A Virtual Reality set-up for the combined assessment of behavioral and psychophysiological cue-reactivity in gambling disorder

Luca R. Bruder*, Lisa Scharer, Jan Peters

Department of Psychology, Biological Psychology, University of Cologne, Germany

*** Corresponding author**

Contact: lbruder@uni-koeln.de, jan.peters@uni-koeln.de

Acknowledgments

We thank Mohsen Shaverdy and Diego Saldivar for the implementation of the VR environments and task programming and all members of the Peters Lab at the University of Cologne for helpful discussions. This work was supported by Deutsche Forschungsgemeinschaft (DFG, grant PE1627/5-1 to J.P.).

Author Contributions

Conceptualization: Jan Peters, Luca Bruder.

Data curation: Luca Bruder.

Formal analysis: Luca Bruder.

Funding acquisition: Jan Peters.

Investigation: Luca Bruder, Lisa Scharer

Methodology: Jan Peters, Luca Bruder.

Project administration: Luca Bruder.

Writing – original draft: Luca Bruder.

Writing – review & editing: Jan Peters, Luca Bruder, Lisa Scharer.

Abstract

Increased reactivity to addiction related cues (cue-reactivity) plays a critical role in the maintenance of addiction. Studies assessing cue-reactivity in gambling disorder often suffer from low ecological validity due to the usage of picture stimuli in a neutral lab environment. Here we describe a novel virtual reality (VR) set-up for the concurrent assessment of behavioral and psychophysiological cue-reactivity in gambling disorder. On two days, thirty-four healthy non-gambling participants explored two rich and navigable VR-environments (neutral: cafe vs. gambling-related: casino and sports-betting facility), while their electrodermal activity was measured using remote sensors. In addition, participants completed a temporal discounting task implemented in each VR environment. On a third day, participants performed the task in a standard lab testing context. We then used comprehensive computational modeling using both standard softmax and drift diffusion model (DDM) choice rules to assess the reliability of discounting model parameters assessed in VR. Test-retest reliability estimates were good to excellent for the discount rate log(k), whereas they were poor to moderate for additional DDM parameters. Differences in model parameters between standard lab testing and VR, reflecting reactivity to the different environments, were mostly numerically small and of inconclusive directionality. Finally, while exposure to VR generally increased tonic skin conductance, this effect was not modulated by the neutral vs. gambling-related VR-environment. Taken together this proof-of-concept study in non-gambling participants demonstrates the feasibility assessing both physiological and behavioral cue-reactivity in VR. We show that temporal discounting measures obtained in VR are reliable, suggesting that VR is a promising tool for an ecologically valid assessment of cue-reactivity in gambling disorder.

Introduction

A core characteristic of many psychiatric and neurological disorders is a detrimental change in decision-making processes. This is especially evident in addiction related disorders such as substance abuse[1–3] or pathological gambling[4–6]. A tool to study such changes in decision making is computational psychiatry[7], which employs theoretically grounded mathematical models to examine cognitive performance in relation to psychiatric disorders. Such a model-based approach allows a better quantification of the underlying latent processes[8].

One process that has been implicated in a range of psychiatric disorders is the discounting of reward value over time (temporal discounting), as both steep and shallow discounting is associated with different psychiatric conditions[1]. In temporal discounting tasks, participants make repeated choices between a fixed immediate reward and larger but temporally delayed rewards[9]. Based on binary choices and/or response time (RT) distributions, the degree to which participants discount the value of future rewards based on the temporal delay provides a measure of individual impulsivity. Increased temporal discounting is thought to be a trans-diagnostic marker with relevance for a range of psychiatric disorders[1], with addictions and related disorders being prominent examples[10,11].

There is preliminary evidence that temporal discounting might to be more pronounced when addiction related cues are present. Pathological gamblers for instance tend to exhibit steeper discounting when gambling related cues are present[4,12]. These findings resonate with theories of drug addiction such as incentive sensitization theory[13] which emphasize a prominent role for addiction-related cues in the maintenance of drug addiction (see below). Identifying the mechanisms underlying such behavioral patterns and how they are modulated by addiction-related cues is essential to the planning and execution of successful interventions that aim to reverse these changes in decision-making[14,15].

Accordingly, the concept of cue reactivity plays a prominent role in research on drug addiction[16], but has more recently also been investigated in behavioral addictions such as gambling disorder[17]. Cue reactivity refers to conditioned responses to addiction-related cues in the environment and is thought to play a major role in the maintenance of addiction. Cue reactivity can manifest in behavioral measures, as described above for temporal discounting, but also in subjective reports and/or in physiological measures[16]. Incentive-Sensitization Theory[13,18] states that neural circuits mediating the incentive motivation to obtain a reward become over sensitized to addiction related cues, mediating craving towards the addictive stimulus. These motivational changes are thought to be mediated by dopaminergic pathways of

the mesocorticolimbic system[19–21]. In line with this, craving following cue exposure correlates with a modulation of striatal value signals during temporal discounting[4], and exposure to drug-related cues increases dopamine release in striatal circuits in humans[21]. While studying these mechanisms in substance addiction is certainly of value, it is also problematic because substances might have direct effects on the underlying neural substrates. Behavioral addictions, such as pathological gambling, however, might offer a somewhat less perturbed view on the underlying mechanisms.

Studies probing cue reactivity in pathological gamblers have typically either used picture stimuli[4,5,30,22–29] or real-life gambling environments (i.e. gambling facilities)[12]. Both methods come with advantages and disadvantages. While presenting pictures in a controlled lab environment enables researchers to minimize the influence of noise factors and eases the measuring of physiological variables, it lacks the ecological validity of real-life stimuli. Vice versa, a field study in a real gambling outlet has an arguably high ecological validity but lacks the control of confounding factors and makes it difficult to obtain physiological measures.

Recent research has exploited the development of high-performance virtual reality (VR) technology to increase the ecological validity of stimuli presented in studies of cue-exposure[31], counterconditioning[32] and equilibrium training[33]. By equipping participants with head-mounted VR-glasses and sufficient space to navigate within the VR-environment, a strong sense of immersion can be created, which in turn generates more realistic stimulation. In this way VR also offers a potential solution for the problem of ecologically valid addiction-related stimuli for studies in the field of cue-reactivity. For example, Bouchard et al.[34] developed a VR-design that is built to provide ecologically valid stimuli for pathological gamblers by placing them in a virtual casino. The design can be used in treatment in order to test reactions and learned cognitive strategies in a secure environment. The present study builds upon this idea to create a design that allows assessment of behavioral, subjective and physiological cue-reactivity in VR-environments. Participants are immersed in two rich and navigable VR environments that either represent a (neutral) café environment or a gambling-related casino environment. Within these environments, behavioral cue reactivity can be measured via behavioral tasks implemented in VR. Given that immersion in the virtual environment takes place in a controlled lab setting, the measurement of physiological variables like skin conductance[35] and heart rate, as indicators of physiological cue reactivity[16,17], is also easily accommodated.

Studies using computational modeling to asses latent processes underlying learning and decision-making increasingly include not only binary decisions, but also response times (RTs) associated with these decisions, e.g. via sequential sampling models such as the drift diffusion model (DDM)[36]. This approach has several potential advantages. First, leveraging the information contained in the full RT distributions can improve the stability of parameter estimates[37,38]. Second, by conceiving decision making as a dynamic diffusion process, a more detailed picture of the underlying latent processes emerges[39–43]. Recent studies, for instance, applied these techniques to temporal discounting, where they revealed novel insights into effects of pharmacological manipulation of the dopamine system on choice dynamics[42]. Likewise, we applied these techniques to examine the processes underlying reinforcement learning impairments in gambling disorder[44] and decision-making alterations following medial orbitofrontal cortex lesions[41]. Importantly, most standard lab-based testing settings use keyboards, button boxes and computer screens to record responses and display stimuli during behavioral tasks. In contrast, in the present study we used VR-controllers in a 3D virtual space. In particular in the context of RT-based modeling, a crucial question is therefore whether responses obtained via VR-controllers allow for a comprehensive RT-based modeling, as previously done using standard approaches. Therefore, we also explored the applicability of drift diffusion modeling in the context of behavioral data obtained in VR.

Besides validating our VR-design with a healthy cohort of participants, the study at hand investigated the stability of parameters derived from behavioral tasks, in particular the temporal discount rate *log(k)*. Recently, the reliability of behavioral tasks as trait indicators of impulsivity and cognitive control has been called into question[45,46], in particular when compared to questionnaire-based measures of self-control[45]. It has been argued that the inherent property that makes behavioral tasks attractive for group-based comparisons renders them less reliable as trait markers[47]. Specifically, Hedge et al.[47] argue that tasks having a low between participant variability produce robust group effects in experimental studies and are therefore employed frequently. However, some of these tasks suffer from reduced test-retest-reliability for individual participants due to their low between-participants variability. Notably, Enkavi et al.[45] reported a reliability of .65 for the discount rate *k*, the highest of all behavioral tasks examined in that study, and comparable to the reliability estimates of the questionnaire-based measures. This is in line with previous studies on the reliability of *k*, which provided estimates ranging from .7 to .77[48,49]. Importantly, as outlined above, both the actual response mode and the contextual setting of VR-based experiments differ substantially from standard lab-based testing situations employed in previous reliability studies of e.g. temporal discounting[45,48–51].

Therefore, it is an open question whether behavioral measures obtained in VR exhibit a reliability comparable to the standard lab-based tests that are typically used in psychology.

Taken together, by examining healthy non-gambling participants on different days and under different conditions (neutral vs. gambling-related VR environment, standard lab-based testing situation), we examined the reliability of temporal discounting. We furthermore explored the feasibility of applying the drift diffusion model in the context of RTs obtained via VR-compatible controllers. Finally, we also examined physiological arousal during exploration of the virtual environments. Importantly, while our design is ultimately aimed to examine these processes in gambling disorder (e.g. the setup includes a gambling-related and a neutral VR environment), the present study constitutes an initial proof-of-concept study examining the reliability of the obtained measures and the technical feasibility in a group of healthy non-gambling control participants.

Methods

Participants. Thirty-four healthy participants (25 female), aged between 18 and 44 (mean = 26.41, std = 6.44), were invited to the lab on three different occasions. Participants were recruited via flyers at the University of Cologne and via postings in local internet forums. No participant indicated a history of traumatic brain injury, psychiatric or neurological disorders or severe motion sickness. Participants were additionally screened for gambling behavior using the questionnaire Kurzfragebogen zum Glückspielverhalten (KFG)[52]. The KFG fulfills the psychometric properties of a reliable and valid screening instrument. No participant showed a high level (>15 points on the KFG) of gambling affinity (mean = 1.56, std = 2.61, range: 0 to 13).

Participants provided informed written consent prior to their participation, and the study procedure was approved by the Ethics Board of the Germany Psychological Society.

VR-Setup. The VR-environments were presented using a wireless HTC VIVE head-mounted display (HMD). The setup provided a 110° field of view, a 90 Hz refresh rate and a resolution of 1440 x 1600 Pixel per eye. Participants had an area of about $6m^2$ open space to navigate the virtual environment. For the execution of the behavioral tasks and additional movement control participants held one VR-controller in their dominant hand. The VR-software was run on a PC with the following specifications: CPU: Intel Core (TM) i7-3600, Memory: 32.0 GB RAM, Windows 10, GPU: NVIDIA GeForce GTX 1080 (Ti). The VR-environments themselves were designed in Unity. Auditory stimuli were presented using on-ear headphones.

VR-Environments. The two VR-environments both consisted of a starting area and an experimental area. The starting area was the same for both VR-environments. It consisted of a small rural shopping street and a small park. Participants heard low street noises. The area was designed for familiarization with the VR-setup and the initial exploration phase. The experimental area of the environments differed for the two environments. For the $VR_{neutral}$ environment it contained a small café with a buffet (Figure 1 a, b and c). Participants could hear low conversations and music. The gambling-related environment ($VR_{gambling}$) contained a small casino with slot machines and a sports betting area (Figure 1 d, e and f). The audio backdrop was the sound of slot machines and sports. The floorplan of both of these experimental areas was identical but mirrored for the café (Figure 1 a and d). Both experimental areas additionally included eight animated human avatars. These avatars performed steady and non-repetitive behaviors like gambling and ordering food for the gambling-related and neutral environments,

respectively. Both experimental areas (café and casino) had entrances located at the same position within the starting area of the VR-environments, which were marked by corresponding signs.

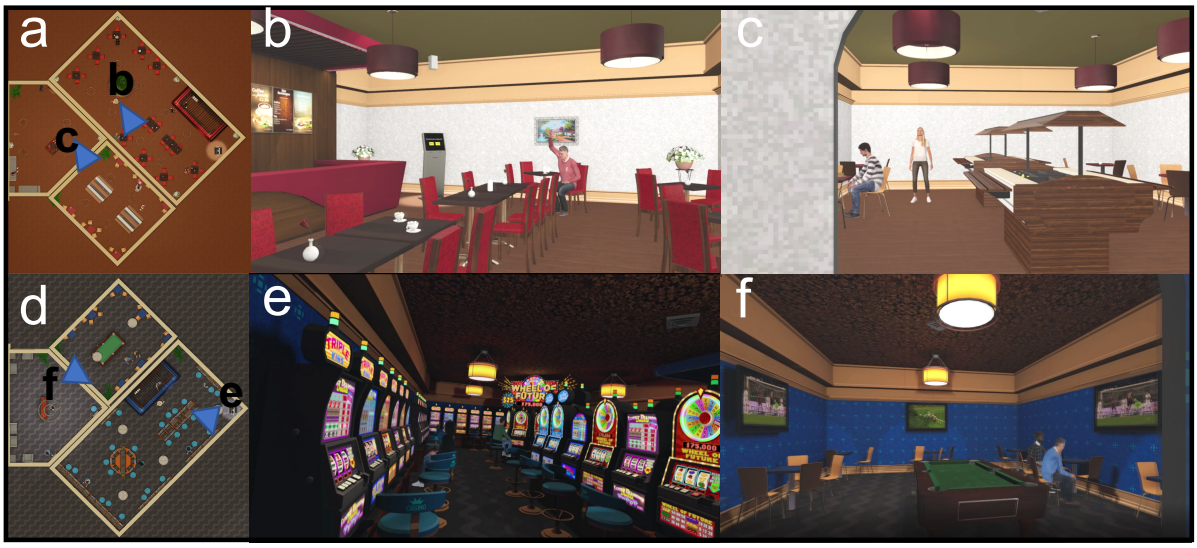


Figure 1. Experimental areas of the VR-environments: a) Floorplan of the café within the VR-neutral environment b) View of the main room of the café c) View of the buffet area of the café d) Floorplan of the casino within the VR-gambling environment e) View of the main room of the casino f) View of the sports bar within the casino

Experimental procedure. Participants were invited to the VR lab for three different sessions on three different days. The time between the sessions was between one day and most days (mean = 3.85, std = 3.36). During the three sessions participants either explored one of two different VR environments (VR-sessions) followed by the completion of two behavioral tasks, or simply performed the same two behavioral tasks in a standard lab-testing context (Lab-session). If the session was a VR-session, electrodermal activity (EDA)[35] was measured during a non-VR baseline period and the exploration of the VR-environments. The order of the sessions was pseudorandomized. At the first session, not depending on if VR was applied or not, participants arrived at the lab and the behavioral tasks were explained in detail. If the session was a Lab-session, participants proceeded with the two behavioral tasks. If the session was the first of the VR-sessions, participants were subsequently familiarized with the VR-equipment and handling. Participants were seated and a five-minute EDA baseline was measured (baseline phase). For both VR-sessions participants were then helped to apply the VR-equipment and entered the VR-environments. Within the VR-environments participants first explored the starting area for 5 minutes (first exploration phase). After these five minutes participants were asked to enter the experimental area of the environment (either the café or the casino) (Figure 1). Participants were instructed to explore the interior experimental area for five minutes (second exploration phase). Each of the three phases was later binned into five one-minute intervals and labeled as B (1 to 5) for the baseline phase, F (1 to 5) for the first

exploration phase and S (1 to 5) for the second exploration phase. During the exploration the experimenter closely monitored the participants and alerted them if they were about to leave the designated physical VR-space. After the second exploration phase participants were asked to proceed to a terminal within the VR-environment on which the behavioral tasks were presented.

Physiological measurements. EDA was measured using a BioNomadix-PPGED wireless remote sensor together with a Biopac MP160 data acquisition system (Biopac Systems, Santa Barbara, CA, USA). A GSR100C amplifier module with a gain of 5V, low pass filter of 10 Hz and a high pass filter DC were included in the recording system. The system was connected to the acquisition computer running the AcqKnowledge software. Triggers for the events within the VR-environments were send to the acquisition PC via digital channels from the VR-PC. Disposable Ag/AgCl electrodes were attached to the thenar and hypothenar eminences of the non-dominant palm. Isotonic paste (Biopac Gel 101) was used to ensure optimal signal transmission. The signal was measured in micro-Siemens units (mS).

Behavioral Tasks. Participants performed the same two behavioral tasks with slightly varied rewards and choices in each of the three sessions: a temporal discounting task[9] and a 2-step sequential decision-making task[53,54]. Results from the 2-step task will be reported separately. In the temporal discounting task participants had to repeatedly choose between an immediately available (smaller-but-sooner, SS) monetary reward of 20 Euros and larger-but-later (LL) temporally delayed monetary rewards. The LL options were multiples of the SS option (range 1.025 to 3.85) combined with different temporal delays (range 1 to122 days). We constructed three sets of six delays and 16 LL options. Each set had the same mean delay and the same mean LL option. Combining each delay with every LL option within each set resulted in three sets of 96 trials. The order of presentation of the trial sets was counter balanced across participants and sessions. All temporal discounting decisions were hypothetical[55,56]. In the VR-version of the task two yellow squares were presented to the participants (Figure 2). One depicted the smaller offer of 20 Euros now, while the other depicted the delayed larger offer. For the lab-based testing session were presented in the same way except that the color scheme was white writing on a black background. Offers were randomly assigned to the left/right side of the display and presented until a decision was made. The next trial started .5 to 1 seconds after the decision. Participants indicated their choice either by aiming the VR-controller at the preferred option and pulling the trigger (VR-sessions) or by pressing the corresponding arrow key on the keyboard (Lab-session).

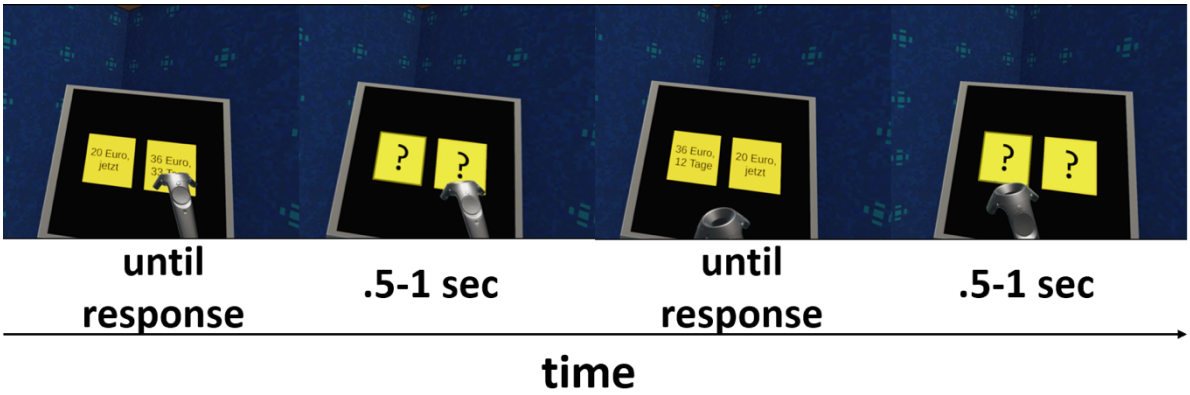


Figure 2. Presentation of the temporal discounting task in VR. Participants had to repeatedly decide between a small but immediate reward (SS) and larger but temporally delayed rewards (LL). Amounts and delays were presented in yellow squares. During the inter-trial intervals (.5-1 sec.) these squares contained only question marks. Participants indicated their choice by pointing the VR-controller at one of the yellow squares and pulling the trigger.

Model-free discounting data analysis. The behavioral data from the temporal discounting task was analyzed using several complementary approaches. First, we used a model-free approach that involved no a priori hypotheses about the mathematical shape of the discounting function. For each delay, we estimated the LL reward magnitudes at which the subjective value of the LL reward was equal to the SS (indifference point). This was done by fitting logistic functions to the choices of the participants, separately for each delay. Subsequently, these indifference points were plotted against the corresponding delays, and the area under the resulting curve (AUC) was calculated using standard procedures[57]. AUC values were derived for each participant and testing session, and further analyzed with the intra-class correlation (ICC) and the Friedman Test, a non-parametric equivalent of the repeated measures ANOVA model.

Computational modeling. Previous research on the effects of the delay of a reward on its valuation proposed a hyperbolic nature of devaluation[58,59]. Therefore, the rate of discounting for each participant was also determined employing a cognitive modeling approach using hierarchical Bayesian modeling[8]. A hierarchical model was fit to the data of all participants, separately for each session (see below). We applied a hyperbolic discounting model (eq. 1):

$$SV(LL_t) = \frac{A_t}{(1+\exp(k)*D_t)} \quad (eq.1)$$

Here, SV(LL) denotes the subjective (discounted) value of the LL. A and D represent the amount and the delay of the LL, respectively. The parameter k governs the steepness of the value decay over time, with higher values of k indicating steeper discounting of value over

time. As the distribution of the discount rate k is highly skewed, we estimated the parameter in log-space (log[k]), which avoids numerical instability in estimates close to 0.

The hyperbolic model was then combined with two different choice rules, a softmax action selection rule[60] and the drift diffusion model[40]. For softmax action selection, the probability of choosing the LL option on trial *t* is given by eq. 2.

$$P(LL_t) = \frac{\exp(SV_{LL_t} * ß)}{\exp(SV_{SS_t} * ß) + \exp(SV_{LL_t} * ß)} \quad \text{(eq. 2)}$$

Here, the ß-parameter determines the stochasticity of choices with respect to a given valuation model. A ß of 0 would indicate that choices are random, whereas higher ß values indicate a higher dependency of choices on option values. The resulting best fitting parameter estimates were used to test the ICC and systematic session effects via comparison of the posterior probabilities of group parameters.

Next, we incorporated response times (RTs) into the model by replacing the softmax choice rule with the drift diffusion model (DDM)[39–42]. The DDM models choices between two options as a noisy evidence accumulation that terminates as soon as the accumulated evidence exceeds one of two boundaries. In this analysis the upper boundary was set to represent LL choices, and the lower boundary SS choices. RTs for choices of the immediate reward were multiplied by -1 prior to model estimation. To prevent outliers in the RT data from negatively impacting model fit, the 2,5% slowest and fastest trials of each participant were excluded from the analysis[40,41]. In the DDM the RT on trial *t* is distributed according to Wiener first passage time (wfpt) (eq. 3).

$$RT_t \sim \text{wfpt}\,(\alpha, \tau, z, v) \quad \text{(eq. 3)}$$

Here α represents the boundary separation modeling the tradeoff between speed and accuracy. τ represents the non-decision time, reflecting perception and response preparation times. The starting value of the diffusion process is given by z, which therefore models a potential bias towards one of the boundaries. Finally, rate of evidence accumulation is given by the drift-rate v.

We first fit a null model (DDM_0), where the value difference between the two options was not included, such that DDM parameters were constant across trials[41,42]. We then used two different temporal discounting DDMs, in which the value difference between options modulated trial-wise drift rates. This was done using either a linear (DDM_L) or a non-linear sigmoid (DDM_S) linking function[43]. In the DDM_L, the drift-rate v in each trial is linearly

dependent on the trial-wise scaled value difference between the LL and the SS options (eq. 4) [40]. The parameter v_{coeff} maps the value differences onto v and scales them to the DDM:

$$v_t = v_{coeff} * (SV\ (LL_t) - SV(SS_t)) \text{ (eq. 4)}$$

One drawback of a linear representation of the relationship between the drift-rate v and trial-wise value differences is that *v* might increase infinitely with high value differences, which can lead the model to under-predict RTs for high value differences[41]. In line with previous work [41,42] we thus included a third version of the DDM, that assumes a non-linear sigmoidal mapping from trial-wise value differences to drift rates (eq. 5 and 6)[39]:

$$v_t = S\ (v_{coeff} * (SV(LL_t) - SV(SS_t))) \text{ (eq. 5)}$$

$$S(m) = \frac{2 * v_{max}}{1 + \exp(-m)} - v_{max} \text{ (eq. 6)}$$

Here, the linear mapping function from the DDM_L is additionally passed through a sigmoid function S with the asymptote v_{max}, causing the relationship between v and the scaled trail-wise value difference *m* to asymptote at v_{max}.

We have previously reported detailed parameter recovery analyses for the DDM_S in the context of value-based decision-making tasks such as temporal discounting[41], which revealed that both subject-level and group-level parameters recovered well.

Hierarchical Bayesian Models. All models were fit to the data of all participants in a hierarchical Bayesian estimation scheme, separately for each session, resulting in independent estimates for each participant per session. Participant-level parameters were assumed to be drawn from group-level Gaussian distributions, the means and precisions of which were again estimated from the data. Posterior distributions were estimated via Markov Chain Monte Carlo in the R programming language[61] using the JAGS software package[62]. For the DDM's the Wiener module for JAGS was used[63]. For the group-level means, uniform priors over numerically plausible parameter ranges were chosen (Table 1). Priors for the precision of the group-level distribution were Gamma distributed (0.001, 0.001). The convergence of chains was determined by the R-hat statistic[64]. Values between 1 and 1.01 were considered acceptable. Comparisons of relative model fit were performed using the Deviance Information Criterion (DIC), where lower values reflect a superior model fit[65].

Table 1. Ranges for the uniform priors of group-level parameter means. Ranges were chosen to cover numerically plausible values. Parameters included in multiple models are only listed once.

Parameter	Min.	Max.
log(k)	-20	3
softmax ß	0	10
$\boldsymbol{v}$	-100	100
$\boldsymbol{\tau}$	.1	6
$\boldsymbol{\alpha}$	.01	5
z	.1	.9
$\boldsymbol{v}_{\mathbf{coeff}}$	-100	100
$\boldsymbol{v}_{\mathbf{max}}$	0	100

Systematic session effects on model parameters. Potential systematic session effects on group level posterior distributions of parameters of interest were analyzed by overlaying the posterior distributions of each group level parameter for the different sessions. Here we report the mean of the posteriors of the estimated group level parameters and the difference distributions between them, the 95% highest density intervals (HDI) for both of these as well as directional Bayes Factors (dBF) which quantity the degree of evidence for reductions vs. increases in a parameter. Because the priors for the group effects are symmetric, this dBF can simply be defined as the ratio of the posterior mass of the difference distributions above zero to the posterior mass below zero[66]. Here directional Bayes Factors above 3 are interpreted as moderate evidence in favor of a positive effect, while Bayes Factors above 12 are interpreted as strong evidence for a positive effect[67]. Specifically, a dBF of 3 would imply that a positive directional effect is three times more likely than a negative directional effect. Bayes Factors below 0.33 are likewise interpreted as moderate evidence in favor of the alternative model with reverse directionality. A dBF above 100 is considered extreme evidence[67]. The cutoffs used here are liberal in this context, because they are usually used if the test is against a H_0 implying an effect of 0. In addition, we report the effect size (Cohen's *d*) based on the mean posterior distributions of the session means, the pooled standard deviations across sessions and the correlation between sessions.

ICC analysis. The test-retest reliability of the best fitting parameter values between the three sessions was analyzed using the intra-class correlation coefficient (ICC). The ICC-analysis was done in the R programming language[61] and was based on a mean-rating of three raters, absolute agreement and a two-way mixed model. ICC values below .5 are an indication of poor test-retest reliability, whereas values in the range between .5 and .75 indicate a moderate test-retest reliability[68]. Higher values between .75 and .9 indicate a good reliability, while values above .9 suggest an excellent test-retest reliability.

Analysis of physiological data. General physiological arousal is thought to be reflected in the slowly-varying skin conductance level (SCL)[35]. Thus, the SCL was extracted from the EDA signal using continuous decomposition analysis (CDA) via the Ledalab toolbox[69] for Matlab (The MathWorks ©). For the deconvolution, default settings were used. The resulting signal was then transformed into percentage change from the mean signal of the five minutes baseline phase at the beginning of the experiment. Subsequently, five one-minute bins were constructed for each phase of the VR-session (baseline phase, the first exploration phase and the second exploration phase). An alternative way of classifying general arousal can be the number of spontaneous phasic responses (SCR) in the EDA signal[69]. Again, the signal was divided in one-minute bins and the number of spontaneous SCRs during each bin was calculated from the phasic component of the deconvoluted EDA signal using the Ledalab toolbox. The resulting values were similarly transformed into percentage change from the mean number of SCRs during the five baseline bins. To test whether entering the VR-environments had a general effect on physiological arousal, we compared the values for the last time point of the base line phase (B5) with the first time point of the first exploration phase (F1) for both sessions using a non-parametric Wilcoxon Signed-Rank Test. To test whether there was a differential effect of entering the different experimental areas of the VR-environments on physiological arousal, for both measures the differences between the last time point of the first exploration phase (F5) and the first time point of the second exploration phase (S1) were compared across VR-sessions using a non-parametric Wilcoxon Signed-Tanks Test[70]. Effect sizes are given as r[71], computed as the statistic Z divided by the square-root of N. Effect sizes between 0 and .3 are considered small and effect sizes between .3 and .5 are considered medium and r values > .5 are considered large effects.

Data and code availability. Raw behavioral and physiological data as well as JAGS model code is available on the Open Science Framework (https://osf.io/xkp7c/files/).

Results

Temporal discounting AUC. The analysis of the AUC values revealed no significant session effect on a group level (Friedman Test: Chi-Squared = 1.235 df =2 p =.539). Furthermore, the ICC value was .93 (95% confidence interval (CI): .89 - .96) (p<.001) indicating an excellent test-retest reliability of temporal discounting AUC values over the three sessions (Table 2). Pairwise correlations between all sessions can be found in the supplementary materials (Supplementary Figure S1).

Softmax choice rule. For the hyperbolic model with softmax choice rule, the group level posteriors showed little evidence for systematic effects of the different sessions on log(k) (all BFs < 3 or >.33) (Figure 3a and c and Table 2). In contrast, the softmax β parameter was higher (reflecting higher consistency) in the $VR_{neutral}$ session compared to the other sessions (vs. Lab: dBF = .01 and vs. $VR_{gambiling}$: dBF = .048) (Figure 3b and d, Table 2). This indicates that a higher β in the $VR_{neutral}$ session was approximately 100 (Lab) or 20 ($VR_{gambling}$) times more likely than a lower β. There was little evidence for a systematic effect between the Lab and $VR_{gambling}$ sessions (dBF = .446).

Table 2. 95% HDIs for the two parameters of the hyperbolic discounting model. HDIs are described by the min. value first and the max value second. Directional Bayes Factors (dBF) are calculated as BF = i/(1-i), with i being the probability mass of the difference distributions above zero. Effect sizes are given as Cohen's d.

	Log(k)					**ß**				
Session	**Mean**	**HDI**		**dBF**	**d**	**Mean**	**HDI**		**dBF**	**d**
Lab	-4.083	-4.643	-3.530	-	-	.417	.355	.489	-	-
$VR_{neutral}$	-4.348	-4.912	-3.797	-	-	.577	.461	.714	-	-
$VR_{gambling}$	-4.274	-4.882	-3.687	-	-	.448	.363	.547	-	-
Lab-$VR_{neutral}$	.266	-.520	1.054	2.712	.38	-.16	-.31	-.024	.01	.9
Lab-$VR_{gambling}$	.191	-.620	1.01	2.162	.3	-.03	-.148	.081	.446	.18
$VR_{gambling}$-$VR_{geutral}$	.074	-.746	.885	1.264	.1	-.129	-.29	.023	.048	.56

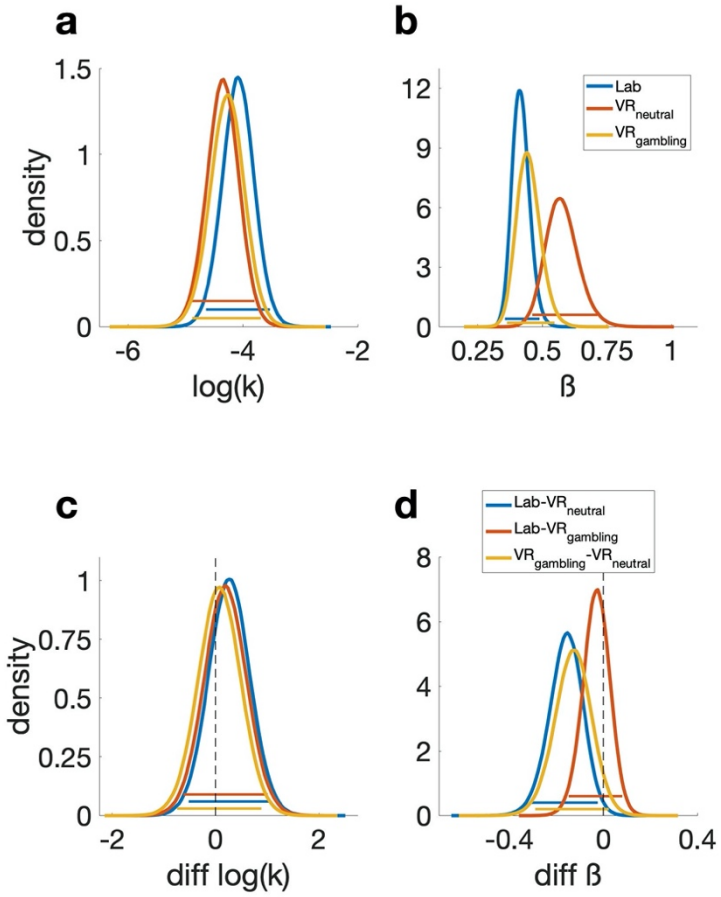


Figure 3. Posterior distributions of the parameters of the hyperbolic discounting model. Colored bars represent the corresponding 95% HDIs. a) Posterior distribution of the log(k) parameter (reflecting the degree of temporal discounting) for all three sessions. b) Posterior distribution of the β or inverse temperature parameter (reflecting decision noise). c) Pairwise difference distributions between the posteriors of the log(k) parameters of all three sessions. d) Pairwise difference distributions between the posteriors of the β parameters of all three sessions.

The ICC value for the log(k) parameter indicated an excellent test-retest reliability of .91 (CI: .86 - .96) (p<.001) (Table 3). For the β-parameter of the softmax choice rule the ICC value was .34 (CI: .17 - .53) (p<.001) indicating a poor test-retest reliability (Table 3). The pairwise correlations of estimated parameter values between all sessions can be found in the Supplement (Supplementary Figure S2 and S3). Pairwise correlations between all sessions for both parameters can be found in the supplementary materials (Supplementary Figure S2 and S3).

Table 3. Summary of the results of the ICC analysis for the AUC values as well as the two parameters of the hyperbolic discounting model with a softmax choice rule. Lower and upper bound describe the 95% confidence interval.

Parameter	ICC	p	Lower bound	Upper Bound
AUC	.93	<.001	.89	.96
log(k)	.91	<.001	.86	.95
β	.34	<.001	.17	.53

Table 4. Summary of the DICs of all DDM models in all sessions. Ranks are based on the lowest DIC in all sessions.

Model	Lab	$VR_{neutral}$	$VR_{gambling}$	Rank
DDM_0	9275.7	9569.8	9225.7	3
DDM_L	7558.9	7921.4	7663.0	2
DDM_S	6992.3	7327.2	7033.1	1

Drift diffusion model choice rule. Model comparison revealed that the DDM_S had the lowest DIC in all conditions (Table 4) replicating previous work [41,42,44]. Consequently, further analyses of session effects and reliability focused on this model. For the log(k) parameter, the 95% HDIs showed a high overlap between all sessions indicating no systematic session effects, however the BFs showed moderate evidence for a reduced log(k) in the $VR_{neutral}$-session (Figure 4a and d, Table 5). A lower value in the $VR_{neutral}$-session was about seven (Lab-session dBF = 6.756) or four times ($VR_{gambling}$ dBF = 3.86) more likely than a lower value. Similarly, the posterior distributions of v_{max}, v_{coeff} and α were highly overlapping, whereas some of the dBFs gave moderate evidence for systematic directional effects within these parameters (Figure 4 b,c,d and f, Figure 5 b and f, Table 5). v_{coeff}, mapping trial-wise value difference onto the drift rate, was lowest in the Lab-session and highest in $VR_{neutral}$ (Lab-$VR_{neutral}$ dBF = .074, Lab-$VR_{gambling}$ = .2, $VR_{gambling}$-$VR_{neutral}$ = .228). Thus, an increase in v_{coeff} in $VR_{neutral}$ compared to the Lab-session was approximately thirteen times more likely than a decrease. Likewise, it was approximately five times more likely that there was an increase in the $VR_{neutral}$ compared to the $VR_{gambling}$-session. For v_{max}, the upper boundary for the value difference's influence on the drift rate, the dBFs indicated that a positive shift from $VR_{gambling}$ to $VR_{neutral}$ was five times more likely than a negative shift (dBF = .203) but there was only very little indication of a systematic difference between both of them and the Lab-session. Finally, a reduction of the boundary separation parameter α was five times more likely than an increase when comparing the $VR_{neutral}$ to the Lab-session (dBF = .255). There was little evidence for any other systematic differences. The bias parameter z displayed high overlap in HDIs and little evidence for any systematic effects between sessions (all dBFs >.33 or <3) (Figure 5 c and f, Table 5). For the non-decision time parameter τ there was extreme evidence for an increase in the VR-sessions compared to the Lab-session (both dBFs >100), reflecting prolonged motor and/or perceptual components of the RT that was more than 100 times more likely than a shortening of these components.

Table 5. Directional Bayes Factors (dBF) and effect sizes (Cohen's d) for all between session comparisons for all parameters of the DDMs. Means and HDIs of the posteriors and difference distributions are summarized in the supplementary materials (Supplementary Table S1). BFs are calculated as BF = i/(1-i), with i being the probability mass of the difference distributions above zero.

Contrast	**log(k)**		v_{coeff}		v_{max}		τ		α		**z**	
	dBF	d	dBF	d	dBF	d	dBF	d	dBF	d	dBF	d
Lab-VR$_{neutral}$	6.756	.37	.074	.37	.377	.2	>100	1.2	.255	.224	.530	.2
Lab-VR$_{gambling}$	1.679	.19	.200	.59	1.573	.09	>100	1.5	.358	.160	1.118	.04
VR$_{gambling}$-VR$_{neutral}$	3.860	.29	.228	.27	.203	.34	3.413	.17	.629	.070	.458	.2

The ICC value for the log(k) parameter was .7 (CI: .56 - .8) indicating a moderate test-retest-reliability (Table 5). For the other DDMs parameters, ICC values were substantially lower (Table 6). Pairwise correlations between all sessions for all parameters can be found in the supplementary materials (Supplementary Figure S4-S9).

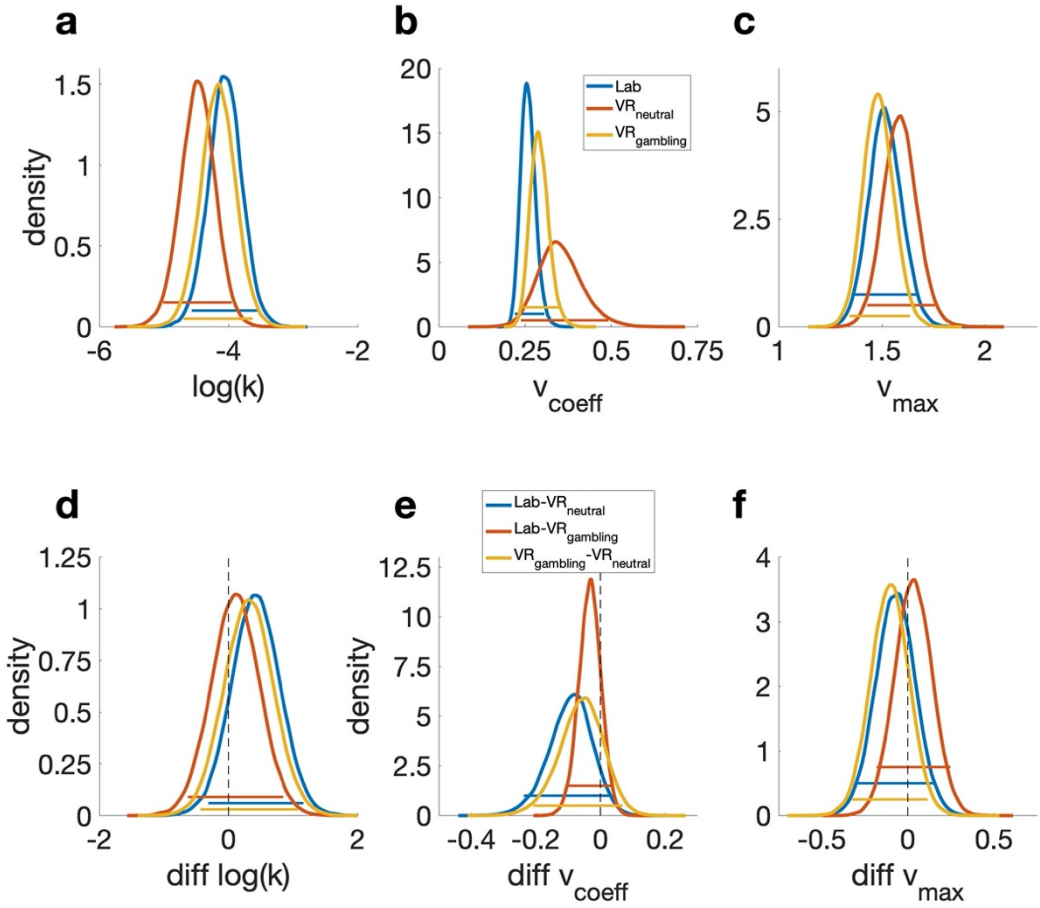


Figure 4. Posterior distributions of the parameters of the DDM_s model. Colored bars represent the corresponding 95% HDIs. a) Posterior distributions of the log(k) parameter for all three sessions. b) Posterior distributions of the v_{coeff} parameter (mapping the drift rate onto the trial wise value difference). c) Posterior distributions of the v_{max} parameter (setting an asymptote for the relation between the trial wise value difference and the drift rate). d) Pairwise difference distributions between the posterior distributions of the log(k) parameters of the three sessions. e) Pairwise difference distributions between the posterior distributions of the v_{coeff} parameters of the three sessions. f) Pairwise difference distributions between the posterior distributions of the v_{max} parameters of the three sessions.

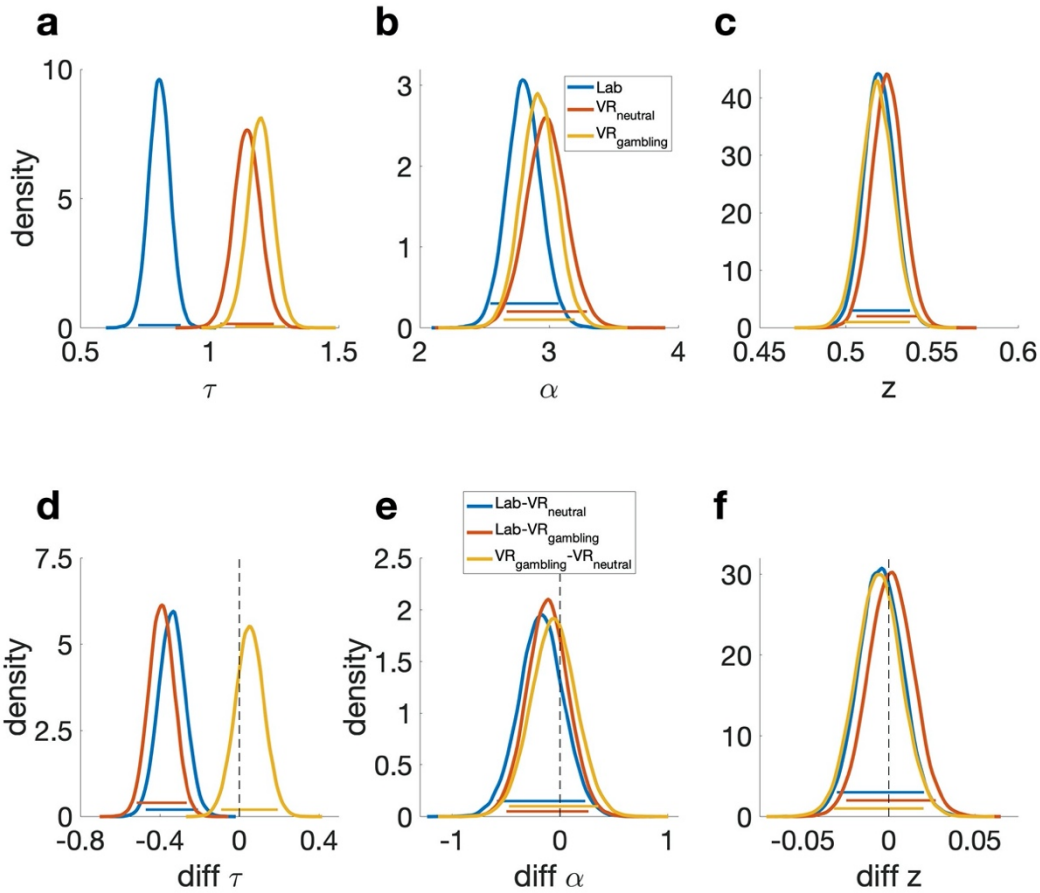


Figure 6. Posterior distributions of the remaining parameters of the DDM_s model. Colored bars represent the corresponding 95% HDIs. a) Posterior distributions of the τ parameter (non-decision time) for all three sessions. b) Posterior distributions of the α parameter (separation between decision boundaries). c) Posterior distributions of the z parameter (bias towards one decision option). d) Pairwise difference distributions between the posterior distributions of the τ parameters of the three sessions. e) Pairwise difference distributions between the posterior distributions of the α parameters of the three sessions. f) Pairwise difference distributions between the posterior distributions of the z parameters of the three sessions.

Table 5. Summary of the results of the ICC analysis of the DDM_s parameters.

Parameter	ICC	p	Lower bound	Upper Bound
log(k)	.7	<.001	.56	.8
ν_{coeff}	.11	.14	-.053	.3
ν_{max}	.33	<.001	.16	.52
τ	.19	.033	.019	.38
α	.42	<.001	.24	.59
z	.4	<.001	.22	.58

Electrodermal activity (EDA). The data of 8 of the 34 participants had to be excluded from the EDA analysis, due to technical problems or missing data during one of the testing sessions. Physiological reactivity in the remaining 26 (18 female) participants was analyzed by converting the SCL signal as well as the nSCRs into percent change from the mean level during the base line phase. Both signals were then binned into five one-minute intervals for each of the three phases (baseline, first exploration and second exploration phase). All comparisons were tested with the Wilcoxon Signed Rank Test. Entering the VR-environments (comparing bin B5 to bin F1 for both environments individually) resulted in a significant increase in the SCL values for both VR-environments ($VR_{neutral}$: $Z = -3.67$, $p < .001$, $r = .72$; $VR_{gambling}$: $Z = -3.543$, $p = .002$, $r = .695$) (Figure 6 b and d). The effect was large in both sessions ($r > .5$). However, for the number of spontaneous SCRs (nSCRs), this effect was only significant in the neutral VR-environment (neutral: $Z = -2.623$, $p = .009$, $r = .515$; gambling: $Z = -.013$, $p = .99$, $r =.002$). There was no significant difference between the two sessions, but the effect was of medium size ($Z = -1.7652$, $p = .078$, $r =.346$) (Figure 6 a and b). To test whether entering the specific experimental areas of the two VR-environments (virtual café vs. virtual casino) had differential effects on physiological responses, the increase in physiological arousal from the end of the first exploration phase to the start of the second exploration phase was examined (comparing bin F5 to bin S1, see Figure 6b and 6d). The SCL (neutral: $Z = -0.7238$, $p = -.469$, $r = .142$; gambling: $Z = -.089$, $p = .929$, $r = .017$) as well as the nSCRs (neutral: $Z = -1.943$, $p = .052$, $r = .381$; gambling: $Z = .982$, $p = .326$, $r = .193$) assessed for each session individually showed no significant effect. The effect size was medium ($r = .381$) for the nSCRs of the $VR_{neutral}$-session and small for all other comparisons ($r < .3$). Furthermore, the Wilcoxon Signed-Ranks test indicated no significant differences between the two experimental areas on both physiological arousal measures (SCL: $Z = -.572$, $p = .381$, $r = .11$; nSCRs.: $Z = -1.7652$, $p = .078$, $r = .346$). For the nSCRs however, the effect was of a medium size ($r = .346$) (Figure 6b).

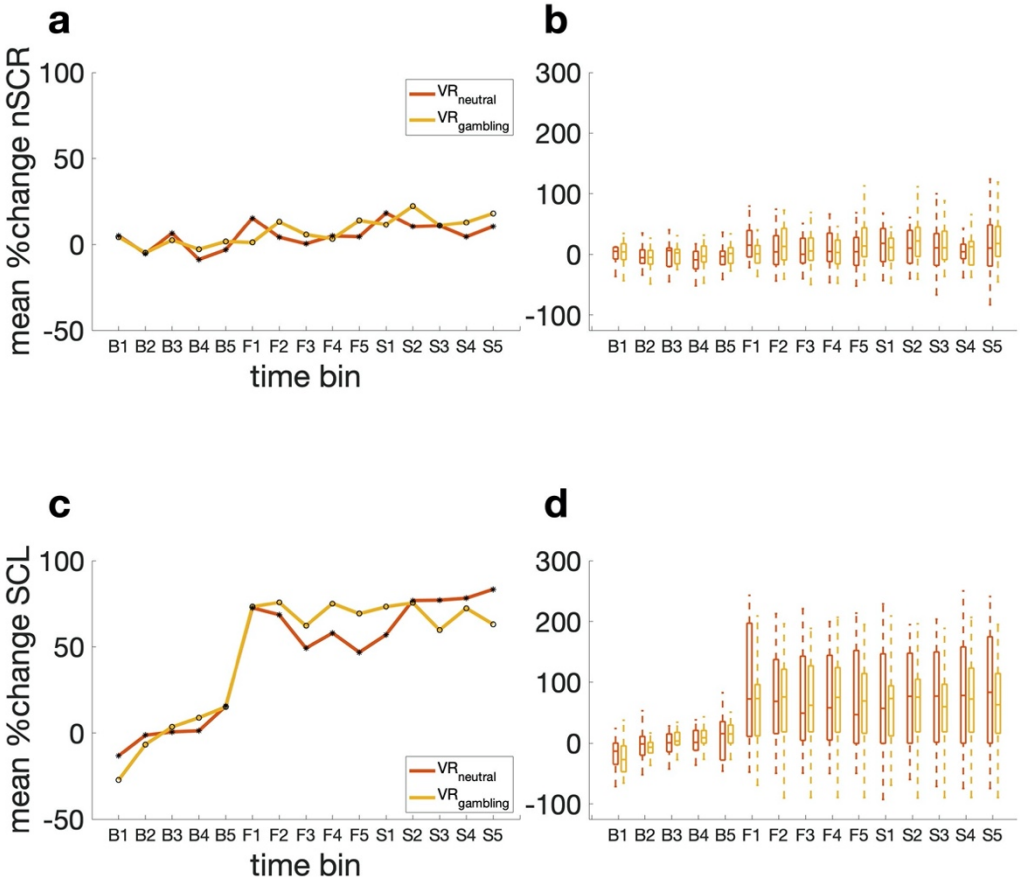


Figure 6. Results of the EDA measurements divided into 15 time points over the course of the baseline phase, measured before participants entered the VR-environments, and the first and second exploration phases. Each of the three phases is divided into five one-minute bins (B1-5: pre-VR baseline, F1-5: first exploration phase in VR, S1-5: second exploration phase VR). a: Median percent change from baseline mean for no. of spontaneous SCRs over all participants. b: Boxplot of percentage change from baseline mean for no. spontaneous SCRs over all participants. c: Median percent change from baseline mean of SCL over all participants. d: Boxplots of percentage change from base line mean of SCL over all participants.

Discussion

Here we investigated the feasibility of a virtual reality (VR) approach that combines the assessment of physiological and behavioral cue-reactivity for potential applications in gambling disorder. In a proof-of-concept study, the approach was tested in a group of healthy young non-gambling control participants, with the aims of investigating technical feasibility, as well as test-retest reliability of temporal discounting measured in VR as a trait-index of impulsivity[51]. Participants performed a temporal discounting task within two different VR-environments ($VR_{neutral}$ vs. $VR_{gambling}$) as well as in a standard computer-based lab testing session. Immersion in VR generally increased physiological arousal as assessed via electrodermal activity (EDA) but these effects were not differentially modulated by the different VR environments. Finally, results revealed good to excellent test-retest reliability of temporal discounting assessed in VR.

To test the behavioral effects of cue-reactivity on impulsivity, we implemented temporal discounting as a measure of choice impulsivity. Healthy controls displayed little evidence for systematic differences in choice preferences between the Lab-session and the VR-sessions. This was observed for model-free measures (AUC), as well as the *log(k)* parameter of the hyperbolic discounting model with the softmax choice rule and the drift diffusion model with non-linear drift rate scaling (DDM_s). Model comparison revealed that the DDM_s accounted for the data best, confirming previous findings[39,41,42,44]. Although generally, discount rates assessed in the three sessions were of similar magnitude, in the DDM_s there was moderate evidence for reduced discounting (i.e. smaller values of *log(k)*) in the $VR_{neutral}$ session. The reasons for this could be manifold. One possibility is that environmental novelty plays a role, such that perceived novelty of the $VR_{neutral}$ session might have been lower than for the $VR_{gambling}$ and Lab-sessions. Exposure to novelty can stimulated dopamine release[72], which is known to impact temporal discounting[73]. Nonetheless, effect sizes were medium (.37 and .29) and the dBFs revealed only moderate evidence. Numerically, the mean *log(k)*'s of the softmax model showed the same tendency, but here effects were less pronounced. One possibility is that the inclusion of additional latent variables in the DDM_s might have increased sensitivity to detect this effect. There was also evidence for a session effect on the scaling parameter (v_{coeff}). Here, the impact of trial-wise value differences on the drift rate was attenuated in the Lab-session, with dBFs revealing strong ($VR_{neutral}$) or moderate evidence ($VR_{gambling}$) for a reduction in v_{coeff} in the Lab-session. Again, effect sizes were medium. Nevertheless, the data suggest increased sensitivity to value differences in VR. This effect might be due to the option presentation in the Lab-session compared to the VR-sessions. The presentation of options within VR might have been somewhat more salient, which might have increased attention allocated to the value differences

568 within the VR-sessions. However, this remains speculative until further research reproduces
569 and further assesses these specific effects on the DDM parameters. Boundary separation (α),
570 drift rate asymptote (v_{max}) and starting point (z) showed little evidence for systematic
571 differences between sessions. The only DDM_s parameter showing extreme evidence for a
572 systematic difference between the lab- and VR-sessions was the non-decision time (τ). This
573 effect is unsurprising, as it describes RT components attributable to perception and/or motor
574 execution. Given that indicating a response with a controller in three-dimensional space takes
575 longer than a simple button press, this leads to substantial increases in τ during VR testing.
576 Finally, the good test-retest reliability of *log(k)* from the DDM_s furthermore indicates that RTs
577 obtained in VR can meaningfully be modeled using the DDM. The potential utility of this
578 modeling approach in the context of gambling disorder is illustrated by a recent study that
579 reported reduced boundary separation (α) in pathological gamblers compared to healthy
580 controls in a reinforcement learning task[44]. Given that there are mixed results when it comes
581 to the effect of addiction related cues on RTs[74–76], the effects of these cues on the latent
582 decision variables included in the DDM could provide additional insights. Taken together, these
583 results show that VR immersion in general does not influence participants inter-temporal
584 preferences in a systematic fashion and might open up a road to more ecologically valid lab
585 experiments measuring behavioral cue reactivity. This is in line with other results showing the
586 superiority of VR compared to classical laboratory experiments in other research lines[77].

587 The present data add to the discussion concerning the reliability of behavioral tasks[1,46–
588 49,51] in particular in the context of computational psychiatry[7,78]. To examine test-retest
589 reliability, the three sessions were performed on different days and with a mean interval of 3.85
590 days between sessions. The test-retest reliability for the AUC and the *log(k)* parameter of the
591 hyperbolic discounting model with softmax choice rule were both excellent. For the *log(k)* of
592 the DDM_s the ICC was good, but slightly lower than for AUC and softmax. This might be
593 related to other factors influencing RTs (fatigue, motivation, etc.) that might vary
594 unsystematically across sessions. Nevertheless, the discount rate *log(k)* was overall stable
595 regardless of the analytical approach. The ICC of .7 observed for the DDM_S was comparable
596 to earlier studies on temporal discounting reliability[48,49]. Kirby and colleagues[48] for instance
597 demonstrated a reliability of .77 for a five-week interval and .71 for one year. This shows that
598 at least over shorter periods from days to weeks, temporal discounting performed in VR has a
599 reliability comparable to standard lab-based testing. Enkavi and colleagues[45] stress that in
600 particular difference scores between conditions (e.g. Stroop, Go-NoGo etc.), show
601 unsatisfactory reliability due to the low between participants variation created by commonly

602 used behavioral tasks. Assessment of difference scores was not applicable in the present study.
603 Nevertheless, there was no positive evidence for systematic effects on *log(k)* (with the
604 exception of the potential novelty effects discussed above), and the test-retest reliability
605 between all conditions was at least good across analysis schemes, indicating short-term stability
606 of temporal discounting measured in VR. It is worth noting, however, that temporal discounting
607 shares some similarities with questionnaire-based measures. As in questionnaires, in temporal
608 discounting tasks participants are explicitly instructed to indicate their preferences. This might
609 be one reason why the reliability of temporal discounting is often substantially higher than that
610 of other behavioral tasks[45,48,49,51]. Other parameters of the DDM_S showed lower levels of test-
611 retest reliability. Especially the v_{coeff} parameter were less reliable, at least when estimated
612 jointly with v_{max}. In the DDM_L, which does not suffer from potential trade-offs between these
613 different drift rate components, the ICC of v_{coeff} was good (Supplementary Table S2). Similarly,
614 here *log(k)* also showed an excellent ICC.

615 The present experimental set-up might be a viable tool to improve the ecological validity
616 of cue-reactivity research in gambling disorder. VR has previously been used to study cue-
617 reactivity in gambling disorder[34,79,80], but also smokers[81] and alcoholics[31]. Our design
618 extends these previous approaches in several ways. First, we included both a neutral and a
619 gambling-related environment, which allows us to disentangle general VR effects from specific
620 contextual effects. Second, our reliability checks for temporal discounting show that model-
621 based constructs with clinical relevance for addiction[10,14] can be reliably assessed during
622 immersion in VR. Together, these advances might yield additional insights into the mechanisms
623 underlying cue-reactivity in behavioral addictions.

624 Understanding how addictions manifest on a computational and physiological level is
625 important to further the understanding the mechanisms underlying maladaptive decision-
626 making. Although alterations in neural reward circuits, in particular in ventral striatum and
627 ventromedial prefrontal cortex, are frequently observed in gambling disorder, there is
628 considerable heterogeneity in the directionality of these effects[82]. Gambling-related visual
629 cues interfere with striatal valuation signals in gamblers, and might thereby increase temporal
630 discounting[4]. In the present work, assessment of physiological cue-reactivity was limited to
631 electrodermal activity (EDA). EDA is an index of autonomic sympathetic arousal, which is in
632 turn related to the emotional response to addiction related cues[35,83–85]. The skin conductance
633 level (SCL) is increased in drug addicts in response to drug related cues[83]. Additionally, it has
634 been shown that addiction related cues in VR can elicit SCR responses in teen[84] and adult[85]
635 smokers. In our study, the number of spontaneous responses in the EDA signal (nSCRs), the

increase upon immersion in VR (B5 vs F1) was only significant in the $VR_{neutral}$ environment. The effect size for the difference between both environments was medium. Given that the two starting areas of the VR-environments were identical, this difference might have been caused by random fluctuations. However, an increase in the number of spontaneous SCRs during VR immersion has been reported previously[33] and thus warrants further investigation. The SCL, on the other hand, increased substantially upon immersion in VR, as indicated by a significant increase between the last minute of baseline recording (B5) and the first minute of the first exploration phase (F1). The effect sizes indicated a large effect. SCL then remained elevated throughout both exploration phases (F1 to S5) but did not increase further when the virtual café/casino area was entered. These results suggest that immersion in VR increases physiological arousal as measured with SCL in healthy controls participants independent of the presented VR environment.

There are several limitations that need to be acknowledged. First, there was considerable variability in test-retest intervals across participants. While most of the sessions were conducted within a week, in some participants this interval was up to three weeks, reducing the precision of conclusions regarding temporal stability of discounting in VR. Other studies, however, have used intervals ranging from five to fifty-seven weeks[48] or three months[49], and have reported comparable reliabilities. Moreover, there is evidence for a heritability of temporal discounting of around 30 and 50 percent at the ages of 12 and 14 years respectively[86]. This increases the confidence in the results obtained here. Nevertheless, a more systematic assessment of how long these trait indicators remain stable in VR would be desirable and could be addressed by future research. Second, the sample size was lower compared to larger studies conducted online[45], and the majority of participants was female. Both factors limit the generalizability of our results. However, large-scale online studies have shortcomings of their own, including test batteries that take multiple hours and/or multiple sessions to complete[45,46], potentially increasing participants' fatigue, and which might have detrimental effects on data quality. We also note that the present sample size was sufficiently large to reveal stable parameter estimates, showing that in our design participants performed the task adequately. Thirdly, the immersion in VR might have been reduced by the available physical lab space. To ensure safety, the experimenter had to at times instruct participants to stay within the designated VR-zone. This distraction might have reduced the effects caused by the VR-environments, because participants were not able to fully ignore the actual physical surroundings. Additionally, it might have influenced the EDA measurements in an unpredictable way. Future research would benefit from the implementation of markers within the VR-environments in order to ensure safety without

breaking immersion. Moreover, participants had to spend about thirty minutes in the full VR-setup. The behavioral tasks were presented after the exploration phase, such that participants might have been fatigued or experienced discomfort during task completion. Finally, cue-reactivity in gambling disorder is determined by many individual factors[28]. The VR-design presented here is designed for slot machine and sports betting players, and thus not applicable for other forms of gambling.

Overall, our results demonstrate the methodological feasibility of a VR-based approach to behavioral and physiological cue-reactivity. Healthy non-gambling control participants showed little systematic behavioral and physiological effects of the two VR environments. Moreover, our data show that temporal discounting is reliable behavioral marker, even if tested in very different experimental settings (e.g. standard lab testing vs. VR). It remains to be seen if such gambling-related environments produce elicit cue-reactivity in gambling disorder. However, results from similar applications have been encouraging[34,80]. These results show the feasibility of VR applications jointly assessing of behavioral and physiological cue-reactivity in addiction science.

References

1. Amlung, M. *et al.* Delay Discounting as a Transdiagnostic Process in Psychiatric Disorders: A Meta-analysis. *JAMA psychiatry* (2019) doi:10.1001/jamapsychiatry.2019.2102.
2. Kirby, K. N. & Petry, N. M. Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction* **99**, 461–471 (2004).
3. Peters, J. *et al.* Lower ventral striatal activation during reward anticipation in adolescent smokers. *Am. J. Psychiatry* **168**, 540–549 (2011).
4. Miedl, S. F., Büchel, C. & Peters, J. Cue-induced craving increases impulsivity via changes in striatal value signals in problem gamblers. *J. Neurosci.* **34**, 4750–4755 (2014).
5. Potenza, M. N. Review. The neurobiology of pathological gambling and drug addiction: An overview and new findings. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **363**, 3181–3189 (2008).
6. Wiehler, A. & Peters, J. Reward-based decision making in pathological gambling: The roles of risk and delay. *Neurosci. Res.* **90**, 3–14 (2015).
7. Huys, Q. J. M., Maia, T. V & Frank, M. J. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat. Neurosci.* **19**, 404–413 (2016).
8. Farrell, S. & Lewandowsky, S. *Computational modeling of cognition and behavior*. *Computational Modeling of Cognition and Behavior* (Cambridge University Press, 2018). doi:10.1017/CBO9781316272503.
9. Miedl, S. F., Peters, J. & Büchel, C. Altered neural reward representations in pathological gamblers revealed by delay and probability discounting. *Arch. Gen. Psychiatry* **69**, 177–186 (2012).
10. Bickel, W. K., Koffarnus, M. N., Moody, L. & Wilson, A. G. The behavioral- and neuro-economic process of temporal discounting: A candidate behavioral marker of addiction. *Neuropharmacology* **76**, 518–527 (2014).
11. Lempert, K. M., Steinglass, J. E., Pinto, A., Kable, J. W. & Simpson, H. B. Can delay discounting deliver on the promise of RDoC? *Psychol. Med.* **49**, 190–199 (2019).
12. Dixon, M. R., Jacobs, E. A., Sanders, S. & Carr, J. E. Contextual Control of Delay Discounting by Pathological Gamblers. *J. Appl. Behav. Anal.* **39**, 413–422 (2006).
13. Robinson, T. E. & Berridge, K. C. The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Res. Rev.* **18**, 247–291 (1993).
14. Bickel, W. K., Yi, R., Landes, R. D., Hill, P. F. & Baxter, C. Remember the future: Working memory training decreases delay discounting among stimulant addicts. *Biol. Psychiatry* **69**, 260–265 (2011).
15. Bickel, W. K., Moody, L. & Quisenberry, A. Computerized Working-Memory Training as a Candidate Adjunctive Treatment for Addiction. *Alcohol Res. Curr. Rev.* (2014).
16. Carter, B. L. & Tiffany, S. T. Meta--analysis of cue--reactivity in addiction research. *Addiction* (1999).
17. Starcke, K., Antons, S., Trotzke, P. & Brand, M. Cue-reactivity in behavioral addictions: A meta-analysis and methodological considerations. *J. Behav. Addict.* **7**, 227–238 (2018).
18. Berridge, K. C. & Robinson, T. E. Liking, wanting, and the incentive-sensitization theory of addiction. *Am. Psychol.* **71**, 670–679 (2016).
19. Anselme, P. Motivational control of sign-tracking behaviour: A theoretical framework. *Neurosci. Biobehav. Rev.* **65**, 1–20 (2016).
20. Berridge, K. C. From prediction error to incentive salience: Mesolimbic computation of

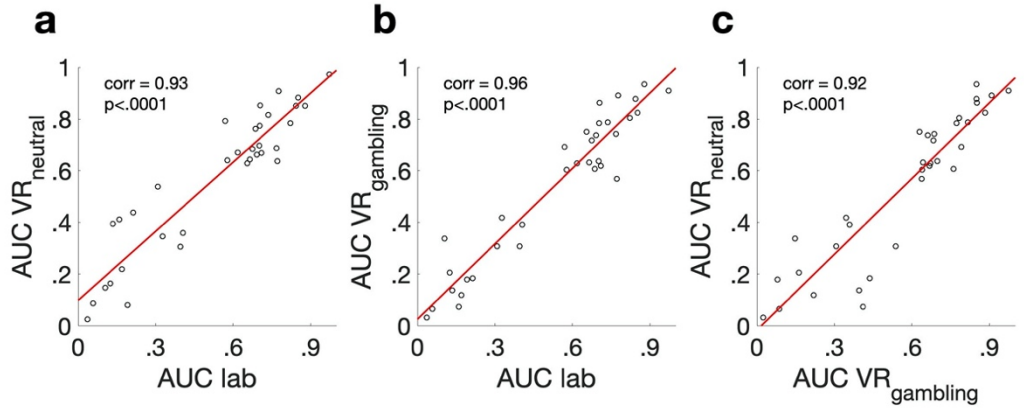
reward motivation. *Eur. J. Neurosci.* **35**, 1124–1143 (2012).
21. Volkow, N. D. *et al.* Cocaine cues and dopamine in dorsal striatum: Mechanism of craving in cocaine addiction. *J. Neurosci.* **26**, 6583–6588 (2006).
22. Brevers, D., He, Q., Keller, B., Noël, X. & Bechara, A. Neural correlates of proactive and reactive motor response inhibition of gambling stimuli in frequent gamblers. *Sci. Rep.* **7**, 1–11 (2017).
23. Brevers, D., Sescousse, G., Maurage, P. & Billieux, J. Examining Neural Reactivity to Gambling Cues in the Age of Online Betting. *Curr. Behav. Neurosci. Reports* **6**, 59–71 (2019).
24. Crockford, D. N., Goodyear, B., Edwards, J., Quickfall, J. & El-Guebaly, N. Cue-induced brain activity in pathological gamblers. *Biol. Psychiatry* **58**, 787–795 (2005).
25. Genauck, A. *et al.* Cue-induced effects on decision-making distinguish subjects with gambling disorder from healthy controls. *Addict. Biol.* 1–10 (2019) doi:10.1111/adb.12841.
26. Goudriaan, A. E., De Ruiter, M. B., Van Den Brink, W., Oosterlaan, J. & Veltman, D. J. Brain activation patterns associated with cue reactivity and craving in abstinent problem gamblers, heavy smokers and healthy controls: An fMRI study. *Addict. Biol.* **15**, 491–503 (2010).
27. Kober, H. *et al.* Brain Activity during Cocaine Craving and Gambling Urges: An fMRI Study. *Neuropsychopharmacology* **41**, 628–637 (2016).
28. Limbrick-Oldfield, E. H. *et al.* Neural substrates of cue reactivity and craving in gambling disorder. *Transl. Psychiatry* **7**, e992 (2017).
29. Potenza, M. N. *et al.* Gambling Urges in Pathological Gambling. *Arch. Gen. Psychiatry* **60**, 828 (2003).
30. van Holst, R. J., van Holstein, M., van den Brink, W., Veltman, D. J. & Goudriaan, A. E. Response inhibition during cue reactivity in problem gamblers: An fmri study. *PLoS One* **7**, 1–10 (2012).
31. Ghiţă, A. *et al.* Cue-Elicited Anxiety and Alcohol Craving as Indicators of the Validity of ALCO-VR Software: A Virtual Reality Study. *J. Clin. Med.* **8**, 1153 (2019).
32. Wang, Y. guang, Liu, M. hui & Shen, Z. hua. A virtual reality counterconditioning procedure to reduce methamphetamine cue-induced craving. *J. Psychiatr. Res.* **116**, 88–94 (2019).
33. Peterson, S. M., Furuichi, E. & Ferris, D. P. Effects of virtual reality high heights exposure during beam-walking on physiological stress and cognitive loading. *PLoS One* **13**, 1–17 (2018).
34. Bouchard, S. *et al.* Using Virtual Reality in the Treatment of Gambling Disorder: The Development of a New Tool for Cognitive Behavior Therapy. *Front. psychiatry* **8**, (2017).
35. Braithwaite, J. J., Watson, D. G., Jones, R. & Rowe, M. A Guide for Analysing Electrodermal Activity (EDA)& Skin Conductance Responses (SCRs)for Psychological Experiments. *Psychophysiology* **2013**, (2013).
36. Forstmann, B. U., Ratcliff, R. & Wagenmakers, E.-J. Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions. *Annu. Rev. Psychol.* **67**, 641–666 (2016).
37. Ballard, I. C. & McClure, S. M. Joint modeling of reaction times and choice improves parameter identifiability in reinforcement learning models. *J. Neurosci. Methods* **317**, 37–44 (2019).
38. Shahar, N. *et al.* Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. *PLoS Comput. Biol.* **15**, e1006803 (2019).
39. Fontanesi, L., Gluth, S., Spektor, M. S. & Rieskamp, J. A reinforcement learning

diffusion decision model for value-based decisions. *Psychon. Bull. Rev.* **26**, 1099–1121 (2019).
40. Pedersen, M. L., Frank, M. J. & Biele, G. The drift diffusion model as the choice rule in reinforcement learning. *Psychon. Bull. Rev.* **24**, 1234–1251 (2017).
41. Peters, J. & D'Esposito, M. The drift diffusion model as the choice rule in inter-temporal and risky choice: A case study in medial orbitofrontal cortex lesion patients and controls. *PLoS Comput. Biol.* **16**, 1–26 (2020).
42. Wagner, B., Clos, M., Sommer, T. & Peters, J. Dopaminergic modulation of human inter-temporal choice: a diffusion model analysis using the D2-receptor-antagonist haloperidol. *bioRxiv* (2020).
43. Miletić, S., Boag, R. J. & Forstmann, B. U. Mutual benefits: Combining reinforcement learning with sequential sampling models. *Neuropsychologia* **136**, (2020).
44. Wiehler, A. & Peters, J. Diffusion modeling reveals reinforcement learning impairments in gambling disorder that are linked to attenuated ventromedial prefrontal cortex value representations. *bioRxiv* 2020.06.03.131359 (2020) doi:10.1101/2020.06.03.131359.
45. Enkavi, A. Z. *et al.* Large-scale analysis of test-retest reliabilities of self-regulation measures. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 5472–5477 (2019).
46. Eisenberg, I. W. *et al.* Applying novel technologies and methods to inform the ontology of self-regulation. *Behav. Res. Ther.* **101**, 46–57 (2018).
47. Hedge, C., Powell, G. & Sumner, P. The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behav. Res. Methods* **50**, 1166–1186 (2018).
48. Kirby, K. N. One-year temporal stability of delay-discount rates. *Psychon. Bull. Rev.* **16**, 457–462 (2009).
49. Ohmura, Y., Takahashi, T., Kitamura, N. & Wehr, P. Three-month stability of delay and probability discounting measures. *Exp. Clin. Psychopharmacol.* **14**, 318–328 (2006).
50. Peters, J. & Büchel, C. Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. *J. Neurosci.* **29**, 15727–15734 (2009).
51. Odum, A. L. Delay discounting: Trait variable? *Behav. Processes* **87**, 1–9 (2011).
52. Petry, J. *Psychotherapie der Glücksspielsucht*. (Psychologie Verlags Union, 1996).
53. Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
54. Kool, W., Cushman, F. A. & Gershman, S. J. When Does Model-Based Control Pay Off? *PLoS Comput. Biol.* **12**, e1005090 (2016).
55. Johnson, M. W. & Bickel, W. K. Within-Subject Comparison of Real and Hypothetical Money Rewards in Delay Discounting. *J. Exp. Anal. Behav.* **77**, 129–146 (2002).
56. Bickel, W. K., Pitcock, J. A., Yi, R. & Angtuaco, E. J. C. Congruence of BOLD response across intertemporal choice conditions: Fictive and real money gains and losses. *J. Neurosci.* **29**, 8839–8846 (2009).
57. Myerson, J., Green, L. & Warusawitharana, M. Area Under the Curve As a Measure of Discounting. *J. Exp. Anal. Behav.* **76**, 235–243 (2001).
58. Green, L., Myerson, J. & Macaux, E. W. Temporal discounting when the choice is between two delayed rewards. *J. Exp. Psychol. Learn. Mem. Cogn.* **31**, 1121–1133 (2005).
59. Mazur, J. E. An adjusting procedure for studying delayed reinforcement. **1987**, 55–73 (1987).
60. Sutton & Barto. *Reinforcement Learning: An Introduction*. (MIT Press, 1998).

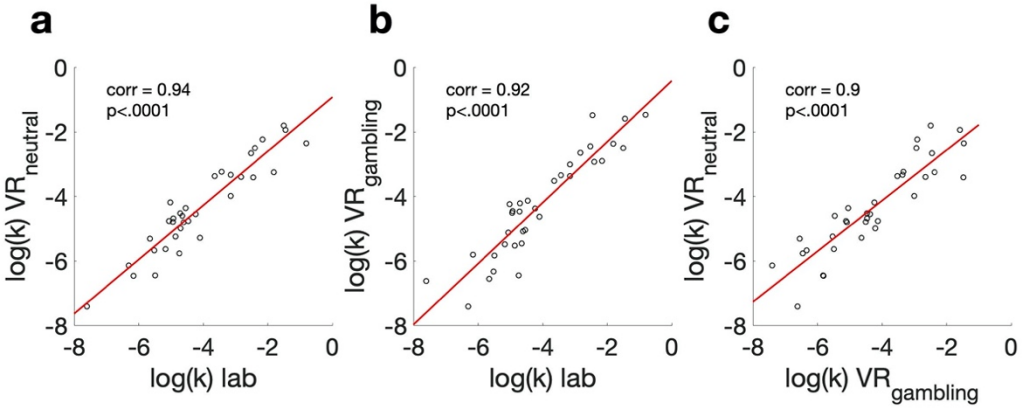
61. R Core Team. R: A Language and Environment for Statistical Computing. (2013).
62. Plummer, M. A Program for analysis of Bayesian graphical models. *Work. Pap.* (2003).
63. Wabersich, D. & Vandekerckhove, J. Extending JAGS: A tutorial on adding custom distributions to JAGS (with a diffusion model example). *Behav. Res. Methods* **46**, 15–28 (2014).
64. Gelman, A. & Rubin, D. B. Inference from Iterative Simulation Using Multiple Sequences. *Stat. Sci. 7* **4**, 457–472 (1992).
65. Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Van Der Linde, A. Bayesian measures of model complexity and fit. *J.R. Stat. Soc. Ser. B Stat. Methodol.* **64**, 583–639 (2002).
66. Marsman, M. & Wagenmakers, E. J. Three Insights from a Bayesian Interpretation of the One-Sided P Value. *Educ. Psychol. Meas.* **77**, 529–539 (2017).
67. Beard, E., Dienes, Z., Muirhead, C. & West, R. Using Bayes factors for testing hypotheses about intervention effectiveness in addictions research. *Addiction* **111**, 2230–2247 (2016).
68. Koo, T. K. & Li, M. Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **15**, 155–163 (2016).
69. Benedek, M. & Kaernbach, C. A continuous measure of phasic electrodermal activity. *J. Neurosci. Methods* **190**, 80–91 (2010).
70. Kerby, D. The Simple Difference Formula: An Approach to Teaching Nonparametric Correlation. *Compr. Psychol.* **3**, (2014).
71. Rosenthal, R., Cooper, H. & Hedges, L. Parametric measures of effect size. *Handb. Res. Synth.* **621(2)**, 231–244 (1994).
72. Duszkiewicz, A. J., McNamara, C. G., Takeuchi, T. & Genzel, L. Novelty and Dopaminergic Modulation of Memory Persistence: A Tale of Two Systems. *Trends Neurosci.* **42**, 102–114 (2019).
73. D'Amour-Horvat, V. & Leyton, M. Impulsive actions and choices in laboratory animals and humans: Effects of high vs. low dopamine states produced by systemic treatments given to neurologically intact subjects. *Front. Behav. Neurosci.* **8**, 1–20 (2014).
74. Juliano, L. M. & Brandon, T. H. Reactivity to instructed smoking availability and environmental cues: evidence with urge and reaction time. *Exp. Clin. Psychopharmacol.* **6(1)**, (1998).
75. Sayette, M. A. *et al.* The effects of cue exposure on reaction time in male alcoholics. *J. Stud. Alcohol* 629–633 (1994).
76. Vollstädt-Klein, S. *et al.* Validating incentive salience with functional magnetic resonance imaging: Association between mesolimbic cue reactivity and attentional bias in alcohol-dependent patients. *Addict. Biol.* **17**, 807–816 (2012).
77. Rubo, M. & Gamer, M. Stronger reactivity to social gaze in virtual reality compared to a classical laboratory environment. *Br. J. Psychol.* 1–14 (2020) doi:10.1111/bjop.12453.
78. Hedge, C., Bompas, A. & Sumner, P. Task reliability considerations in computational psychiatry. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* (2020).
79. Bouchard, S., Loranger, C., Giroux, I., Jacques, C. & Robillard, G. Using Virtual Reality to Provide a Naturalistic Setting for the Treatment of Pathological Gambling. in *The Thousand Faces of Virtual Reality* (ed. Sik-Lanyi, C.) (InTech, 2014). doi:10.5772/59240.
80. Giroux, I. *et al.* Gambling exposure in virtual reality and modification of urge to gamble. *Cyberpsychol. Behav. Soc. Netw.* **16**, 224–231 (2013).
81. Gamito, P. *et al.* Eliciting nicotine craving with virtual smoking cues. *Cyberpsychology, Behav. Soc. Netw.* **17**, 556–561 (2014).

82. Clark, L., Boileau, I. & Zack, M. Neuroimaging of reward mechanisms in Gambling disorder: an integrative review. *Mol. Psychiatry* **24**, 674–693 (2019).
83. Havermans, R. C., Mulkens, S., Nederkoorn, C. & Jansen, A. The efficacy of cue exposure with response prevention in extinguishing drug and alcohol cue reactivity. ,. *Behav. Interv. Pract. Resid. Community-Based Clin. Programs* **22(2)**, 121–135 (2007).
84. Bordnick, P. S., Traylor, A. C., Graap, K. M., Copp, H. L. & Brooks, J. Virtual reality cue reactivity assessment: A case study in a teen smoker. *Appl. Psychophysiol. Biofeedback* **30**, 187–193 (2005).
85. Choi, J. S. *et al.* The effect of repeated virtual nicotine cue exposure therapy on the psychophysiological responses: A preliminary study. *Psychiatry Investig.* **8**, 155–160 (2011).
86. Anokhin, A. P., Golosheykin, S., Grant, J. D. & Heath, A. C. Heritability of delay discounting in adolescence: A longitudinal twin study. *Behav. Genet.* **41**, 175–183 (2011).

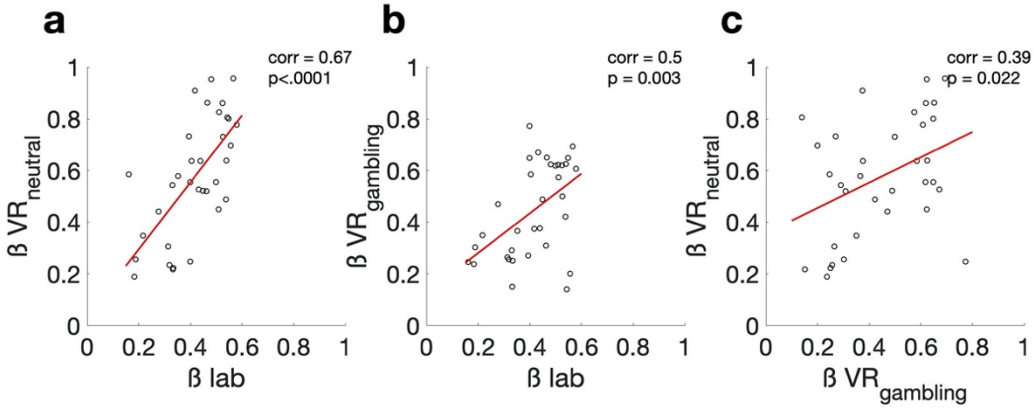
Supplementary Materials



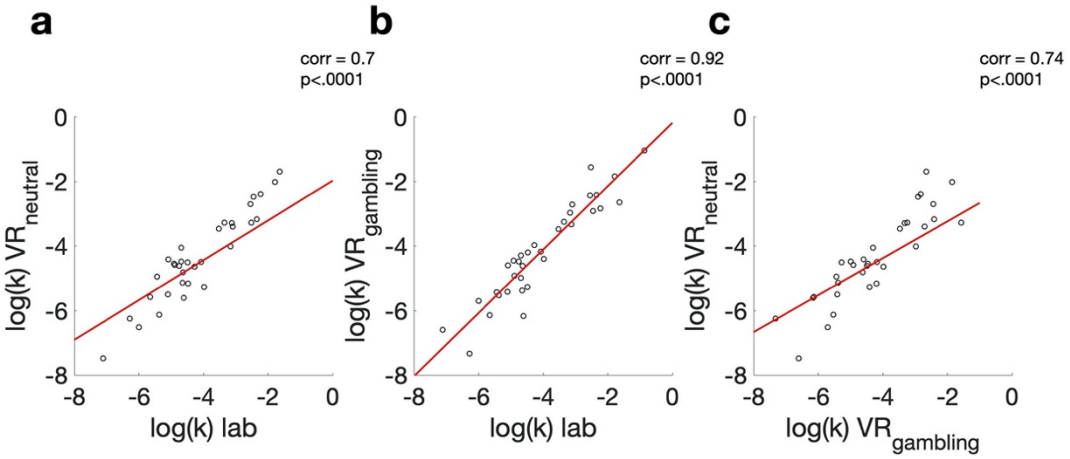
Supplementary Figure S1. Scatterplots of the individual participants AUC values. a) lab vs $VR_{neutral}$ b) lab vs $VR_{gambling}$ c) $VR_{gambling}$ vs $VR_{neutral}$..



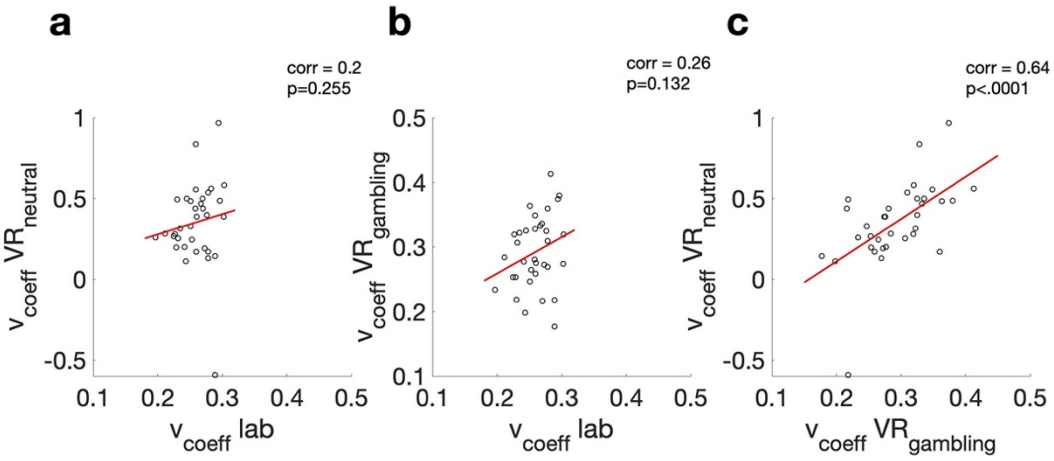
Supplementary Figure S2. Scatterplots of the mean of the individual participants parameter posterior distributions for the *log(k)* parameter of the hyperbolic discounting model with softmax choice rule (see equation X). a) lab vs $VR_{neutral}$ b) lab vs $VR_{gambling}$ c) $VR_{gambling}$ vs $VR_{neutral}$.



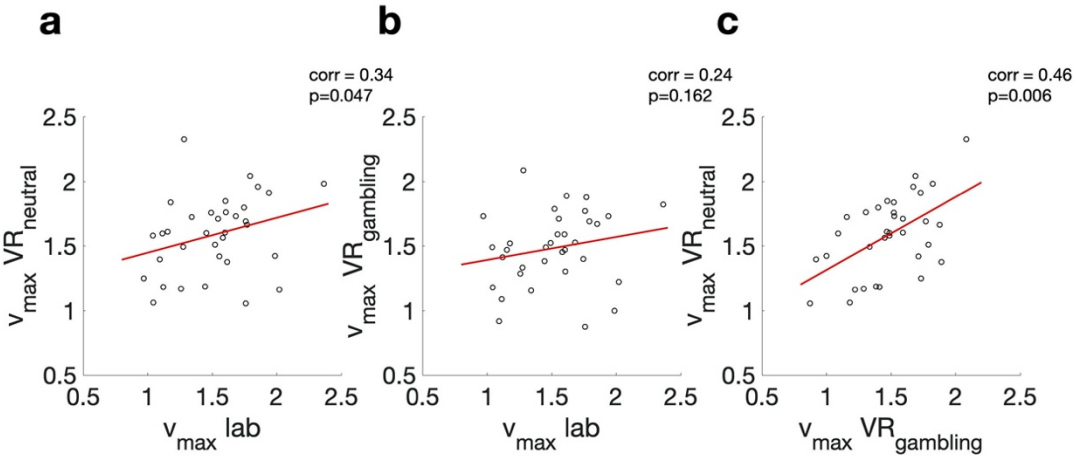
Supplementary Figure S3. Scatterplots of the mean of the individual participants parameter posterior distributions for the softmax ß parameter of the hyperbolic discounting model. a) lab vs $VR_{neutral}$ b) lab vs $VR_{gambling}$ c) $VR_{gambling}$ vs $VR_{neutral}$.



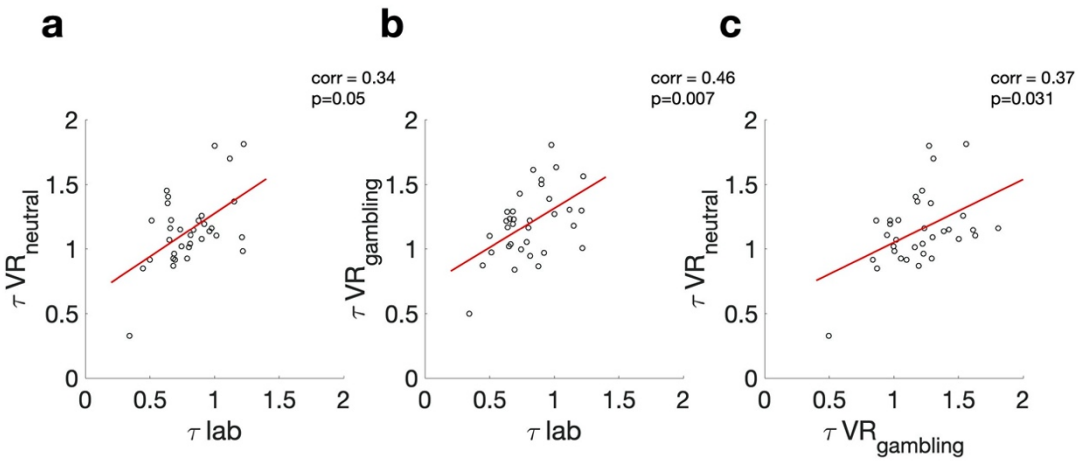
Supplementary Figure S4. Scatterplots of the mean of the individual participants parameter posterior distributions for the *log(k)* parameter of the DDM_S temporal discounting model. a) lab vs $VR_{neutral}$ b) lab vs $VR_{gambling}$ c) $VR_{gambling}$ vs $VR_{neutral}$.



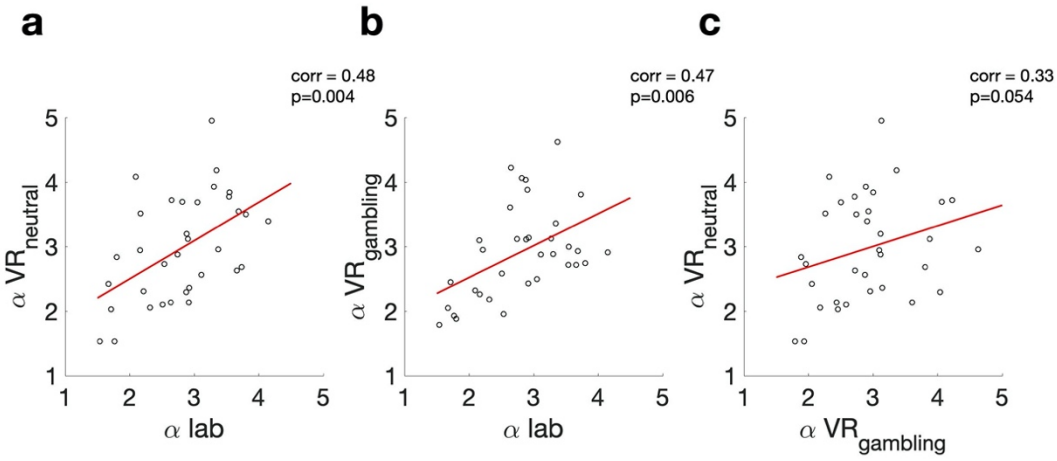
Supplementary Figure S5. Scatterplots of the mean of the individual participants parameter posterior distributions for the v_{coeff} parameter of the DDM_S model. a) lab vs $VR_{neutral}$ b) lab vs $VR_{gambling}$ c) $VR_{gambling}$ vs $VR_{neutral}$.



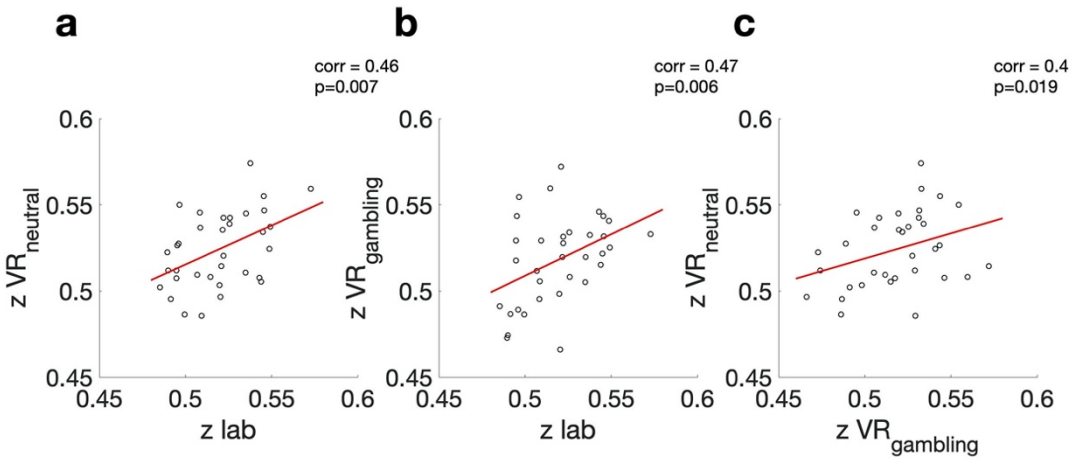
Supplementary Figure S6. Scatterplots of the mean of the individual participants parameter posterior distributions for the v_{max} parameter of the DDM_S model. a) lab vs $VR_{neutral}$ b) lab vs $VR_{gambling}$ c) $VR_{gambling}$ vs $VR_{neutral}$.



Supplementary Figure S7. Scatterplots of the mean of the individual participants parameter posterior distributions for the tau parameter of the DDM_S model. a) lab vs $VR_{neutral}$ b) lab vs $VR_{gambling}$ c) $VR_{gambling}$ vs $VR_{neutral}$.



Supplementary Figure S8. Scatterplots of the mean of the individual participants parameter posterior distributions for the alpha parameter of the DDM_S model. a) lab vs $VR_{neutral}$ b) lab vs $VR_{gambling}$ c) $VR_{gambling}$ vs $VR_{neutral}$.



Supplementary Figure S9. Scatterplots of the mean of the individual participants parameter posterior distributions for the beta parameter of the DDM_S model. a) lab vs $VR_{neutral}$ b) lab vs $VR_{gambling}$ c) $VR_{gambling}$ vs $VR_{neutral}$.

Supplementary Table S1. Means and 95% HDIs of the posteriors of the parameters from the DDM_s model.

	Log(k)			vcoeff			vmax			τ			α			z		
Session	**Mean**	**HDI**		**Mean**	**HDI**		**Mean**	**HDI**		**Mean**	**HDI**		**Mean**	**HDI**		**Mean**	**HDI**	
Lab	-4.051	-4.57	-3.54	.26	.22	.306	1.516	1.362	1.677	.806	.724	.889	2.806	2.545	3.073	.519	.502	.537
$VR_{neutral}$	-4.474	-5.012	-3.95	.352	.238	.491	1.588	1.428	1.756	1.145	1.041	1.249	2.978	2.671	3.293	0.524	.506	.542
$VR_{gambling}$	-4.16	-4.698	-3.626	.292	.244	.351	1.484	1.342	1.636	1.197	1.099	1.294	2.921	2.648	3.199	.518	.499	.537
Lab-$VR_{neutral}$	.422	-.309	1.16	-.093	-.237	.0305	-.072	-.3	.157	-.339	-.471	-.207	-.173	-.585	0.236	-.005	-.03	.02
Lab-$VR_{gambling}$	.108	-.628	.849	-.033	-.104	.034	.032	-.181	.247	-.391	-.516	-.265	-.115	-.497	0.266	.001	-.025	.027
$VR_{gambling}$-$VR_{neutral}$	.314	-.434	1.072	-.06	-.208	.069	-.104	-.324	.115	.052	-.091	.194	-0.058	-.473	.357	-.006	-.032	.02

Supplementary Table S2. Summary of the results of the ICC analysis of the DDM_L parameters.

Parameter	ICC	p	Lower bound	Upper Bound
log(k)	.92	<.001	.88	.95
$\boldsymbol{v}_{\mathbf{coeff}}$	.65	<.001	.5	.77
$\boldsymbol{\tau}$	.15	.067	-.014	.35
$\boldsymbol{\alpha}$	.36	<.001	.19	.55
$\boldsymbol{z}$	.60	<.001	.45	.74