Supplementary Information

Supplementary Methods

fMRI Data Collection and Pre-Processing

A Phillips 3T MRI scanner was used to obtain anatomical and functional images of participants completing the task. High-resolution structural images were collected using a multi-echo MP-RAGE sequence with the following parameters: TR/TE/TI = 2300/2.74/900 ms, flip angle = 8°, FOV = 256x256 mm, Slab thickness = 176, Voxel size =1x1x1 mm, Number of echos = 4, Pixel bandwidth =650 Hz, Total scan time = 6 min. These structural scans were used for alignmemnt of images. fMRI scans were obtained by setting repetition time (TR) to 2000 ms, echo time (TE) to 25ms, voxel size to 3.4 mm × 3.4 mm × 4.0 mm, flip angle to 90°, and slice number to 37. The functional scans were preprocessed using the statistical parametric mapping software package (SPM12, Wellcome Department of Imaging Neuroscience;) and by implementing time correction, co-registration, and normalization with resampled voxel size of 2mm × 2mm × 2mm and smoothing with an 8mm Gaussian kernel.

Model- Free Reinforcement Learning

In our model-free learning, reward prediction error (RPE) θ_i, incorporates current utility and the discounted value of an observed next offer:

$$\theta_i = U(r_i, f_i) + \gamma \max\big(Q(s_{i+1}, 1), Q(s_{i+1}, 0)\big) - Q_i(s_i, a_i)$$

Here, the value of an observed next offer is calculated as the difference between the actual reward of either accepting $Q(s_{i+1}, 1)$ or rejecting $Q(s_{i+1}, 0)$ the offer and the predicted reward at the *i*th trial given a certain action $Q_i(s_i, a_i)$. The value of an observed next offer assumes a deterministic greedy choice at each trial, and thus selects the maximum value between accepting or rejecting.

The subsequent value of an offer given a certain action is in turn updated with the RPE and a learning rate of ζ ($0 \leq \zeta \leq 1$).

$$Q_{i+1}(s_i, a_i) = Q_i(s_i, a_i) + \zeta\theta_i$$

Internal valuation ΔQ_i, or the difference in the value of accepting and rejecting an offer, is similarly inputted in a softmax function as shown above for FT.

$\Delta Q_i = Q(s_i, 1) - Q(s_i, 0)1.$

Risky Decision-making Task and Risk Aversion Model (online sample)

A part of a larger study, online participants (n= 219) completed a risky decision-making task [21]. Risk aversion parameters were extracted from this task to evaluate the extent to which risk aversion contributes to differences in total and medium rejection rates between smokers and non-smokers (see below for details on the model). In the task, participants made 30 choices between two gamble options where one option always had a larger difference between high- and low- potential payoffs (i.e., riskier). Each pair of gamble options had the same high- and low- payoff probabilities. We first generated eight unique lottery payoff menus with eight paired gambles (from 30% to 100%) and selected 30 unique pairs out of all possible pairs to reduce the number of choices without sacrificing the task sensitivity in capturing individuals' risk preferences (see **Table S10** for the full gamble pairs). The positions of the safe and risky gambles were randomly swapped and the gamble pairs were presented in a pseudorandom sequence.

Per Expected Utility Theory, we used a power utility function, where its concavity, rho, captures an individual's risk preference:

$$U(x) = x^{rho}, \text{ where } \begin{array}{l} risk\ aversion: rho < 1 \\ risk\ neutral: rho = 1 \\ risk\ seeking: rho > 1 \end{array}$$

Individual-level risk preference parameter, as well as additional value sensitivity parameter in softmax decision rule, was estimated from individuals' choices using maximum likelihood

fitting. The parameter estimation was conducted with custom MATLAB scripts and the fminsearch function in MATLAB with multiple initial values.

Consequently, we conducted additional analyses to examine the extent to which risk aversion contributes to differences in delta values between smokers and non-smokers. We constructed general linear models (GLMs) with risk aversion parameter values as an independent variable to predict rejection rates (**Tables S8- S9**).

Supplementary Results

Order Effects

We also examined the possibility of order effects on the key computational parameter of interest, delta, which represented the mentally estimated controllability, across both fMRI and online studies. We found no evidence of condition order having an impact on delta:

- **Order effect in fMRI smokers:** Delta scores for participants that started with the controllable (n=14, 1.536 ± 0.595) vs. uncontrollable condition (n= 11, 1.218 ± 0.710) are non-significant with a non-parametric bootstrapping p= 0.2011.

- **Order effects in fMRI non-smokers:** Delta scores for participants that started with the controllable (n= 10, 0.992 ± 1.239) vs. uncontrollable condition (n= 7, -0.562 ± 1.549, are non-significant with a non-parametric bootstrapping p= 0.117.

- **Order effect in all fMRI subjects (smokers & non-smokers combined):** Delta scores for participants that started with the controllable vs. uncontrollable condition are non-significant with a non-parametric bootstrapping p= 0.1532.

- **Order effect online smokers:** Delta scores for participants that started with the controllable (n= 40, 1.03±1.39) vs. uncontrollable condition (n= 32, 1.24 ± 0.59) are non-significant with a non-parametric bootstrapping p= 0.364.

- **Order effect in online non-smokers:** Delta scores for participants that started with the controllable (n=75, 1.29±1.05) vs. uncontrollable condition (n=72, 1.41 ± 0.33) are non-significant with a bootstrapping non-parametric p= 0.4331.

- **Order effects in all online subjects (smokers & non-smokers combined):** Delta scores for participants that started with the controllable (n= 115, 1.20±1.17) vs. uncontrollable condition (n= 104, 1.36 ± 0.41) are non-significant with a non-parametric bootstrapping p= 0.256.

Supplementary Tables

Table S1. Demographic and clinical variables of participants in the fMRI study. Statistical tests show no significant difference between these variables. Education is measured on a scale from 1-6 (1= Completed Middle School, 2= High School diploma or equivalent, 3= Some college, 4= Bachelor's degree, 5= Master's degree, 6= Ph.D.) A 2-sample t-test is used to match age, chi-square test to match sex and handedness, and a Mann-Whitney U test to match education and income. Additionally, race, ethnicity, and daily cigarette consumption reported by nicotine users is included in the last rows.

	Smokers (n=17)	**Non-smokers** (n=25)	**Statistical test**
Sex			
Male/ Female	14/3	14/11	$\chi 2$= 3.16
			p= 0.75
Handedness			
Right/Left	15/2	24/1	$\chi 2$= 0.92
			p= 0.34
Education			
Mean (SD)	2.88 (0.70)	3.36 (1.25)	z= 1.755
			p= 0.078
Age			
Mean (SD)	36.88(10.44)	31.16(11.08)	t= 1.42
			p= 0.16
Daily Cigarette Consumption			
Mean (SD)	18.6(6.92)	N/A	N/A

Table S2. Demographic and clinical variables of non-smoker and smoker participants for an independent larger online sample. 147 healthy participants were matched to 72 nicotine users based on sex, handedness, education, and age. Statistical tests show no significant difference between these variables. Education is measured on a scale from 1-6 (1= Completed Middle School, 2= High School diploma or equivalent, 3= Some college, 4= Bachelor's degree, 5= Master's degree, 6= Ph.D.) Annual income level is reported as 1= less than $10,000; 2=$10,000- $19,999; 3= $20,000-$29,999; 4= $30,000-39,999; 5= $40,000-$49,999; 6= $50,000-$59,999; 7= $60,000-$69,999; 8= $70,000-$79,999; 9= $80,000-$89,999; 10= $90,000-$99,999; 11= $100,000-$149,999; 12=more than $150,000. A 2-sample t-test is used to match age, chi-square test to match sex and handedness, and a Mann-Whitney U test to match education and annual income. Additionally, race, ethnicity, daily cigarette consumption, and percentage of craving reported by nicotine users is included in the last rows.

	Smokers (n=72)	**Non-smokers** (n=147)	**Statistical test**
Sex			
Male/ Female/Other	41/31	93/54	$\chi2$= 0.813 p= 0.367
Handedness			
Right/Left/Ambidextrous	68/4	129/7/11	$\chi2$= 0.523 p= 0.768
Education			
Mean (SD)	4.58 (1.01)	4.38(0.67)	z= 0.757 p= 0.447
Annual Income			
Mean (SD)	6.43(8.98)	6.74(11.29)	z= 1.137 p= 0.254
Age			
Mean (SD)	38.79(12.27)	35.81(13.51)	t= -1.581 p= 0.058
Race			
Asian/Black/White/Multiracial/Other	4/2/4/59/3	12/12/7/102/11	N/A
Ethnicity			
Hispanic or Latino/ Other	8/64	12/135	N/A
Daily Cigarette Consumption			
Mean (SD)	9.34(7.59)	N/A	N/A
Craving Scores			
Mean (SD)	64.51(26.90)	N/A	N/A

Table S3. Smokers and non-smokers implement a model of forward-thinking (FT) with a 2-step planning horizon The forward thinking (FT) models best accounted for non-smokers' choices compared to the model-free learning. Specifically, the 2-step FT model generated the lowest total Deviance Information Criterion (DIC) score for smokers, non-smokers and for both groups combined, and was thus used for further analyses. Behavioral data was fitted beyond the 2-step model (3-step and 4-step FT models), but not all parameters were recoverable for these models (see Table S4-S5).

Deviation Information Criteria

	Model Free	0-Step	1-Step	2-Step
Non-smokers (n=25)	604.14	353.43	324.65	274.16
Smokers (n=17)	284.77	138.06	133.26	119.13
Total	**888.91**	**491.49**	**457.91**	**393.29**

Table S4. Parameter recovery for smokers' data in the controllable condition. Parameters were fully recoverable for the 2-step forward model, which was not the case for the other models listed. The 2-step model for the smoker cohort was selected for further analysis.

		Inverse Temperature β	Sensitivity to Norm Violation α	Initial Norm μ	Adaptation Rate ε	Estimated Controllability δ
Model Free	r=	0.332036	0.142153	0.314758	-0.38634	
	p=	0.192897	0.586273	0.218508	0.125579	
0-Step	r=	0.866509	0.372229	0.238859	0.482958	
	p=	6.83E-06	0.141205	0.355857	0.04956	
1-Step	r=	0.313313	0.906318	0.775207	0.549993	0.803568
	p=	0.220744	5.41E-07	0.000257	0.022175	0.000102
2-Step	r=	0.969942	0.601287	0.664848	0.87666	0.992108
	p=	1.29E-10	0.010676	0.003593	3.89E-06	6.07E-15
3-Step	r=	0.691575	0.242479	0.753357	-0.11139	0.906134
	p=	0.002103	0.348391	0.00048	0.670372	5.48E-07
4-Step	r=	0.742083	-0.26433	0.710864	0.245351	0.841218
	p=	0.000648	0.305263	0.00138	0.342534	2.32E-05

Table S5. Parameter recovery for non-smokers' data in the controllable condition.
Parameters were fully recoverable for the 2-step forward thinking model, but not for any other model listed. The 2-step FT model for the non-smoker cohort was selected for further analysis.

		Inverse Temperature β	Sensitivity to Norm Violation α	Initial Norm μ	Adaptation Rate ε	Estimated Controllability δ
Model Free	r=	0.29830485	0.146045138	0.34691062	0.00478418	
	p=	0.14751065	0.486059444	0.08932368	0.98189253	
0-Step	r=	0.50784388	0.395816192	0.35446278	0.22892989	
	p=	0.00955184	0.050160161	0.08211607	0.27100684	
1-Step	r=	0.71934991	0.242538656	0.60372172	0.37453873	0.6805286
	p=	5.07E-05	0.242745551	0.00139639	0.06509328	0.00018145
2-Step	r=	0.77288398	0.763419019	0.85288249	0.77245707	0.79877273
	p=	5.94E-06	9.03E-06	6.13E-08	6.06E-06	1.70E-06
3-Step	r=	0.70667704	0.65055199	0.86935432	0.32990004	0.93765286
	p=	7.86E-05	0.000429989	1.70E-08	0.10728636	4.85E-12
4-Step	r=	0.67223779	0.551904816	0.75040641	0.37621427	0.84747376
	p=	0.00023258	0.004233153	1.56E-05	0.06380603	9.03E-08

Table S6. General linear model of negative mood (measured by Beck Depression Inventory (BDI) -II) on model estimated controllability (online sample). Delta ~ intercept + group (HC/smoker) + BDI + group x BDI. The overall regression was not statistically significant ($R^2 = 0.02$, $F(3, 215) = 1.47$, $p = 0.23$). The group predictor ($\beta = -0.23$, $p = 0.07$), negative mood predictor measured by z-scored BDI ($\beta = -0.01$, $p = 0.91$), and interaction ($\beta = 0.12$, $p = 0.34$) were not significant in this model. BDI was normalized across all subjects. The lack of significance of the group predictor here may not necessarily conflict with the significant two sample t test result presented in the main text, and the total variance explained is now shared between group and BDI, and their interactions.

Delta ~ Intercept + Group (HC/Smoker) + BDI + Group x BDI

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	1.350498	0.074156	18.212	<2e-16
Group	-0.231576	0.129331	-1.791	0.0748
BDI	-0.008616	0.074410	-0.116	0.9079
Group x BDI	0.124282	0.130086	0.955	0.3405
Residuals Standard Error	0.8991 on 215 degrees of freedom			
Multiple R-squared	0.02003, Adjusted R-squared: 0.006357			
F-statistic	1.465 on 3 and 215 DF, p-value= 0.2251			

Table S7. General linear model of impulsivity (measured by the Barratt Impulsivity Scale, BIS) on model estimated controllability (online sample). Delta ~ intercept + group (HC/smoker) + BIS total + group x BIS. The overall regression was not statistically significant ($R^2 = 0.03$, $F(3, 94) = 0.87$, $p = 0.46$). The group predictor ($\beta = 0.28$, $p = 0.26$) and impulsivity predictor measured by BIS ($\beta = -0.07$, $p = 0.49$) were not significant in this model. BIS was normalized across all subjects. The lack of significance of the group predictor here may not necessarily conflict with the significant two sample t test result presented in the main text, and the total variance explained is now shared between group and BIS, and their interactions.

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
Intercept	1.21743	0.11106	10.962	<2e-16
Group	0.28341	0.24809	1.142	0.256
BIS	-0.07487	0.10827	-0.692	0.491
Group x BIS	-0.26936	0.26708	-1.009	0.316
Residuals Standard Error	0.915 on 94 degrees of freedom			
Multiple R-squared	0.02699, Adjusted R-squared: -0.004064			
F-statistic	0.8691 on 3 and 94 DF, p-value= 0.46			

Table S8. General Linear Model of Risk Aversion on Rejection Rate (online sample): Rejection rate (Medium OFFERs) ~ intercept + group + risk aversion + group x risk aversion. There is no significant effect of risk aversion, group, or interaction on the rejection rate in the medium offer range. The lack of significance of the group predictor here may not necessarily conflict with the significant two sample t test result presented in the main text, and the total variance explained is now shared between group and risk aversion, and their interactions.

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0.65349	0.02694	24.259	<2e-16 ***
Group	-0.06252	0.04748	-1.317	0.189
Risk Aversion	0.04156	0.06075	0.684	0.495
Group x Risk Aversion	-0.10590	0.08597	-1.232	0.219
Residuals Standard Error	0.2797 on 211 degrees of freedom			
Multiple R-squared	0.03141, Adjusted R-squared: 0.01764			
F-statistic	2.281 on 3 and 211 DF, p-value= 0.08033			

Table S9. Payoffs and probabilities of paired gambles used in the risky decision-making task. Both safer and riskier options have a larger difference between high- and low- potential payoffs (i.e., riskier). Each pair of gamble options had the same high- and low- payoff probabilities ("probability of earning high payoff (%))". We generated eight unique lottery payoff menus with eight paired gambles (from 30% to 100%) and selected 30 unique pairs out of all possible pairs (indexed from 1 to 30).

	Safer gamble		Riskier gamble		
Gamble index	High payoff	Low payoff	High payoff	Low payoff	Probability of earning high payoff (%)
1	33.2	23.1	56.8	1.7	50
2	33.2	23.1	56.8	1.7	60
3	33.2	23.1	56.8	1.7	70
4	33.2	23.1	56.8	1.7	90
5	20.8	15.2	37.4	1.1	30
6	20.8	15.2	37.4	1.1	40
7	20.8	15.2	37.4	1.1	50
8	20.8	15.2	37.4	1.1	100
9	19.6	18.0	38.6	0.9	40
10	19.6	18.0	38.6	0.9	70
11	19.6	18.0	38.6	0.9	80
12	19.6	18.0	38.6	0.9	90
13	25.5	24.9	50.8	1.3	30
14	25.5	24.9	50.8	1.3	50
15	25.5	24.9	50.8	1.3	70
16	25.5	24.9	50.8	1.3	90

17	24.4	23.0	51.1	1.2	40
18	24.4	23.0	51.1	1.2	50
19	24.4	23.0	51.1	1.2	80
20	24.4	23.0	51.1	1.2	90
21	26.7	21.4	51.6	1.4	40
22	26.7	21.4	51.6	1.4	60
23	26.7	21.4	51.6	1.4	70
24	26.7	21.4	51.6	1.4	100
25	26.5	25.2	55.3	1.3	30
26	26.5	25.2	55.3	1.3	60
27	26.5	25.2	55.3	1.3	80
28	28.3	26.6	55.6	1.6	30
29	28.3	26.6	55.6	1.6	60
30	28.3	26.6	55.6	1.6	80

Supplementary Figure

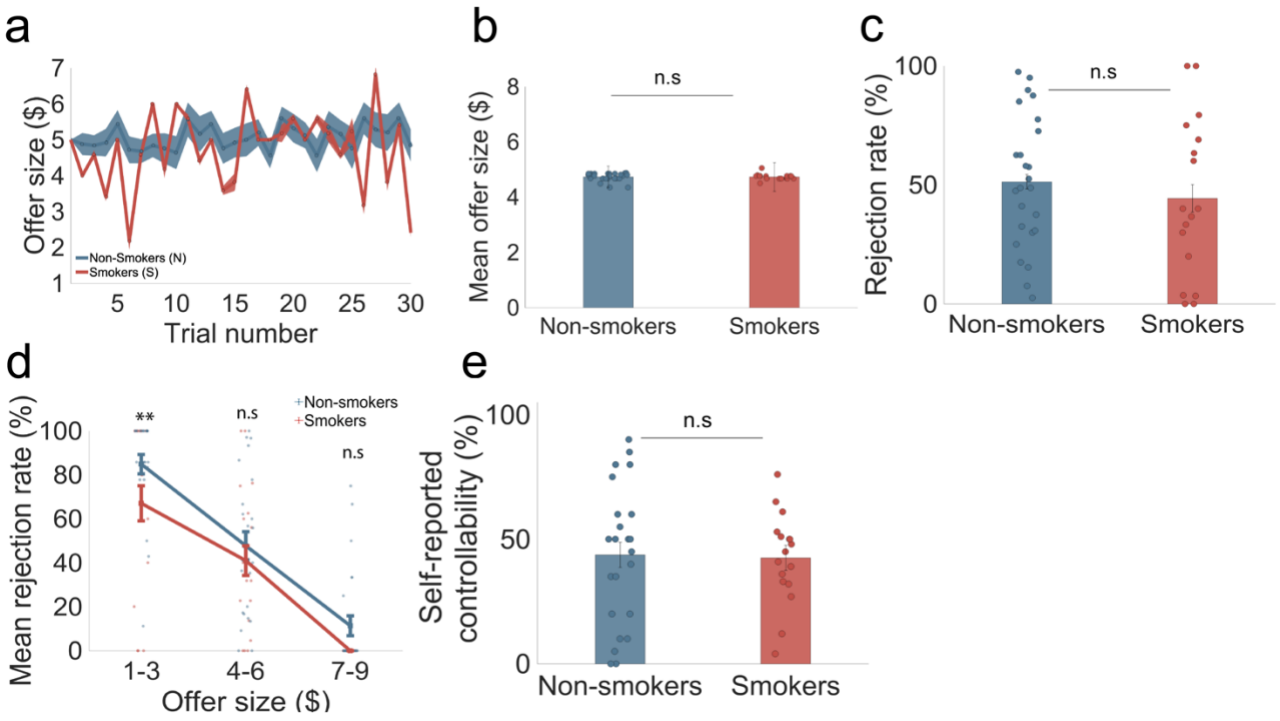

Figure S1. Behavioral results from the "uncontrollable" condition (i.e. standard ultimatum game). In the uncontrollable condition of the task, offers were randomly drawn from a Gaussian distribution ($\mu = 5$, $\sigma = 1.2$, rounded to the nearest integer, max = 8, min = 2). **a)** Nicotine smokers' and non-smokers' offer sizes do not follow a clear pattern as trial number increases. **b)** A two-sampled t-test reveals that individual mean offer sizes are not significantly different between smokers ($4.73±0.11) and non-smokers ($5.13±0.15), where t(40)= 0.076, p= 0.47. **c)** Percentage of total rejection is not significant different between smokers (44.32%±0.33) and non-smokers (51.29%±0.29). **d)** When rejection rates were divided and categorized by low ($1-$3), medium ($4-$6) and high ($7-$9) offers, a two-sample t-test revealed that smokers had a significantly lower rejection rate for low offer sizes (67.06%± 0.40) compared to non-smokers (84.86%±0.18), where t(40)= 1.86, p=0.036. **e)** Perceived controllability rated on a scale of 1% to 100% after each condition of the task was not significantly different for smokers (42.53%±18.05) compared to non-smokers (43.70%±29.38), where t(40)=0.15, p= 0.441.