = \{e_{c_i} \mid \text{sim}(e_q, e_{c_i}) \geq \theta\} \tag{1.4}$$

Knowledge Graphs

Knowledge graphs structure relational data to represent entities and their relationships explicitly. Neo4j is an online graph database to store the extracts knowledge from textual data as entity-relationship triples [33]. This structured representation facilitates retrieval and embedding of domain-specific information into the model in later stages.

Multimodal Fusion

Multimodal fusion integrates heterogeneous data types. It allows the model to handle different data dimensionalities and distributions within the datasets (e.g. the in-game extracted measures all represent different behaviors within different ranges). Early fusion (feature-level fusion) was used to concatenate modality-specific embeddings into a unified embedding space. This approach facilitates prediction through cosine similarity.[34].

2 Preliminary Research

This chapter outlines research on neural mechanisms of addiction, psychometric prediction with machine learning, and VR- and non-VR-based cognitive assessments. These elements are the foundation for the ideation of this project.

2.1 Neural Pathways & Addiction

In this section we present findings linking neurophysiological responses and addiction, focusing on impaired cognitive control, risk-taking behavior, and altered reinforcement learning mechanism. We will present few studies using psychophysiological and neuroimaging methods to reveal dysfunctions in key brain regions such as the ventromedial prefrontal cortex (vmPFC), amygdala, striatum, and right inferior frontal gyrus (rIFG), which contribute to maladaptive decision-making patterns in individuals with substance use disorders (SUD) [35, 36, 37, 38, 14, 39, 40, 41, 42].

2.1.1 Neurophysiological Responses in Addiction

Dysfunctional neural mechanisms contribute to altered impulse control, reward sensitivity, and emotional regulation in context of addiction. The vmPFC and amygdala play a crucial role in processing risk and reward. Dysfunctions in these regions lead to impaired decision-making and excessive risk-taking behavior [43, 36, 35, 14, 39, 40, 41, 44]. Studies indicate that individuals with SUD fail to generate anticipatory Skin Conductance Response (SCRs) [43, 45]. This reflects impaired emotional foresight during high-risk decision-making tasks. Dysfunction in the rIFG, which is linked to response inhibition [46, 38], causing difficulty in suppressing impulsive actions [37, 47].

Computerized Gambling Task CGT

The Computerized Gambling Task CGT showed immediate trigger of punishment SCR (Electrodermal response following monetary loss) after a card selection resulting in penalty.[48] Whereas, anticipatory SCR (Electrodermal response before making a decision) occurs before selecting a card [43]. As observed in individuals with SUD, SCR indicates failure to emotionally anticipate losses and weakened conditioning effects

[21]. Additionally Studies conducted on Iowa Gambling Task (IGT) 2.3 and Game of Dice Task (GDT) 2.3 show a deficit in long-term reward evaluation and impulse control consistent with the vmPFC dysfunction observed during CGT [37, 14, 39, 40, 44].

Go/No-Go Task

The Go/No-Go Task (GNG) is a standard task to examine neural correlates of cognitive control. Participants see a sequence of letters appearing one at a time on a screen. They must press a button as quickly as possible whenever any letter appears ("go" trials). However, if the letter "X" appears ("no-go" trial), they must withhold their response. Yet the "X" appears in 25% of trials, making response inhibition challenging. Higher no-go errors (pressing the button for "X") indicate weakened impulse control and correlate with compulsive behavior[49]. Functional imaging studies have linked the "no-go errors" to reduced activation in rIFG (impaired response inhibition in SUD individuals) [46, 36, 35, 37, 47, 38].

2.1.2 Oculomotor Markers of Addiction

Eye-tracking-related research in addiction demonstrate that oculomotor control is influenced by reward salience and impulsivity traits [50]. Individuals with SUD and behavioral addictions exhibit faster engagement with reward-related cues (attentional bias) with slower disengagement from high-salience cues (impaired cognitive flexibility), and altered anticipatory saccades showcasing impairment in decision-making [51, 9].

Predictive Visual Pursuit Task PVPT

During this task, a moving visual target follows a predictable trajectory. Participants must track the target with their eyes as accurately as possible. This task includes standard trials (target moves as expected) and catch trials (target does not reappear). The PVPT assesses anticipatory saccades and smooth pursuit eye movements in response to a predictable motion stimulus. This task evaluates impulsivity and predictive control to correlate latency and velocity with UPPS scores. They found that higher UPPS impulsivity scores are reflected in shorter pursuit latency, increased anticipatory errors and lower pursuit gain [50].

Reward-Driven Attentional Bias Task (RD-ABT)

RD-ABT consists of two phases. Training Phase where participants complete a visual search task, learning that two target colors are linked to monetary rewards (high vs. low). Then a test Phase where the previously rewarded colors act as irrelevant distractors in a shape-based search task without rewards.This task measures visual working memory capacity and attentional bias by showing that response time slows in the presence of reward-associated distractors. This task also shows that, reward history influences attention allocation. This task measures visual working memory capacity and attentional bias by showing that response time slows in the presence of reward-associated distractors. This task also shows that, reward history influences attention allocation.

2.2 Machine Learning In Psychometric Related Research

Machine learning is increasingly applied in psychometric research to improve assessment accuracy, automate scoring, and extract latent behavioral patterns.

2.2.1 Multi Model Approach

Given the complexity of the psychometric scores, extracting meaningful patterns requires often multiple models [52]. Ensemble learning techniques (e.g., random forests, gradient boosting) enhance reliability by integrating weak learners [53]. Deep learning architectures combined with traditional psychometric models (e.g., IRT, factor analysis) improve predictive accuracy for impulsivity and personality traits. By assigning models to distinct behavioral markers, the Multi-model mitigates overfitting. For example, support vector machines (SVMs) are effective for binary classifications (e.g., risky vs. safe behaviors), while deep neural networks (DNNs) capture non-linear relationships in data [54, 55].

2.2.2 Domain Adaptation

ML models in psychometric are generally trained on small labeled datasets of one population, which delivers unsatisfactory results when applied to different demographic, linguistic and cognitive variations [56]. To maintain reliability of the model transfer learning techniques (e.g fine-tuning pretrained models) should be applied. Applying domain-specific embeddings allows the ML model to maintain prediction accuracy even when assessing populations with different sociocultural backgrounds [53, 52].

2.3 Comparison of VR & Non-VR Solutions

In this section we compare few existing VR and non-VR solutions designed for addiction research and addiction behavioral treatment and assessment. Please note this is not a Meta-analysis study nor a Systematic review study, to keep this chapter's size reasonable yet include useful information, the tasks which led to many findings mentioned above were omitted from this section. Additionally in this section we only include a brief description that is not reflective of the immense work nor the research conducted using those solutions. Therefore we ask you to please refer to the references for these solution or a more detailed yet still brief explanation of these few solutions in the supplementary materials.

Iowa Gambling Task (IGT)

The IGT assesses decision-making under implicit risk conditions. Participants are presented with four virtual card decks labeled A, B, C, and D. Each selection results in either a monetary gain or a monetary loss. Two decks (A & B) are disadvantageous (high rewards with high long-term loss). Two decks (C & D) are advantageous (smaller rewards with favorable long-term outcome). The task starts with a monetary sum and participants repeatedly select cards. Participants must adapt their choices based on experienced wins and losses. SUD individuals consistently selecting more cards from the disadvantageous decks (A & B), failing to adapt to long-term negative consequences. The IGT does not isolate specific cognitive processes, making it unclear whether impairments result from impulsivity, reward sensitivity, working memory deficits, or general risk-taking tendencies. The structure of IGT may not reflect real-world risk-based decision-making.

Game of Dice Task (GDT)

The GDT assesses decision-making under explicit risk conditions. Probabilities of gains and losses are known beforehand. This task includes: High-Risk Scenario: Large monetary rewards with low probability of success (e.g., betting on a single number with a 1/6 chance of winning). Moderate-Risk Scenario: Moderate rewards with medium probability of success (e.g., betting on two numbers with a 2/6 chance). Low-Risk Scenario: Smaller rewards with higher probability of success (e.g., betting on three or four numbers with a 3/6 or 4/6 chance). Minimal-Risk Scenario: Small but consistent wins with the highest probability of success. Participants are given a virtual monetary amount and they must make calculated decisions on how to bet for each dice roll. SUD individuals consistently selected disadvantageous high-risk options, despite experiencing repeated losses. The GDT does not capture real-world impulsivity, as SUD individuals often struggle with ambiguous rather than explicit risks.

AlcoVR

AlcoVR is a cue-exposure therapy solution designed for AUD patients. AlcoVR recreates alcohol-related settings such as bars and social gatherings. Its strength manifests in replicating personalized alcohol-associated environments, while also integrating peer influence via interactive NPCs (Non-Playable Characters) [57]. This approach allows for customization of exposure intensity [58, 59, 60]. However such environment face ethical concerns regarding induced cravings giving their graphic and explicit nature.

(a) Lounge-bar environment

(b) Neutral Cafe environment

(c) Bar environment

(d) Neutral house environment

Figure 2.1: Images representing different environments included in AlcoVR

Virtual Alcohol Approach-Avoidance Training Task (VAAAT)

VAAAT is a VR-based intervention designed to influence avoidance biases toward alcohol-related stimuli. The participants are set in either a bar depicting social drinking scenarios, or a cafe (neutral stimuli). A red or green cue appears after scene exposure and participant must either approach (pull) or avoid (push). The environment dynamically adjusts exposure intensity based on user performance and compliance. VAAAT serves as a counter-conditioning mechanism with high ecological validity, allowing controlled exposure to alcohol-related cues [61]. It effectively alters automatic approach tendencies, but longitudinal studies are required to confirm its real-world impact on sustained behavioral change.

Playmancer

Playmancer is a serious game aiming for emotional regulation. The game is composed of multiple mini-games with adaptive difficulty based on user arousal state[62]. The solution trains self-regulation through interactive challenges and controlled breathing. Its strength is showcased through the gamified approach which unarguably enhances engagement, its adaptive feedback which fosters self-control and regulates not only frustration but also anxiety [63]. As described this solution does not directly target addiction triggers so it might not correlate directly and its effectiveness requires ongoing user participation for sustained results.

Risk-Taking Computerized Task (RTT)

RTT is a non-VR computerized tool assessing decision-making under varying risk levels. It measures impulsivity and cognitive control in high-stakes contexts. It provides quantitative decision-making metrics by identifies risk-taking tendencies in controlled settings [64]. On the down side it lacks immersive engagement and it does not provide real-time intervention possibilities.

Methamphetamine Craving Virtual Environment (METH-VR)

METH-VR is a virtual reality-based exposure therapy designed for methamphetamine users. The environment is modeled after real-world meth-use settings, integrating user-reported experiences to enhance realism. The virtual environment consists of an apartment where meth use and transactions occur. Animated avatars engage in meth use (smoking, injecting, snorting) with drug-related paraphernalia and auditory cues such as music commonly associated with meth use. A neutral-apartment without any drug-related cues, with neutral auditory stimuli such as Latin jazz. Participants were free to navigate in both apartments. METH-VR is used to measure craving through physiological and behavioral tracking. Then each participant engagement type with the stimuli is analyzed to personalize intervention [65]. While its high-fidelity design allows for detailed assessment, ethical concerns remain regarding potential craving induction.

Implicit Association VR Task (IAT-VR)

IAT-VR evaluates subconscious biases linked to addiction through structured scenarios such as medical assessments and social confrontations. It deploys eye-tracking and

Figure 2.2: Images representing the modeled environment with its different facets

Stroop tests to measure craving intensity [66]. This solution only measures implicit biases influencing addictive behaviors in addition to behavioral and physiological assessment tools. On the contrary this solution seems inapplicable outside laboratory conditions.

Figure 2.3: Images displaying elements of the environment of IAT-VR

Wheel of Fortune Task (WoF)

WoF is a non-VR cognitive task analyzing risk-based decision-making by simulating counterfactual outcomes such as regret and relief. It assesses cognitive-emotional interplay in risky decisions which is useful for identifying biases in addiction-related choices [67]. Participants are presented with two gamble options, each displayed as a wheel divided into two colored sectors: green (associated with the best outcome) and red (associated with the worst outcome). The probabilities of winning or losing are visually represented by the relative sector sizes. Participants must choose between the two wheels within a fixed decision window (4.5s). Once a selection is made, the chosen wheel is highlighted, and an arrow begins spinning before stopping on a final outcome. Depending on the task condition: Counterfactual (CF) Condition where both selected and unselected gamble outcomes are revealed. Allowing participants to compare their actual result with what they could have won or lost; Partial Feedback (PF) Condition where only the chosen gamble's outcome is revealed. Preventing counterfactual comparison. In the CF condition, participants experience relief if their choice resulted in a better outcome than the unselected option and regret if it led to a worse outcome. The PF condition isolates satisfaction or disappointment without counterfactual comparison. These emotional responses impacted their future decision-making, revealing biases in risk assessment. Despite its ability to measure decision-related emotions, WoF lacks immersive engagement and real-time intervention possibilities compared to IAT-VR, AlcoVR, and METH-VR.

Cognitive Impulsivity Suite (CIS)

CIS is an online tool designed to measure impulsivity through three tasks each targeting different cognitive mechanism. Bounty Hunter (Attentional Control) to measure inhibition. Participants must "shoot" the bandit while refraining from responding to the sheriff. Caravan Spotter (Information Gathering) to measure impulsivity. Participants must identify an object from a gradually unpixelating image. Prospector's Gamble (Feedback Monitoring/Shifting) to measure cognitive flexibility. Participants select one of two gold prospectors. On Prospector is consistently "lucky" but periodically reverses without warning. Participants must detect and adapt to shifting reward probabilities. CIS provides a unified quantitative measure of impulsivity with high test-retest reliability. Unlike VR-based tasks or traditional lab-based studies, CIS does not track physiological automatic responses corresponding to impulsive decisions [68].

Solution Name	Key Attributes	Limitations	Addiction Type	Target Platform
AlcoVR	Personalized alcohol-related environments, peer interaction, customizable exposure	Limited realism, ethical concerns regarding induced cravings	Alcohol Use Disorder (AUD)	Virtual Reality (VR)
Playmancer	Gamified self-regulation, adaptive difficulty, emotional regulation training	Not addiction-specific, requires continued engagement	General Emotional Regulation	Non-VR Serious Game
RTT	Quantitative decision-making metrics, controlled risk assessment	Lacks immersion, no real-time interventions	Risk-Taking Behavior	Non-VR Computerized Task
METH-VR	High-fidelity methamphetamine-related environment, multi-sensory engagement	High computational needs, ethical concerns	Methamphetamine Addiction	Virtual Reality (VR)
IAT-VR	Measures implicit addiction biases, integrates behavioral & physiological assessment	Influenced by external cognitive factors, limited generalizability	Implicit Alcohol Addiction Bias	Virtual Reality (VR)
WoF	Assesses cognitive-emotional decision making, identifies biases in addiction choices	No real-time interventions, lacks VR engagement	Gambling & Risk-Based Addiction	Non-VR Cognitive Task

Table 2.1: Compiled findings of existing VR and non VR solutions as basis for our novel solution.

3 Ideation & Reasoning

Given the limitations of each of the previous solutions, such as either the inability to track real-time impulsive decisions, or reliance on self-reports, or lack of an immersive environment, we decided to create a solution encompassing all these points. The approach systematically incorporates existing reliable and validated games within the research field on addiction and impulsivity as well as elements from the questionnaires (BIS-11, UPPS, and BFI-2). In this chapter we will showcase the proposed project and argument the validity of the design and correlation of our extracted measurements.

3.1 State Of The Art Solution

This project presents a novel approach to behavioral assessment. It offers a customized VR unbiased decision-making trial. Given the new VR game we aim to extract certain values. The extracted data is then fed to a prediction model, which will make sense of these values and map them to psychometric scores. Unlike traditional methods, which rely on self-reported measures which may be prompt to biases, or are only significant in clinical setup, we deviated from a choice based scenario or directly interpreting the data. As a starting point we opted to predict the exact same scores of the well-established German version of the questionnaires BIS-11 30 items, BFI-2 and UPPS. After predicting these scores we will try to minimize the difference between the predicted scores and the actual scores answered by the participants. Henceforth, the model will be refined using different techniques, then checked for improvements at later stages. Therefore we partitioned our project into phases. Phase one focusing on data collection and predicting the scores with the model as-is. Phase two focusing on refining the predictive model aiming to surpass existing assessment tools in accuracy and profiling. In figure 4.2 we showcase the four parts constituting our project. The VR game responsible for extracting relevant data to feed the prediction model. The prediction model to predict thirty different scores. Please refer to table 5.1 for the extracted scores. The Digitilized questionnaire app, to collect the actual scores from the questionnaires BFI-2, UPPS, BIS-11 German versions and finally the refinement of the model.

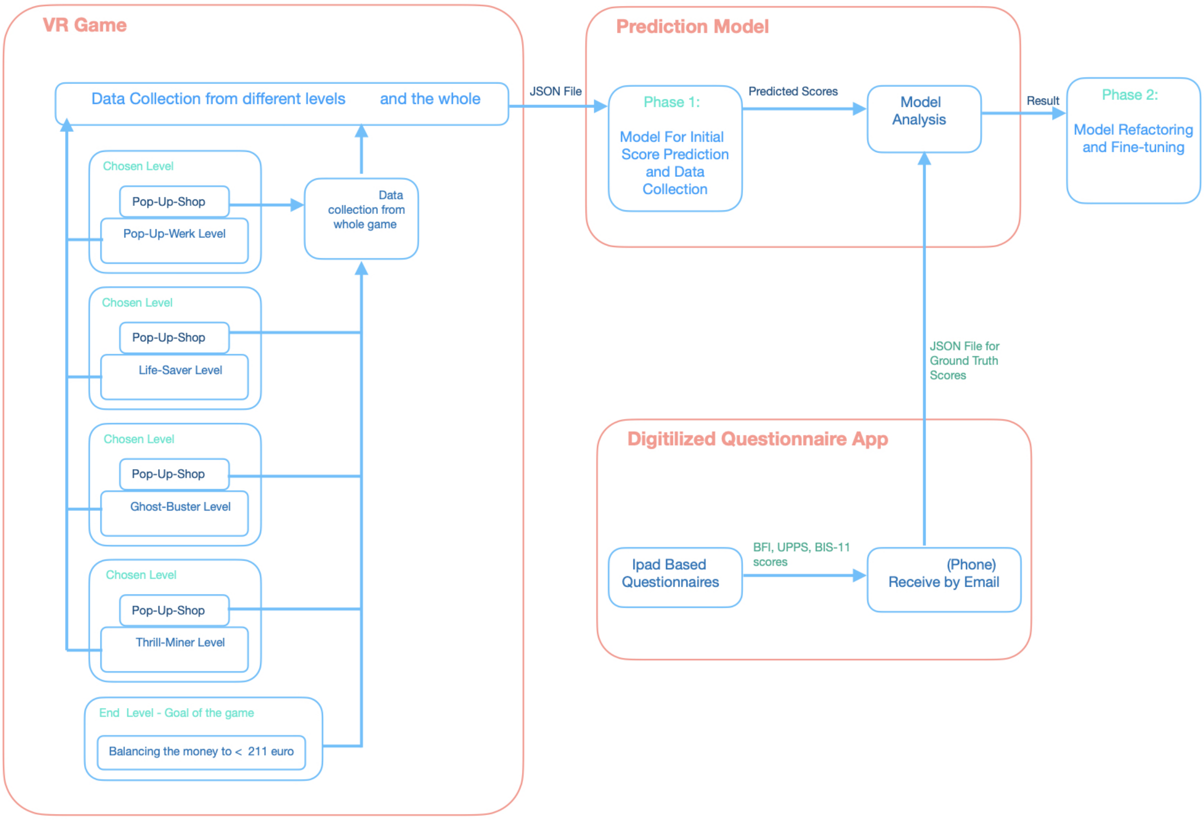


Figure 3.1: Caption

3.1.1 Validity of Chosen Measure Clusters

Each task was designed based on correlated variables from previous studies on impulsivity, decision-making and risk-taking [69, 70]. Data such as fixation time, reaction latency, gaze shifts, decision patterns and strategy adaptation during the game are general clusters of the measured variables. Fixation time has been repeatedly studied to assess attentional bias and reward sensitivity. Prolonged fixation on reward cues correlates with heightened craving and relapse risk. Reaction latency correlates to cognitive processing speed and impulse control [71]. Delayed reaction times often linked to impaired executive function in highly impulsive individuals. Gaze shifts between reward and non-reward stimuli correlates to decision-making biases[72, 25]. Individuals with SUD or gambling tendencies, when faced with high-risk choices and reward elements exhibit faster engagement and slower disengagement [9]. Decision patterns reveal rationality under uncertainty. Impulsive individuals exhibit sub-optimal choices even when given extensive decision-making time. Strategy adaptation is a metric for cognitive flexibility and learning factor assessment. Poor strategy shifts correlate with rigid or compulsive behaviors [24, 25]. These points highlight the validity of our extracted measure variables for further processing.

3.1.2 Prediction Model

Due to the novelty of this approach, no pre-existing dataset was available for training. Thus, in this first step, we rely on **gpt-4o-mini** to calculate initial correlations between measured variables and psychometric questionnaire scores. Then we train a correlation learner on similar research papers for domain adaptation [73], our extracted in-game measures and the scores of the output. At the end we readjust the initially predicted scores given the learned correlations. This approach was chosen over other methods which could be found in [53, 52, 74, 75] as initial testing phases led to always zero-value predictions due to the lack of large labeled dataset. Furthermore our approach align with research on eye-tracking and cognitive processing emphasizing the integration of multi-step prediction [76, 77].

3.2 General Game Idea Validity

The overall game is designed to assess cognitive flexibility, rational long-term decision-making [69, 24, 25], adaptation to ambiguity [78] and strategic planning under constraint. The trial measures their ability to manage short-term gains vs long-term balance [79, 80]. Aiming to extract values which could directly correlate to their capacity for delayed gratification, risk mitigation, and cognitive complexity [43]. To avoid cognitive fatigue and boredom the trial includes different short tasks [81]. The trial has a global goal mimicking complex tasks, and small immediate goals mimicking simple tasks [82, 83]. Measure of performance is derived from both these goals as individuals with higher modularity perform worse on complex tasks yet better on simple tasks [84]. Additionally we introduce a leader-board for the competitive element. Aiming to extract values which could directly correlate to social influences on choices [85], risk-taking tendencies, and the ability to remain rational under external pressure.

3.2.1 Logical Setup

This trial is structured as a competition, denoted by a leader-board ranking all winning participants. A single player have the possibility to claim their accumulated in-game monetary sum given certain conditions. Each player starts with a balance of two thousand euros (2000€). The game consists of multiple timed tasks denoted as levels. After completing a level a shop will appear for a limited time, offering different items to purchase. Each level is designed to evaluate certain behavioral traits, decision-making strategies, and risk management skills. Please refer to the respective sections below for detailed description of these levels. At the start of each level its respective objective is stated once, with explicit rewards and penalties but also with ambiguous statements hinting to implicit penalties. Players must fulfill the goal of each level while mentally keeping track of their game sum balance. The ultimate goal is to end this trial with the highest monetary sum among all participants that is under two hundred eleven euros (211€).

3.2.2 Pop-Up-Shop Validity

The Pop-Up-Shop serves as a controlled environment to measure impulsive spending tendencies and reward sensitivity [25]. Players are presented with a limited-time purchasing scenario where they must choose between alcoholic and non-alcoholic beverages [86]. Both options are priced equally, removing cost as a decision factor. Aiming to analyze preference-based tendencies. The speed of purchase is also tracked. Aiming to extract variables which could correlate directly to reward-driven behaviors and real-life impulsivity traits[79].

3.2.3 Derived Measurements

Please refer at the end of this chapter to tables 3.3, 3.4, 3.7 and 3.6 for extracted values from the respective levels. Here in table 3.1 we showcase the additional values that are being measured consistently from the start of the trial until the end. For additional information regarding exact scene layouts, logical elements and specific extracted measurements, please refer to the Supplementary Material.

3.3 Pop-Up-Werk Level Validity

In Pop-Up-Werk, we track reward sensitivity through unconscious attentional fixation on alcoholic versus non-alcoholic stimuli [86]. Sensation seeking and risk preference are inferred from the selection frequency of high-reward versus lower-risk slot machines [19, 25, 87]. Decision-making tendencies are assessed by analyzing spending behavior and impulse-driven purchases [69, 36, 71, 70]. Impulsivity and adaptation are measured through slot machine persistence, money tracking, and response to penalties and gains [46, 47, 80]

Table 3.1: Measured Variables throughout the whole game

Variable Name	Type	Description
InstructionPanelOpen CountPerLevel	List<int>	Number of times the instruction panel was opened during the game.
TimeRequiredTo Purchase	List<float>	Reaction time to buy items from the pop-up shop (Each time).
SafeAccountBalance PerLevel	List<float>	Accumulated monetary amount in each level denoted as the safe account balance.
AlcoholicItems Purchased	int	Number of alcoholic items purchased during the game
NonAlcoholicItems Purchased	int	Number of non-alcoholic items purchased during the game.
GameAccountBalance	float	The final monetary balance if it is under 211€.

3.3.1 Goal & Scene Setup

This level is a VR version of the traditional CGT, used to assess decision-making in addiction related research [37, 79, 69, 70]. Our objective is to simulate real-world gambling behaviors. The goal of this level is to accumulate and subtract reasonable monetary amounts through slot machine playing and pop-up-shop purchase. Players must purchase coins deducted from their final balance automatically to play any slot machine. This level includes four different types of slot machines. Players must use coins to play any slot machine. Different types of slot machines cost different amounts of coins and present different winning probabilities and winning sums.
The Diamond machine offers a one-hundred percent (100%) winning probability of one thousand euro (1000€) and costs only two coins per spin.
The Bill machine offers a twenty-five percent (25%) chance of winning fifty euro (50€), twelve-point-five percent (12.5%) chance of winning two hundred euro (200€), twelve-point-five percent (12.5%) chance of winning ten euro (10€), twelve-point-five percent (12.5%) chance of winning twenty euro (20€) and thirty-seven-point-five (37.5%) chance of winning nothing and costs two coins per spin.
The Cake machine offers a seventy percent (70%) chance of winning eight euro (8€) and costs five coins per spin.
The Mystery machine offers a seventy-five percent (75%) chance to randomly playing

one of the other slot machines (Diamond, Bill, or Cake) with their respective winning probabilities and payouts, twenty-five percent (25%) of no winnings and costs one coin per.Please refer to table 3.2 for a compiled feature description of our slot machines and which scenario in addition to the type of player is being targeted by each of them.
Finally this level also includes alcoholic beverage ATMs to track the subconscious gaze deviation to reward vs non-reward elements [72]. Gaze fixation on reward-associated stimuli can indicate addictive behaviors and attentional biases in at-risk individuals.
The scene is inspired by traditional German sport's bar aesthetics. It is composed of animated jackpot screens on arcade machines, mainly alcoholic drink ATMs and a bar area. The atmosphere is designed to subtly encourage impulsivity while still allowing room for strategic purchase. Audio cues, such as jackpot sounds, coin drops [82] and ambient German chatter were included to reinforce reward anticipation, emotional investment and familiarity. In Figure 3.2, we showcase some of the described elements within the scene.

Figure 3.2: Image showcasing the different slot machines and the alcoholic beverage ATM of the level Pop-Up-Werk.

3.3.2 Key Elements

The key elements in the Pop-Up-Werk level include: Slot machine preferences, where players must decide which slot machine to engage with based on expected payouts and personal risk tolerance. Attention towards specific visual cues, such as alcoholic beverage distribution machines or carbonated drinks distribution machines is tracked. Aiming to extract values which correlate directly to attentional bias in AUD patients and cue-reactivity.. Two distinct penalties in form of a stated and a hidden penalty. The bar deducts house fees where the player anticipates money reduction. The other hidden penalty is when the total accumulated money within this level exceeds a certain amount and the player can notice it by keeping track of the money counter displayed at all time. Aiming to measure variables that could correlate directly to impulsive reaction to losses and gains, where players either continue uncontrollably spinning the slot machines or adapt.

Table 3.2: Feature comparison of slot machines and their targeted player behaviors.

Slot-Machine	Scenario	Designed to Evaluate	Targeted Player Type
Diamond.	High reward, low cost, low risk	Risk-avoidance and preference for guaranteed rewards	Players seeking high returns with minimal investment
Bill.	Consistent but low reward, low cost, cumulative success rate	Preference for steady gains and strategic planning	Players with a conservative strategy
Cake.	Moderate risk and reward	Willingness to take controlled risks	Players balancing between risk-taking and safe choices
Mystery.	Unpredictable outcomes, extreme losses or gains	Thrill-seeking and tolerance for randomness	Players drawn to uncertainty and gambling tendencies

3.3.3 Derived Measurements

Please refer to 3.7 at the end of this chapter for Pop-Up-Werk level extracted behavioral measures. For additional information about the exact scene layout and exact measurements that has been extracted please refer to the Supplementary Material.

3.4 Life Saver Level Validity

In Life Saver, we measure decision-making through the prioritization of rescue cases under time and resource constraints [86]. Empathy and emotional responses are assessed based on the selection of highly critical or lower-priority cases and the type of case. Stress management and emotional control are analyzed through performance under time pressure [88]. Cognitive flexibility and rationality is measured through resource allocation efficiency across different cases [89]. Awareness is measured given the strategy pattern [83], distinguishing between emotional and strategic decision-making [90]. Avoidance behavior and apathy is inferred through skipped cases..

3.4.1 Goal & Scene Setup

The goal of this level is to distribute limited resources and non limited resources within a limited time frame, if the player wishes to. The player is presented with eight case, four human and four animal cases requiring different levels of medical attention. Each case has a survival timer indicating the urgency of intervention. The player must make quick yet rational decisions on how to allocate replenishable oxygen, non-replenishable antidotes, replenishable blood supply and speak soothing words into microphone. A player can assign only one of each required resources within thirty seconds. Each time a single resource is allocated an extra ten seconds is added to the timer of the specific case. Each case has a unique combination of needed resources. Deliberately, it was made impossible to save all the eight cases given the limited resources. A rational decision may differ in definition from individual to another [90, 89]. Therefor the player can follow different prioritization strategies in form of combination. Here we state all the possible strategies:

- Not attending to any case
- Maximizing over critical animal cases
- Maximizing over critical human cases
- Maximizing over non critical human cases
- Maximizing over non critical animal cases
- Maximizing over non critical animal while also maximizing over critical human cases
- Maximizing over non critical human while also maximizing over critical animal cases
- Optimizing resources over human cases
- Optimizing resources over animal cases
- Optimizing resources over critical human cases
- Optimizing resources over critical animal cases
- Optimizing resources over non critical human cases

- Optimizing resources over non critical animal cases
- Optimizing resources over non critical human and critical animal cases
- Optimizing resources over non critical animal and critical human cases

The scene is set in a bunked-like medical experimental facility. Each case is presented as a cell with UI-Menu for interaction and displaying relevant informations. A dark atmosphere, talking personnel and emergency cues at the end amplifying the sense of urgency. Countdown timers are present to increase stress and force intuitive decision-making.

3.4.2 Key Elements

The key elements in the Life Saver level include: Resource Allocation Strategy, where players must determine how to distribute limited medical supplies effectively across multiple cases. Aiming to extract values which correlate directly to distress and pressure tolerance.. Prioritization of case types under time pressure, where time-to-live timers are present. Aiming to extract values which correlate directly to decision-making and real-time adaptability. Players must remember the order of which they attended to previous cases. Aiming in extracting values which correlate directly to cognitive Load Assessment.
In Figure 3.3, we showcase some of the described elements within the scene.

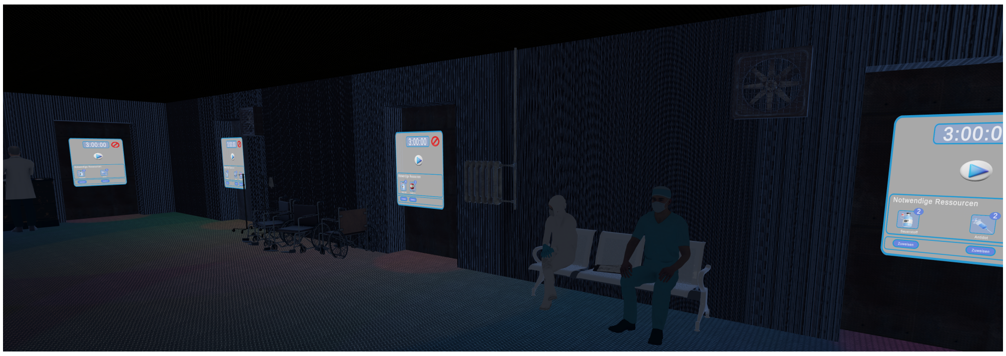


Figure 3.3: Image showcasing the different cases and their UI panel with required resources of the Life-Saver level.

3.4.3 Derived Measurements

Please refer to 3.6 at the end of this chapter for an example of Life Saver level extracted behavioral measures. Please note that this level has eight different cases so the example of the extracted variables is consistent throughout all the eight cases within the game. By CA or CH we denote Case Human or Case Animal. The strategy contains eight aspects please refer to Chapter **??** for details about how the strategy was derived. For additional information about the exact scene layout and exact measurements that has been extracted please refer to the Supplementary Material.

3.5 Ghost Buster Level Validity

In Ghost Buster, we measure attention and concentration through reaction times when identifying and shooting the correct ghosts. Cognitive flexibility is assessed through adaptation to changing rules across the different phases denoted as quests and reaction times to different ghost types. Sensitivity to distractions is also analyzed based on the impact of random stimuli on accuracy. Impulsivity and control mechanisms are reflected in incorrect ghost eliminations or rushed decisions [70, 69]. Emotional responses such as stress or agitation under time pressure are measured based on performance. Learning ability and strategy adaptation are evaluated through improvements in accuracy and reaction time across different rules.

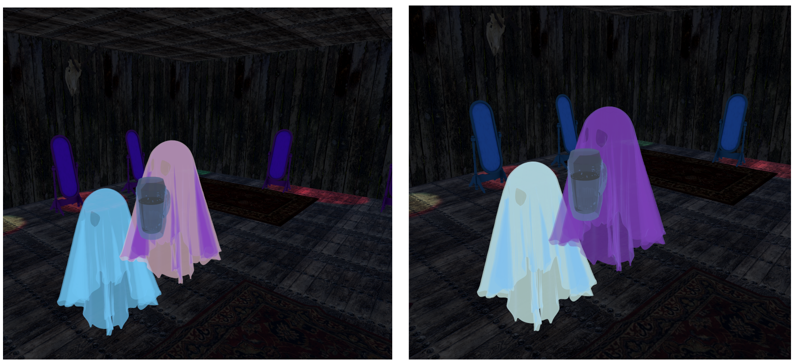
3.5.1 Goal & Scene Setup

The goal of this level is to eliminate correct targets under progressively increasing distractions and where previously relevant informations become irrelevant. This task is a VR version of the classical color-dot trials. Throughout the whole level, four different types of ghosts appear randomly. The ghosts rotate inside the enclosed room and disappear after a full rotation is completed or when it was busted by the player. This level is composed of four phases denoted as quests. At the start of each phase instructions are said regarding which type of ghost the player must bust. During the first and second phases, in addition to the auditory instructions elements inside the room will also indicate which ghosts to bust. When a player busts a correct ghost a sound indicating money has been added is played [82]. On the other hand when they bust the wrong ghost a sound indicating failure is played [82]. However during the third and fourth phases, the auditory cues of either correctly or wrongly busting a ghost and the elements within the room will become irrelevant. These sounds and element's colors will change randomly. The time interval of changing these features will decrease from phase three to phase four. Evidently the appearance rate of the ghosts will also gradually increase from phase one to phase four.

The scene is designed to invoke distress with its dark ambient feeling. Although invoking fear was the main distraction and distress factor was were oping for, we had to account for different phobias (Nyctophobia, Claustrophobia, Teratophobia, Thanatophobia). To that regard the degree of creepiness was adjusted. Additionally, to keep the focus of the player on the targets only, the ghosts were designed with cartoony and neon colored aesthetics. Ghosts are classified as either purple or blue and either holding an alcoholic beverage or not. The addition of the alcoholic beverages is merely a distraction and do not contribute in any shape or form to our extracted values regarding attraction to reward elements.

3.5.2 Key Elements

The key elements in Ghost Buster include: Target identification under distraction, where players must distinguish between correct and incorrect ghosts under switching visual and auditory conditions. Reaction times to busting a correct ghost, and if this average value gets progressively better or worst during a certain and consecutive phase, are values recorded during this level. Aiming to have values which could directly correlate to decision making speed, adaptability and error-based learning. First two phases use direct feedback while the last two use randomized feedback. Aiming to extract values which could correlate to cognitive resilience. Eye-tracking data, reaction consistency, and movement stability are analyzed throughout the whole level. Aiming to extract values which could correlate to impulsivity and frustration tolerance. In the figure 3.4, we demonstrate key scene elements of the level Ghost Buster.



(a) Gaze fixed on the purple ghost holding beverage and room elements indicating target is the purple ghost

(b) Gaze fixed on the blue ghost not holding beverage and room elements indicating target is the blue ghost

Figure 3.4: Images showcasing different ghosts inside the scene of the Ghost-Buster Level.

3.5.3 Derived Measurements

Please refer to the end of this chapter for the table 3.4 and table 3.5 containing the extracted values from this level during the game. For additional information about the exact scene layout, the exact measurements that has been extracted please refer to the Supplementary Material.

3.6 Thrill Miner Level Validity

In Thrill Miner, spontaneity and risk-taking are assessed through the player's choices between high-reward, high-risk paths and medium-reward medium-risk routes [18, 69, 86, 87]. Novelty-seeking is evaluated by tracking the selection of unknown paths and mystery rewards [13, 17]. Perseverance is measured based on the player's ability to complete a chosen path despite increasing difficulty or repeating it after getting familiar with it [45]. Rationality and decision-making are analyzed through the collection of rewards along a path, the ability to chose a different path if his trials were not fruitful and the time spent on already visited paths (detect if there was a learning factor)[20]. Emotional responses such as excitement, pressure,and frustration are inferred from performance across a chosen path, the tendency to either repeat, abandon or partially switch a path under limited time [89]. In the figure below we illustrate all the possible paths of this level. We denote a level as node R_xP_x where R is the reward grad ranging from 1 to 4 and P denotes the penalty grad ranging from 1 to 4, with 1 denoting least amount and 4 denoting maximum amount. Different paths connecting nodes denoted as arrows $R_1P_1 \rightarrow R_4P_4$.

3.6.1 Goal & Scene Setup

The goal of this level is to reach an end-level we denote as node R_4P_4 and survive it during the limited time trial. To reach this node the player can chose different paths. The length of each path, hence the number of nodes between first and end-node is unknown to the player. The player will only be notified ones they reach the end node. Different rewards, penalties and surprise elements are along these paths. The difficulty of the path is determined by the size of its reward and penalty or by the drive for the unknown namely a path with mystery rewards. The goal of this level is still consistent with the general game goal which is collecting reasonable amount of money to balance their account after the trial is over. This level has different grades of penalties. The first penalty is in the form of death traps along a path which forces the player to repeat a single level. The second penalty is forcing the player to repeat from the starting point (the cave) if they fails a single level thrice. The third and last penalty is stripping off the whole monetary amount from the account if the player fails the last level.

The scene is designed as a crossover between the Tarzan game 1999 and the Mario Bros game 1983 but as a VR version. The player starts in a cave with transparent walls where they can see possible upcoming paths. This was a deliberate choice to induce excitement and thrill. The player will have to choose between mystery-level-of-difficulty paths or gradual increase of reward and difficulty path or fast increase of reward and penalty path. The players will be faced with the choice of different paths next to each other to arrive to a teleport way-point leading to the next path. Placing the path next to each other was also deliberate so the player can make an informed decision if they fail this path and therefore we would be able to test if there was a learning factor.

Natural environment elements (trees, rocks, snowy rocks, bridges), glowing transparent materials, floating diamonds, hanging ropes and sound cues[82] are used to create a dynamic setting that builds suspense and influences impulsive behaviors.

3.6.2 Key Elements

The key elements in the Thrill Miner level include: Risky vs. Safe Paths mimicking high-risk to high-reward gambling games. Stable and gradual increase of penalty to reward ratio paths. Aiming to extract values which could directly tests risk-taking behavior. Paths with uncertain outcomes that could either increase or decrease the total sum. Aiming to extract values which could directly assess novelty-seeking behavior. All the different paths progressively increases in difficulty with obstacles that require rapid responses. Aiming to extract values which could directly evaluate perseverance under pressure. Due to time constraints and the player's uncertainty about the final node's distance, each path require rapid decision making. Aiming to extract values which could directly measure impulsivity and real-time decision-making skills. Players must at all time keep count of the collected rewards, to balance gains and losses strategically.

3.6.3 Derived Measurements

The Thrill Miner level extracted behavioral measures.
By node we denote R_xP_x and path the sequence of $R_xP_x \rightarrow R_xP_x \rightarrow R_xP_x \rightarrow R_4P_4$ where R_4P_4 is the last level with highest reward and penalty grad. At The end of this chapter please refer to 3.3 for insights about the extracted values during this level. For additional information about the exact scene layout and exact measurements that has been extracted please refer to the Supplementary Material.

Table 3.3: Measured Variables in Thrill Miner

Variable Name	Type	Description
FullyChosenPath	string	The full path selected when a player reaches the final node without abandoning previous choices (including failed trials).
DroppedPath	string	Path where a player initially chose but later abandoned before reaching the final node.
NotVisitedPath	string	Paths that were available but not selected by the player.
PercentageOfProgress Path	float	The percentage of completion along a chosen path.
IncreaseOrDecrease LearningFactorPath	float	The ratio of time of each consecutive node, along the visited nodes within a single path
VisitCount	int	The number of times a specific path was visited by the player.
avgSpentTime	float	The average time spent per node in a given path.
GetTimeChange Ratio	float	The ratio of time variation between consecutive attempts on the same node.
remainingTrials	int	Number of attempts left before a node is locked out.

Table 3.4: Measured Variables in Ghost Buster

Variable Name	Type	Description
q1AvgReactionTimeToBustAnyGhost	float	Average reaction time to bust any ghost in phase 1.
q1AvgReactionTimeToBustCorrect Ghost	float	Average reaction time to correctly bust the target ghost in phase 1.
q1AvgReactionTimeToBustWrong Ghost	float	Average reaction time when mistakenly busting the wrong ghost in phase 1.
q1AvgTimeGazeOnCorrectGhost	float	Average gaze time on the correct ghost in phase 1.
q1AvgTimeGazeOnWrongGhost	float	Average gaze time on the wrong ghost in phase 1.
q1AvgNumberCorrectBusted Ghosts	float	Number of correctly busted ghosts in phase 1.
q1AvgNumberWrongBusted Ghosts	float	Number of mistakenly busted ghosts in phase 1.
q12AvgReactionTimeToBust AnyGhost	float	Combined average reaction time tobust any ghost across phase 1 and 2.
q12AvgReactionTimeToBustCorrect Ghost	float	Combined average reaction time to bust correct ghosts across phase 1 and 2.
q12AvgReactionTimeToBustWrong Ghost	float	Combined average reaction time to bust wrong ghosts across phase 1 and 2.
q12AvgTimeGazeOnCorrectGhost	float	Combined average gaze time on correct ghosts across phase 1 and 2.
q12AvgTimeGazeOnWrongGhost	float	Combined average gaze time on wrong ghosts across phase 1 and 2.
incDecMarginReactionTimeAny Q3To21	float	Change in reaction time to bust any ghost from phase 3 to 2-1 average.
incDecMarginReactionTimeCorrect Q3To21	float	Change in reaction time to correctly bust a ghost from phase 3 to 2-1 average.
incDecMarginReactionTimeWrong Q3To21	float	Change in reaction time when busting the wrong ghost from phase 3 to 2-1 average.

Table 3.5: Measured Variables in Ghost Buster

Variable Name	Type	Description
incDecMarginNumberBustAny Q3To21	float	Change in number of any ghost busts from phase 3 to 2-1 average.
incDecMarginNumberBustCorrect Q3To21	float	Change in number of correct ghost busts from phase 3 to 2-1 average.
incDecMarginNumberBustWrong Q3To21	float	Change in number of wrong ghost busts from phase 3 to 2-1 average.
incDecMarginReactionTimeAny Q4To3	float	Change in reaction time to bust any ghost from phase 4 to phase 3.
incDecMarginReactionTimeCorrect Q4To3	float	Change in reaction time to correctly bust a ghost from phase 4 to phase 3.
incDecMarginReactionTimeWrong Q4To3	float	Change in reaction time when busting the wrong ghost from phase 4 to phase 3.
incDecMarginNumberBustAny Q4To3	float	Change in number of any ghost busts from phase 4 to phase 3.
incDecMarginNumberBustCorrect Q4To3	float	Change in number of correct ghost busts from phase 4 to phase 3.
incDecMarginNumberBustWrong Q4To3	float	Change in number of wrong ghost busts from phase 4 to phase 3.

Table 3.6: Measured Variables in Life Saver

ariable Name	Type	Description
ProgressOfSavingWithinCH4	float	Percentage of completion for saving the critical human case 4.
TimeSpentOnCH4	float	Total time spent on critical human case 4.
CH4DiscoveredAndSaved?	string	Indicates whether critical human case 4 was discovered and saved.
CH4DiscoveredAndKilled?	string	Indicates whether critical human case 4 was discovered but not saved.
CH4WatchedVideo	string	Indicates whether the instructional video for CH4 was watched.
ProgressOfSavingWithinAStrategy	float	Average percentage of completion across all cases.
AvgOfProgressOfHumanCases	float	Average completion percentage for all human cases.
AvgOfProgressOfAnimalCases	float	Average completion percentage for all animal cases.
Strategy	string	The player's overall strategy in managing resources and prioritizing cases.

Table 3.7: Measured Variables in Pop-Up-Werk

Variable Name	Type	Description
CoinsQuantity	float	Total number of coins collected by the player.
CoinsBoughtCount	float	Number of times the player purchased additional coins.
MysterySlotPlaysPerMinute	float	Frequency of playing the mystery slot machine per minute.
DiamondSlotPlaysPerMinute	float	Frequency of playing the diamond slot machine per minute.
BillSlotPlaysPerMinute	float	Frequency of playing the bill slot machine per minute.
AvgFixationTimeAlcoholic Displays	float	Average fixation time on alcoholic-related displays.
AvgFixationTimeAlcoholicVs Non-Alcoholic	float	Ratio of fixation time between alcoholic and non-alcoholic displays.
DiamondSlotPlayCount	float	Total number of times the player used the diamond slot machine.
BillSlotPlayCount	float	Total number of times the player used the bill slot machine.
CakeSlotPlayCount	float	Total number of times the player used the cake slot machine.
MysterySlotPlayCount	float	Total number of times the player used the mystery slot machine.
AvgTimePlayingDiamondMachine	float	Average duration spent playing the diamond slot machine.
AvgTimePlayingBillMachine	float	Average duration spent playing the bill slot machine.
AvgTimePlayingMysteryMachine	float	Average duration spent playing the mystery slot machine.
AvgTimePlayingCakeMachine	float	Average duration spent playing the cake slot machine.
AvgStagnantTime	float	Average idle time when the player is not engaging with any slot machine.

4 VR Game Implementation

In the previous chapter, we presented the idea behind each task and how each task (Level) is connected together to deliver the whole game. The following chapter outlines in detail how we implemented the presented idea, the chosen solution and algorithms given best practices and keeping performance as our most important goal.

4.1 Hardware

To implement the proposed solution we used the following hardware, please refer to chapter 7 where we discuss why our choice of VR-Headset was not the most suitable given our solution. Razer Blade 14 (RZ09-0370) – Windows 11 Home Edition (Version 10.0.26100, Build 26100), AMD Ryzen 9 5900HX (3.3 GHz, 8 cores), 16GB RAM, and NVIDIA GeForce RTX 3060 GPU; Pico Neo 3 Eye Pro VR headset – XR2 Qualcomm Snapdragon processor, 6GB RAM, and Adreno 650 GPU;

4.2 Software & Patterns

The game was created using: Unity 2022.3.54f1 LTS; Recommended Android SDK and NDK installed with Unity; XR Plugin Management 4.4.0; XR Legacy Input Helpers 2.1.11; XR Interaction Toolkit 3.0.3; XR Core Utilities 2.3.0; OpenXR Plugin 1.13.1; Input System 1.11.2; PICO Integration 2.5.3;
To implement a seamless experience the following best practices were used: Observer pattern (UnityAction<>.invoke not UnityEvent) for sending tracked values to be measured and handling every communication between all the different scripts; Singletons for every script handling the participant, overall game and data calculation; Scriptable objects for most of the prefabs including ghosts and Pop-Up-Shop items; Factory pattern for every scriptable object; Task-Based asynchronous programming and utility class (helper) pattern for calculating the measured values and writing them to the exported JSON file; Coroutine-based state management instead of Update() function to have a more controlled per-frame behavior handling timed tasks (works better with invokes than simple update). Procedural scene generation with additive scene loading to reduce overall rendering load.

4.3 Whole Game & Pop-Up-Shop

The game's architecture is designed to measure participant specific behaviors across all levels, purchases and game progress. The implementation follows all the above mentioned best practices.

Dynamic Level Customization & Participant Assignment

The architecture supports participant-specific settings, levels order and each level duration. Please refer to supplementary material for screenshots and videos regarding the familiarization scene and all the customization-setup screens.

(a) Role choosing screen

(b) Participant choosing screen

Figure 4.1: Images from the initial familiarization level and setup panels.

Monetary Balance & Purchase Management

The Money Manager and Pop-Up-Shop tracks all in-game purchases and monetary changes. The end level is designed so that participants reduce or increase their game amount to under 211€given their wins from different levels.

Data Persistence & File Management

The architecture serializes and stores the progress and purchases into JSON format. Each JSON file is automatically named using a unique identifier assigned to each participant. To ensure anonymity the filename follows the pattern: **CGXXX.json** with CG or AD as the prefix to specify either Control Group or SUD individuals. XXX are the three-digit unique identifier assigned sequentially per participant (CG001.json, CG002.json, AD001.json). The output JSON contains **10 columns**. Please refer to Figure 4.2 for an overview of the structure of the output file.

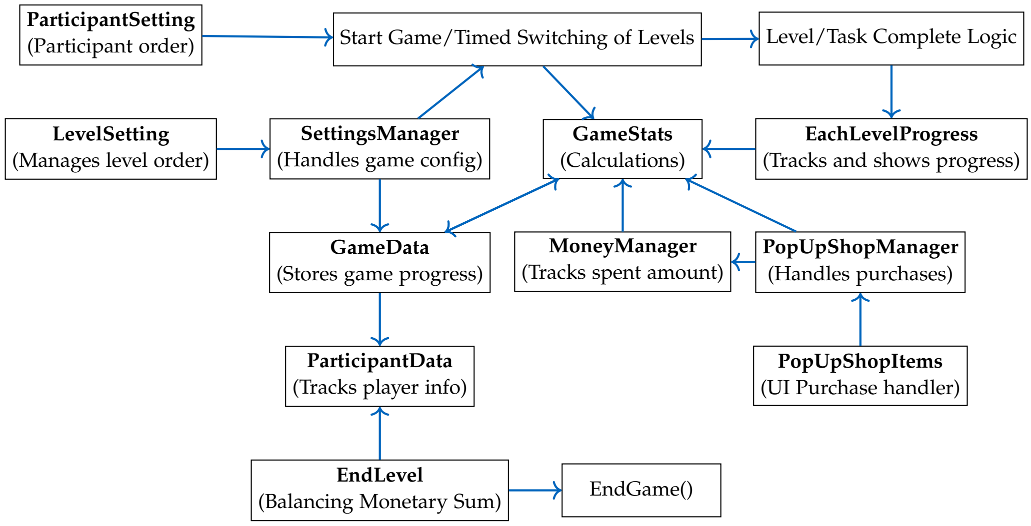


Figure 4.3: Example of persistent scripts workflow from start to end of the game.

4.3.1 Eye-Tracked Components

Eye-tracked components are consistent throughout Ghost Buster level and Pop-Up-Werk level. These components detect, measure and save participant gaze, fixation durations, and interactions with objects classified as either reward-based or non-reward-based.

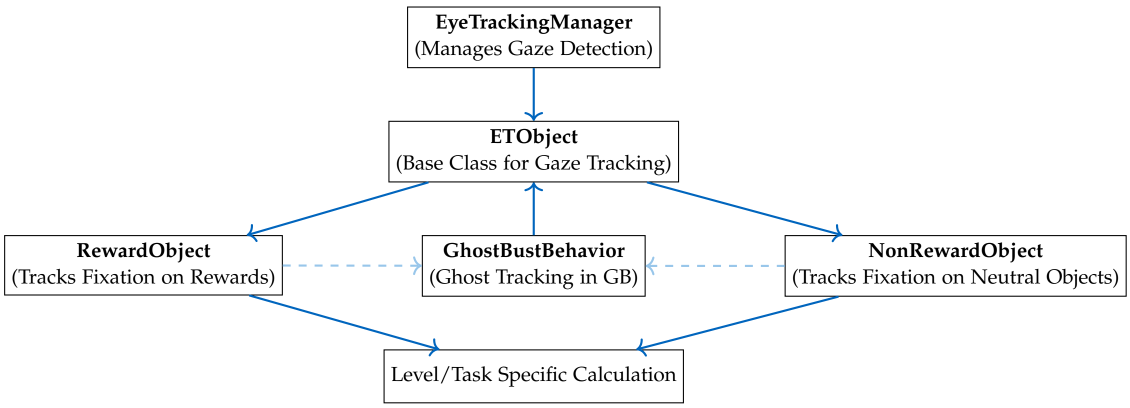


Figure 4.4: Eye tracking scripts workflow

4.4 Pop-Up-Werk

Pop-Up-Werk level is designed to measure participant's behavior when engaging with slot (gambling) machines and involuntary gaze-shift toward alcoholic beverage distributing machines. The Key point of this level is capturing different slot machine interactions while avoiding redundancy and including the possibility of integration of further slot machines.
Therefore we detached the logic of the gameplay of the level from its composing components and detached the data calculation from those components. This logic is applied to each and every level. At the end, we have different scripts each handling a specific slot machine and eye-tracked components. Each script sends tracked values to designated scripts for further calculation. A collective script handling start/end of the level and other information.

- `PUWStats.cs` Static class: computing statistics related to slot usage, play frequency, and fixation times without redundant calculations.
- `PUWData.cs` Persistent singleton listener: updating values collected for each participant and triggering the necessary event listeners for output file.
- Individual slot machine scripts (`BillsSlotBehavior.cs`, `DiamondSlotBehavior.cs`, etc.) managing participant interactions, animations, different probabilities and rewards per slot type.
- Individual eye-tracked scripts objects
- `PopUpWerkManager.cs` Singleton: handling the entire process and level from synchronization, updates, hidden or stated penalties and monetary event triggering.

The collected data is then serialized into the JSON file. Please refer to Table 3.7 for an overview of these variables.
Please refer to figure 4.5 for an illustration regarding the workflow of these scripts.

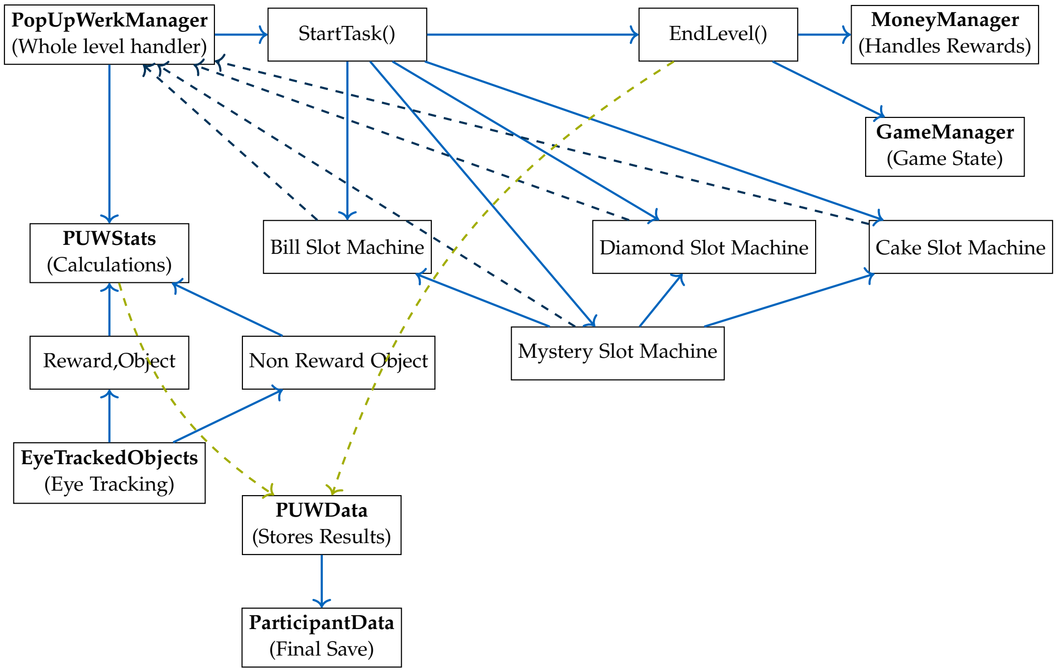


Figure 4.5: Pop-Up-Werk level scripts workflow

4.5 Ghost Buster

Ghost Buster level is designed to measure adaptability and progress of reaction time to fast changing conditions. The level is sequence of four short sub-tasks denoted as quests. Q1 $\rightarrow$ Q2 $\rightarrow$ Q3 $\rightarrow$ Q4. Where each Q_i represents a quest. Each quest requires player to bust a different ghost:

- **Quest 1 (Q1)**: Bust purple ghosts. Elements in room illuminated at the start in purple. Slow spawn rate of ghosts. Incorrect action triggers a sound cue. Correct action triggers a sound cue.
- **Quest 2 (Q2)**: Bust blue ghost. Elements in room illuminated at the start in blue. faster spawn rate of ghosts. Incorrect action triggers a sound cue. Correct action triggers a sound cue.auditory penalty.
- **Quest 3 (Q3)**: Bust ghosts not holding a drink. Element's color changes between purple and blue during task period randomly. Faster spawn rate of ghosts. Sound cues are randomly triggered.

- **Quest 4 (Q4)**: Bust blue ghosts holding a drink. Faster random elements' color change. Faster spawn rate of ghosts. Faster random sound cues triggers.

Following the same architecture as previous level, this level based on event-triggers with gameplay logic decoupled from gameplay element's behaviors and data handling:

- `GBStats.cs` Static class: computing each quest statistics and learning effects for consecutive quests.
- `GBData.cs` Persistent singleton: updating values collected for output file.
- `GhostBustBehavior.cs`: managing participant interactions and ghost animations.
- Individual eye-tracked scripts objects
- `GhostBusterManager.cs` Singleton: handling the entire process and level from synchronization, updates, hidden or stated penalties and monetary event triggering.

The collected data is then serialized into the JSON file. Please refer to **??** for an overview of these variables.
Please refer to figure 4.6 for an illustration regarding the workflow of these scripts.

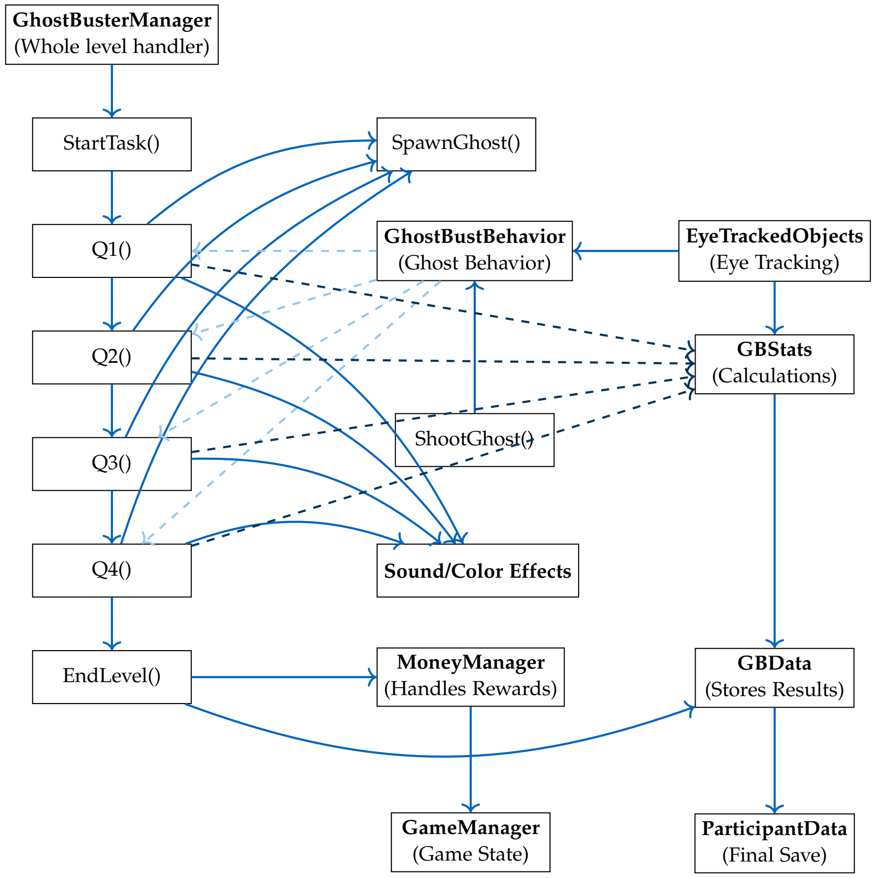


Figure 4.6: Ghost Buster level scripts workflow

4.6 Life-Saver

Life-Saver level is designed to measure the rational prioritization under constraints. It presents a decision-making problem under uncertainty, where players must allocate limited resources to save as many experimental cases as possible within a fixed time frame while considering the impact of their decisions on the other cases. Classical approaches in optimization, such as integer programming or combinatorial search can model this scenario perfectly but are computationally expensive, making them not suitable for a real-time game.
Instead of solving an optimization problem (creating a computational overhead) we found a simpler solution based on **first-order logical equations and classification**.

Formal Strategy Definitions

Each participant's behavior can be partitioned into multiple Clusters. We distinguish between three different clusters.
Optimization-type cluster. Case preference cluster. Prioritization cluster.
Each cluster is composed of multiple sub-strategies.
Each sub-strategy is formulated as a first-order logical equation.
Each selected and visited case belongs to the **set S** of all possible cases which can be explored in each specific sub-strategy. Note with CH_i: we denote Case Human from 1 to 4 and with CA_i we denote Animal Case from 1 to 4:

· **Critical Cases Optimization:** Cases with the shortest time-to-survive.

$$
\begin{aligned}
S_c = \big\{(CH_4 \wedge CA_4) \wedge \big[\\
&\big(CH_1 \wedge (CH_2 \vee CH_3)\big) \\
&\vee \big((CA_3 \wedge CA_2) \wedge CH_1\big) \\
&\vee \big(CA_1 \wedge (CA_3 \vee CA_2)\big)\big]\big\}
\end{aligned}
\tag{4.1}
$$

Formulated output: *"Individual generally maximized cases with least time to survive."*

· **Non-Critical Cases Optimization:** Cases with the longest time-to-survive.

$$
\begin{aligned}
S_{nc} = \big\{(CH_4 \wedge CA_4) \wedge \big[\\
&\big((CH_3 \wedge CH_2) \wedge (CA_3 \vee CA_2)\big) \\
&\vee \big((CA_3 \wedge CA_2) \wedge (CH_3 \vee CH_2)\big)\big]\big\}
\end{aligned}
\tag{4.2}
$$

Formulated output: *" Individual generally maximized cases with longest time to survive."*

· **Resource Optimization:** Maximizing used resources and saved cases count.

$$
\begin{aligned}
S_{op} = \big\{(CH_4 \wedge CA_4) \wedge \big[\\
&\big((CA_3 \wedge CA_2) \wedge (CH_1 \vee CH_2 \vee CH_3)\big) \\
&\vee \big(CA_1 \wedge (CA_3 \vee CA_2)\big) \\
&\vee \big((CH_3 \wedge CH_2) \wedge (CA_2 \vee CA_3)\big)\big]\big\}
\end{aligned}
\tag{4.3}
$$

Formulated output: *"Individual optimized the allocation of resources to cover most of cases."*

· **No optimization:**

Formulated output: *"Unstructured! Individual did not optimize the allocation of resources to cover most of cases. Individual did not optimize on longest time to survive of cases. Individual did not optimize on least time to survive of cases."*

- **Animal Cases Preference:** Exploring on average more animal cases.

  Formulated output: *"Individual generally prioritized Animal cases over Human cases."*

- **Human Cases Preference:** Exploring on average more human cases.

  Formulated output: *"Individual generally prioritized Human cases over Animals cases."*

- **No Preference:**

  Formulated output: *"Individual did not distinguish nor prefer Animal cases from human cases."*

- **Animal over Human Prioritization:** Exploring critical animal cases with less critical human cases.

$$
\begin{aligned}
S_{cha} = \big\{ &(CH_4 \land CA_4 \land CA_3 \land CA_1) \\
&\lor (CH_4 \land CA_4 \land CA_2 \land CA_1) \\
&\lor (CH_4 \land CA_4 \land CA_3 \land CH_3 \land CH_2) \\
&\lor (CH_4 \land CA_4 \land CA_2 \land CH_3 \land CH_2) \big\}
\end{aligned}
\tag{4.4}
$$

  Formulated output: *"Individual did chose longest time to survive for human case and least time to survive for animal cases"*

- **Human over Animal Prioritization:** Exploring critical human cases with less critical animal cases.

$$
\begin{aligned}
S_{cah} = \big\{ &(CH_4 \land CA_4 \land CA_3 \land CA_2 \land CH_1) \\
&\lor (CH_4 \land CA_4 \land CH_2 \land CH_3) \big\}
\end{aligned}
\tag{4.5}
$$

  Formulated output: *"Individual did choose longest time to survive for Animal case and least time to survive for human cases"*

Then the final strategy output is a textual output **containing** (by progressively adding) each followed sub-strategy's **formulated-output** into the final string.

$$
S_{\text{final}} = \bigcup_{i=1}^{n} S_i \tag{4.6}
$$

where each S_i represents a chosen sub-strategy. In a similar manner the logic of the level is decoupled from the components composing it:

- `LSStats.cs` Static class: calculating statistics related to progress per case and chosen strategy.
- `LSData.cs` Persistent singleton: updating values for the output file.
- Case and different resource types handling scripts (`Case.cs`, `ResourceElementCase.cs`, `ResourceElementParticipant.cs`etc.) managing participant interactions, resource replenishing and tracking state of progress withing each case.
- `LifeSaverManager.cs` Singleton: handling the entire process from data synchronization, updates, hidden and stated penalties and monetary event triggering.

The collected data is then serialized into the JSON file. Please refer to 3.6 for an example overview of these variables.
Please refer to figure 4.7 for an illustration regarding the workflow of these scripts.

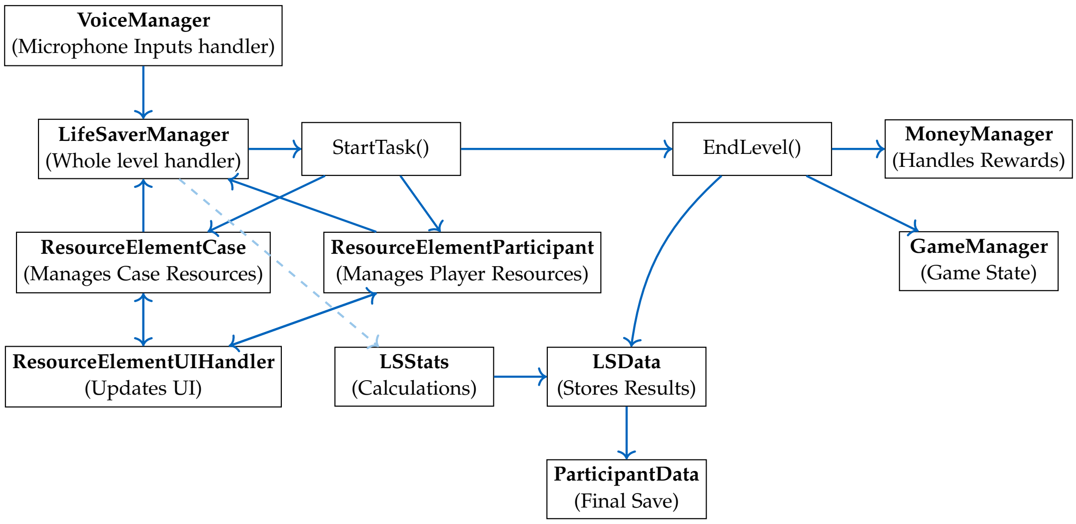


Figure 4.7: Life Save level scripts workflow

4.7 Thrill Miner

Thrill Miner level is designed to measure learning effects to increasing rewards and risks from chosen decision. It presents a procedurally populated and instantiated (not generated) environment. This approach was the best solution given the design of this level in chapter 3. Each path is distinguished with its presented rewards and penalties. Players must learn by trial and error. Unlike traditional deterministic decision-making models, each edge in the graph represents both a risk and a reward-level, and each node represents the player's progress and learned behavior. The progression through the level can be represented by a **directed graph G** $= (N, E)$, where:

- **N** Nodes: Represent distinct decision points denoted as quests. Each quest categorized by the risk-reward degree $R_i P_j$, where R denotes the reward degree (low to high: R_1, R_2, R_3, R_4) and P denotes the penalty degree (low to high: P_1, P_2, P_3, P_4).
- **E** Edges: Represent transitions between nodes denoted as paths. Each path categorized by the risk-reward-degree.
- **Risks**: Are in form of obstacles, unknown collectables (unknown rewards) and increased difficulty to cross the path.
- **Rewards**: Are in form of amount and/or type of collectables and remaining nodes (quests) to reach the end node (end quest).
- **Collectables**: Are in form of blue Diamonds (1000). Pink Diamonds (1000). Pink Coins (20). Blue Coins (20) or mystery (Random amount in the $[-100; 100]$ range).

Please refer to Figure 4.7 showcasing the different nodes:

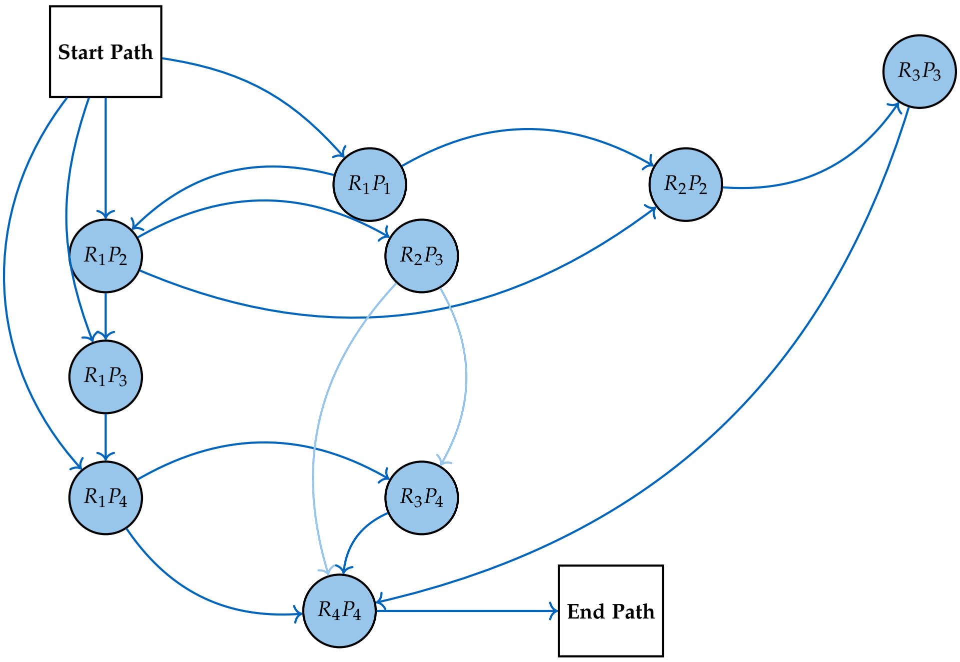


Figure 4.8: Path progression of Thrill Miner level, where Start Path: The initial entry point (Cave) where the player begins. Starting Nodes: R_1P_1, R_1P_2, R_1P_3, and R_1P_4. Ending Node: R_4P_4. End Path is the exiting of this level. Edges indicate the different paths within a node.

The paths are recursively calculated and the player's final passed, or started but dropped paths are saved as our measured values. $R_4P_4 \leftarrow R_3P_4 \leftarrow R_2P_3 \leftarrow R_1P_2$ In a similar manner to previously explained levels, all scripts interact through events. The logic of the level is decoupled from the components composing it:

- `TMStats.cs` Static class: calculating statistics of progress per node and chosen path.
- `TMData.cs` Persistent singleton: updating values for output file.
- Path and different obstacle handling scripts (`PathSetting.cs`, `TMObjects.cs` etc.) managing participant interactions, instantiating and behavior of each component composing the Thrill Miner level.
- `ThrillMinerManager.cs` Singleton: handling the entire process from data synchronization, updates, hidden and stated penalties and monetary event triggering.

The collected data is then serialized into the JSON file. Please refer to 3.3 for an example overview of these variables.
Please refer to figure 4.9 for an illustration regarding the workflow of these scripts.

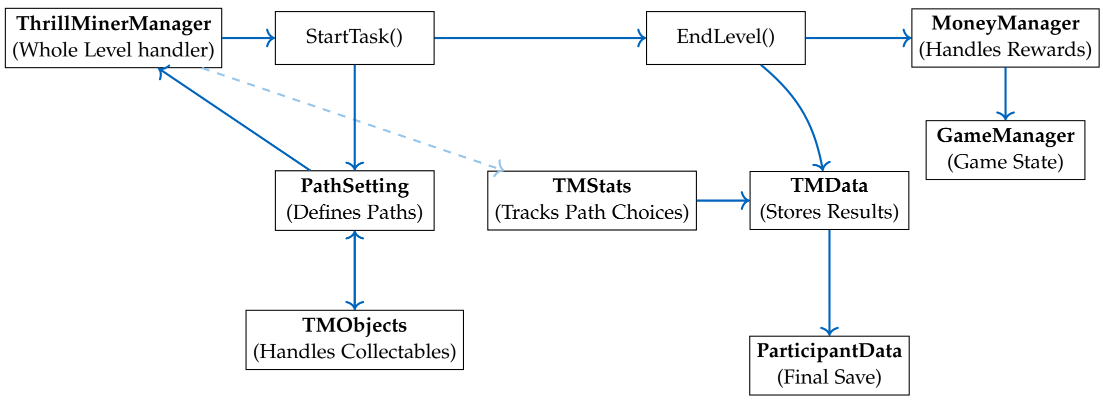


Figure 4.9: Thrill Miner Level scripts workflow

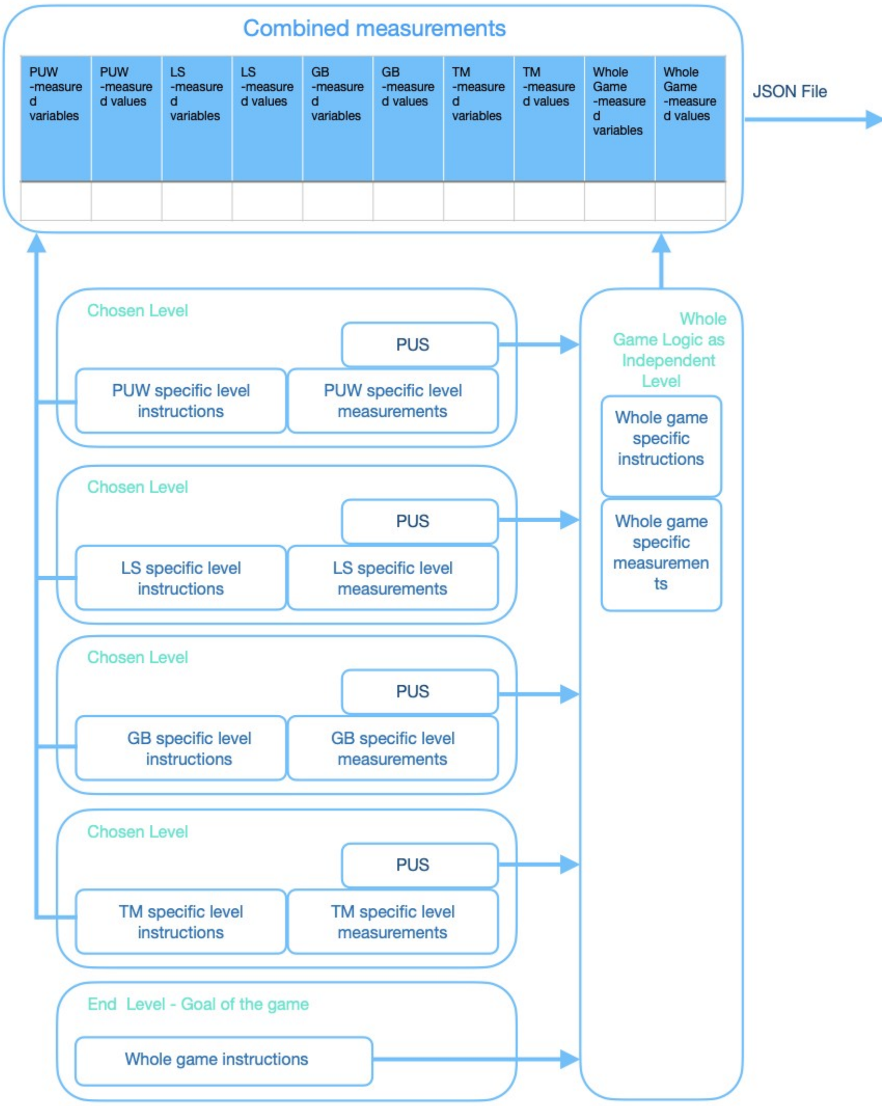


Figure 4.2: Diagram showcasing the different levels information sending to the final JSON file output

5 Digitilized Questionnaire Application

For the sake of completeness, we included this chapter to present how we digitized the three questionnaires to compute directly the scores and send them securely for the final comparison with the predicted values.

5.1 Questionnaire Application

Each question is displayed sequentially. Participants select their responses via a five-point or four-point Likert scale. The questions are ordered into groups of five. The transition between groups is automatically managed and prevents progression until all required questions within a group are answered. Responses are saved in dictionaries, calculated then stored in a structured JSON file sent via email to a private email address. Each participant receives a unique identifier to name the JSON file, ensuring anonymous data collection, corresponding to the identifier given during the VR game.

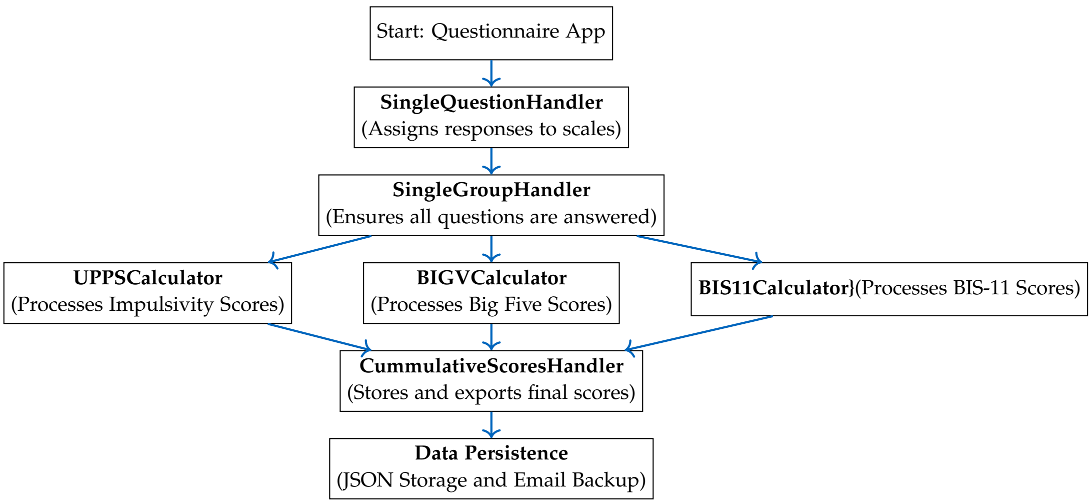


Figure 5.1: Questionnaire application scripts workflow.

Table 5.1: Psychometric scores, their ranges, and computation methods based on UPPS, BFI-2, and BIS-11 questionnaires. With R we denote the reverse value.

Score Name	**Range**	**Computation Method**
Urgency	(11, 44)	2R, 6R, 10R, 14R, 18R, 24R, 28R, 32R, 36R, 41R, 43, 45R (UPPS)
Lack of premeditation	(10, 40)	1, 5, 9, 13, 17, 23, 27, 31, 35, 39, 40 (UPPS)
Lack of perseverance	(10, 40)	4, 8R, 12, 16, 20, 22, 26, 30, 34, 38R (UPPS)
Sensation seeking	(12, 48)	3R, 7R, 11R, 15R, 19R, 21R, 25R, 29R, 33R, 37R, 42R, 44R (UPPS)
Extraversion	(12, 60)	1, 6, 11R, 16R, 21, 26R, 31R, 36R, 41, 46, 51R, 56 (BFI-2)
Verträglichkeit	(12, 60)	2, 7, 12R, 17R, 22R, 27, 32, 37R, 42R, 47R, 52, 57 (BFI-2)
Gewissenhaftigkeit	(12, 60)	3, 8, 13, 18, 23R, 28, 33R, 38R, 43R, 48, 53R, 58 (BFI-2)
Negative Emotionalität	(12, 60)	4R, 9R, 14R, 19R, 24, 29R, 34, 39, 44R, 49, 54, 59 (BFI-2)
Offenheit	(12, 60)	5, 10, 15R, 20, 25, 30, 35, 40, 45, 50, 55, 60 (BFI-2)
Geselligkeit	(4, 20)	1, 6, 11R, 16R (BFI-2 Facet)
Durchsetzungsfähigkeit	(4, 20)	21, 26R, 31R, 36R (BFI-2 Facet)
Aktivität	(4, 20)	41, 46, 51R, 56 (BFI-2 Facet)
Mitgefühl	(4, 20)	2, 7, 12R, 17R (BFI-2 Facet)
Höflichkeit	(4, 20)	22R, 27, 32, 37R (BFI-2 Facet)
Zwischenmenschliches Vertrauen	(4, 20)	42R, 47R, 52, 57 (BFI-2 Facet)
Ordnungsliebe	(4, 20)	3, 8, 13, 18 (BFI-2 Facet)
Fleiß	(4, 20)	23R, 28, 33R, 38R (BFI-2 Facet)
Verlässlichkeit	(4, 20)	43R, 48, 53R, 58 (BFI-2 Facet)
Ängstlichkeit	(4, 20)	4R, 9R, 14R, 19R (BFI-2 Facet)
Niedergeschlagenheit	(4, 20)	24, 29R, 34, 39 (BFI-2 Facet)
Unbeständigkeit der Gefühle	(4, 20)	44R, 49, 54, 59 (BFI-2 Facet)
Ästhetisches Empfinden	(4, 20)	5, 10, 15R, 20 (BFI-2 Facet)
Intellektuelle Neugierde	(4, 20)	25, 30, 35, 40 (BFI-2 Facet)
Kreativer Einfallsreichtum	(4, 20)	45, 50, 55, 60 (BFI-2 Facet)
Aufmerksamkeit	(0, 20)	11, 28, 5, 9R, 20R (BIS-11)
Kognitive Instabilität	(0, 12)	26, 6, 24 (BIS-11)
Motorische Impulsivität	(0, 28)	17, 19, 22, 3, 2, 25, 4 (BIS-11)
Beharrlichkeit	(0, 16)	21, 16, 30R, 23 (BIS-11)
Selbst Kontrolle	(0, 24)	12R, 1R, 8R, 7R, 13R, 14 (BIS-11)
Kognitive Komplexität	(0, 20)	15R, 29R, 10R, 27, 18 (BIS-11)

6 Model Implementation

The following chapter presents the initial models built based on prior research [53, 52], followed by an explanation of why these models were discarded given the constraints of this study. Then we go into details of our final implemented model used for predictions. In Table 5.1 we presented the 30 psychometric scores we are predicting using our model.

6.1 Testing Models

The two initial models were selected based from research with similar contexts [53, 52]. These models perform well given abundant training data and well-defined feature distributions. However, the extracted measures from the game are completely new variables therefore no pre-existing labeled dataset is available. The lack of direct training data makes any supervised learning approach impossible for our purposes.

6.1.1 ML Based Model

The ML-based model consists of a neural network trained using an adaptive optimization strategy. The architecture was tested with multiple configurations of hidden layers and activation functions. The optimization is performed using **Adam** with a learning rate of **0.01** and a maximum of **500** epochs, **2000** epochs and **1000** epochs. These values were validated in similar research focusing on machine learning applications involving psychometrics [73, 75]. The output layer delivers scores within predefined ranges using **mean squared error (MSE)** as the loss function. To account for non-linearities in the initial predictions, a secondary regressor is applied to model residual errors. In table 6.1, we present the workflow of this model as pseudocode.

6.1.2 XGBoost-Based Model

The XGBoost-based model consists of embedding generation from textual inputs using **SBERT embeddings** to capture high-dimensional semantic relationships [76]. These embeddings are then concatenated with numerical and categorical features. The concatenated data is then passed into an **XGBoost regressor**. The XGBoost regression

Table 6.1: Pseudocode for the ML-Based model. Where `data` = input features; `labels` = target values; `hidden_layers` = network depth; `epochs` = training iterations (**500**, **1000**, **2000**); `learning_rate` = step size (**0.01**); `optimizer` = **Adam**; `loss_fn` = **MSE**; `regressor` = secondary model for residual correction.

```
function train_ML_model(data, labels, hidden_layers, epochs, learning_rate):

  X, y = preprocess(data, labels)
  X_train, X_test, y_train, y_test = split(X, y, test_size=0.2)
  model = NeuralNetwork(input_size = X.columns, hidden_layers = hidden_layers, output_size = y.columns)
  optimizer = Adam(model.parameters, learning_rate)
  loss_fn = MeanSquaredError()

  for epoch in range(epochs):
    optimizer.zero_grad()
    predictions = model(X_train)
    loss = loss_fn(predictions, y_train)
    loss.backward()
    optimizer.step()

  residuals = compute_residuals(predictions, y_train)
  regressor = fit_secondary_regressor(X_train, residuals)
  final_predictions = model(X_test) + regressor.predict(X_test)

  return final_predictions
```

model optimizes feature importance weighting during training. The final score predictions are then refined using **mean-centered shifting** and **random noise** to reduce skewness in score distributions. In table 6.2, we present the workflow of this model as pseudocode.

6.2 Chosen Model Pipeline

As stated before the absence of training data makes any supervised learning approach impossible for our purposes. Therefore our model begins with an initial prediction phase, where a prompting pipeline generates psychometric scores from extracted measurements. These scores serve as preliminary estimates before refinement. Then we implement the RAG pipeline for the knowledge enhancement step. In this step we extract and structure information from academic sources (143 files ranging from papers

Table 6.2: Pseudocode for the XGBoost-Based model. Where `sbert_model` = **SBERT embeddings**; `XGBoostRegressor` = model optimizing feature weighting; post-processing includes **mean-centered shifting** and **random perturbations** instead of quantile regression.

```
function train_XGBoost_model(data, labels, sbert_model):

  X, y = preprocess(data, labels)
  text_features = extract_text(X)
  text_embeddings = encode(text_features, model=sbert_model)
  categorical_features, numerical_features = extract_structured(X)
  X_processed = concatenate(text_embeddings, categorical_features, numerical_features)

  model = XGBoostRegressor(n_estimators=200, max_depth=6, learning_rate=0.05)
  model.fit(X_processed, y)
  predictions = model.predict(X_processed)

  for i in range(len(predictions)):
    mean_val = (SCORE_RANGES[i][0] + SCORE_RANGES[i][1]) / 2
    predictions[i] = clip(predictions[i] + mean_val * 0.5, SCORE_RANGES[i][0],
SCORE_RANGES[i][1])
    predictions[i] = predictions[i] + random_noise(-3, 3)

  return predictions
```

to journals). Then Named Entity Recognition (NER) is applied on the extracted data to identify relevant constructs. Then a **Bayesian network** is used to map causal relationships between variables. Additionally the Knowledge representations are converted into **SciBERT embedding** to ensuring alignment with domain-specific semantics [74, 76]. The training step composed of the **correlation learner**. The correlation learner is a neural network that optimizes feature-score relationships (by feature we denote our in-game extracted measurements, and by score we denote the 30 psychometric scores). The learner also incorporates statistical dependencies identified in the Bayesian model. This structure ensures that psychometric estimates are constrained by both empirical measurements and structured domain knowledge [73, 75]. The processing step where the initial predictions are refined. In this step, the extracted measurements are mapped to predicted scores, and correction factors derived from the correlation learner adjusts the outputs. This step applies Bayesian-TF-IDF approach to ensure the reliability of the estimated scores. Given this approach, the final output accounts for both observed measurements and inferred relations.

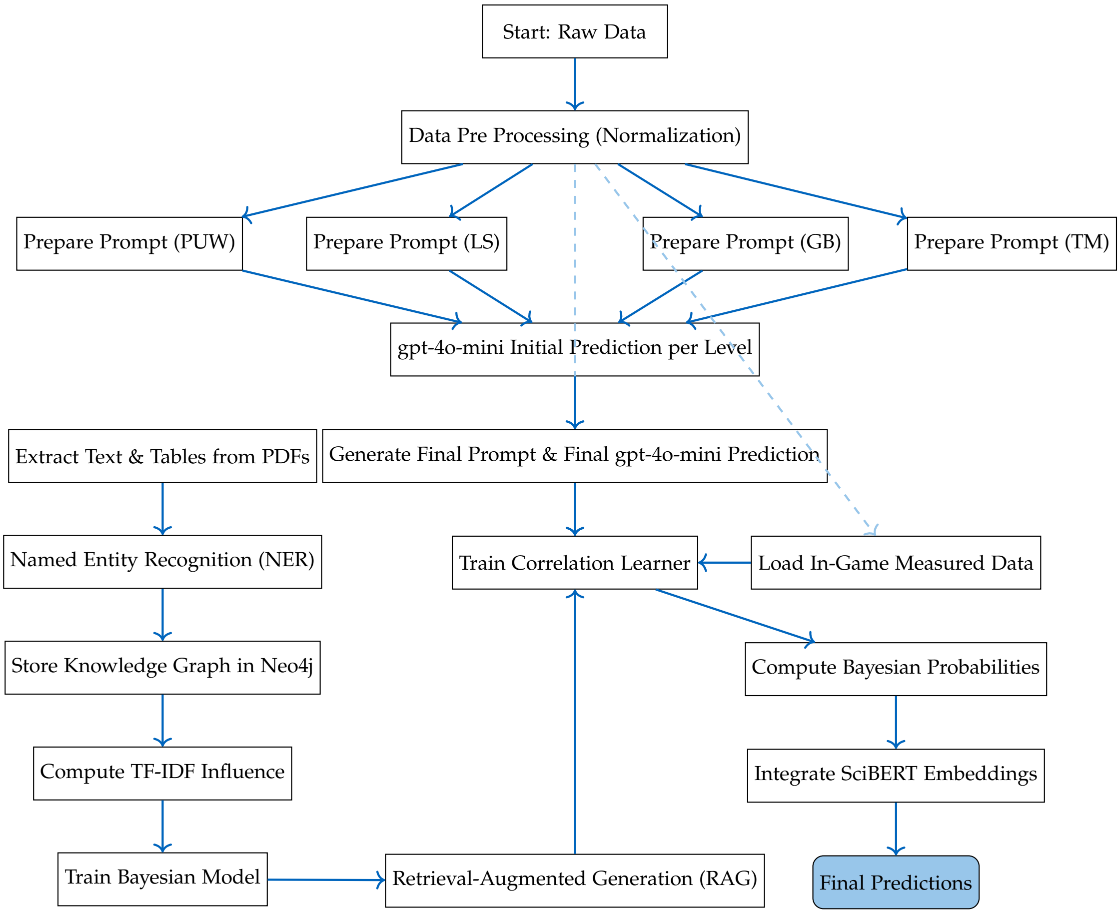


Figure 6.1: Workflow of the Prediction Model. In-Game Measured Data are 3.6 3.1 3.7 3.4 3.5 3.3 and Predictions are 5.1.

6.2.1 Measurements Normalization

To ensure our extracted in-game measurements remain statistically comparable and relax outliers introduced by scale variations, each variable is normalized. We use different normalization techniques as follows.

Min-Max Scaling is used for bounded variables such as coin quantities, slot play counts, and drinks purchased, so that values range between 0 and 1 while preserving relative magnitudes. Preventing extreme values from influencing the model.

Z-score normalization is used for rate-based variables such as plays per minute.

Subtracting the mean and dividing by the standard deviation, making comparisons across distributions more reliable.

Max normalization is used for fixation times and time-based progress measures. Each observation is scaled as a proportion of the longest recorded duration, maintaining interpretability across participants.

6.2.2 Initial Prediction

The initial prediction step maps extracted in-game measurements to psychometric scores. To avoid bias caused by a single measure, predictions are processed separately per level. We distinguish four distinct prompts, each targeting a specific level **PUW-Level Prompt**, **LS-Level Prompt**, **GB-Level Prompt**, **TM-Level Prompt**. Each prompt generates independent 30 scores. The outputs from these four levels are then fed into a fifth prompt to ensure these scores contribute proportionally to the final prediction.
The model uses **gpt-4o-mini**, configured with **Temperature = 0.2** to minimize variability and enforce consistency; **Max tokens = 3000**; **Zero-shot** prompting to bound predictions within each score range.

6.2.3 Knowledge Enhancement

The knowledge enhancement step, integrates structured academic relevant research information into the prediction pipeline. This step prevents direct data-inference biases by embedding domain-specific relationships and probabilistic structures. The process consists of: Retrieval-augmented generation (RAG) for extracting relevant information; Named Entity Recognition (NER) for structuring extracted knowledge; Graph-based storage for relational modeling; Bayesian networks for causal learning, and SciBERT fine-tuning to encode extracted knowledge into domain-specific vector representations.

RAG Pipeline

Academic PDFs are processed using OCR-based text extraction via **pdfplumber** and table conversion using **LayoutLMv3**. Extracted text is then normalized and concatenated for consistent knowledge representation. In table 6.3, we present the workflow of this model as pseudocode.

Table 6.3: Pseudocode for the RAG Pipeline. Where `pdf_folder` = folder containing PDFs; `text` = extracted text; `tables` = raw table data; `tables_text` = tables converted to text.

```
function RAG_Pipeline (pdf_folder):

  knowledge_base = ""
  for each pdf_file in pdf_folder:
    text = extract_text_from_pdf(pdf_file)
    tables = extract_tables_from_pdf(pdf_file)
    tables_text = convert_tables_to_text(tables)
    knowledge_base = knowledge_base + text + tables_text

  return knowledge_base
```

NER Extraction from Previous Research

Named Entity Recognition (NER) to extract domain-specific constructs. As the academic papers used here are in German and English we used two pre-trained models **"mschiesser/ner-bert-german"** for German text and **"dbmdz/bert-large-cased-finetuned-conll03-english"** for English text. In table 6.4, we present the workflow of this model as pseudocode.

Table 6.4: Pseudocode for extracting entities. Where `text` = input text; language is detected; model selection is based on language; and entities are parsed from the NER results.

```
function extract_entities(text):

  language = detect_language(text)
  if language == "de":
    ner_results = run_NER(text, model="mschiesser/ner-bert-german")
  else:
    ner_results = run_NER(text, model="dbmdz/bert-large-cased-finetuned-conll03-
english")
  entities = parse_ner_results(ner_results)

  return entities
```

Neo4j Graph-Based Storage

Extracted knowledge is stored in Neo4j, a graph database. Nodes represent concepts (e.g., psychological traits), and edges define their relationships.

Bayesian Networks for Causal Learning

The Bayesian network to model probabilistic dependencies between extracted entities. The structure is inferred from the knowledge graph using **Maximum Likelihood Estimation (MLE)**. Then the Bayesian network is trained with **pgmpy**'s MaximumLikelihoodEstimator to understand dependencies within the data.

SciBERT Fine-tuning on Knowledge

The **"allenai/scibert_scivocab_uncased"** model is fine-tuned on the extracted knowledge to generate domain-specific text embeddings. This helps encode extracted knowledge into numerical representations for predictions. The fine-tuning is optimized with **Batch size = 16**; **Learning rate = 2e-5**, and **Max sequence length = 512**.

6.2.4 Training phase: Correlation Learner

The correlation learner is a neural network to estimate dependencies between in-game measured features and the 30 psychometric scores. The model maps an input feature vector $x \in \mathbb{R}^d$ to an output vector $\hat{y}$, for both numerical and learned textual embeddings. The input vector is constructed by concatenating: In-game measured data; TF-IDF representations of the extracted domain knowledge and fine-tuned SciBERT embeddings.
The model is optimized using **Mean Squared Error (MSE) loss**, **Adam optimizer** with **Learning rate = 0.01**; **Maximum epochs= 500**, and **Early stopping threshold = 0.05**. In table 6.5, we present the workflow of this model as pseudocode.

6.2.5 Processing Phase: Scores Adjustment

In the processing step each predicted file is re-evaluated via Bayesian-TF-IDF Score Adjustment. In table 6.8, we present the workflow of this step as pseudocode. The correlation learner outputs adjustment factors, which are combined with TF-IDF-based influences and Bayesian causals. The adjusted scores are again bounded to the defined range. In table 6.6, we present the workflow of this model as pseudocode. TF-IDF ensures the relevance of extracted knowledge to each score. In table 6.7, we present the workflow of this function as pseudocode.

6.2.6 Evaluation

As stated before, due to the lack of participants and ground truth data, we will only hypothesis regarding the validity of our model. In the following table 6.9 we showcase the deviation per score given five German fellow student volunteers.

Table 6.5: Pseudocode for the correlation learner. Where `measured_values` = raw measured data; `predicted_values` = initial predictions; `knowledge_data` = TF-IDF vector representation; `scibert_embeddings` = domain-specific embeddings from SciBERT; `epochs` = maximum training iterations (500); `threshold` = early stopping loss threshold (0.05); `learning_rate` = optimizer step size (0.01).

```
function train_correlation_learner(measured_values, predicted_values, knowledge_data,
scibert_embeddings, epochs = 500, threshold = 0.05):

  measured_tensor = tensor(measured_values, dtype=float32)
  knowledge_tensor = tensor(knowledge_data, dtype=float32)
  knowledge_tensor = repeat(knowledge_tensor, measured_tensor.rows)
  scibert_tensor = tensor(scibert_embeddings, dtype=float32)
  scibert_tensor = repeat(scibert_tensor, measured_tensor.rows)
  input_tensor = concatenate(measured_tensor, knowledge_tensor, scibert_tensor)
  model = CorrelationLearner(input_size = input_tensor.columns, output_size =
length(predicted_values))
  optimizer = Adam(model.parameters, learning_rate=0.01)
  loss_fn = MeanSquaredError()

  for epoch in range(epochs):
    optimizer.zero_grad()
    output = model(input_tensor)
    loss = loss_fn(output, tensor(predicted_values, dtype=float32))
    for condition: if loss < threshold:
      for action: break
    loss.backward()
    optimizer.step()

  return model
```

Table 6.6: Pseudocode for the Bayesian-TF-IDF Score Adjustment. Where `predicted_scores` = initial model outputs; `measured_values` = empirical data; `model` = trained Correlation Learner; `knowledge_base` = extracted academic insights; `bayesian_model` = causal dependency structure.

```
function apply_rlaif(predicted_scores, measured_values, model, knowledge_base,
bayesian_model):

    adjusted_scores =
    numeric_features = convert_measured_values(measured_values)
    measured_tensor = tensor(numeric_features, dtype=float32)
    predicted_adjustments = model(measured_tensor).detach().numpy().flatten()

    for i, (score, value) in enumerate(predicted_scores.items()):
        adjustment = predicted_adjustments[i] if i < len(predicted_adjustments) else 0.0
        influence = compute_combined_influence(score, knowledge_base, bayesian_model)
        adjusted_value = value + adjustment * influence
        adjusted_scores[score] = normalize_score(adjusted_value, define_ranges[score])

    return adjusted_scores
```

Table 6.7: Pseudocode for computing TF-IDF influence. Where `score` = psychometric target; `knowledge_base` = extracted academic text; `vectorizer` = TF-IDF implementation.

```
function compute_tfidf_influence(score, knowledge_base):

    vectorizer = TfidfVectorizer()
    tfidf_matrix = vectorizer.fit_transform(knowledge_base.split("
textbackslash n"))
    feature_names = vectorizer.get_feature_names_out()
    if score in feature_names:
        return tfidf_matrix[:, feature_names.tolist().index(score)].sum()
    return 0
```

Table 6.8: Pseudocode for processing phase. Where `knowledge_base` = extracted domain knowledge; `bayesian_model` = causal structure; `measured_matrix` = stored empirical data; `correlation_model` = trained score adjustment network.

```
function processing_phase(knowledge_base, bayesian_model, measured_matrix, correlation_model):

    predicted_matrix = []
    enhanced_matrix = []
    header_row = ["Filename"] + list(mapping.values())
    predicted_matrix.append(header_row)
    enhanced_matrix.append(header_row)

    for filename in os.listdir(predicted_folder):
        predicted_data = load_predicted_json(filename)
        measured_values = retrieve_measured_values(filename, measured_matrix)
        adjusted_scores = apply_rlaif(predicted_data, measured_values, correlation_model,
knowledge_base, bayesian_model)
        predicted_matrix.append([filename] + list(predicted_data.values()))
        enhanced_matrix.append([filename] + list(adjusted_scores.values()))

    visualize_measured_data(enhanced_matrix)
```

Table 6.9: Evaluation of psychometric scores based on mean deviation, standard deviation, and range.

Score Name	Range	Mean Deviation	Standard Deviation
Urgency	(11, 44)	5.67	8.08
Lack of premeditation	(10, 40)	3.00	3.46
Lack of perseverance	(10, 40)	3.67	2.52
Sensation seeking	(12, 48)	16.33	8.02
Extraversion	(12, 60)	2.00	1.73
Verträglichkeit	(12, 60)	4.67	4.73
Gewissenhaftigkeit	(12, 60)	9.33	1.53
Negative Emotionalität	(12, 60)	14.67	8.96
Offenheit	(12, 60)	15.33	7.09
Geselligkeit	(4, 20)	2.33	1.53
Durchsetzungsfähigkeit	(4, 20)	1.67	2.89
Aktivität	(4, 20)	1.67	1.15
Mitgefühl	(4, 20)	2.00	1.00
Höflichkeit	(4, 20)	3.00	2.65
Zwischenmenschliches Vertrauen	(4, 20)	2.33	2.08
Ordnungsliebe	(4, 20)	4.00	1.73
Fleiß	(4, 20)	4.67	2.89
Verlässlichkeit	(4, 20)	1.67	0.58
Ängstlichkeit	(4, 20)	2.67	0.58
Niedergeschlagenheit	(4, 20)	4.33	1.53
Unbeständigkeit der Gefühle	(4, 20)	6.00	2.65
Ästhetisches Empfinden	(4, 20)	2.00	1.73
Intellektuelle Neugierde	(4, 20)	1.67	1.15
Kreativer Einfallsreichtum	(4, 20)	4.00	2.65
Aufmerksamkeit	(0, 20)	4.33	0.58
Kognitive Instabilität	(0, 12)	3.33	0.58
Motorische Impulsivität	(0, 28)	5.67	2.08
Beharrlichkeit	(0, 16)	1.00	0.00
Selbst Kontrolle	(0, 24)	3.67	3.21
Kognitive Komplexität	(0, 20)	3.33	2.08

7 Challenges & Furtue work

In this chapter, we state the challenges encountered during the development and propose recommendations for future work. The proposed next steps are based on the current market's availability of hardware, off-the-shelf VR headsets and specifically the state of development of eye-tracking within those head-mounted displays (HMDs).

7.1 Rendering & Performance

The Pop-Up-Werk level was designed to mimic a realistic and immersive gambling environment. Multiple animated machines, alcohol dispensers and NPCs introduce potential performance drop rates. Undoubtedly before integrating any asset, its mesh size was reduced significantly in addition to the number of elements to be incorporated to the scene. To enhance the visual appeal and realism we of course chose high detailed textures and dimmed the lights. Additionally we opted for fully baked light maps to counter the real time lighting issue. However the latter was not possible and we had to use mixed light settings due to the version of Unity used. During testing the level ran seamlessly on the used computer with real-time lighting and initial assets integration. During Build on the Pico Neo 3 Eye Pro however, the scene was partially black and it required significant time to draw it. Therefore we had to reduce the number of NPCs and number of beverage dispensers and yet we could not mitigate the issue. The scene was left as-is for now as these elements are crucial for our measurements.

7.2 Eyetracking & Accuracy

The target device is Pico Neo 3 Pro Eye Business edition headset. During the development, we encountered many fallbacks caused by the device's eye-tracking drivers, sensors and software, despite the headset's advertised support for eye-tracking features. The ergonomics of the headset were the main cause for failed sensors captures. The Pico Neo 3 Pro Eye is designed for the Chinese demographic, leading to a misfit for users with different craniofacial structures, specifically Caucasian skull [91]. The Pico Software Development Kit (SDK) version 2.5.3 lacked support for global deployment, necessitating the use of an outdated Unity version, limiting the use of advanced XR features and relying on unstable deprecated packages. The Pico SDK also lacked editor support for run-time environment, hindering the testing process significantly. Consequently, we had to implement seven distinct methods to capture eye position and rotation. These methods range from actual measurements of a single eye to approximation of the glabella; However, none of the actual (not approximated methods) yielded to results, as extracting per single eye movement was not yet available even though variables to contain these information were predefined and supported by the Pico SDK and the Pico headset. The challenges encountered with the Pico Neo 3 Pro Eye highlights the need for careful evaluation when selecting hardware. It is primordial to ensure that devices are suitable for the target demographic and provide accurate robust tracking data.

7.3 Motion Sickness & Proficiency

Another faced limitation, is the feasibility of the Thrill Miner level. This level involves dynamic movements such as jumping and climbing, which requires familiarity with VR controls and highly susceptible to induce motion sickness. Taking into consideration the age range of our target group we assumed unfamiliarity with the VR controls; Given the improper head fit, the Pico Neo 3 Pro Eye induced stronger motion sickness than other VR headsets, even in relatively static scenes. Feelings of dizziness, nausea, and disorientation hinder the ability to complete the tasks therefore the Thrill Miner level was omitted. A computer or a CAVE (Cave Automatic Virtual Environment) setup could have mitigated the motion sickness and familiarity issue.

7.4 Apple vs Windows & Android

During the implementation of the digitized questionnaire app, we encountered a critical limitation: the inability to store responses locally on the iPad. This issue stems from restrictions within iOS regarding file storage. App sandboxing prevents certain applications from saving data outside of designated secure areas. The workaround was to send questionnaire scores via email. While sending computed ground truth scores via email was initially not planned, this approach offers privacy and security. Currently for testing purposes, the questionnaire app includes all three questionnaires UPPS, BFI and BIS-11 30 items. In the future we will divide this app into three distinct apps.

7.5 Prediction Model Implementation & Accuracy

Since no large-scale validation dataset exists as stated in previous chapters, the model's accuracy remains theoretical for now. However, based on known psychometric distribution properties, adjustments and the 5 tester to compare their predicted scores vs their ground truth, we can hypothesize that the initial prediction model expected error range is ±1-20 points per psychometric score. Please note that this deviation is proportional to the range of the single psychometric score 5.1.

7.6 Compiled Recommendation

Given the restrictions identified and explained in previous sections, we recommend the following: Two distinct target platforms of this project for two distinct purposes.

Research and Progress Measuring Purposes

We recommend using a a large, high-resolution screen with a dedicated eye-tracking device. A high resolution screen attached to a powerful computer to bypass the encountered rendering issues and therefore give more room to integrate high-end graphics and objects to achieve a more realistic scene setup. Devices such as the Tobii Pro Glasses 3 or the EyeLink 1000 Plus for eye-tracking. As stated above, the most significant challenge is the current limitations of eye-tracking in commercially available VR headsets, where devices lack the precision and robustness, particularly in tracking isolated eye movements (without turning the head) and pupillometry [7]. The Tobii Pro Glasses 3 deliver accurate and robust gaze data, while the EyeLink 1000 Plus is a high-precision, video-based eye tracker sampling binocularly at up to 2000 Hz. Additionally we recommend incorporating a surrounding sound system and a scent generator, such as Olorama Technology's Professorial scents generators, to further enhance the sense of presence and immersion. This approach will not only provide Auditory stimuli but also olfactory stimuli on top of the immersive visual cues.

Behavioral Therapy Purposes

We recommend without compromising the user experience, specifically immersion to keep using a VR headset but a better suited device for the target demographic.

By adopting these recommendations, we aim to create a more controlled and immersive environment that expends the type and accuracy of collected data, to further refine our predictive model making it more versatile and robust.

8 Conclusion

This project introduces a novel approach to predicting psychometric scores from implicitly measuring in-game behavioral data within a virtual environment using a multi-layered model integrating generative AI, Bayesian inference, and retrieval-augmented knowledge processing . The proposed prediction model consists of:

- Initial score predictions via gpt-4o-mini structured prompting across each game level then a final prompt to generate the actual scores.
- Knowledge enhancement using RAG, where 143 academic sources were processed to extract domain-specific embedding.
- Refinement via Bayesian inference and TF-IDF weighting to constrain predictions within theoretical bounds.
- Correlation Learner, a neural network that connects in-game extracted measures to psychometric scores.
- Final Adjustment of initially predicted scores with the learned correlations.

While the model remains theoretical due to the lack of ground truth validation data, the methodology outlined here establishes a solid foundation for future research within this area.

List of Figures

List of Tables

Bibliography

[1] J. Weafer, M. J. Baggott, and H. de Wit. "Test-retest reliability of behavioral measures of impulsive choice, impulsive action, and inattention." In: *Experimental and Clinical Psychopharmacology* 21.6 (Dec. 2013), pp. 475–481. DOI: 10.1037/a0033659.

[2] F. Morales-Vives, U. Lorenzo-Seva, and A. Vigil-Colet. "How response biases affect the factor structure of Big Five personality questionnaires." In: *Anales de Psicología* 33.3 (Oct. 2017), pp. 589–596. ISSN: 0212-9728. DOI: 10.6018/analesps.33.2.254841.

[3] N. Vafaie and H. Kober. "Association of Drug Cues and Craving With Drug Use and Relapse: A Systematic Review and Meta-analysis." In: *JAMA Psychiatry* 79.7 (July 2022), pp. 641–650. ISSN: 2168-622X. DOI: 10.1001/jamapsychiatry.2022.1240. eprint: https://jamanetwork.com/journals/jamapsychiatry/articlepdf/2793124/jamapsychiatry_vafaie_2022_oi_220030_1655850843.27702.pdf.

[4] K. Lewandowski. "Ecological validity in cognitive assessment and treatment." In: *Schizophrenia Research: Cognition* (Dec. 2024), p. 100341. DOI: 10.1016/j.scog.2024.100341.

[5] M. Zhang, J. Ying, G. Song, R. C. Ho, D. S. Fung, and H. Smith. "Attention Bias in Individuals with Addictive Disorders: Systematic Review Protocol." In: *JMIR Research Protocols* 7.2 (Feb. 2018), e41. DOI: 10.2196/resprot.9385.

[6] J. Simon, D. Grogna, M.-C. Rivard, M. Heck, S. Bouchard, and E. Quertemont. "Assessing attentional bias for alcohol-related cues using eye tracking in a virtual reality environment." In: *Frontiers in Virtual Reality* 3 (2022). ISSN: 2673-4192. DOI: 10.3389/frvir.2022.849840.

[7] M. K. Eckstein, B. Guerra-Carrillo, A. T. Miller Singley, and S. A. Bunge. "Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development?" In: *Developmental Cognitive Neuroscience* 25 (2017). Sensitive periods across development, pp. 69–91. ISSN: 1878-9293. DOI: https://doi.org/10.1016/j.dcn.2016.11.001.

[8] K. E. Courtney, W. Liu, G. Andrade, J. Schulze, and N. Doran. "Attentional Bias, Pupillometry, and Spontaneous Blink Rate: Eye Characteristic Assessment Within a Translatable Nicotine Cue Virtual Reality Paradigm." In: *JMIR Serious Games* 12 (June 2024), e54220. ISSN: 2291-9279. DOI: 10.2196/54220.

[9] C. Culbertson, S. Nicolas, I. Zaharovits, E. D. London, R. 2. De La Garza, A. L. Brody, and T. F. Newton. "Methamphetamine craving induced in an online virtual reality environment." In: *Pharmacology Biochemistry and Behavior* 96.4 (2010), pp. 454–460. ISSN: 0091-3057. DOI: 10.1016/j.pbb.2010.07.005.

[10] A. Wolf and K. Ueda. "Contribution of Eye-Tracking to Study Cognitive Impairments Among Clinical Populations." In: *Frontiers in Psychology* 12 (2021). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2021.590986.

[11] V. Branchadell, R. Poy, P. Segarra, P. Ribes-Guardiola, and J. Moltó. "Attentional biases in abstinent patients with cocaine use disorder: rapid orienting or delayed disengagement?" In: *Frontiers in Psychology* 15 (2024). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2024.1290890.

[12] Z. Topbaş, E. Albayrak, N. Şimşek, and E. Usta. "Does Attentional Bias Predict Relapse in Addiction? A Systematic Review of Longitudinal Studies." In: *Brain and Behavior* 15 (Feb. 2025). DOI: 10.1002/brb3.70300.

[13] A. M. Herman, H. D. Critchley, and T. Duka. "Risk-Taking and Impulsivity: The Role of Mood States and Interoception." In: *Frontiers in Psychology* 9 (Aug. 2018), p. 1625. DOI: 10.3389/fpsyg.2018.01625.

[14] L. Fisher-Fox, C. J. Prestigiacomo, and M. A. Cyders. "Urgency Theory in the context of broader emotion theories: a conceptual review." In: *Frontiers in Psychiatry* 15 (2024). ISSN: 1664-0640. DOI: 10.3389/fpsyt.2024.1403639.

[15] Y. Huang, S. Luan, B. Wu, Y. Li, J. Wu, W. Chen, and R. Hertwig. "Impulsivity is a stable, measurable, and predictive psychological trait." In: *Proceedings of the National Academy of Sciences* 121.24 (2024), e2321758121. DOI: 10.1073/pnas.2321758121. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.2321758121.

[16] B. Morgan, J. Suhr, and S. Jones. "C-31The Relationship of Urgency to Impulsive Decision-Making During Heightened Affective States in Problem Drinkers." In: *Archives of Clinical Neuropsychology* 32.6 (Sept. 2017), pp. 667–765.

[17] E. Fehrman, A. K. Muhammad, E. M. Mirkes, V. Egan, and A. N. Gorban. *The Five Factor Model of personality and evaluation of drug consumption risk*. 2017. arXiv: 1506.06297 [stat.AP].

[18] M. A. Bornovalova, M. A. Gwadz, C. Kahler, W. M. Aklin, and C. W. Lejuez. "Sensation seeking and risk-taking propensity as mediators in the relationship between childhood abuse and HIV-related risk behavior." In: *Child Abuse & Neglect* 32.1 (Jan. 2008), pp. 99–109. DOI: 10.1016/j.chiabu.2007.04.009.

[19] A. K. Verma and U. Chivukula. "The effect of an affect, sensation seeking, and premeditation on risky decision-making: Conditional process analysis." In: *PLOS ONE* 18 (Feb. 2023), pp. 1–12. DOI: 10.1371/journal.pone.0281324.

[20] S. Xu, L. Luo, Z. Xiao, K. Zhao, H. Wang, C. Wang, and H. Rao. "High sensation seeking is associated with behavioral and neural insensitivity to increased negative outcomes during decision-making under uncertainty." In: *Cognitive, Affective, & Behavioral Neuroscience* 19.6 (Dec. 2019). Erratum in: Cogn Affect Behav Neurosci. 2020 Feb;20(1):214. doi: 10.3758/s13415-019-00766-4., pp. 1352–1363. DOI: 10.3758/s13415-019-00751-x.

[21] A. Abid, R. J. Mach, M. Middlebrooks, C. Lamm, and M. Judah. "B - 73 Neurocognitive Correlates of Impulsivity: Lack of Perseverance Is Associated with Poorer Cognitive Control." In: *Archives of Clinical Neuropsychology* 39.7 (Sept. 2024), pp. 1168–1168. ISSN: 1873-5843. DOI: 10.1093/arclin/acae067.234. eprint: https://academic.oup.com/acn/article-pdf/39/7/1168/60086307/acae067.234.pdf.

[22] H. L. Soodla and K. Akkermann. "Bottom-up transdiagnostic personality subtypes are associated with state psychopathology: A latent profile analysis." In: *Frontiers in Psychology* 14 (2023). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2023.1043394.

[23] F. Bevione, M. Martini, F. Toppino, P. Longo, G. Abbate-Daga, A. Brustolin, and M. Panero. "Cognitive Impulsivity in Anorexia Nervosa in Correlation with Eating and Obsessive Symptoms: A Comparison with Healthy Controls." In: *Nutrients* 16.8 (2024), p. 1156. ISSN: 2072-6643. DOI: 10.3390/nu16081156.

[24] A. Vigil-Colet and F. Morales-Vives. "How Impulsivity is Related to Intelligence and Academic Achievement." In: *The Spanish journal of psychology* 8 (Dec. 2005), pp. 199–204. DOI: 10.1017/S1138741600005072.

[25] K. Sakurai, H. Li, N. Inamura, N. Masuoka, and T. Hisatsune. "Relationship between elevated impulsivity and cognitive declines in elderly community-dwelling individuals." In: *Scientific Reports* 10.1 (2020), p. 21032. ISSN: 2045-2322. DOI: 10.1038/s41598-020-78124-5.

[26] C. MacCann, A. L. Duckworth, and R. D. Roberts. "Empirical identification of the major facets of Conscientiousness." In: *Learning and Individual Differences* 19.4 (2009), pp. 451–458. ISSN: 1041-6080. DOI: https://doi.org/10.1016/j.lindif.2009.03.007.

[27] S. R. Carlson, A. A. Pritchard, and R. M. Dominelli. "Externalizing behavior, the UPPS-P Impulsive Behavior scale and Reward and Punishment Sensitivity." In: *Personality and Individual Differences* 54.2 (2013), pp. 202–207. ISSN: 0191-8869. DOI: https://doi.org/10.1016/j.paid.2012.08.039.

[28] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. *Attention Is All You Need*. 2023. arXiv: 1706.03762 [cs.CL].

[29] I. Beltagy, K. Lo, and A. Cohan. "SciBERT: A Pretrained Language Model for Scientific Text." In: *EMNLP/IJCNLP (1)*. Ed. by K. Inui, J. Jiang, V. Ng, and X. Wan. Association for Computational Linguistics, 2019, pp. 3613–3618. ISBN: 978-1-950737-90-1.

[30] X. Liu, F. Zhang, Z. Hou, L. Mian, Z. Wang, J. Zhang, and J. Tang. "Self-supervised Learning: Generative or Contrastive." In: *IEEE Transactions on Knowledge and Data Engineering* (2021), pp. 1–1. ISSN: 2326-3865. DOI: 10.1109/tkde.2021.3090866.

[31] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton. *A Simple Framework for Contrastive Learning of Visual Representations*. 2020. arXiv: 2002.05709 [cs.LG].

[32] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, and D. Kiela. *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. 2021. arXiv: 2005.11401 [cs.CL].

[33] S. Xu, M. Chen, and S. Chen. "Enhancing Retrieval-Augmented Generation Models with Knowledge Graphs: Innovative Practices Through a Dual-Pathway Approach." In: *Advanced Intelligent Computing Technology and Applications*. Ed. by D. S. Huang, Z. Si, and W. Chen. Vol. 14880. Lecture Notes in Computer Science. Singapore: Springer, 2024. DOI: 10.1007/978-981-97-5678-0_34.

[34] T. Baltrušaitis, C. Ahuja, and L.-P. Morency. *Multimodal Machine Learning: A Survey and Taxonomy*. 2017. arXiv: 1705.09406 [cs.LG].

[35] D. B. Dorman, A. L. Sampson, P. Sacre, V. Stuphorn, E. Niebur, and S. V. Sarma. "Decomposing Executive Function into Distinct Processes Underlying Human Decision Making." In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2022, pp. 807–811. DOI: 10.1109/EMBC48229.2022.9871849.

[36] C. Korponay, D. Dentico, T.-M. Kral, M. Ly, A. Kruis, R. Goldman, A. Lutz, and R. J. Davidson. “Neurobiological correlates of impulsivity in healthy adults: Lower prefrontal gray matter volume and spontaneous eye-blink rate but greater resting-state functional connectivity in basal ganglia-thalamo-cortical circuitry.” In: *NeuroImage* 157 (Aug. 2017). Erratum in: Neuroimage. 2018 Feb 15;167:505. doi: 10.1016/j.neuroimage.2017.10.042, pp. 288–296. DOI: 10.1016/j.neuroimage.2017.06.015.

[37] A. Lohse, A. Løkkegaard, H. R. Siebner, and D. Meder. “Linking Impulsivity to Activity Levels in Pre-Supplementary Motor Area during Sequential Gambling.” In: *Journal of Neuroscience* 43.8 (Feb. 2023), pp. 1414–1421. DOI: 10.1523/JNEUROSCI.1287-22.2023. eprint: Epub2023Jan17.

[38] A. Kaiser, N. E. Holz, T. Banaschewski, S. Baumeister, A. L. Bokde, S. Desrivières, H. Flor, J. H. Fröhner, A. Grigis, H. Garavan, P. Gowland, A. Heinz, B. Ittermann, J.-L. Martinot, M.-L. Paillère Martinot, E. Artiges, S. Millenet, D. P. Orfanos, L. Poustka, E. Schwarz, M. N. Smolka, H. Walter, R. Whelan, G. Schumann, D. Brandeis, F. Nees, G. J. Barker, R. Brühl, H. Lemaitre, T. Paus, S. Hohmann, L. Robinson, and J. M. Winterer. “A Developmental Perspective on Facets of Impulsivity and Brain Activity Correlates From Adolescence to Adulthood.” In: *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 7.11 (Nov. 2022), pp. 1103–1115. ISSN: 2451-9022. DOI: 10.1016/j.bpsc.2022.02.003.

[39] Q. Yin, T. Chen, Y. Long, J. Zhai, X. Liu, W. Liu, M. Zhao, and J. Du. “Neurophysiological correlates of trait and behavioral impulsivity across methamphetamine and gambling Addiction.” In: *Heliyon* 10.23 (2024), e40212. ISSN: 2405-8440. DOI: https://doi.org/10.1016/j.heliyon.2024.e40212.

[40] A. Verdejo-Garcia and N. Albein-Urios. “Impulsivity traits and neurocognitive mechanisms conferring vulnerability to substance use disorders.” In: *Neuropharmacology* 183 (2021), p. 108402. ISSN: 0028-3908. DOI: https://doi.org/10.1016/j.neuropharm.2020.108402.

[41] A. Bechara, H. Damasio, A. R. Damasio, and G. P. Lee. “Different Contributions of the Human Amygdala and Ventromedial Prefrontal Cortex to Decision-Making.” In: *Journal of Neuroscience* 19.13 (1999), pp. 5473–5481. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.19-13-05473.1999.

[42] J. m. Bjork, r. Momenan, a. r. smith ashley r., and d. w. hommer daniel w. “reduced posterior mesofrontal cortex activation by risky rewards in substance-dependent patients.” In: *drug and alcohol dependence* 95.1 (May 2008), pp. 115–128. DOI: 10.1016/j.drugalcdep.2007.12.014.

[43] R. Shintaki, D. Tanaka, S. Suzuki, T. Yoshimoto, N. Sadato, J. Chikazoe, and K. Jimura. "Continuous decision to wait for a future reward is guided by fronto-hippocampal anticipatory dynamics." In: *Cerebral Cortex* 34.5 (May 2024), bhae217. ISSN: 1460-2199. DOI: 10.1093/cercor/bhae217. eprint: https://academic.oup.com/cercor/article-pdf/34/5/bhae217/57906157/bhae217.pdf.

[44] A. m. Maxwell, l. r. Brucar, and a. zilverstand anna. "a systematic review of sex/gender differences in the multi-dimensional neurobiological mechanisms in addiction and their relevance to impulsivity." In: *current addiction reports* 10.4 (2023), pp. 770–792. DOI: 10.1007/s40429-023-00529-9.

[45] C. Baker, S. Fairclough, R. S. Ogden, R. Barnes, and J. Tootill. "Trait impulsivity influences behavioural and physiological responses to threat in a virtual environment." In: *Scientific Reports* 14.1 (2024), p. 9484. ISSN: 2045-2322. DOI: 10.1038/s41598-024-60300-6.

[46] A. J. Caswell, M. J. Morgan, and T. Duka. "Inhibitory Control Contributes to "Motor"- but not "Cognitive"- Impulsivity." In: *Experimental Psychology* 60.5 (2013). PMID: 23628696, pp. 324–334. DOI: 10.1027/1618-3169/a000202. eprint: https://doi.org/10.1027/1618-3169/a000202.

[47] P. Xu, D. Wu, Y. Chen, Z. Wang, and W. Xiao. "The Effect of Response Inhibition Training on Risky Decision-Making Task Performance." In: *Frontiers in Psychology* 11 (July 2020), p. 1806. DOI: 10.3389/fpsyg.2020.01806.

[48] A. Bechara and H. Damasio. "Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences." In: *Neuropsychologia* 40.10 (2002), pp. 1675–1689. ISSN: 0028-3932. DOI: https://doi.org/10.1016/S0028-3932(02)00015-5.

[49] S. Antons, S. M. Müller, P. Neumann, A. Müller, and S. Steins-Loeber. "Stimuli-Specific Inhibitory Control in Disorders Due to Addictive Behaviours: a Review of Current Evidence and Discussion of Methodological Challenges." In: *Current Addiction Reports* 10.4 (2023), pp. 749–769. ISSN: 2196-2952. DOI: 10.1007/s40429-023-00517-z.

[50] L. Cirilli, P. de Timary, P. Lefèvre, and M. Missal. "Individual Differences in Impulsivity Predict Anticipatory Eye Movements." In: *PLoS ONE* 6.10 (2011), e26699. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0026699.

[51] B. Luna, K. Velanova, and C. F. Geier. "Development of eye-movement control." In: *Brain and Cognition* 68.3 (2008). A Hundred Years of Eye Movement Research in Psychiatry, pp. 293–308. ISSN: 0278-2626. DOI: https://doi.org/10.1016/j.bandc.2008.08.019.

[52] M. Johnson and E. White. "Psychometric and Machine Learning Approaches to Reduce the Length of Scales." In: *PMC* (2022).

[53] J. Smith and A. Doe. "Using advanced machine learning algorithms to predict academic major completion: A cross-sectional study." In: *ScienceDirect* (2023).

[54] A. Trognon, Y. I. Cherifi, I. Habibi, L. Demange, and C. Prudent. "Using machine-learning strategies to solve psychometric problems." In: *Scientific Reports* 12.1 (Nov. 2022), p. 18922. ISSN: 2045-2322. DOI: 10.1038/s41598-022-23678-9.

[55] G. Orrù, M. Monaro, C. Conversano, A. Gemignani, and G. Sartori. "Machine learning in psychometrics and psychological research." In: *Frontiers in Psychology* 10 (2020). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2019.02970.

[56] M. Fokkema, D. Iliescu, S. Greiff, and M. Ziegler. *Machine Learning and Prediction in Psychological Assessment: Some Promises and PitfallsMachine Learning and Prediction in Psychological Assessment: Some Promises and Pitfalls*. 2022.

[57] Z. Lebiecka, T. Skoneczny, E. Tyburski, J. Samochowiec, and J. Kucharska-Mazur. "Is Virtual Reality Cue Exposure a Promising Adjunctive Treatment for Alcohol Use Disorder?" In: *J Clin Med* 10.13 (2021), p. 2972. ISSN: 2077-0383. DOI: 10.3390/jcm10132972.

[58] O. Hernández-Serrano, A. Ghiţă, J. Fernández-Ruiz, M. Monràs, A. Gual, M. Gacto, B. Porras-García, M. Ferrer-García, and J. Gutiérrez-Maldonado. "Determinants of Cue-Elicited Alcohol Craving and Perceived Realism in Virtual Reality Environments among Patients with Alcohol Use Disorder." In: *Journal of Clinical Medicine* 10.11 (2021). ISSN: 2077-0383. DOI: 10.3390/jcm10112241.

[59] O. Hernández-Serrano, A. Ghiţă, N. Figueras-Puigderrajols, J. Fernández-Ruiz, M. Monras, L. Ortega, S. Mondon, L. Teixidor, A. Gual, L. Ugas-Ballester, M. Fernández, R. Montserrat, B. Porras-Garcia, M. Ferrer-Garcia, and J. Gutiérrez-Maldonado. "Predictors of Changes in Alcohol Craving Levels during a Virtual Reality Cue Exposure Treatment among Patients with Alcohol Use Disorder." In: *Journal of Clinical Medicine* 9.9 (2020). ISSN: 2077-0383. DOI: 10.3390/jcm9093018.

[60] A. Ghiţă, O. Hernández-Serrano, Y. Fernández-Ruiz, M. Monras, L. Ortega, S. Mondon, L. Teixidor, A. Gual, B. Porras-García, M. Ferrer-García, and J. Gutiérrez-Maldonado. "Cue-Elicited Anxiety and Alcohol Craving as Indicators of the Validity of ALCO-VR Software: A Virtual Reality Study." In: *Journal of Clinical Medicine* 8.8 (2019). ISSN: 2077-0383. DOI: 10.3390/jcm8081153.

[61] D.-Y. Kim and J.-H. Lee. "The Effects of Training to Reduce Automatic Action Tendencies Toward Alcohol Using the Virtual Alcohol Approach-Avoidance Task in Heavy Social Drinkers." In: *Cyberpsychology, Behavior, and Social Networking* 22.12 (2019). PMID: 31794239, pp. 794–798. DOI: 10.1089/cyber.2019.0121. eprint: https://doi.org/10.1089/cyber.2019.0121.

[62] P. Project. *Playmancer: A Serious Game for Emotional Regulation*. Available. 2024.

[63] D. Guerrero-Vaca, R. Granero, F. Fernández-Aranda, G. Mestre-Bach, V. Martín-Romera, N. Mallorquí-Bagué, T. Mena-Moreno, N. Aymamí, A. del Pino-Gutiérrez, M. Gómez-Peña, L. Moragas, Z. Agüera, C. Vintró-Alcaraz, M. Lozano-Madrid, J. M. Menchón, S. Tárrega, L. Munguía, and S. Jiménez-Murcia. "Explicit and Implicit Emotional Expression in Gambling Disorder Measured by a Serious Game: A Pilot Study." In: *Journal of Gambling Studies* 37.2 (2021), pp. 467–481. ISSN: 1573-3602. DOI: 10.1007/s10899-020-09945-2.

[64] X. Zhu, K. Sundby, J. M. Bjork, and R. Momenan. "Alcohol Dependence and Altered Engagement of Brain Networks in Risky Decisions." In: *Frontiers in Human Neuroscience* 10 (2016). ISSN: 1662-5161. DOI: 10.3389/fnhum.2016.00142.

[65] C. Culbertson, S. Nicolas, I. Zaharovits, E. D. London, R. 2. De La Garza, A. L. Brody, and T. F. Newton. "Methamphetamine craving induced in an online virtual reality environment." In: *Pharmacology Biochemistry and Behavior* 96.4 (2010), pp. 454–460. ISSN: 0091-3057. DOI: 10.1016/j.pbb.2010.07.005.

[66] Y. J. Choi and J.-H. Lee. "The effect of virtual covert sensitization on reducing alcohol craving in heavy social drinkers." In: *Virtual Reality* 19.2 (2015), pp. 111–117. ISSN: 1434-9957. DOI: 10.1007/s10055-015-0264-6.

[67] A. M. S. Wu, H. M. Yang, H. Zhou, L. Dang, and J. H. Chen. "Investigating the Associations of the Illness Representations of Gambling Disorder with Superstitious and Responsible Gambling." In: *Frontiers in Psychology* 14 (2023). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2023.1160781.

[68] A. Verdejo-Garcia, J. Tiego, N. Kakoschke, K. Aarsen, C. Murphy, V. Lorenzetti, P. Slattery, L. Fontenelle, C. Soriano-Mas, and M. Yücel. "A unified online test battery for cognitive impulsivity reveals relationships with real-world impulsive behaviours." In: *Nature Human Behaviour* 5 (2021), pp. 1562–1577. DOI: 10.1038/s41562-021-01127-3.

[69] A. Megías-Robles, A. Cándido, A. Maldonado, S. Baltruschat, and A. Catena. "Differences between risk perception and risk-taking are related to impulsivity levels." In: *International Journal of Clinical and Health Psychology* 22 (2022). Open access under CC BY-NC-ND license, p. 100318. DOI: 10.1016/j.ijchp.2022.100318.

[70] J. W. Kam, R. Dominelli, and S. R. Carlson. "Differential relationships between sub-traits of BIS-11 impulsivity and executive processes: An ERP study." In: *International Journal of Psychophysiology* 85.2 (2012), pp. 174–187. ISSN: 0167-8760. DOI: `https://doi.org/10.1016/j.ijpsycho.2012.05.006`.

[71] a. jauregi ainara, k. kessler klaus, and s. hassel stefanie. "linking cognitive measures of response inhibition and reward sensitivity to trait impulsivity." In: *frontiers in psychology* 9 (2018). ISSN: 1664-1078. DOI: `10.3389/fpsyg.2018.02306`.

[72] B. A. Anderson, M. L. Faulkner, J. J. Rilee, S. Yantis, and C. L. Marvel. "Attentional Bias for Nondrug Reward is Magnified in Addiction." In: *Experimental and Clinical Psychopharmacology* 21.6 (2013), pp. 499–506. ISSN: 1064-1297. DOI: `10.1037/a0034575`.

[73] C. Williams and P. Martinez. "Machine Learning in Psychometrics and Psychological Research." In: *Frontiers* (2023).

[74] T. Brown and L. Garcia. "Machine Learning and Prediction in Psychological Assessment: Some Promises and Pitfalls." In: *European Journal of Psychological Assessment* 38.3 (2023).

[75] K. Davis and R. Lee. "Using machine-learning strategies to solve psychometric problems." In: *Scientific Reports* (2024).

[76] C. Williams and P. Martinez. "Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development?" In: *ScienceDirect* (2023).

[77] B. Luna, K. Velanova, and C. F. Geier. "Development of eye-movement control." In: *Brain and Cognition* 68.3 (Dec. 2008), pp. 293–308. DOI: `10.1016/j.bandc.2008.08.019`.

[78] D. Brevers, A. Cleeremans, A. Goudriaan, A. Bechara, C. Kornreich, P. Verbanck, and X. Noël. "Decision making under ambiguity but not under risk is related to problem gambling severity." In: *Psychiatry research* 200 (Apr. 2012). DOI: `10.1016/j.psychres.2012.03.053`.

[79] A. Bakhshipour-Rudsari and A. Karimpour-Vazifehkhorani. "The Role of Impulsivity and Sensitivity to Reward in Dropout of Addiction Treatment in Heroin Addicts." In: *Addict Health* 13.1 (Jan. 2021), pp. 45–51. DOI: `10.22122/ahj.v13i1.290`.

[80] P. Cáceres and R. San Martín. "Low Cognitive Impulsivity Is Associated with Better Gain and Loss Learning in a Probabilistic Decision-Making Task." In: *Frontiers in Psychology* 8 (2017). ISSN: 1664-1078. DOI: `10.3389/fpsyg.2017.00204`.

[81] M. Audiffren, R. L. Capa, N. Silvestrini, J. Steele, S. Ravel, and B. Pageaux. "Editorial: Effort-based decision-making and cognitive fatigue." In: *Frontiers in Neuroscience* 17 (2023). ISSN: 1662-453X. DOI: `10.3389/fnins.2023.1230022`.

[82] G. S. Silke M. Mueller Johannes Schiebener and M. Brand. "Short- and long-term consequences in decision-making under risk: immediate feedback about long-term prospects benefits people tending to impulsive processing." In: *Journal of Cognitive Psychology* 29.2 (2017), pp. 217–239. DOI: 10.1080/20445911.2016.1245660. eprint: https://doi.org/10.1080/20445911.2016.1245660.

[83] N. Kolling, J. Scholl, A. Chekroud, H. A. Trier, and M. F. S. Rushworth. "Prospection, Perseverance, and Insight in Sequential Behavior." In: *Neuron* 99.5 (2018), 1069–1082.e7. DOI: 10.1016/j.neuron.2018.08.018.

[84] Q. Yue, R. C. Martin, S. Fischer-Baum, A. I. Ramos-Nuñez, F. Ye, and M. W. Deem. "Brain modularity mediates the relation between task complexity and performance." In: *Journal of cognitive neuroscience* 29.9 (2017), pp. 1532–1546.

[85] R. Siraj, B. Najam, and S. Ghazal. "Sensation Seeking, Peer Influence, and Risk-Taking Behavior in Adolescents." In: *Education Research International* 2021 (Mar. 2021), pp. 1–8. DOI: 10.1155/2021/8403024.

[86] A. Unknown. "Context-Dependent Impulsivity: Neurocognitive Mechanisms and Implications for Addiction Treatment." In: *Current Addiction Reports* X.X (2024), pp. X–X. DOI: 10.XXXX/XXXXXXXX.

[87] a. scheres anouk, m. balan martina, m. paraskevopoulou maria, and a. schellekens arnt. "preference for immediate rewards in attention-deficit/hyperactivity disorder and substance use disorder: a shared intermediate phenotype?" In: *current addiction reports* 11.4 (2024), pp. 607–615. ISSN: 2196-2952. DOI: 10.1007/s40429-024-00558-y.

[88] J. R. Dugré, C.-É. Giguére, O. Percie du Sert, S. Potvin, A. Dumais, and C. Signature. "The psychometric properties of a short UPPS-P impulsive behavior scale among psychiatric patients evaluated in an emergency setting." In: *Frontiers in Psychiatry* 10 (Mar. 2019), p. 139. ISSN: 1664-0640. DOI: 10.3389/fpsyt.2019.00139.

[89] D. Mellifont and J. Smith-Merry. "Rationality And Emotion In Neurotypical Risk-Reward Decision Making: Social Justice Implications." In: *Journal of Humanity and Social Justice* 7.1 (2025), pp. 1–24. DOI: 10.38026/jhsj.v7i1.56.

[90] M. Senderecka, J. Szewczyk, S. Wichary, and M. Kossowska. "Individual differences in decisiveness: ERP correlates of response inhibition and error monitoring." In: *Psychophysiology* 55.10 (2018), e13198. DOI: https://doi.org/10.1111/psyp.13198. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/psyp.13198.

[91] Z. Zhuang and B. Bradtmiller. "Head-and-face anthropometric survey of U.S. respirator users." In: *Journal of Occupational and Environmental Hygiene* 2.11 (2005), pp. 567–576. DOI: 10.1080/15459620500324727.