Article

An Immersive Virtual Reality Game for Predicting Risk Taking through the Use of Implicit Measures

Carla de-Juan-Ripoll *, José Llanes-Jurado, Irene Alice Chicchi Giglioli [ID], Javier Marín-Morales [ID] and Mariano Alcañiz [ID]

Instituto de Investigación e Innovación en Bioingeniería (i3B), Universitat Politècnica de València, 46022 Valencia, Spain; jllajur@i3b.upv.es (J.L.-J.); alicechicchi@i3b.upv.es (I.A.C.G.); jamarmo@i3b.upv.es (J.M.-M.); malcaniz@i3b.upv.es (M.A.)
* Correspondence: cardejua@i3b.upv.es; Tel.: +34-617626946

Featured Application: The tool presented in this article can be applied as an ecological measure for evaluating decision-making processes in risky situations. It can be used in different contexts from both Occupational Safety and Health practices and for research purposes.

Abstract: Risk taking (RT) measurement constitutes a challenge for researchers and practitioners and has been addressed from different perspectives. Personality traits and temperamental aspects such as sensation seeking and impulsivity influence the individual's approach to RT, prompting risk-seeking or risk-aversion behaviors. Virtual reality has emerged as a suitable tool for RT measurement, since it enables the exposure of a person to realistic risks, allowing embodied interactions, the application of stealth assessment techniques and physiological real-time measurement. In this article, we present the assessment on decision making in risk environments (AEMIN) tool, as an enhanced version of the spheres and shield maze task, a previous tool developed by the authors. The main aim of this article is to study whether it is possible is to discriminate participants with high versus low scores in the measures of personality, sensation seeking and impulsivity, through their behaviors and physiological responses during playing AEMIN. Applying machine learning methods to the dataset we explored: (a) if through these data it is possible to discriminate between the two populations in each variable; and (b) which parameters better discriminate between the two populations in each variable. The results support the use of AEMIN as an ecological assessment tool to measure RT, since it brings to light behaviors that allow to classify the subjects into high/low risk-related psychological constructs. Regarding physiological measures, galvanic skin response seems to be less salient in prediction models.

Keywords: virtual reality; risk taking; personality; sensation seeking; impulsivity; eye tracking; galvanic skin response; implicit measures

1. Introduction

Risk taking (RT) is a component of the decision-making process in uncertain situations, in which the subject rationally knows the probability of each outcome [1,2]. The decision-making process is influenced by three main elements [3,4]: decision features, which are the characteristics of the decision itself, such as the ordering of the choice options [5]; situational factors, which refer to the context of the decision, for example, time pressure [6]; and individual differences, which have been identified as the perception of benefits, the perception of risks and risk attitude in the field of RT [7]. In the first stage of RT process, the subject thinks about the possible positive/negative outcomes of his/her actions before acting [8]. During this process, emotional states have an influence on the weighting of cost-benefit assessment [9], and its relation with RT has been widely studied. On the first hand, it has been suggested that people experiencing positive emotions tend

to maintain this positive state [10] and protect themselves from the potential negative outcomes of a decision [11], which leads to risk avoidance. On the other hand, positive emotions can be associated with greater risk tolerance [12], promoting RT. These results suggest that the relation between emotional states and RT relies upon the decision-making context. Considering this, in the present paper we are focusing on a cross-domain trait, risk proneness, understood as the propensity to be attracted to potentially risky activities [13], which is related to personality traits and temperamental aspects such as sensation seeking and impulsivity, which influence the individual's approach to RT, prompting risk-seeking or risk-aversion behaviors [8].

Personality, Sensation Seeking, Impulsivity and RT

Among the numerous personality models developed in psychology research, the Big Five factorial model of personality—composed of neuroticism, extraversion, openness, agreeableness and conscientiousness factors [14]—seems to be the most generally recognized in terms of the study of the relation between personality and RT. On the first hand, neuroticism, which is connected to sensitivity to punishment and negative affect [15]; agreeableness, which is characterized by trust, straightforwardness and compliance; and conscientiousness, understood as a need for compliance under conditions of conformity and control; have been related to risk aversion in most domains [16–19]. On the other hand, extraversion, as a generalized need for stimulation; and openness to experience, which relates to cognitive risk seeking, acceptance of experimentation, and tolerance of uncertainty, change and innovation; have been related to risk seeking [20].

Sensation seeking has been defined as "the seeking of varied, novel, complex and intense sensations and experiences, and the willingness to take physical, social, legal, and financial risks for the sake of such experience" [21] (p. 27) and has been positively related to RT in several domains, such as recreation, health, career, finance, safety, social life and sex [18,22]. Zuckerman [16] identified four dimensions of the sensation seeking trait: thrill and adventure seeking, which reflects a desire to engage in physical activities that provide unusual sensations [21]; experience seeking, which has been related to lower sensitivity to aversive stimulation [23]; disinhibition, which appears as significant predictor of RT in several domains, as in rule-breaking behaviors and violations of societal norms [22,24,25] and boredom susceptibility, which is connected to intolerance for routine and repetitive activities [26].

Impulsivity has been defined as the "predisposition toward rapid, unplanned reactions to internal or external stimuli without regard to the negative consequences of these reactions to the impulsive individual or to others" [27] (p. 1784) and has been associated with RT in terms of drug use, drink driving and seatbelt use, among others [28,29]. Some authors have also demonstrated its connection with emotional self-control, inhibition and, especially, the management of frustrating situations [30,31]. Whiteside and Lynam [32] argued that impulsivity is made up of a set of five impulse-related traits: negative urgency, lack of premeditation, lack of perseverance, sensation seeking and positive urgency. According to [32,33], negative and positive urgency traits relate to the tendency to exhibit impulsive behaviors when facing negative/positive situations. Lack of premeditation relates to thoughtless behaviors and to the tendency to choose alternatives with short-term rewards, rather than options that might lead to more valuable but delayed rewards. Lack of perseverance reflects an absence of focus on a boring or difficult activity. Sensation seeking is an attraction toward exciting, new and potentially dangerous experiences.

2. Measurement of RT

RT measurement constitutes a challenge for researchers and practitioners and has been addressed from different perspectives. To date, most of the theoretical constructs used in RT assessment are based on explicit measures such as self-reports, although these measures have been applied from different points of view. While some authors employ self-reported measures to assess risky-related psychological constructs, such as personality,

impulsivity and sensation seeking [34–36]; other authors use self-reported daily habits as a measure of RT [8,37]. Alternatively, ref. [38] developed a measure of the tendency to engage in real-life RT behaviors in different domains: ethic, financial, health, recreational and social.

However, self-reported measures present some limitations. On the first hand, with the use of these instruments it is assumed that humans are able to think and verbalize accurately about their attitudes, emotions and behaviors, while it has been demonstrated that most of the brain processes that regulate attitudes, emotions and behaviors are not conscious, and consequently, cannot be verbalized [39–41]. On the other hand, questionnaires have an important intrinsic bias since individuals need to remind past situations or imagine future experiences to answer, rather than actually undergoing the experiences that the researchers wish to analyze [42].

To overcome these limitations, the approach of "stealth assessment" [43] emerged focusing on the study of how psycho-cognitive states can be assessed in an ecological, non-intrusive, non-biased way. Studies under this paradigm record subjects' performance during a serious game, and then conclusions are drawn about individual competencies based on the data [44,45]. In the field of RT, the Bechara gambling task (BGT) [46] and the balloon analogue risk task (BART) [34] could be considered the most used measures that aim to assess RT under this methodology. In BGT, participants are given four decks of cards and are asked to choose a card from any one of the four decks. Once a card is chosen, it is turned over, and the amount of money won or lost for choosing that card is revealed. This is repeated for 100 times, and the player is never told the distribution of wins and losses associated with each deck, and instead the distributions are learned from experience. In BART, a balloon is presented in the middle of a screen, and subjects are asked to pump it as much as possible, knowing that it could exploit at any time. At the beginning of the task participants are told that they will obtain a financial reward the more they could inflate each balloon without breaking it. Although the reliability of these tools has been retested [47,48], it has been proved that the correspondence between performance in neuropsychological tests and real-life behaviors is very weak [49–51].

2.1. Virtual Reality for RT Assessment

Conversely, virtual reality (VR) provides the capacity of simulate real experiences in which subjects can interact as if they were in the real world [52], and there is empirical evidence demonstrating similarities between the neural mechanisms that subjects experience when immersed in a virtual environment and in those real life [53,54]. VR allows to record the behavioral responses of the users while they are interacting with a virtual environment [55], making VR an innovative, effective, active, engaging and adaptive tool that has been applied in numerous fields of human behavior research e.g., [52], providing better results than 2D solutions [56].

VR has emerged as a suitable tool for RT measurement, since it enables the exposure of a person to realistic risks, allowing embodied interactions, the application of stealth assessment techniques and physiological real-time measurement [57]. On the basis of this, we developed the spheres and shield maze task (SSMT) [58], a virtual environment for RT measurement. It consisted in an out-of-context maze, through which participants had to pass from start to finish before three minutes, accumulating as much "karma" as possible by collecting spheres down the road. Participants could lose "karma" if they were attacked by a risk. Furthermore, participants had the option of activating a shield, which protected them from the risks. This virtual environment supposed a first approach for the measurement of the risk-related constructs sensation seeking and impulsivity, although it presented two main limitations. First, the practice session was too short and insufficient. Second, it measured only three variables: "karma", distance covered and shield use, ignoring real-time behavioral and psychophysiological measures. In this article, we propose an enhanced version of the SSMT, by which the authors intend to overcome these issues.

2.2. Implicit Measures in VR

The interactions of the users with the virtual environment can be also studied by the analysis of their gaze movements, which have shown to be related to information processing in risky decisions [59] and problem solving [60]. The eye tracking (ET) measure can be integrated into a VR set-up, in order to record fixations and eye movements during an experience in a virtual environment. This technology has been applied in combination with VR for the study of the influence of contextual elements in human behavior, such as in street robbery [61], identifying if the presence of particular components of a physical space can influence in decision-making. Furthermore, ET has been employed to study whether if exists a relationship among gaze patterns and human behavior [62,63], or even if these gaze patterns could contribute to predict humans' decisions [64].

In the field of RT, ET has been employed as a reliable indicator of information processing patterns in risky decisions. On the first hand, greater number of fixations, longer fixations and larger quantity of available information fixed have been related to deeper pre-decision processes, which lead to risk aversion [65–68]. On the other hand, in a study with construction workers, lower dwell time was connected with a higher risk perception [69]. The authors interpreted this result as follows: participants with higher risk perception identified the hazards rapidly, so they could spend their time searching other possible hazards present in the situation.

In addition to behavioral measures, physiological measures have been proposed as implicit measures of human behavior [42]. Galvanic skin response (GSR) has been successfully used as a measure of implicit processes such emotional arousal [70], which plays a decisive role in the decision-making process. GSR has been employed in combination with VR to evaluate the stress generated by changes in contextual aspects, such as architectural stimuli [71], as predictor of anxiety level [72] and as a measure to discriminate between Autism Spectrum Disorder and typical development populations [73], among others.

In the field of RT, high physiological arousal acts as a "warning signal" in risky situations and tends to lead to safe decisions [1]. This relationship has been demonstrated to be mediated by emotional intelligence, such a way that, low emotional intelligence may lead to maladaptive decision-making, due to an impaired interpretation of physiological arousal [74]. Additionally, situational factors, such as time pressure, have an influence on the relationship between GSR and RT. In an experiment with two kind of decisions (time pressure and time delay), the relationship between GSR and RT was positive in situations under time pressure, and negative in situations under time delay [75].

Despite these measures having been widely adopted in VR-based experiments, to our knowledge, ET and GSR have not been employed in combination with VR to evaluate RT.

2.3. The Current Study

Starting from these premises, we present the assessment on decision making in risk environments (AEMIN) tool, as a new interactive virtual environment for RT measurement. Compared to the SSMT, AEMIN has longer duration, which allows a wider and enriched recording of information from the subjects, and contains more elements along the maze, such as spheres of different colors and a pause button. Furthermore, features in AEMIN were rated depending on whether the subject was in a risk zone or in a no risk zone, to provide further information about the subjects´ behavior depending on the situation. Additionally, the appearance and characteristics of the risks have been improved, in order to provide a more natural experience and consequently, more natural behaviors. A detailed description of AEMIN is provided in the Materials and Methods section.

The main aim of this study is to discriminate participants with high versus low scores in the measures of neuroticism, extraversion, openness to experience, agreeableness, conscientiousness, sensation seeking and impulsivity, through their behaviors and physiological responses during playing AEMIN. Applying machine learning (ML) methods to the dataset we explored: (a) if through these data it is possible to discriminate RT domains, sensation seeking and impulsivity, allowing to qualitatively determine a general level of RT for

each subject; and (b) which parameters better discriminate between the two populations in each variable.

3. Materials and Methods

3.1. Participants

A group of 98 subjects was recruited to participate in the experiment. They were balanced in terms of gender (56 men and 55 women) and age (35% under 30, 35% among 30–45, 30% above 45; mean age = 37.08, SD = 10.91). Prior to their participation, they received documentary information on the study and gave their written consent for their involvement. The responses were anonymized and randomized to ensure the privacy of the information. The study obtained the ethical approval of the Ethical Committee of the Polytechnic University of Valencia (P4_18_06_19).

3.2. Self-Reported Measures

The risk-related constructs were measured by means of a battery of self-reported measures:

Personality: Spanish version of the NEO five-factor inventory (NEO-FFI). This comprises 60 items and is composed by the factors neuroticism, extraversion, openness, conscientiousness and agreeableness [76,77]. The reliability coefficients' Cronbach's alphas ranged from 0.75 to 0.83. The internal consistency of the scales in the present study was: neuroticism α = 0.77; extraversion α = 0.85; openness α = 0.79; agreeableness α = 0.75; conscientiousness α = 0.84.

Sensation seeking: Spanish version of the 40-item Sensation Seeking Scale-V (SSS-V) [78,79]. This includes subscales for thrill and adventure seeking, experience seeking, disinhibition and boredom susceptibility, and a total sensation seeking score. The reliability coefficients' Cronbach's alphas ranged between 0.67 and 0.81, which suggests the subscales have acceptable internal consistency. The internal consistency of the scale in the present study was 0.77.

Impulsivity: Short Spanish version of the UPPS-P impulsive behavior scale [32,80]. Composed of 20 items, this measures five impulsivity traits: negative urgency, lack of premeditation, lack of perseverance, sensation seeking and positive urgency. The Cronbach's alphas coefficients ranged from 0.66 to 0.81. The internal consistency of the scales in the present study was: negative urgency α = 0.72; lack of premeditation α = 0.77; lack of perseverance α = 0.78; sensation seeking α = 0.79; positive urgency α = 0.60.

As a measure of the sense of presence in the virtual environment, participants responded the Sense of Presence Inventory, which is composed by the dimensions of spatial presence, engagement, ecological validity and negative effects (ITC-SOPI) [81]. Cronbach´s alphas coefficients in ITC-SOPI ranged from 0.76 to 0.94. The internal consistency of the scales in the present study was: spatial presence α = 0.91; engagement α = 0.84; ecological validity α = 0.77; negative effects α = 0.86.

3.3. The Virtual Environment

We present the assessment on decision making in risk environments (AEMIN) tool, as a new interactive virtual environment for RT measurement. As an extension of the SSMT [58], AEMIN is an interactive virtual environment that is composed by two mazes that participants must pass through from start to finish before the allocated time expires without (virtually) hurting themselves (see Figure 1a). One of the mazes must be solved individually, while in the other one the subject is accompanied by four avatars. The avatars are represented by robots (see Figure 1b), which can express basic emotions through a screen located on their faces.

Participants have 10 min to negotiate each maze and they are instructed to accumulate as much energy as possible, since it is the source of life of their avatar. If a robot is poor of energy, it shows dying breathing and its movements are slower, which implies a waste of time to find the exit of the maze. There are green spheres distributed throughout the maze,

which earn participants´ energy if they collect them. Furthermore, participants can lose energy if they are attacked by a risk. These risks are also distributed throughout the maze and are of four types: bridges, swarms of insects, storms and haunted rooms. Some spheres are close to hazards and others are located in no-risk zones. Participants have the option of activating a shield, which protects them from the risks. When the shield is active, the user's speed is reduced and (s)he cannot collect any spheres. The shield is a finite resource that subjects need to optimize. While passing through the maze, the participants have information about the remaining time (orange circle in Figure 1), their level of energy (green circle in Figure 1), and the battery life of the shield (blue circle in Figure 1). Table 1 shows a brief description of each risk and the consequences of each one for the robots. In addition, there are some purple spheres hidden in some endless roads. Catching one of these purple elements can take uncertain effects, such as simplifying the route or subtracting 10 s to the participant. The game can be paused by the participant at any time, so that (s)he is moved to a virtual relaxing room, until (s)he is ready to return to the game. The reason why we included this virtual room is that the use of it by the participant can be considered as an inhibition strategy, and as an indicator of emotional self-control. The navigation metaphor is indirect walking, in which pushing down on the controller's integrated touchpad moves the user´s avatar in the direction (s)he is facing at 2 m/s (speeds above 3 m/s. can increase cybersickness symptoms [82]). Before undertaking the AEMIN game, the participants underwent a guided practice session in which they learned how to travel through the virtual environment, how to collect spheres and how to activate the shield.
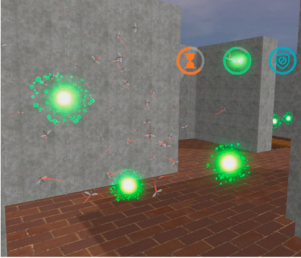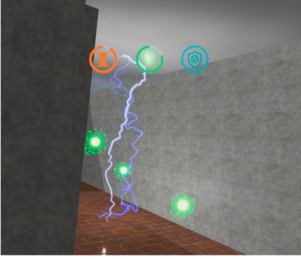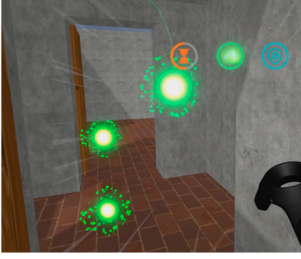

(**a**)　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 1. (**a**) AEMIN maze; (**b**) Avatars.

The virtual environment was developed in Unity (version 2018.4.1f1) using c# as programming language. Participants performed the AEMIN game using the HTC Vive Pro-eye head mounted display[1], with 2880×1600 pixels (1440×1600 per eye), a field of view of $110°$ degrees, working at 90 Hz refresh rate. The ET data were obtained from the Unity VR through the ET SDK (SRanipal), with a maximum frequency of 120 Hz and an accuracy of $0.5°-1.1°$.

Galvanic Skin Response (GSR) is also recorded in the experimentation. Data was collected with the Shimmer3 GSR sensor2, sampled at 128 Hz. We measured skin conductance between two reusable electrodes attached to human fingers.

The computer used was an Intel Core i7-770 CPU 3.60 GHz with an NVIDIA GeForce GTX 1070.

Table 1. Description of the risks.

Risk	Description	Consequences
Bridge	Walkway that allows the robots to cross from one side to another to continue the path. Participants can cross it as many times as they like in both directions.	If a robot falls into the pit, it will lose part of the battery of its shield. Additionally, the robot reappears at the beginning of the bridge, and this supposes a little time to cross again.
Swarm of insects	Swarm of flying insects that flits over an area of the maze.	In an insect bites a robot, it will suffer blurred vision a few seconds later, which supposes a little time to recover the normal vision. Furthermore, this makes the robot to lose energy.
Storm	In some areas of the maze the weather is stormy.	If a lightning strikes a robot, it will suffer a large loss of energy.
Haunted room	Room that becomes increasingly smaller when someone enters it. The room has an enter and an exit door and participants can cross them as many times as they like, in both directions. Participants are asked to catch the key inside the room to open the doors.	Opening the doors is an investment of time.

3.4. Experimental Procedure

Each participant responded to the self-report questionnaires on a personal computer. The process took approximately 30 min, and was completed in an experimental room, supervised by a research assistant. The subject was thereafter conducted to a second experimental room where (s)he received a brief contextualization of the VR game. Consecutively, the research assistant equipped the participant with the GSR device and the HMD system in the correct position. After a calibration process of the eye tracking apparatus, the subject was asked to sit and relax during 90 s in order to record a GSR baseline. During this period, the subject listened to a relaxing audio to create a common state of calm. After that, the subject stood up and completed the practice session, which included a brief presentation of the avatars. Hereafter, the participant solved the two mazes (50% of the participants began

by the individual scene, and the other 50% started by the group of avatars). Finally, the subjects responded to the presence questionnaires in a personal computer.

3.5. Data Processing

The virtual environment (VE) is divided in two areas: risk zone and no risk zone. The defined risk zone areas correspond to the situation where the subject is inside a risk such as bridge, swarm of insects, storm and haunted room. The no risk zone is defined for the situations where the subject is not inside of a risk zone. According with this division, we analyzed two groups of variables: (a) measures in risk zones; and (b) measures in no risk zones. The features were divided depending on the source of data where have been computed. Three different sources of data were established: VR, ET and GSR. Table 2 summarizes the complete set of features that was used from each source.

Table 2. Description of the set of features obtained by data source.

Data Source		Risk Zone		No Risk Zone	
		Features	[1]N	Features	[1]N
VR	Navigation	Time spent Visits to each risk Distance covered Time walking Velocity Acceleration	28	Time spent - Distance covered Time walking Velocity Acceleration	20
	Interactions	Green spheres caught - Pause button use Shield use Total interactions		Green spheres caught Purple spheres caught Pause button use Shield use Total interactions	
ET		Time to first fixation Number of fixations Fixation duration Number of objects seen Number of saccades Angular saccade distance Velocity of saccades Distance in saccades	37	Time to first fixation Number of fixations Fixation duration Number of objects seen Number of saccades Angular saccade distance Velocity of saccades Distance in saccades	34
GSR		Mean, std and median signal Phasic and tonic value Number of phasic peaks Skewness of phasic signal Kurtosis of phasic signal Entropy of phasic signal	18	Mean, std and median signal Phasic and tonic value Number of phasic peaks Skewness of phasic signal Kurtosis of phasic signal Entropy of phasic signal	18

[1] Number of features.

Features study from VR data are divided between navigation and interaction features. The navigation part obtains a set of features related with the trajectory of the subject in the maze whereas the interaction features counts the number of times that the subject uses or touch some element in the maze.

ET data was processed in order to obtain a classification between fixations and saccades, using the dispersion threshold (DT) algorithm with $1°$ as a dispersion threshold and 0.25 s as a time window threshold [83]. A complete set of features was obtained from the classification between fixation and saccade.

Before the obtainment of features from GSR, two previous steps were done. The first of them was the manual cleaning of the signal. Commonly, GSR signal could suffer from different types of noises that hide correlations between the signal of the subject and its level of stress [84]. The manual correction was done using *Ledalab*[3] software in MATLAB.

The second step was the division of the signal into phasic and tonic components using continuous decomposition analysis (CDA) [85]. After this pre-processing, a set of features was obtained from the raw signal and the phasic and tonic components including time and non-linear domain analysis [86].

In order to approach a classification problem, the target variables were divided in two groups: high score and low score. The division was done according with the normality of the distribution of each target variable. If the distribution was normal, the target was segmented by the mean target value, whereas if the target distribution was not normal, the target was segmented by the median. The significance level between groups in each target variable was checked through a statistical *t*-test in features with normal distribution and Mann–Whitney for features without a normal distribution.

3.6. Statistical Analysis

Firstly, a multivariate outlier detection was performed by group of variables (VR, ET and GSR). Mahalanobis distance between every subject and the probability that it belongs to a Chi-square distribution was calculated. Subjects that belonged to the most extreme 1% of the data distribution were defined as outliers.

Some pre-processing steps were done before the modelling study. The variables with a Pearson-correlation higher than 0.95 in absolute value were removed. After that, no-normal feature distributions were transformed using logarithms. The variables which after this transformation were normal distributed keep the transformation, the ones that were not normal were not transformed.

A ML method was applied to find the best possible selection of features that classify whether the subject have a high or low score in the studied target variables. The used model was a support vector machine (SVM) [87]. The pipeline for the modeling of the data is equal for every target.

The pipeline is designed to find the best possible features to explore the importance of each one in combination with the rest of features. To address this goal, the ML pipeline removes iteratively the feature which achieves the lowest accuracy for each model in the iteration. Iteration k, computes the mean accuracy in a cross-validation (CV) of 10 folds and 2 repetitions. After that, a backward feature selection (BFS) [88] method removes one feature selecting the set of k-1 features with highest accuracy. This method also uses a CV with 10 folds and 2 repetitions. The process ended-up when only one feature remains. The set of features with highest accuracy are selected. After that, an hyperparameter tunning is performed to the SVM. Finally, the model is validated in a CV of 10 folds with 4 repetitions. The average and standard deviation of the metrics accuracy, kappa, true positive ratio (TPR) and true negative ratio (TNR) were reported. Moreover, the experiment explored the importance of each group of features using four different sets based on the source. Three datasets including VR, GSR and ET features respectively were created. Moreover, an additional dataset which is called ALL that joins all the features was included.

To check the overfitting of the ML pipeline, the obtained results are compared against the ones obtained from a generated random target. The unique condition imposed to the generation of this random target is that it must have a coincidence in its labels lower than a 67.5%, compared with the rest of the real targets, in order to avoid a random target very similar to a real one. The objective is to compare, according with a one-way ANOVA test, the statistical distribution of the set of accuracies obtained from the last CV of the ML pipeline of the random target and each real target. Six random targets are generated to extend the number of accuracy samples from the random targets. Figure 2 shows a scheme of the ML pipeline used and the overfitting check method exposed. If the comparison between both distributions shows a statistical difference (p-value < 0.05), it supports that the ML pipeline is over the chance level. Finally, the dataset with highest accuracy is reported as the best classification model.

Regarding the presence questionnaires, mean and standard deviation for each dimension were calculated.
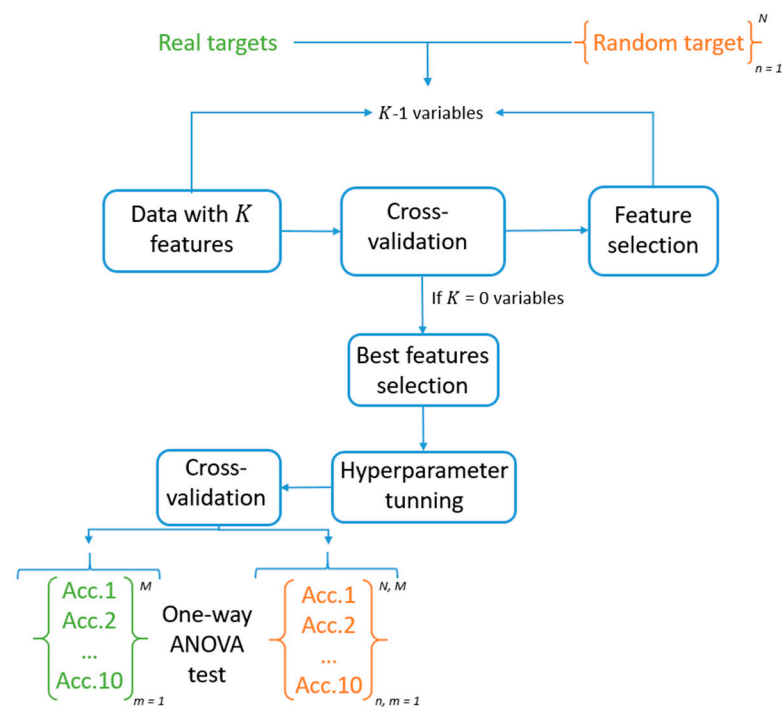
Figure 2. Scheme of ML pipeline and its overfitting check using a set of generated random targets. *M* indicates the number of the CV repetitions set to 4, and *N* the number of random targets generated, equal to 6.

4. Results

From the initial 98 set of subjects, 10 of them were removed due to the not properly collection of data. A total of 88 subjects were processed properly (43 women, 45 men, mean age= 35.33 and SD = 10.50) (for further details, please see Supplementary Materials). Outlier studies were performed by data source, dividing between VR, ET and GSR. Finally, 3 outliers were found for the data-source of ET, whereas any outlier was found for VR and GSR. The final dataset, without outliers, had in total 85 subjects (42 women, 43 men, mean age = 35.49 and SD = 10.64).

The 93.33% of the target variables were normal distributed whereas, only one target variable, Thrill and adventure seeking, which represents the 6.67%, was not. All the target variables present statistical differences between the high and low groups. Table 3 shows the statistical description of every subscale.

A total of 4 features were removed due to its no variation between subjects. 31 features (20.53%) were correlated above 0.95 in Pearson coefficient. These variables were removed from the dataset. Moreover, 16 variables were transformed using logarithms. The final dataset ended-up with a total of 120 features were 42 belong to the VR, 60 to ET and 18 to GSR.

Table 4 presents the best models obtained by the ML pipeline, according with the dataset used, the balance of the sample, the significant level between the target variable and the generated random distribution of target variables and four different metrics such as accuracy, kappa, TPR and TNR.

According with Table 4, neuroticism, extraversion, openness, thrill and adventure seeking, experience seeking, disinhibition, boredom susceptibility, negative urgency, lack of premeditation and positive urgency, have been well recognized since their accuracy shows statistical differences with random models. On the other hand, agreeableness, conscientiousness, sensation seeking (overall), lack of perseverance and sensation seeking have not been recognized over the chance level. The data source ALL appears 10 (66.67%) times as the data source with highest accuracy, VR data source 4 (26.67%) times and ET once (6.67%). Table 5 summarizes the selected features for each model.

Table 3. Statistical description of each target variable.

Dimension	Target Variable	Mean	[1] Std	Median	Number of Highs	Number of Lows	[2] St. Sig.
Personality	Neuroticism	20.92	7.20	21	45	43	***
	Extraversion	32.64	7.32	32.5	44	44	***
	Openness	32.28	6.65	33	45	43	***
	Agreeableness	31.35	6.12	31.5	44	44	***
	Conscientiousness	32.60	7.57	33.5	49	39	***
Sensation seeking	Thrill and adventure seeking	6.75	2.84	8	34	54	***
	Experience seeking	3.68	1.08	4	48	40	***
	Disinhibition	4.31	2.18	4	41	47	***
	Boredom susceptibility	3.89	1.87	4	51	37	***
	Sensation seeking (overall score)	18.63	5.59	19	47	41	***
Impulsivity	Negative urgency	9.35	2.51	9	43	45	***
	Lack of premeditation	5.58	1.59	5.5	44	44	***
	Lack of perseverance	6.82	230	7	48	40	***
	Sensation seeking	10.32	2.69	11	45	43	***
	Positive urgency	9.98	2.08	10	52	36	***

[1] Standard deviation. [2] Statistical significance of the feature between high and low groups: *** *p*-value < 0.001.

Table 4. Results of the models obtained with highest accuracy for every target variable.

Dimension	Subscale	Data s.	[1] St. Sig.	Accuracy	Kappa	[2] TPR	[3] TNR
Personality	Neuroticism	ALL	*	0.73 (0.14)	0.45 (0.29)	0.74 (0.14)	0.72 (0.24)
	Extraversion	ALL	**	0.75 (0.15)	0.51 (0.31)	0.81 (0.18)	0.71 (0.19)
	Openness	ET	***	0.71 (0.16)	0.40 (0.34)	0.78 (0.17)	0.63 (0.30)
	Agreeableness	ALL	-	0.72 (0.16)	0.43 (0.32)	0.82 (0.17)	0.6 (0.21)
	Conscientiousness	ALL	-	0.71 (0.10)	0.38 (0.21)	0.83 (0.12)	0.54 (0.27)
Sensation seeking	Thrill and adventure seeking	ALL	*	0.73 (0.13)	0.40 (0.27)	0.94 (0.13)	0.43 (0.19)
	Experience seeking	ALL	*	0.73 (0.19)	0.46 (0.39)	0.78 (0.17)	0.68 (0.28)
	Disinhibition	ALL	*	0.72 (0.15)	0.43 (0.31)	0.58 (0.32)	0.86 (0.13)
	Boredom susceptibility	VR	***	0.73 (0.08)	0.31 (0.24)	0.30 (0.25)	0.98 (0.05)
	Sensation seeking (overall score)	ALL	-	0.67 (0.14)	0.35 (0.28)	0.69 (0.17)	0.67 (0.26)
Impulsivity	Negative urgency	VR	***	0.78 (0.14)	0.55 (0.28)	0.70 (0.19)	0.86 (0.16)
	Lack of premeditation	ALL	**	0.75 (0.10)	0.50 (0.20)	0.72 (0.25)	0.79 (0.21)
	Lack of perseverance	VR	-	0.67 (0.17)	0.33 (0.32)	0.68 (0.29)	0.65 (0.23)
	Sensation seeking	ALL	-	0.68 (0.21)	0.36 (0.42)	0.68 (0.26)	0.67 (0.25)
	Positive urgency	VR	**	0.71 (0.16)	0.34 (0.37)	0.92 (0.10)	0.41 (0.31)

[1] * *p* < 0.05, ** *p* < 0.01 and *** *p* < 0.001; [2] true positive rate; [3] true negative rate.

Regarding the presence questionnaire, the results for the ITC-SOPI were (mean, SD): spatial presence 3.79, 0.53; engagement 3.99, 0.5; ecological validity 3.26, 0.75; and negative effects 2.36, 0.87.

Table 5. Selected features for each classification model.

Dim.	Subscale (n Features)	Risk Zone			No Risk Zone		
		VR	ET	GSR	VR	ET	GSR
Pers.	Neuroticism (6)	Time spent	Visits to keys Visits to green spheres	-	Total interactions	Fixation duration Visits to purple spheres	-
	Extraversion (4)	Green spheres caught	Distance in saccades	-	-	Fixation duration Velocity of saccades	-
	Openness (5)	-	Number of objects seen Angular saccade distance	-	-	Distance in saccades	-
Sens. Seek.	Thrill and adventure seeking (5)	Time spent Green spheres caught Visits to each risk Distance covered	-	-	Pause button use	-	-
	Experience seeking (9)	Acceleration	Fixation duration Visits to keys	-	Time spent Purple spheres caught Total interactions	Fixation duration	Number of phasic peaks Phasic value
	Disinhibition (11)	Velocity Distance covered	Distance in saccades	Skewness of phasic signal	Purple spheres caught Velocity	Number of fixations Velocity in saccades	Number of phasic peaks Phasic value Skewness of phasic signal
	Boredom susceptibility (11)	Shield use Total interactions	Fixation duration Distance in saccades	-	-	Angular saccade distance Velocity and distance in saccades	Kurtosis of phasic signal Phasic value
Imp.	Negative urgency (4)	Time spent Pause button use	-	-	Total interactions	-	-
	Lack of prem. (5)	Velocity	Velocity of saccades	-	Purple spheres caught Time spent	Visits to green spheres	Phasic value
	Positive urgency (10)	Time spent Visits to each risk Shield use	-	-	Distance covered Pause button use Shield use	-	-

5. Discussion

In this article, we present the assessment on decision making in risk environments (AEMIN) tool, as a new interactive virtual environment for RT measurement. The main aim of this study is to discriminate participants with high versus low scores in the measures of personality, sensation seeking and impulsivity, through their behaviors and physiological responses during playing AEMIN. Applying ML methods to the dataset we explored: (a) if through these data it is possible to discriminate between RT domains, allowing to qualitatively determinate a general level of RT for each subject; and (b) which parameters better discriminate between the two populations in each variable.

The results are discussed by sections: (1) accuracy of the models to discriminate RT domains; (2) the influence of the features used in each model selected; (3) limitations and further studies; (4) conclusion.

5.1. Accuracy of the Models to Discriminate RT Domains

5.1.1. Personality Recognition

Regarding the final models on personality recognition, the dimensions of neuroticism, extraversion and openness to experience have been properly recognized. The validation set using 88 subjects achieved 72.6% accuracy (kappa: 0.447), 75.4% accuracy (kappa: 0.506) and 70.8% accuracy (kappa 0.402) respectively. The selected models for predicting agreeableness and conscientiousness have not overcome the chance level.

Interestingly, these results show that neuroticism, extraversion and openness to experience are the better predicted personality dimensions. On the first hand, neuroticism has been related to negative affect and sensitivity to punishment [15], but its relationship with RT seems to be more complex and context-related. Therefore, although high levels of neuroticism may lead to risk aversion in most domains, as a way of avoiding guilt or anxiety about negative outcomes, the relation between neuroticism and RT seems to be inverse in the health domain [18], in which some studies identified a tendency to take risks to alleviate anxiety and other emotions in subjects with high neuroticism [89]. On the other hand, high extraversion and openness to experience have been related to risk approach across domains, due to a generalized need for stimulation and cognitive risk seeking, acceptance of experimentation, tolerance of uncertainty, change and innovation [14,20]. In the light of these findings, we could conclude that suitably our tool brings out the personality dimensions most context-dependent and related to the approach to risk, and not so much those related to general risk avoidance.

5.1.2. Sensation Seeking Recognition

Regarding the final models on sensation seeking recognition, the dimensions of experience seeking, thrill and adventure seeking, boredom susceptibility and disinhibition were predicted with robust models. The validation set achieved 73.3% accuracy (kappa: 0.456), 72.6% accuracy (kappa: 0.311), 72.1% accuracy (kappa: 0.425) and 73.1% accuracy (kappa: 0.402), respectively. The selected model for predicting overall sensation seeking score seemed to be overfitted.

These results demonstrate that AEMIN is a suitable tool to measure sensation seeking. As mentioned in previous sections, there is a great consensus in the literature regarding the influence of each of the sensation seeking subdimensions on RT [21–26], so we consider that AEMIN meets the expectations in this regard.

5.1.3. Impulsivity Recognition

Regarding the final models on impulsivity recognition, the dimensions of negative urgency, lack of premeditation and positive urgency were predicted with robust models. The validation set achieved 77.5% accuracy (kappa: 0.553), 75.1% accuracy (kappa: 0.5) and 70.8% accuracy (kappa: 0.341), respectively. The selected models for predicting lack of perseverance and sensation seeking seemed to be overfitted.

In this case, three of the five subdimensions of impulsivity were well predicted. Interestingly, negative and positive urgency, which are related to context-related behaviors—when facing negative/positive situations [32]—are included. This result suggests that with AEMIN we could identify RT behaviors in widely varying situations, encompassing negative and positive contexts.

5.2. Influence of the Features Used in Each Model Selected

5.2.1. Influence of VR Features

Regarding VR variables, the results show that navigation variables, which are related to the movements of the subject in the virtual environment, seem to be more meaningful in risky zones; while interaction variables, which are related to the interactions of the subjects with the different elements of the virtual environment (buttons and virtual elements), seem to be more relevant in no risk zones. Results of the presence questionnaires are similar, or even better, to those obtained in other works [90,91].

Starting with the navigation variables in risk zone, the results show that the time spent in risk zone has a strong influence on the prediction of variables of personality, impulsivity and sensation seeking. Longer time spent in a risk zone could mean either that the subject has passed through these areas slower, or that (s)he has passed through them a greater number of times. In any case, this may be related to the higher/lower susceptibility to punishment or to negative consequences. The neuroticism and negative urgency variables, in which time spent in risk zone appears as important predictor, are related to the sensitivity to punishment or to negative stimuli [15,32], so this would explain the relationship with the time spent in risk zone in our virtual environment. On the other hand, it also appears as important for the classification of the subjects in thrill and adventure seeking and positive urgency variables, together with the number of visits to each risk. In these cases, it is possible that subjects with greater interest in risky physical activities or with impulsive behaviors when facing situations perceived as positive decide to experiment and spend more time in these risk areas, to see what the consequences are.

The variable of distance covered in risk zone refers to the length of the subject´s trajectory in the risk zones of the maze. This variable appears as important in the classification of the subjects in the variables of thrill and adventure seeking and disinhibition, both belonging to the dimension of sensation seeking. This variable was also measured in the previous version of AEMIN [58], and significant correlations were obtained with almost all sensation seeking subdimensions, so our results in both articles seem to be consistent. Covering a greater distance in AEMIN could be interpreted as a greater interest in exploring different areas of the maze, which could be related to the variables thrill and adventure seeking and disinhibition, since both of them are reflected in high engagement in activities that generate new sensations and in rule-breaking behaviors [21].

Velocity and acceleration in risk zones appear as important variables in predicting lack of premeditation, experience seeking, and disinhibition. The action of quickly passing through the risk areas, without stopping to pick up spheres could have a double interpretation. On the one hand, it can be understood as an unpremeditated or risky action. Instead, it could also be interpreted as an intention to pass something bad as quickly as possible, avoiding the possible damage that could be caused by passing through a risk area.

As for the interaction variables in risk zones, the number of green spheres collected in risk zones helps to classify subjects into extraversion and thrill and adventure seeking subdimensions. Picking up spheres that are in risk areas can pose a risk, since the subject must pass through these areas without the protection of the shield to pick them up. Therefore, the decision to take a sphere that is in a risk zone may be related to excitement seeking—which is characterized by an interest for shinny colors and noisy environments [92] —a common feature of the extraversion and thrill and adventure seeking dimensions [93].

The use of the shield only influences in risk zones, to classify subjects in terms of boredom susceptibility. This result seems surprising, since it was expected that the use of the shield would be a somewhat more revealing variable in terms of the behaviors

of the subjects, which would add richness to the predictive models of a greater number of variables. In the SSMT [58], the use of the shield was related to subdimensions of impulsivity and sensation seeking, for which similar results were expected in the predictive models of the present article. One possible reason could be that participants did not fully understand the mechanics of the shield and did not use it enough to reflect certain behavioral patterns. This will be considered as one of the limitations of this research, and we will work to improve the understanding of the shield element in enhanced versions of AEMIN.

The use of the pause button in risk situations appears as an important variable in the prediction of negative urgency. The use of the pause button in risky situations may reflect a strategy of psychological distancing from negative stimuli, while the non-use of this resource may be due to thoughtless reactions to risky situations. This could have a strong relationship with the negative urgency variable, defined as the tendency to show impulsive behaviors in negative situations [33].

Regarding the interaction variables in no risk zone, the number of purple spheres collected appears as significant to classify the subjects in high/low lack of premeditation, experience seeking and disinhibition. These purple spheres were included in the virtual environment as elements that generate uncertainty, so collecting these spheres is clearly a risky behavior, which can be taken due to a lack of premeditation, or due to the voluntary search for new experiences or sensations.

The use of the pause button in no risk zone is meaningful for the prediction of thrill and adventure seeking. The use of this button in non-risk areas may be related to wanting to rest from the experience in general or to being curious to try it, and not so much to applying a specific coping technique in a specific moment of stress as occurs in the risk areas.

Total interactions with elements in no risk zones appears as an important variable for predicting neuroticism, negative urgency and experience seeking. A greater or lesser number of interactions with the elements of the virtual environment can be related to very different behaviors or decisions, since in the virtual environment there are very different elements, from the shield, to the spheres or the pause button. What the total interactions variable can be an indicator of, is perhaps a greater or lesser involvement of the subject within the virtual environment, as well as a better understanding of the mechanics of the game. On the other hand, it can also be related to anxious or impulsive behaviors, as well as the search for different experiences and the desire to explore the virtual environment.

5.2.2. Influence of ET Features

Our results show that ET variables have a strong influence in most of the classification models, so we could say that it is an important measure in combination with those variables of the VR dataset, both in risk zones and in no risk zones. The variables that provide the most relevant information to classify the subjects in terms of risk-related dimensions are: fixation duration, number of fixations in no risk zones, visits to keys, green spheres and purple spheres, angular saccade distance, velocity in saccades and distance in saccades.

As mentioned in previous sections, fixations duration could be an indicator of depth of processing [94], and a good predictor of perception [95] and risk aversion [65]. The fixations duration in risk zones appeared as a meaningful variable to classify subjects in boredom susceptibility. This result suggests that participants with high boredom susceptibility show different information processing patterns in risk zones than those with low boredom susceptibility, since these areas arouse a different interest in them, taking them out of the routine of the game. Conversely, fixations duration in no risk zones was an important variable in the classification models of neuroticism, extraversion and experience seeking. In these cases, a deeper processing of information in no risk areas can be interpreted as a state of alert, waiting for something bad to happen, or as a search for new or different elements in areas that apparently are simpler and show a smaller number of stimuli than risk zones.

Regarding the number of fixations and visits to concrete objects, these variables are an indicator of interest in concrete elements [94] and have been related to risk aversion, as a strategy to collect information in the analytical pre-decision process [65–67]. Spheres and keys that open the doors are the most important elements when studying the number of fixations in AEMIN. The number of fixations in green spheres located in risk zones appears as an important variable in the classification of subjects regarding neuroticism, while the number of fixations in green spheres in no risk zones is related to lack of premeditation. On the other hand, the number of fixations on keys is a fundamental variable in the classification of the subjects in neuroticism and experience seeking. Finally, the number of fixations in purple spheres is important to classify as high or low neuroticism. In light of these results, we could understand that the elements that can be captured or collected during the game (green spheres, purple spheres and keys) are the most relevant when analyzing the number of fixations. On the other hand, other elements of the game that seem more visually striking and of greater interest within the game, such as risks, do not appear as meaningful variables from the point of view of the number of fixations. Interestingly, these results could help game designers to guide the user's attention to specific elements of the virtual environment, incorporating the interaction of "collecting", as a guarantee that visual patterns related to personality, impulsivity and sensation seeking will come to light.

Angular saccade distance and distance in saccades could discriminate global and focal visual search strategies [94,96]. These variables appear as relevant for classification in risk zones, for predicting extraversion, openness, boredom susceptibility and disinhibition. These variables are also meaningful for the classification in boredom susceptibility when the subject is in a no risk zone. These results could indicate that subjects´ visual search patterns in risk zones can help to classify them in high or low extraversion, openness and disinhibition. On the other hand, to differentiate subjects with high or low boredom susceptibility, it is necessary to study their visual search strategy throughout the experience, both in risk areas and in no risk areas.

Finally, the velocity in saccades is an indicator, together with the number and duration of fixations, of an adaptive attention process, depending on the uncertainty or the perceived difficulty of each situation, so that slower saccades have been related to information acquisition processes in situations perceived as uncertain or difficult [68,97,98]. In our study, the velocity in saccades appears as an important variable in risk areas for classifying subjects in lack of premeditation, while it is meaningful in no risk areas for classifying subjects in extraversion, openness, boredom susceptibility and disinhibition. This result could be interpreted as follows: the velocity in saccades in no risk areas, as a behavior dependent on the perception of difficulty or uncertainty of a situation, is an indicator of the subjects´ interpretation of the no risk zones, based on their level of extraversion, openness, boredom susceptibility, and disinhibition. Thus, it is possible that some participants were in a high alert state while passing through these areas, since they identified them as of low certainty, while other subjects crossed these areas with the feeling of being in a safe place.

5.2.3. Influence of GSR Features

The variables obtained from GSR while the subject was in the no risk areas were relevant in the final models, while they were not relevant when the subject was in risk areas. All the GSR variables selected in the final models correspond to metrics of the phasic component of the signal, which is characterized by rapid and event-related changes, so it takes less time to show changes [99]. The GSR signal usually peaks between 2 and 10 s after stimulation and recovers at approximately the same rate [99]. Ayata et al. [100] found that a 3-s time window in the phasic signal is the most optimal for the prediction of valence and arousal. Since the periods of time in which subjects usually remain in risk areas are short (between 1 and 8 s) except in rooms, where they can spend more time, it is possible that changes in the phasic signal, which can be interpreted as a "warn" in risky situations [1], are reflected few seconds after the subject has left the risk areas.

Another possible interpretation of these results is that changes in the phasic component are meaningful to differentiate subjects with high/low extraversion, boredom susceptibility and disinhibition based on their level of activation in the no risk zones. In this regard, decisions in no risk zones present less time pressure, and they are of the type: selection of paths or decision of whether to take spheres or not. Since time pressure has been raised as one of the influential factors in the relationship between GSR and RT [75], it is possible that, in these decisions in which there are no situational biases, decision-making is more guided by individuals´ personality and temperamental factors than in risk zones.

5.3. Limitations and Further Studies

We acknowledge that this study presents some methodological limitations. First, the sample size was not large. Second, we built the high/low target variables basing on the mean or median results of the responses from this study, so it may not be extrapolated to the rest of the population. Third, it could be possible that participants did not fully understand the mechanics of the shield and did not use it enough to reflect certain behavioral patterns. For future investigations, we will recruit a larger sample of participants, we will look for validated reference scales to label the subjects and we will work to improve the understanding of the shield element in enhanced versions of AEMIN.

5.4. Conclusions

Concerning the features that better predict each dimension, we could conclude that behavioral measures—interaction with the virtual environment and ET—provide the core information in the classification models. Therefore, the results support the use of AEMIN as an ecological assessment tool to measure RT, since it brings to light behaviors that allow to classify the subjects into high/low risk-related psychological constructs. Regarding physiological measures, GSR seems to be less salient in prediction models.

References

1. Bechara, A.; Damasio, H.; Tranel, D.; Damasio, A. The Iowa Gambling Task and the somatic marker hypothesis: Some questions and answers. *Trends Cogn. Sci.* **2005**, *9*, 159–162. [CrossRef] [PubMed]
2. Krain, A.L.; Wilson, A.M.; Arbuckle, R.; Castellanos, F.X.; Milham, M.P. Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making. *NeuroImage* **2006**, *32*, 477–484. [CrossRef] [PubMed]
3. Einhorn, H.J. The use of nonlinear, noncompensatory models in decision making. *Psychol. Bull.* **1970**, *73*, 221–230. [CrossRef]
4. Hunt, R.G.; Krzystofiak, F.J.; Meindl, J.R.; Yousry, A.M. Cognitive style and decision making. *Organ. Behav. Hum. Decis. Process.* **1989**, *44*, 436–453. [CrossRef]
5. Appelt, K.C.; Milch, K.F.; Handgraaf, M.J.J.; Weber, E.U. The decision making individual differences inventory and guidelines for the study of individual differences in judgment and decision-making research. *Judgm. Decis. Mak.* **2011**, *6*, 252–262.
6. Dror, I.E.; Basola, B.; Busemeyer, J.R. Decision making under time pressure: An independent test of sequential sampling models. *Mem. Cogn.* **1999**, *27*, 713–725. [CrossRef]
7. Figner, B.; Weber, E.U. Who Takes Risks When and Why? *Curr. Dir. Psychol. Sci.* **2011**, *20*, 211–216. [CrossRef]
8. Zuckerman, M.; Kuhlman, D.M. Personality and Risk-Taking: Common Bisocial Factors. *J. Pers.* **2000**, *68*, 999–1029. [CrossRef]
9. Fessler, D.M.T. Emotions and cost–benefit assessment: The role of shame and self-esteem in risk taking. In *Bounded Rationality: The Adaptive Toolbox*; Gigerenzer, G., Selten, R., Eds.; The MIT Press: Cambridge, MA, USA, 2001; pp. 191–214.
10. Isen, A.M.; Simmonds, S.F. The Effect of Feeling Good on a Helping Task that is Incompatible with Good Mood. *Soc. Psychol.* **1978**, *41*, 346–349. [CrossRef]
11. Arkes, H.R.; Herren, L.T.; Isen, A.M. The role of potential loss in the influence of affect on risk-taking behavior. *Organ. Behav. Hum. Decis. Process.* **1988**, *42*, 181–193. [CrossRef]
12. Nguyen, Y.; Noussair, C.N. Risk Aversion and Emotions. *Pac. Econ. Rev.* **2014**, *19*, 296–312. [CrossRef]
13. Raffaelli, M.; Crockett, L.J. Sexual Risk Taking in Adolescence: The Role of Self-Regulation and Attraction to Risk. *Dev. Psychol.* **2003**, *39*, 1036–1046. [CrossRef] [PubMed]
14. McCrae, R.R.; Costa, P.T., Jr. Conceptions and correlates of openness to experience. In *Handbook of Personality Psychology*; Academic Press: Cambridge, MA, USA, 1997; pp. 825–847. [CrossRef]
15. Elliot, A.J.; Thrash, T.M. Approach and Avoidance Temperament as Basic Dimensions of Personality. *J. Pers.* **2010**, *78*, 865–906. [CrossRef] [PubMed]
16. Gullone, E.; Moore, S. Adolescent risk-taking and the five-factor model of personality. *J. Adolesc.* **2000**, *23*, 393–407. [CrossRef] [PubMed]
17. Hoyle, R.H.; Fejfar, M.C.; Miller, J.D. Personality and Sexual Risk Taking: A Quantitative Review. *J. Pers.* **2000**, *68*, 1203–1231. [CrossRef] [PubMed]
18. Nicholson, N.; Soane, E.; Fenton-O'Creevy, M.; Willman, P. Personality and domain-specific risk taking. *J. Risk Res.* **2005**, *8*, 157–176. [CrossRef]
19. Schwebel, D.C.; Severson, J.; Ball, K.K.; Rizzo, M. Individual difference factors in risky driving: The roles of anger/hostility, conscientiousness, and sensation-seeking. *Accid. Anal. Prev.* **2006**, *38*, 801–810. [CrossRef]
20. Lauriola, M.; Levin, I.P. Personality traits and risky decision-making in a controlled experimental task: An exploratory study. *Pers. Individ. Differ.* **2001**, *31*, 215–226. [CrossRef]
21. Zuckerman, M. *Behavioral Expressions and Biosocial Bases of Sensation Seeking*; Cambridge University Press: Cambridge, UK, 1994.
22. Donohew, L.; Zimmerman, R.; Cupp, P.S.; Novak, S.; Colon, S.; Abell, R. Sensation seeking, impulsive decision-making, and risky sex: Implications for risk-taking and design of interventions. *Pers. Individ. Differ.* **2000**, *28*, 1079–1091. [CrossRef]
23. Netter, P.; Hennig, J.; Roed, I. Serotonin and Dopamine as Mediators of Sensation Seeking Behavior. *Neuropsychobiology* **1996**, *34*, 155–165. [CrossRef]
24. Roberti, J.W. A review of behavioral and biological correlates of sensation seeking. *J. Res. Pers.* **2004**, *38*, 256–279. [CrossRef]
25. De Vries, R.E.; de Vries, A.; Feij, J.A. Sensation seeking, risk-taking, and the HEXACO model of personality. *Pers. Individ. Differ.* **2009**, *47*, 536–540. [CrossRef]
26. Zuckerman, M. Biosocial bases of sensation seeking. In *Biology of Personality and Individual Differences*; The Guilford Press: New York, NY, USA, 2006; pp. 37–59.
27. Moeller, F.G.; Barratt, E.S.; Dougherty, D.M.; Schmitz, J.M.; Swann, A.C. Psychiatric Aspects of Impulsivity. *Am. J. Psychiatry* **2001**, *158*, 1783–1793. [CrossRef] [PubMed]
28. De Wit, H. Impulsivity as a determinant and consequence of drug use: A review of underlying processes. *Addict. Biol.* **2008**, *14*, 22–31. [CrossRef] [PubMed]
29. Stanford, M.S.; Greve, K.W.; Boudreaux, J.K.; Mathias, C.W.; Brumbelow, J.L. Impulsiveness and risk-taking behavior: Comparison of high-school and college students using the Barratt Impulsiveness Scale. *Pers. Individ. Differ.* **1996**, *21*, 1073–1075. [CrossRef]
30. Cooper, M.L.; Agocha, V.B.; Sheldon, M.S. A Motivational Perspective on Risky Behaviors: The Role of Personality and Affect Regulatory Processes. *J. Pers.* **2000**, *68*, 1059–1088. [CrossRef]
31. Boyer, T.W. The development of risk-taking: A multi-perspective review. *Dev. Rev.* **2006**, *26*, 291–345. [CrossRef]
32. Whiteside, S.P.; Lynam, D.R. The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Pers. Individ. Differ.* **2001**, *30*, 669–689. [CrossRef]

33. Whiteside, S.P.; Lynam, D.R.; Miller, J.D.; Reynolds, S.K. Validation of the UPPS impulsive behaviour scale: A four-factor model of impulsivity. *Eur. J. Pers.* **2005**, *19*, 559–574. [CrossRef]

34. Lejuez, C.W.; Read, J.P.; Kahler, C.W.; Richards, J.B.; Ramsey, S.E.; Stuart, G.L.; Strong, D.R.; Brown, R.A. Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *J. Exp. Psychol. Appl.* **2002**, *8*, 75–84. [CrossRef]

35. Skeel, R.L.; Neudecker, J.; Pilarski, C.; Pytlak, K. The utility of personality variables and behaviorally-based measures in the prediction of risk-taking behavior. *Pers. Individ. Differ.* **2007**, *43*, 203–214. [CrossRef]

36. Horvath, P.; Zuckerman, M. Sensation seeking, risk appraisal, and risky behavior. *Pers. Individ. Differ.* **1993**, *14*, 41–52. [CrossRef]

37. Lejuez, C.W.; Aklin, W.M.; Zvolensky, M.J.; Pedulla, C.M. Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. *J. Adolesc.* **2003**, *26*, 475–479. [CrossRef]

38. Blais, A.R.; Weber, E.U. A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgm. Decis. Mak.* **2006**, *1*, 37–47.

39. Barsade, S.G.; Ramarajan, L.; Westen, D. Implicit affect in organizations. *Res. Organ. Behav.* **2009**, *29*, 135–162. [CrossRef]

40. Becker, W.J.; Cropanzano, R.; Sanfey, A.G. Organizational Neuroscience: Taking Organizational Theory Inside the Neural Black Box. *J. Manag.* **2011**, *37*, 933–961. [CrossRef]

41. George, J.M. The Illusion of Will in Organizational Behavior Research: Nonconscious Processes and Job Design. *J. Manag.* **2009**, *35*, 1318–1339. [CrossRef]

42. Kivikangas, J.M.; Chanel, G.; Cowley, B.; Ekman, I.; Salminen, M.; Järvelä, S.; Ravaja, N. A review of the use of psychophysiological methods in game research. *J. Gaming Virtual Worlds* **2011**, *3*, 181–199. [CrossRef]

43. Shute, V.J. Stealth assessment in computer-based games to support learning. *Comput. Games Instr.* **2011**, *55*, 503–524.

44. Shute, V.J.; Wang, L.; Greiff, S.; Zhao, W.; Moore, G.R. Measuring problem solving skills via stealth assessment in an engaging video game. *Comput. Hum. Behav.* **2016**, *63*, 106–117. [CrossRef]

45. Mayer, I.; Bekebrede, G.; Warmelink, H.; Zhou, Q. A brief methodology for researching and evaluating serious games and game-based learning. In *Psychology, Pedagogy, and Assessment in Serious Games*; IGI Global: Hershey, PA, USA, 2014; pp. 357–393. [CrossRef]

46. Bechara, A.; Damasio, A.R.; Damasio, H.; Anderson, S.W. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* **1994**, *50*, 7–15. [CrossRef]

47. Buelow, M.T.; Suhr, J.A. Construct Validity of the Iowa Gambling Task. *Neuropsychol. Rev.* **2009**, *19*, 102–114. [CrossRef] [PubMed]

48. White, T.L.; Lejuez, C.W.; De Wit, H. Test-retest characteristics of the Balloon Analogue Risk Task (BART). *Exp. Clin. Psychopharmacol.* **2008**, *16*, 565–570. [CrossRef] [PubMed]

49. Manchester, D.; Priestley, N.; Jackson, H. The assessment of executive functions: Coming out of the office. *Brain Inj.* **2004**, *18*, 1067–1081. [CrossRef] [PubMed]

50. Sbordone, R.J. Ecological validity of neuropsychological testing: Critical issues. In *Neuropsychology Handbook*; Springer: New York, NY, USA, 2008; pp. 367–394.

51. Bottari, C.; Dassa, C.; Rainville, C.; Dutil, E. The factorial validity and internal consistency of the Instrumental Activities of Daily Living Profile in individuals with a traumatic brain injury. *Neuropsychol. Rehabil.* **2009**, *19*, 177–207. [CrossRef] [PubMed]

52. Alcañiz, M.; Lozano, J.A.; Rey, B. Technological background of VR. *Stud. Health Technol. Inform.* **2003**, *99*, 199–214.

53. Tarr, M.J.; Warren, W.H. Virtual reality in behavioral neuroscience and beyond. *Nat. Neurosci.* **2002**, *5*, 1089–1092. [CrossRef]

54. Alcañiz, M.; Rey, B.; Tembl, J.; Parkhutik, V. A Neuroscience Approach to Virtual Reality Experience Using Transcranial Doppler Monitoring. *Presence Teleoperators Virtual Environ.* **2009**, *18*, 97–111. [CrossRef]

55. Parsons, T. Virtual Reality for Enhanced Ecological Validity and Experimental Control in the Clinical, Affective and Social Neurosciences. *Front. Hum. Neurosci.* **2015**, *9*, 660. [CrossRef]

56. Giglioli, I.A.C.; de-Juan-Ripoll, C.; Parra, E.; Raya, M.A. Are 3D virtual environments better than 2D interfaces in serious games performance? An explorative study for the assessment of executive functions. *Appl. Neuropsychol. Adult* **2019**, 1–10. [CrossRef]

57. De-Juan-Ripoll, C.; Soler-Domínguez, J.L.; Guixeres, J.; Contero, M.; Gutiérrez Álvarez, N.; Alcañiz, M. Virtual Reality as a New Approach for Risk Taking Assessment. *Front. Psychol.* **2018**, *9*, 2532. [CrossRef] [PubMed]

58. De-Juan-Ripoll, C.; Soler-Domínguez, J.L.; Giglioli, I.A.C.; Contero, M.; Alcañiz, M. The Spheres & Shield Maze Task: A Virtual Reality Serious Game for the Assessment of Risk Taking in Decision Making. *Cyberpsychol. Behav. Soc. Netw.* **2020**, *23*, 773–781. [CrossRef] [PubMed]

59. Glöckner, A.; Herbold, A.-K. An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *J. Behav. Decis. Mak.* **2010**, *24*, 71–98. [CrossRef]

60. Knoblich, G.; Ohlsson, S.; Raney, G.E. An eye movement study of insight problem solving. *Mem. Cogn.* **2001**, *29*, 1000–1009. [CrossRef] [PubMed]

61. Yang, J.W.; Kim, D.; Jung, S. Using Eye-Tracking Technology to Measure Environmental Factors Affecting Street Robbery Decision-Making in Virtual Environments. *Sustainability* **2020**, *12*, 7419. [CrossRef]

62. Pettersson, J.; Albo, A.; Eriksson, J.; Larsson, P.; Falkman, K.W.; Falkman, P. Cognitive Ability Evaluation Using Virtual Reality and Eye Tracking. In Proceedings of the 2018 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications, Ottawa, ON, Canada, 12–13 June 2018; pp. 1–6. [CrossRef]

63. Porras-Garcia, B.; Ferrer-Garcia, M.; Ghita, A.; Moreno, M.; López-Jiménez, L.; Vallvé-Romeu, A.; Serrano-Troncoso, E.; Gutiérrez-Maldonado, J. The influence of gender and body dissatisfaction on body-related attentional bias: An eye-tracking and virtual reality study. *Int. J. Eat. Disord.* **2019**, *52*, 1181–1190. [CrossRef] [PubMed]

64. Rojas, J.-C.; Marín-Morales, J.; Azofra, J.M.A.; Contero, M. Recognizing Decision-Making Using Eye Movement: A Case Study with Children. *Front. Psychol.* **2020**, *11*, 2542. [CrossRef]

65. Kwak, Y.; Payne, J.W.; Cohen, A.L.; Huettel, S.A. The rational adolescent: Strategic information processing during decision making revealed by eye tracking. *Cogn. Dev.* **2015**, *36*, 20–30. [CrossRef]

66. Su, Y.; Rao, L.-L.; Sun, H.-Y.; Du, X.-L.; Li, X.; Li, S. Is making a risky choice based on a weighting and adding process? An eye-tracking investigation. *J. Exp. Psychol. Learn. Mem. Cogn.* **2013**, *39*, 1765–1780. [CrossRef]

67. Payne, J.W.; Braunstein, M.L. Risky choice: An examination of information acquisition behavior. *Mem. Cogn.* **1978**, *6*, 554–561. [CrossRef]

68. Velichkovsky, B.M.; Rothert, A.; Kopf, M.; Dornhöfer, S.M.; Joos, M. Towards an express-diagnostics for level of processing and hazard perception. *Transp. Res. Part F Traffic Psychol. Behav.* **2002**, *5*, 145–156. [CrossRef]

69. Habibnezhad, M.; Fardhosseini, S.; Vahed, A.M.; Esmaeili, B.; Dodd, M.D. The relationship between construction workers' risk perception and eye movement in hazard identification. In *Construction Research Congress 2016*; American Society of Civil Engineers: Reston, VA, USA, 2016; pp. 2984–2994. [CrossRef]

70. Nourbakhsh, N.; Wang, Y.; Chen, F. GSR and blink features for cognitive load classification. In *Lecture Notes in Computer Science (LNCS)*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8117, pp. 159–166. [CrossRef]

71. Ergan, S.; Radwan, A.; Zou, Z.; Tseng, H.-A.; Han, X. Quantifying Human Experience in Architectural Spaces with Integrated Virtual Reality and Body Sensor Networks. *J. Comput. Civ. Eng.* **2019**, *33*, 04018062. [CrossRef]

72. Šalkevicius, J.; Damaševičius, R.; Maskeliūnas, R.; Laukienė, I. Anxiety Level Recognition for Virtual Reality Therapy System Using Physiological Signals. *Electronics* **2019**, *8*, 1039. [CrossRef]

73. Alcañiz Raya, M.; Giglioli, I.A.C.; Marín-Morales, J.; Higuera-Trujillo, J.L.; Olmos, E.; Minissi, M.E.; Garcia, G.T.; Sirera, M.; Abad, L. Application of Supervised Machine Learning for Behavioral Biomarkers of Autism Spectrum Disorder Based on Electrodermal Activity and Virtual Reality. *Front. Hum. Neurosci.* **2020**, *14*, 90. [CrossRef] [PubMed]

74. Yip, J.A.; Stein, D.H.; Côté, S.; Carney, D.R. Follow your gut? Emotional intelligence moderates the association between physiologically measured somatic markers and risk-taking. *Emotion* **2020**, *20*, 462–472. [CrossRef]

75. Persson, E.; Asutay, E.; Hagman, W.; Västfjäll, D.; Tinghög, G. Affective response predicts risky choice for fast, but not slow, decisions. *J. Neurosci. Psychol. Econ.* **2018**, *11*, 213–227. [CrossRef]

76. Cordero, A.; Pamos, A.; Seisdedos, N. *Inventario de Personalidad Neo Revisado (NEO PI-R). Inventario Neo Reducido de Cinco Factores (NEO-FFI)*; TEA: Madrid, Spain, 1999. (In Spanish)

77. Costa, P.T.; McCrae, R.R. NEO PI/FFI manual supplement for use with the NEO Personality Inventory. In *Psychological Assessment Resources*; PAR Inc.: Odessa, FL, USA, 1989.

78. Pérez, J.; Torrubia, R. Fiabilidad y validez de la versión española de la Escala de Búsqueda de Sensaciones (Forma V). *Rev. Latinoam. Psicol.* **1986**, *18*, 7–22. [CrossRef]

79. Zuckerman, M.; Kolin, E.A.; Price, L.; Zoob, I. Development of a sensation-seeking scale. *J. Consult. Psychol.* **2006**, *28*, 477–482. [CrossRef]

80. Candido, A.; Orduña, E.; Perales, J.C.; Verdejo-García, A.; Billieux, J. Validation of a short Spanish version of the UPPS-P impulsive behaviour scale. *Trastor. Adict.* **2012**, *14*, 73–78. [CrossRef]

81. Lessiter, J.; Freeman, J.; Keogh, E.; Davidoff, J.B. A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence Teleoperators Virtual Environ.* **2001**, *10*, 282–297. [CrossRef]

82. So, R.H.Y.; Lo, W.T.; Ho, A.T.K. Effects of Navigation Speed on Motion Sickness Caused by an Immersive Virtual Environment. *Hum. Factors J. Hum. Factors Ergon. Soc.* **2001**, *43*, 452–461. [CrossRef] [PubMed]

83. Llanes-Jurado, J.; Marín-Morales, J.; Guixeres, J.; Alcañiz, M. Development and Calibration of an Eye-Tracking Fixation Identification Algorithm for Immersive Virtual Reality. *Sensors* **2020**, *20*, 4956. [CrossRef] [PubMed]

84. Shukla, J.; Barreda-Ángeles, M.; Oliver, J.; Puig, D. Efficient wavelet-based artifact removal for electrodermal activity in real-world applications. *Biomed. Signal Process. Control* **2018**, *42*, 45–52. [CrossRef]

85. Benedek, M.; Kaernbach, C. A continuous measure of phasic electrodermal activity. *J. Neurosci. Methods* **2010**, *190*, 80–91. [CrossRef] [PubMed]

86. Taylor, S.; Jaques, N.; Chen, W.; Fedor, S.; Sano, A.; Picard, R. Automatic Identification of Artifacts in Electrodermal Activity Data. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, Italy, 25–29 August 2015; pp. 1934–1937. [CrossRef]

87. Schoelkopf, B.; Smola, A.J.; Williamson, R.C.; Bartlett, P.L. New Support Vector Algorithms. *Neural Comput.* **2000**, *12*, 1207–1245. [CrossRef]

88. Doak, J. An evaluation of feature selection methods and their application to computer security. In *Techninal Report CSE-92-18*; University of California: Berkeley, CA, USA, 1992.

89. Vollrath, M.E.; Torgersen, S. Who takes health risks? A probe into eight personality types. *Pers. Individ. Differ.* **2002**, *32*, 1185–1197. [CrossRef]

90. De Leo, G.; Diggs, L.A.; Radici, E.; Mastaglio, T.W. Measuring Sense of Presence and User Characteristics to Predict Effective Training in an Online Simulated Virtual Environment. *Simul. Healthc. J. Soc. Simul. Healthc.* **2014**, *9*, 1–6. [CrossRef]

91. Piccione, J.; Collett, J.; De Foe, A. Virtual skills training: The role of presence and agency. *Heliyon* **2019**, *5*, e02583. [CrossRef]

92. Costa, P.T.; McCrae, R.R. NEO-PI-R. Professional manual. In *Psychological Assessment Resources*; PAR Inc.: Odessa, FL, USA, 1992.

93. Aluja, A.; García, Ó.; García, L.F. Relationships among extraversion, openness to experience, and sensation seeking. *Pers. Individ. Differ.* **2003**, *35*, 671–680. [CrossRef]

94. Al-Moteri, M.O.; Symmons, M.; Plummer, V.; Cooper, S. Eye tracking to investigate cue processing in medical decision-making: A scoping review. *Comput. Hum. Behav.* **2017**, *66*, 52–66. [CrossRef]

95. Krupinski, E.A.; Nodine, C.F.; Kundel, H.L. Enhancing recognition of lesions in radiographic images using perceptual feedback. *Opt. Eng.* **1998**, *37*, 813–819. [CrossRef]

96. Kundel, H.L.; Nodine, C.F.; Conant, E.; Weinstein, S.P. Holistic Component of Image Perception in Mammogram Interpretation: Gaze-tracking Study. *Radiology* **2007**, *242*, 396–402. [CrossRef] [PubMed]

97. Brunyé, T.T.; Gardony, A.L. Eye tracking measures of uncertainty during perceptual decision making. *Int. J. Psychophysiol.* **2017**, *120*, 60–68. [CrossRef] [PubMed]

98. Heekeren, H.R.; Marrett, S.; Ungerleider, L.G. The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* **2008**, *9*, 467–479. [CrossRef]

99. Sharma, M.; Kacker, S.; Sharma, M. A Brief Introduction and Review on Galvanic Skin Response. *Int. J. Med. Res. Prof.* **2016**, *2*, 13–17. [CrossRef]

100. Ayata, D.; Yaslan, Y.; Kamaşak, M. Emotion Recognition via Random Forest and Galvanic Skin Response: Comparison of Time Based Feature Sets, Window Sizes and Wavelet Approaches. In Proceedings of the 2016 Medical Technologies National Congress, Antalya, Turkey, 27–29 October 2016; pp. 1–4. [CrossRef]