nature portfolio

Peer Review File

Aberrant neural computation of social controllability in nicotine-dependent humans

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

This is an elegant computational study using a modified social valuation task to assess behavioral and neural correlates of social-decision making and perceived controllability among smokers (n=17) and non-smokers (n=25). Behavioral (but not fMRI) effects were then also tested in an online convenience sample of 147 individuals. Behavioral computational findings from the in-person and on-line studies largely aligned (although temporal trajectories of were different between the two samples, with smokers in the in-person study reducing offers over time and smokers in the online study increasing offers over time). Neuroimaging analyses in the smaller in-person sample also yielded interesting findings involving reduced tracking of forward projected choice and impaired social prediction errors among smokers. The modeling and analyses are all well done. My primary concern is the overall framing of the study and the absence of any attempt to explore other factors beyond smoking that might account for these results. Specifically, low-level depressive symptoms, including anhedonia, are common among smokers. While individuals were excluded for psychiatric diagnoses, it is very possible that between-group differences in dimensional features such as negative affect, impulsivity, and myriad other factors might in fact account for the observed differences between smokers and non-smokers. Thus, presenting the results as specifically relating to social controllability in smokers (and addiction more generally) is problematic. It would be very surprising if other factors did not differ between groups (consistent with the general phenotypic presentation of individuals with tobacco use), thus failure to explore how these factors might have related to study findings significantly diminishes the overall impact.

Reviewer #2 (Remarks to the Author):

In this paper, Caroline McLaughlin and colleagues analyze how smokers and non-smokers differ – behaviorally and neutrally - in how they play a social game (the Ultimatum Game) where participants have some degree of control over the future offers made by the putative partner. The use of different methods (neuroeconomics task, computational modeling, fMRI) is a strength of the paper, as well as the inclusion of both an in-person and an (online) replication sample for the behavioral/computer modeling part of the paper. I have, however, still a few questions/comments:
1. Were there group differences between the participants that started with the controllable vs. the uncontrollable condition?
2. When analyzing differences in mean offer sizes in the online replication sample, the p-value mentioned is a “bootstrapping p-value”.
a. Could you give the details of what this non-parametric/ bootstrapping method is, and what is the rationale of doing it instead of a simple unpaired t-test (like it seems to have been done for the in-person sample) or a Mann-Whitney U-test/Wilcoxon rank sum test?
b. Were the assumptions necessary to perform a t-test met/not met in either of the samples? Given the larger sample size, I would expect the assumptions of the t-test to be met more easily in this sample than in the in-person sample, although this may of course not be the case.
c. Are the other p-values shown for the online sample also “bootstrapped p-values”?

3. Pre-registration: I applaud the authors' decision to share the data and code for this study, as well as for preregistering studies. In terms of preregistrations, however, the linked preregistrations are more general broad study preregistrations, and do not seem to have any specific hypothesis related to this particular study. Namely, for example, for the in-person sample, even though it says that the beliefs about control are going to be analyzed, there are no specific hypotheses associated with it (e.g., that smokers would have higher/lower beliefs), and no mention that offers are going to be analyzed divided in small, medium, and high. There are also specific analyses mentioned in the preregistration that are not included in the paper (and data from 2 other tasks, but I assume they will be analyzed in different papers). Similarly, for the online sample, the hypothesis and goals do not even mention smoking. As in the previous sample, there are also no specific hypotheses related to this particular study. It is, of course, ok to do exploratory analysis, and they can yield interesting insights, but I do have a few suggestions:
a. In the methods, when you refer to the preregistrations, mention that these are general study/protocol registrations;
b. Mention in the text which analyses were preregistered and which were not (I think many were not, at least in the model-agnostic measures)
c. For future preregistrations, aim to have more specific (often direction) predictions – you do have a few, but those do not seem to be related to this study. Similarly, consider using the OSF template instead of the "as predicted" if you need more space, as the "as predicted" website has a relatively short word limit (you may have already figured this out, as you used the OSF template for the online sample);
d. After analyzing the in-person sample, you could have potentially preregistered the specific analysis you were going to do in the "replication sample", even if the replication sample was already collected, with specific hypotheses but before looking at the data (e.g., that you would see a lower controllability belief; lower rejection for medium offer sizes; …) – that would have made the results even stronger.
4. How much did participants believe that there was an actual other partner playing? Were the authors able to analyze this?

Minor
5. Was the total number of trials in the experiment 60/80 (i.e., 30/40 per condition and 2 conditions – one controllable, one uncontrollable)? Also, were participants told how many trials the game had and which trial number they were in?
6. Discussion: Given the small sample size of the in-person sample, I suggest changing the sentence "demonstrating that addiction is also linked to neural deficits in the midbrain during aberrant norm updating in complex social environments." (line 256) to "suggesting that addiction is also linked to …."
7. In the discussion, you mention that "Futures studies may investigate the relationship between deprivation level and task-based measures by systematically manipulating participants' abstinence." (line 278/279) – this was also one of the analyses preregistered; is there any reason why the analysis wasn't actually done in this paper?

Reviewer #3 (Remarks to the Author):

In this study the authors conduct an investigation of the behavior of current smokers versus non-smokers in a modified dictator game where their acceptance or rejection influences the offers that will be made to them in future trials. This task has been previously developed and published along with the development of a computational model that explains the probability of accepting or rejecting as a function of its projected value (forward thinking) which depends on the current utility of their choice plus a discounted utility of future splits, which is weighed by the agent's belief of the current choice influence. This belief is a free parameter estimated for each subject and is considered to represent "social controllability". The novelty of this study relies on the fact that the task and computational modeling is performed in a "clinical" population and that it is also used for an fMRI model-based investigation.

I generally find the results convincing and the motivation to explore social decision-making in the context of substance use disorders compelling. I do find that the paper could be improved both in clarity/readability and by addressing some issues that could take away from the findings.

Major points:
1. Did you control for socio-economic factors other than education? There are many reasons why one may believe to posses less agency or controllability. For example, perceived social status may have an impact. While it is likely that you don't have direct measures of subjective social status, you may be able to demonstrate that your groups are comparable in income, employment, reliance on social security, etc.
2. One alternative explanation to smokers having a lower inclination to reject medium size offers (which prevents them from getting higher value offers) may be that because the increase is probabilistic and thus rejecting may be perceived as risky. On the other hand, accepting more offers could be due to impatience where the received amount now is much more valuable than the potential larger amount in the future. While how the results of this paper align with temporal discounting is partially mentioned in the discussion, please expand on how the current modeling approach compares to alternative explanations for the behavior observed.
3. The discounting factor was kept fixed in the modeling of forward-thinking value. Please show the results of the modeling when the discounting factor is also estimated. Is there trade-off between this parameter and your controllability parameter? Are there group differences in the discounting factor?
4. Please clarify whether the monetary rewards in the task were hypothetical or real and how payment was implemented.
5. Were participants instructed to consider the impact of their choices on subsequent offers? When was the question about their perceived controllability delivered relative to the task?
Minor points:
1. How many subjects performed the task in the scanner in each group? Please provide more detail about the fMRI task and how it differed from the behavioral task.
2. The paper seems to have been written by multiple different authors with divergent styles. The introduction and results sections lack clarity, cohesiveness, care for style, and is replete with typos,

in comparison to the much better-written discussion. I recommend that the authors re-read the manuscript carefully and improve the writing.
3. Was this a pre-registered study? Please comment on possible limitations if not.

Rebuttal Letter

Reviewer #1 (Remarks to the Author):

This is an elegant computational study using a modified social valuation task to assess behavioral and neural correlates of social-decision making and perceived controllability among smokers (n=17) and non-smokers (n=25). Behavioral (but not fMRI) effects were then also tested in an online convenience sample of 147 individuals. Behavioral computational findings from the in-person and online studies largely aligned (although temporal trajectories of were different between the two samples, with smokers in the in-person study reducing offers over time and smokers in the online study increasing offers over time). Neuroimaging analyses in the smaller in-person sample also yielded interesting findings involving reduced tracking of forward projected choice and impaired social prediction errors among smokers. The modeling and analyses are all well done. My primary concern is the overall framing of the study and the absence of any attempt to explore other factors beyond smoking that might account for these results. Specifically, low-level depressive symptoms, including anhedonia, are common among smokers. While individuals were excluded for psychiatric diagnoses, it is very possible that between-group differences in dimensional features such as negative affect, impulsivity, and myriad other factors might in fact account for the observed differences between smokers and non-smokers. Thus, presenting the results as specifically relating to social controllability in smokers (and addiction more generally) is problematic. It would be very surprising if other factors did not differ between groups (consistent with the general phenotypic presentation of individuals with tobacco use), thus failure to explore how these factors might have related to study findings significantly diminishes the overall impact.

Response: Thank you for your thorough review of our paper and your insightful comment. We agree that there may be additional variables that could influence the outcomes besides smoking consumption, and that it is important to explore those potential effects. Although the study excluded individuals with psychiatric conditions, we recognize the importance of controlling for negative affect, low-level depressive symptoms, and impulsivity factors (other factors that could be confounding the results such as risk aversion and socioeconomic status are also examined below). We have been able to leverage the large online dataset that included these other participant variables (mood symptoms, impulsivity, etc) to conduct multiple general linear models (GLMs) to investigate the degree to which our parameter of interest, delta, can be accounted for by either negative affect, measured by Beck Depression Inventory-II (BDI-II)), and impulsivity, measured by Barratt Impulsivity Scale (BIS). These measures are commonly used to assess negative mood and impulsivity in the literature and were collected at the time of online study using standardized questionnaires. Results demonstrate that neither negative mood nor impulsivity had a significant impact on delta in the current study. Please see below for stats and we have also added these additional results into the manuscript (see revised main text page 8 and **Table S6-S7**, in the revised manuscript).

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

There is no significant effect of group, negative mood, and group and mood interaction on model-estimated controllability. The significant intercept shows that on average healthy subjects have higher model-estimated controllability than smoker subjects. The overall regression was not statistically significant ($R^2 = 0.02$, $F(3, 215) = 1.47$, $p = 0.23$). The group predictor ($\beta = -0.79$, $p = 0.19$), negative mood predictor measured by BDI ($\beta = -0.001$, $p = 0.91$), and interaction ($\beta = 0.02$, $p = 0.34$) were not significant.

	Estimate Std.	Error	t value	Pr(>|t|)
Intercept	1.390122	0.350161	3.970	9.8e-05
Group	-0.786860	0.600108	-1.311	0.191
BDI	-0.01169	0.010100	-0.116	0.908
Group x BDI	0.016915	0.017691	0.956	0.340
Residuals Standard Error	0.9881 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
There is no significant effect of impulsivity measured by BIS, group, and group interaction on model-estimated controllability. Multiple linear regression was used to test if group and impulsivity significantly predicted our model-estimated controllability. The overall regression was not statistically significant ($R^2 = 0.03$, $F(3, 94) = 0.87$, $p = 0.46$). The group predictor ($\beta = 3.07$, $p = 0.29$) and impulsivity predictor measured by BIS ($\beta = -0.01$, $p = 0.49$) were not significant.

	Estimate Std.	Error	t value	Pr(>|t|)
Intercept	1.99312	1.10433	1.805	0.0743
Group	3.07410	2.88935	1.064	0.2901
BIS	-0.01176	0.01700	-0.692	0.4909
Group x BIS	-0.04230	0.04194	-1.009	0.3158
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			

Revised manuscript Results and Methods sections:

Results - lines 169-173: “Furthermore, this effect could not be attributed to mood symptoms (assessed by Beck Depression Inventory-II, **Table S6**) or impulsivity tendencies (assessed by Barratt Impulsivity Scale) commonly associated with addiction (**Table S7**).”

Methods - line 483-488: “We applied additional general linear modeling to further explore the effects of negative mood and impulsivity on the model estimated controllability (**Supplementary Information S6-S7**). Similar regression approach is applied to explore the effect of risk aversion on smokers and non-smokers’ choice behavior measured as rejection rates (**Supplementary Information S8-S9**). Analyses were performed using MATLAB (2020b)[64], R 4.3.1[65], and RStudio 2023.6.0.42[66]. MATLAB was used for data storage. R and RStudio were used for data curation and regression analysis using *lme4* package in R[67].”

Reviewer #2 (Remarks to the Author):

In this paper, Caroline McLaughlin and colleagues analyze how smokers and non-smokers differ – behaviorally and neutrally - in how they play a social game (the Ultimatum Game) where participants have some degree of control over the future offers made by the putative partner. The use of different methods (neuroeconomics task, computational modeling, fMRI) is a strength of the paper, as well as the inclusion of both an in-person and an (online) replication sample for the behavioral/computer modeling part of the paper. I have, however, still a few questions/ comments:

Response: Thank you for taking the time to review our manuscript and provide these highly constructive comments. Please see our point-by-point response below.

1. *Were there group differences between the participants that started with the controllable vs. uncontrollable condition?*

Response: We have previously examined the possibility of order effect in a published study using the same task in healthy controls; this analysis showed no order effect in our key computational parameters [1]. Additionally, to corroborate the absence of order effect in this task study, we have conducted additional analyses to compare delta scores among participants who started with either the controllable condition first or the uncontrollable condition first. We have added these additional analyses to **Supplementary Results**:

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.

** The parametric bootstrapping method was used in all these analyses for consistency and given the unbalanced online sample sizes (also see below).

We have added these additional analyses and results to the revised manuscript –

Main text: line 131. “There was no order effect on delta in either group (*P*s>0.1; see **Supplemental Results**), similar to our previous results[1].”

Main text: lines 169-170. “Similar to our in-person sample and previous study[1], we did not observe any order effect on the estimated controllability parameter (*P*s>0.2; see **Supplemental Results**).”

2. *When analyzing differences in mean offer sizes in the online replication sample, the p-value mentioned is a “bootstrapping p-value.”*
a. *Could you give the details of what this non-parametric/ bootstrapping method is, and what is the rationale of doing it instead of a simple unpaired t-test (like it seems to have been done for the in-person sample) or a Mann-Whitney U-test/Wilcoxon rank sum test?*

Response: For all online comparisons we used a bootstrapping method because: 1) of a highly unbalanced sample where n = 72 for online smokers and n = 147 for online non-smokers; 2)

behavioral measures and estimated parameters from the model are not normally distributed for the online sample, especially among non-smokers. For instance, the total rejection rate for healthy online participants (n=147) exhibits a Kolmogrov-Smirnov (K-S) test statistic (D) of 0.20015 where p is 0.0001, indicating strong evidence that rejection rates are not normally distributed across participants. Similarly, parameter values from the model also deviate from a normal distribution, as demonstrated by delta with a K-S test statistic (D) of 0.21002, $p<0.00001$. Healthy online participants have a delta skewness value of -2.117 and a kurtosis value of 5.002, while online smokers have a skewness delta value of -1.737 and a kurtosis value of 2.858. These differences in skewness and kurtosis suggest that the distributions are not similar in shape. Therefore, using a Mann-Whitney U-Test, which assumes similar distribution shapes, may additionally not yield valid results.

As such, we used a non-parametric bootstrapping method following previous work conducted with similar constraints[2-5] to assess the probability of observing a difference between two groups. The bootstrapping procedure was conducted with 10000 iterations as follows (e.g. the comparison between 72 smokers and 147 non-smokers): (i) 72 subjects were selected randomly as the surrogate smoker group, from the whole group of 219 online subjects including both smokers and non-smokers; (ii) 147 subjects were selected randomly as the surrogate non-smoker group from the whole group of 219 subjects; and (iii) the t-value of the difference between the two surrogate groups was calculated. After 10000 iterations, the distribution of the t-values was obtained. The observed t-value (e.g. between smoker and non-smoker groups) was then calculated and compared along the t distribution. If the probability of obtaining the observed t-value along the permutated distribution of t-value is <5% (one tailed), we considered the difference between the patient and control groups to be significant.

We have now included a short description in the Methods section explaining our rationale for using this statistical method.

Methods: lines 459-481.

"For statistical tests, we first examined if our key measures and parameters of interest met the criteria for standard parametric tests. We found that the total rejection rate for healthy online participants (n=147) exhibited a Kolmogrov-Smirnov (K-S) test statistic (D) of 0.20015 where p is 0.0001, indicating strong evidence that rejection rates were not normally distributed. Similarly, parameter values from the model also deviate from a normal distribution, as demonstrated by delta for all online subjects with a K-S test statistic (D) of 0.21002, $p<0.00001$. In light of this evidence as well as the highly unbalanced sample sizes of the online study, we used a non-parametric bootstrapping method following previous work conducted with similar constraints[5] to assess the probability of observing a difference between two groups. Therefore, a bootstrapping method was employed to compare all online parameters between smokers and non-smokers, while a 2-sample t-test was used to compare fMRI in-person parameters between the two groups.

The bootstrapping procedure was conducted with 10000 iterations as follows (e.g. the comparison between 72 smokers and 147 non-smokers): (i) 72 subjects were selected randomly as the surrogate smoker group, from the whole group of 219 online subjects including both

smokers and non-smokers; (ii) 147 subjects were selected randomly as the surrogate non-smoker group from the whole group of 219 subjects; and (iii) the t-value of the difference between the two surrogate groups was calculated. After 10000 iterations, the distribution of the t-values was obtained. The observed t-value (e.g. between smoker and non-smoker groups) was then calculated and compared along the t distribution. If the probability of obtaining the observed t-value along the permutated distribution of t-value is <5% (one tailed), we considered the difference between the patient and control groups to be significant."

b. *Were the assumptions necessary to perform a t-test met/not met in either of the samples? Given the larger sample size, I would expect the assumptions of the t-test to be met more easily in this sample than in the in-person sample, although this may of course not be the case.*

Response: This is correct (please see our last response to point 2a). The assumption of normality required for a t-test was not satisfied for the online sample. Various parameters, including the following, were found to deviate from a normal distribution based on a Kolmogrov-Smirnov test. Consequently, bootstrapping methods were employed for all analyses involving online subjects.

c. *Are the other p-values shown for the online sample also "bootstrapped p-values"?*

Response: Thank you for requesting this clarification. That is correct - the online sample results are all based on the same bootstrapping method. We have incorporated this clarification in the methods section pages 20-21.

Methods: lines 465-470. "In light of this evidence as well as the highly unbalanced sample sizes of the online study, we used a non-parametric bootstrapping method following previous work conducted with constraints to assess the probability of observing a difference between two groups. Therefore, a bootstrapping method was employed to compare all online parameters between smokers and non-smokers, while a 2-sample t-test was used to compare fMRI in-person parameters between the two groups."

3. *Pre-registration: I applaud the authors' decision to share the data and code for this study, as well as for pre-registering studies. In terms of pre-registrations, however, the linked pre-registrations are more general broad study pre-registrations, and do not seem to have any specific hypothesis related to this particular study. Namely, for example, for the in-person sample, even though it says that the beliefs about control are going to be analyzed, there are no specific hypotheses associated with it (e.g., that smokers would have higher/lower beliefs), and no mention that offers are going to be analyzed divided in small, medium, and high. There are also specific analyses mentioned in the preregistration that are not included in the paper (and data from 2 other tasks, but I assume they will be analyzed in different papers). Similarly, for the online sample, the hypothesis and goals do not even mention smoking. As in the previous sample, there are also no specific hypotheses related to this particular study. It is, of course, ok to do exploratory analysis, and they can yield interesting insights, but I do have a few suggestions:*

a. *In the methods, when you refer to the pre-registrations, mention that these are general study/protocol registrations.*

Response: Thank you for this suggestion. Indeed, both pre-registrations were submitted as a general study. We have now specified in the Methods section that the study was part of a more general pre-registered protocol (see below how we specified this in the manuscript).

Methods: lines 320-321. "The fMRI study was pre-registered as part of a larger fMRI study examining decision-making in smokers (https://osf.io/m9cws)."

Methods: lines 335-337. "This online study was pre-registered as a subcomponent of a larger longitudinal study investigating social decision-making (https://osf.io/8s5mu)."

b. Mention in the text which analyses were pre-registered and which were not (I think many were not, at least in the model-agnostic measures)

Response: Thank you for this suggestion. Group-based comparisons and computational modeling were included in the pre-registration, although we did not specify the statistical tests in the original pre-registrations (as the actual data distributions could be uncertain and might require post hoc choices of specific tests). We originally planned to conduct correlation with craving measures, yet this attempt failed as our carbon monoxide measure indicated that smokers did not follow our instruction of staying abstinent from the midnight before scanning. We have now added these details and specified in the manuscript which analyses were pre-registered.

Methods: lines 347-349. "However, CO levels measured on the day of scanning suggested that smokers were likely not able to stay abstinent as instructed. As such, we were unable to conduct planned analyses on how craving affected social controllability."

Methods: lines 453-455. "The computational models were also pre-registered as a subset of the project examining value-based decision making in nicotine addiction (https://osf.io/m9cws)."

Methods: lines 458-459: "Our analyses focused on group comparisons between smokers and non-smokers, per our pre-registrations."
Methods: lines 520-521. "The ROIs in the analyses were specified in the pre-registered study (https://osf.io/m9cws)."

c. For future pre-registrations, aim to have more specific (often direction) predictions – you do have a few, but those do not seem to be related to this study. Similarly, consider using the OSF template instead of the "as predicted" if you need more space, as the "as predicted" website has a relatively short word limit (you may have already figured this out, as you used the OSF template for the online sample).

Response: We greatly appreciate your advice and fully agree that it is crucial for researchers to incorporate preregistration into their standard laboratory practice, in order to enhance the rigor and reproducibility across our field. We sincerely apologize again for not being able to follow the best practice, despite our effort in pre-registering both the in-person and online studies, as these were amongst the first projects that were preregistered in the lab when the PI and her group lacked experience with preregistrations. However, we can confirm that newer studies in the lab

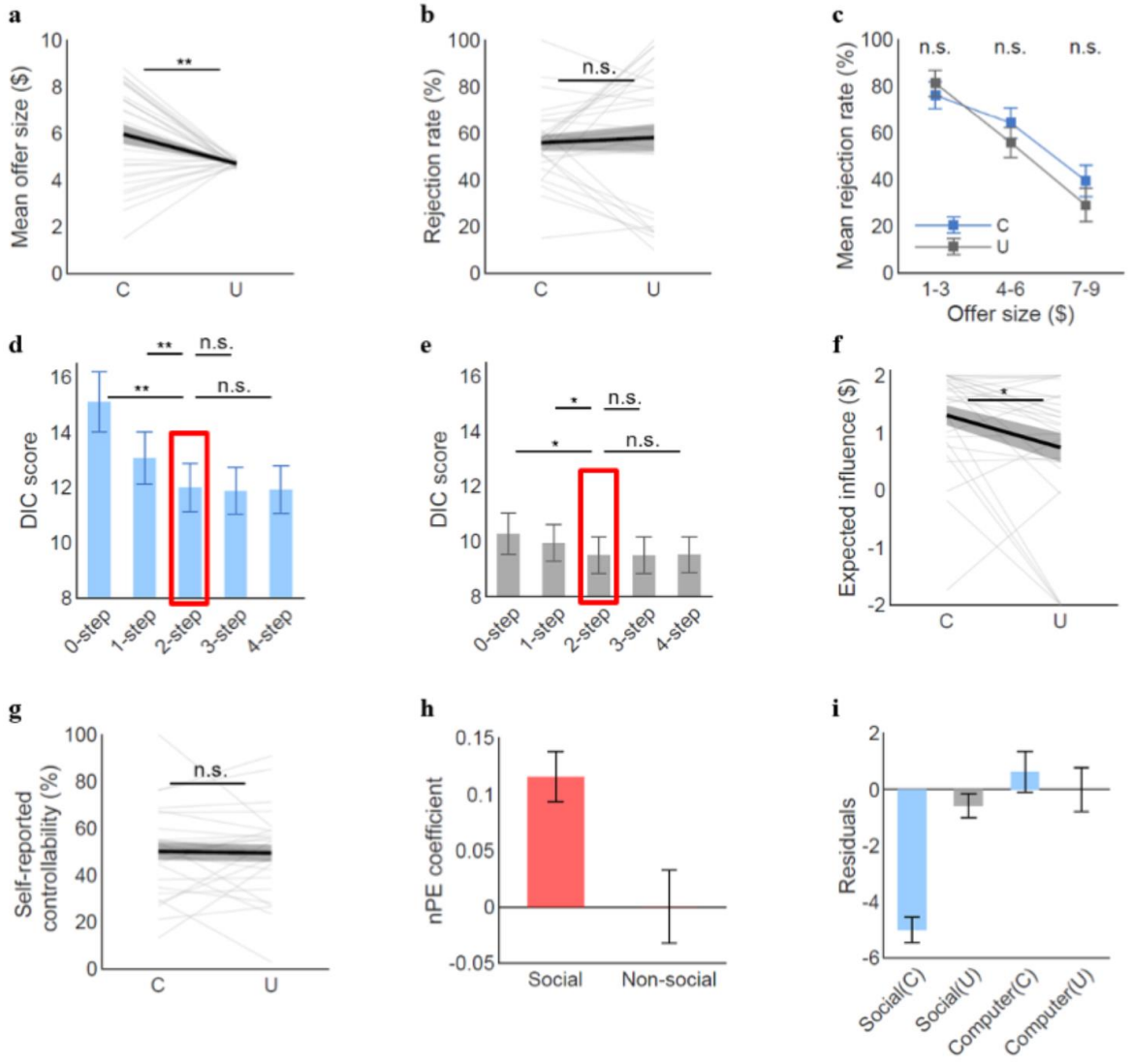
have been preregistered with the standard OSF template and we will follow your suggestions to continue to adhere with OSF templates much more closely for future pre-registrations.

d. After analyzing the in-person sample, you could have potentially preregistered the specific analysis you were going to do in the "replication sample", even if the replication sample was already collected, with specific hypotheses but before looking at the data (e.g., that you would see a lower controllability belief; lower rejection for medium offer sizes; …) – that would have made the results even stronger.

Response: We completely agree with your comment. Unfortunately, we did not finish analyzing the in-person data before pre-registering the replication online study (which took place during Covid in 2020). As such, we did not manage to list all analyses in the online pre-registration based on our in-person analysis and results, other than key components such as the computational model and group-based comparisons. We recognize the importance of following good practice of OSF and have since mandated all lab projects to strictly follow OSF guidelines in terms of how to specify analysis plans.

4. How much did participants believe that there was an actual other partner playing? Were the authors able to analyze this?

Response: Thank you for this insightful comment. We instructed the subjects that they were going to play with simulated partners, as deception was not part of the study design or IRB protocol. We agree that subjects might still develop beliefs about the "humanness" of the other player, which we unfortunately did not measure. However, in a previous study using a similar paradigm [1], we did include a non-social control condition where subjects were told that they were playing with a computer (Na et al., 2021). Notably, there are both similarities and differences in the results from the non-social vs. social condition of the task. Specifically, we found that subjects showed similar choice behaviors in both conditions, with the 2-step model once again emerging as the most favorable model in the non-social context as well (see figure below, panel a-f). There are also noticeable differences when subjects played with a computer, such as they reported a lower sense of control even in the Controllable condition (panel g) and that norm prediction errors did not affect how they felt (panel h-i).



Author Response Figure 1 from Na et al., 2021[1]: Behavioral results of a non-social controllability task. To investigate whether our results are specific to the social domain, we ran another batch of the task in which 27 out of the 48 original participants were re-contacted with a 14- to 24-month temporal gap and played the same game with the instruction of "playing with computer" instead of "playing with virtual human partners". Overall, we found choice patterns (**a-f**) similar to those in the social task while the subjective states (i.e., self-reported controllability (**g**) and the impact of the norm prediction error on the emotion ratings (**Table S1**) differed from the social task. (**a**) Similar to the results of the social task, offers ($mean_C$ = 6.0, $mean_U$ = 4.7, $t(26.23)$ =3.03, $P < 0.01$) were higher for the Controllable than the Uncontrollable. (**b**) Overall rejection rates ($mean_C$ = 55.9%, $mean_U$ = 58.1%, $t(40.76)$ = -0.33, P = 0.74) or (**c**) any of the binned rejection rates were not significantly different between the two conditions (paired t-test: *low ($1-3)* $mean_C$ = 76%, $mean_U$ = 81%, $t(12)$ = 1.54, P = 0.15, *middle ($4-6)* $mean_C$ = 64%, $mean_U$ = 56%, $t(26)$ = 1.74, P = 0.09, *high ($7-9)* $mean_C$ = 39%, $mean_U$ = 29%, $t(19)$ = 0.80, P = 0.44). (**d,e**) The DIC scores showed a similar pattern to the social task, with the elbow point at the 2-step FT model for both conditions. Paired t-tests confirmed that the 2-step

model's DIC scores were significantly lower than the 0-step model (Controllable: $t(26) = -3.16$, $P < 0.01$; Uncontrollable: $t(26) = -2.38$, $P < 0.05$) and the 1-step model (Controllable: $t(26) = -3.02$, $P < 0.01$; Uncontrollable: $t(26) = -2.31$, $P < 0.05$), whereas the DIC scores were not significantly different between the 2-step model and the 3-step model (Controllable: $t(26) = -1.23$, $P = 0.23$; Uncontrollable: $t(26) = 0.20$, $P = 0.84$) or the 4-step model (Controllable: $t(26) = 0.68$, $P = 0.50$; Uncontrollable: $t(26) = -0.13$, $P = 0.90$). (**f**) Expected influence was significantly higher for the Controllable than the Uncontrollable condition ($mean_C = 1.31$, $mean_U = 0.75$, $t(26) = 2.54$, $P < 0.05$). (**g**) In contrast to the social task, self-reported controllability was not different between the two conditions when individuals played the game with a computer ($mean_C = 62.7$, $mean_U = 56.9$, $t(25) = 0.78$, $P = 0.44$). (**h**) To unpack the norm prediction error × social interaction effect in **Supplementary file 1a**, we used the regression coefficients from the original mixed-effect regression ('emotion rating ~ offer + norm prediction error + condition + task + task*(offer + norm prediction error + condition) + (1 + offer + norm prediction error | subject)') and calculated the residual, which should be explained by the differential impact of nPE between social and non-social tasks. Correlation coefficients between the residuals and nPE were plotted for each task condition ($mean_{Social} = 0.151$, $mean_{Non\text{-}social} = 0.005$; $SD_{Social} = 0.023$, $SD_{Non\text{-}social} = 0.025$). Note that the non-social task was coded as the reference group (0 for the group identifier) in our original regression. This result indicates that the impact of nPE was stronger in the Social than in the non-social Computer task. Bars represent the mean of the coefficients and error bars represent the standard deviation. (**i**) To unpack the Controllable × social task interaction effect in **Supplementary file 1a**, similar to (**h**), we used the coefficients from the original mixed-effect regression ('emotion rating ~ offer + norm prediction error + condition + task + task*(offer + norm prediction error + condition) + (1 + offer + norm prediction error | subject)') and calculated the residual by each condition and task as shown in the figure ($mean_{Social(C)} = -5.00$, $mean_{Social(U)} = -0.58$, $mean_{Computer(C)} = 0.63$, $mean_{Computer(U)} = 0.00$; $SEM_{Social(C)} = 0.46$, $SEM_{Social(U)} = 0.43$, $SEM_{Computer(C)} = 0.72$, $SEM_{Computer(U)} = 0.77$). Bars represent the mean of the coefficients and error bars represent SEM. Note that the non-social task and the Uncontrollable condition were coded as the reference group (0 for the group identifiers) in the regression. These results show that the emotion ratings were lower in the Controllable social context compared to the non-social as well as the Uncontrollable social context. We speculate that exerting control over other people – compared to not needing to exert control over other people or playing with computer partners – might be more effortful (as shown by our RT results). Intentionally decreasing other people's portion of money might also induce a sense of guilt. Satterthwaite's approximation was used for the effective degrees of freedom for t-test with unequal variance. The variance significantly differed for the offer and the overall rejection rates. Error bars and shades represent s.e.m. * $P < 0.05$; ** $P < 0.01$; n.s. indicates not significant. For **a**, **b**, **f**, and **g**, each line represents a participant and each bold line represents the mean.

To summarize, based on these results, we speculate that subjects' beliefs regarding if their opponents were real humans or not were less likely to influence their choice behaviors but more likely to influence their emotional ratings (which is not part of the research question in this current study). Nevetheless, we have also added the following as a limitation to Discussion - Main text: lines 294-297. "Lastly, while participants were informed that they were playing with simulated players, we did not track their beliefs about the "humanness" of the other players. Future work could explicitly measure such belief and examine how it might affect participants' social choices."

Minor:

5. Was the total number of trials in the experiment 60/80 (i.e., 30/40 per condition and 2 conditions – one controllable, one uncontrollable)? Also, were participants told how many trials the game had and which trial number they were in?

Response: Thank you for bringing up this question. As you rightly pointed out, there were 30 trials per condition, totaling 60 trials for the fMRI study; and there were 40 trials per condition, totaling 80 trials for the online study. Participants were not explicitly told about the specific number of trials for each condition or the entire task. We have also added this to Methods:

Methods: lines 387-388: "The number of trials of the game was unknown to the participant."

6. Discussion: Given the small sample size of the in-person sample, I suggest changing the sentence "demonstrating that addiction is also linked to neural deficits in the midbrain during aberrant norm updating in complex social environments." (line 256) to "suggesting that addiction is also linked to …."

Response: We appreciate this revision point and have changed the wording, as you suggested.

Line 271: "…suggesting that addiction is also linked to neural deficits in the midbrain…"

7. In the discussion, you mention that "Futures studies may investigate the relationship between deprivation level and task-based measures by systematically manipulating participants' abstinence." (line 278/279) – this was also one of the analyses pre-registered; is there any reason why the analysis wasn't actually done in this paper?

Response: Nicotine abstinence was not implemented in the actual study due to challenges encountered after we pre-registered the study. Specifically (and also explained in a previous point), we detected high levels of exhaled CO in the participants at the time of the scanning session, indicating that our instructions to the smokers of staying abstinent overnight were likely ignored by them. We have added these details in Methods and Discussion -

Methods: lines 289-297: "Furthermore, although we were able to demonstrate group differences between smokers and controls, we were not able to carry out the planned analysis on how craving status might affect social controllability computations, as our attempt of instructing smokers to stay abstinent overnight failed in the experiment (see Methods). Futures studies may investigate the relationship between deprivation level and social decision-making by better experimental designs that can effectively vary participants' abstinence and craving levels."

Methods: lines 347-350: "However, CO levels measured on the day of scanning suggested that smokers were likely to have smoked regardless and failed to stay abstinent as instructed. As such, we were unable to conduct planned analyses on how craving affected social controllability per our preregistration."

Reviewer #3 (Remarks to the Author):

In this study the authors conduct an investigation of the behavior of current smokers versus non-smokers in a modified dictator game where their acceptance or rejection influences the offers that will be made to them in future trials. This task has been previously developed and published along with the development of a computational model that explains the probability of accepting or rejecting as a function of its projected value (forward thinking) which depends on the current utility of their choice plus a discounted utility of future splits, which is weighed by the agent's belief of the current choice influence. This belief is a free parameter estimated for each subject and is considered to represent "social controllability". The novelty of this study relies on the fact that the task and computational modeling is performed in a "clinical" population and that it is also used for an fMRI model-based investigation.

I generally find the results convincing and the motivation to explore social decision-making in the context of substance use disorders compelling. I do find that the paper could be improved both in clarity/readability and by addressing some issues that could take away from the findings.

Response: Thank you for your thorough and constructive review of our manuscript. We have made every effort to respond to your comments and revise the manuscript accordingly (see point-by-point response below). We hope you find the revised manuscript much improved and now suitable for publication.

Major points:

1. Did you control for socio-economic factors other than education? There are many reasons why one may believe to possess less agency or controllability. For example, perceived social status may have an impact. While it is likely that you don't have direct measures of subjective social status, you may be able to demonstrate that your groups are comparable in income, employment, reliance on social security, etc.

Response: Thank you for highlighting additional socioeconomic factors that could potentially influence the results. Apart from education, we have collected income data in our online replication study and confirmed that income levels among smokers and non-smokers are comparable. Implementing parametric bootstrapping, we found no significant difference ($p = 0.295$) in the income level of online smokers (6.43 ± 8.98) and non-smokers (6.74 ± 11.29). Income levels were categorized using brackets of $10,000 and ranged from 1, representing less than $10,000, to 12, representing more than $150,000 (refer to **Table S2** for intermediate values). We incorporated yearly income levels in the demographics table for online participants (**Table S2**). We have made sure that revised **Table S2** clearly lists all relevant socio-economic variables and statistical comparison results.

2. One alternative explanation to smokers having a lower inclination to reject medium size offers (which prevents them from getting higher value offers) may be that because the increase is probabilistic and thus rejecting may be perceived as risky. On the other hand, accepting more offers could be due to impatience where the received amount now is much more valuable than

the potential larger amount in the future. While how the results of this paper align with temporal discounting is partially mentioned in the discussion, please expand on how the current modeling approach compares to alternative explanations for the behavior observed.

Response: We appreciate this insightful comment and in response have included additional analyses and a discussion on temporal discounting and risk aversion. In our online sample and as part of a larger project, risk aversion (or seeking) was assessed using a risky decision-making task[6] analyzed by a computational model of risk preference[7]. We have now added the following additional analysis to address your concern, leveraging this secondary task collected in the same online participants.

Specifically, we asked participants to make 30 choices between two 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 S9** below for the full gamble pairs). The position 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 ($U(x) = x^{\rho}$) where its concavity (ρ) captures an individual's risk preference: $\rho < 1$ indicates risk-aversion, $\rho = 1$ indicates risk neutrality, and $r > 1$ indicates risk-seeking. 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 rejection rates between smokers and non-smokers (for medium sized offers, which was the main statistically significant result). We constructed general linear models with risk aversion parameter values as an independent variable to predict rejection rates.

Table S8. Risk Aversion & Medium Offer Rejection Rate GLM:

Rejection rate (Medium OFFERs) ~ intercept + group + risk aversion + group x risk aversion
There is no significant effect of risk aversion, group, and interaction on the rejection rate in the medium offer range. The significant intercept shows that on average healthy subjects have a higher rejection rate than smoker subjects in the medium offer range. An ANOVA on the regression model shows a significant group effect on the rejection rate but with a non-significant beta coefficient. The overall regression was not statistically significant ($R^2 = 0.03$, $F(3, 211) = 2.28$, $p = 0.08$). The group predictor ($\beta = -0.06$, $p = 0.19$), risk aversion ($\beta = 0.04$, $p = 0.50$), and interaction ($\beta = -0.11$, $p = 0.22$) were not significant.

	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

We have added these results and new methods to the paper:

Main text: lines 155-161. "Leveraging the existence of an independent risky decision-making task in this cohort[6], we then tested whether the group difference in rejection rate observed here were attributable to individual differences in risk aversion (see **Supplemental Methods** for details). To this end, we constructed GLMs with risk aversion parameter values estimated from this task[7] as an independent variable to predict rejection rate. This analysis demonstrated that risk aversion did not underpin the observed differences in rejection rates (for medium-sized offers) between the subject groups (**Table S8**)."

The risk aversion model is described in **Supplementary Information - *Risky Decision-making Task and Risk Aversion Model (online sample)*** with the payoffs of the added risky decision-making task listed in **Table S9**. The new GLM results are included as **Table S8.**

Table S9. Payoffs and probabilities of paired gambles of the risky decision-making task.

	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

3. *The discounting factor was kept fixed in the modeling of forward-thinking value. Please show the results of the modeling when the discounting factor is also estimated. Is there a trade-off between this parameter and your controllability parameter? Are there group differences in the discounting factor?*

Response: Thank you for this comment. Fixing the discounting factor was a deliberate choice when designing the computational model. In a previous study, we examined a version of the model with the temporal discounting parameter as a free parameter[1]. We found that this parameter had poor recoverability: the correlation between actual and recovered parameter value was 0.27 ($p>0.05$). Further, the correlation between the temporal discounting parameter and our key parameter of interest delta was -0.32 ($p=0.026$). As such, we fixed it at 0.8, the empirical mean across the participants from one initial round of estimation, in order to avoid collinearity with the parameter of our interest. The current study and task design were not well suited for questions related to temporal discounting, and we acknowledge this as a limitation of the study as well as added more relevant discussions. We do hope that the added risk aversion analyses presented above can partially address this question.

Discussion, line 238-244: "Our finding aligns with and provides a computational explanation for the observed greater temporal discounting of future rewards in SUD [8-10] , as well as findings suggesting that cognitive strategies involving deliberation and mental imagery of future events can reduce temporal discounting and cigarette consumption [11]. Although the temporal discounting parameter was fixed in our model setup due to technical considerations, our study expands this literature by showing that future-oriented valuation of one's own agency (i.e. calculating the impact of one's action on future events) is altered in smokers".

4. Please clarify whether the monetary rewards in the task were hypothetical or real and how payment was implemented.

Response: The monetary rewards in the task were real. Participants were compensated based on time, in addition to a randomly drawn outcome of the task. These details have been clarified in the methods section.

Methods: lines 309-311. "All participants provided written informed consent before participating in the study and were compensated based on time and task performance (i.e. the outcome of a randomly drawn trial)."

5. Were participants instructed to consider the impact of their choices on subsequent offers? When was the question about their perceived controllability delivered relative to the task?

Response: Participants were not given explicit information about how their actions could influence the offers or which condition they might have influence. Instead, they were instructed that "you may or may not have influence over the offers made by this team" regardless of the condition/team. Perceived controllability was rated at the end of each condition (30/40 trials). We have clarified these detailed in the revised manuscript -

Methods: lines 381- 383. “Subjects were told that they were playing with members of two different teams and were not given explicit information regarding how the two teams might differ. Instead, they were instructed that they “may or may not have influence over the offers made” by the team.”

Main text: lines 388-390. “After completing the task, subjects were asked to rate their perceived influence over their partners’ offers at the end of each condition using a scale from 0 to 100 (“perceived controllability).”

Minor points:

1. How many subjects performed the task in the scanner in each group? Please provide more detail about the fMRI task and how it differed from the behavioral task.

Response: We apologize for this confusion. The number of participants in each group has been included in the figure legends for easier reference. Specifically, there were a total of 17 smokers and 25 non-smokers in the fMRI group. We have added the specific sample size for each study/cohort throughout the manuscript. Additionally, we have revised the methods section to emphasize that the same task was used for both online and fMRI participants, with slightly different numbers of trials between the two versions.

2. The paper seems to have been written by multiple different authors with divergent styles. The introduction and results sections lack clarity, cohesiveness, care for style, and is replete with typos, in comparison to the much better-written discussion. I recommend that the authors re-read the manuscript carefully and improve the writing.

Response: We apologize for the lack of clarity or cohesiveness. We have made thorough revisions to the entire manuscript, including the Introduction and Results sections, to improve the writing style.

3. Was this a pre-registered study? Please comment on possible limitations if not.

Response: Both the fMRI and online studies were pre-registered as a subcomponent of larger projects. We have specified in the manuscript which parts of the study were registered -

Methods: lines 320-321. “The fMRI study was pre-registered as part of a larger fMRI study examining decision-making in smokers (https://osf.io/m9cws).”

Methods: lines 335-337. “This online study was pre-registered as a subcomponent of a larger longitudinal study investigating social decision-making (https://osf.io/8s5mu).”

Methods: lines 453-455. “The computational models were also pre-registered as a subset of the project examining value-based decision making in nicotine addiction (https://osf.io/m9cws).”

Methods: lines 458-459: “Our analyses focused on group comparisons between smokers and non-smokers, per our pre-registrations.”

Methods: lines 520-521. “The ROIs in the analyses were specified in the pre-registered study (https://osf.io/m9cws).”

Rebuttal Letter References

1 Na, S. *et al.* Humans use forward thinking to exploit social controllability. *eLife* **10**, e64983 (2021). https://doi.org:10.7554/eLife.64983

2 Mooney, C. Z. & Duval, R. D. *Bootstrapping: A Nonparametric Approach to Statistical Inference*. (Sage Publications, 1993).

3 Hasson, U., Avidan, G., Deouell, L. Y., Bentin, S. & Malach, R. Face-selective Activation in a Congenital Prosopagnosic Subject. *Journal of Cognitive Neuroscience* **15**, 419-431 (2003). https://doi.org:10.1162/089892903321593135

4 Gu, X. *et al.* Anterior insular cortex is necessary for empathetic pain perception. *Brain* **135**, 2726-2735 (2012). https://doi.org:10.1093/brain/aws199

5 Gu, X. *et al.* Necessary, Yet Dissociable Contributions of the Insular and Ventromedial Prefrontal Cortices to Norm Adaptation: Computational and Lesion Evidence in Humans. *Journal of Neuroscience* **35**, 467-473 (2015). https://doi.org:10.1523/JNEUROSCI.2906-14.2015

6 Chung, D., Christopoulos, G. I., King-Casas, B., Ball, S. B. & Chiu, P. H. Social signals of safety and risk confer utility and have asymmetric effects on observers' choices. *Nature Neuroscience* **18**, 912-916 (2015). https://doi.org:10.1038/nn.4022

7 Bernoulli, D. Exposition of a New Theory on the Measurement of Risk. *Econometrica* **22**, 23-36 (1954). https://doi.org:10.2307/1909829

8 Audrain-McGovern, J. *et al.* Does delay discounting play an etiological role in smoking or is it a consequence of smoking? *Drug and alcohol dependence* **103**, 99 (2009). https://doi.org:10.1016/j.drugalcdep.2008.12.019

9 Johnson, M. W., Bickel, W. K. & Baker, F. Moderate drug use and delay discounting: A comparison of heavy, light, and never smokers. *Experimental and Clinical Psychopharmacology* **15**, 187-194 (2007). https://doi.org:10.1037/1064-1297.15.2.187

10 Reynolds, B., Richards, J. B., Horn, K. & Karraker, K. Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behavioural Processes* **65**, 35-42 (2004). https://doi.org:10.1016/S0376-6357(03)00109-8

11 Stein, J. S., Tegge, A. N., Turner, J. K. & Bickel, W. K. Episodic future thinking reduces delay discounting and cigarette demand: an investigation of the good-subject effect. *Journal of Behavioral Medicine* **41**, 269-276 (2018). https://doi.org:10.1007/s10865-017-9908-1

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

The authors have been very responsive and the manuscript is improved.

Reviewer #2 (Remarks to the Author):

The paper is now much improved, with many clarifications and additional analyses.
I have, however, still a few questions/concerns:

1. For the in-person study, the reported results on sense of control seem to be from a one-tailed t-test ($t(40)=1.93$, $p= 0.031$)? The two-tailed t-test was thus likely non-significant ($p=0.062$). I understand the temptation to do a one-tailed test, especially if the results go in the direction we expected, but I think that in this case the most appropriate would be to do a two-tailed test. The exception would have been if this particular directional hypothesis and analysis had been preregistered, which it wasn't. It is totally ok if the result is non-significant, it should just not be labelled as significant if it isn't, thus the text should be reworded accordingly.
2. Similarly, are the p-value results related to the online sample one-tailed or two-tailed?
3. I appreciate the added analyses answering the other reviewers' comments, namely looking at potential effects of depression and/or impulsivity that could also help explain the results. I do not understand the authors' claim that "The significant intercept shows that on average healthy subjects have higher model-estimated controllability than smoker subjects". Shouldn't the effect of group (non-smokers vs. smokers) instead be captured by the "group" parameter also included in the general linear model? But maybe I am missing something…
4. On that note, the group predictor in both the depression and impulsivity general linear models was actually non-significant, and the whole GLMs seem to explain very little of the total variance in delta. Overall, this suggests that group membership (being a smoker or non-smoker) explains very little of the variation in delta (mentally simulated influence of one's actions on future social outcomes). How do the authors interpret these findings?

Minor
1. I suggest replacing "health controls" with "non-smokers", as there may be healthy smokers and non-health non-smokers.
2. Could the effect sizes for each result be reported?

Reviewer #3 (Remarks to the Author):

I think the authors addressed all the reviewers' comments and suggestions satisfactorily. The manuscript is greatly improved in clarity and detail.

Response Letter

Dear Editor,

Thank you and the reviewers for providing important insight and feedback for our manuscript. We are glad to hear that all three reviewers have found the paper significantly improved and that Reviewers 1 & 3 were both happy with the revised manuscript as it was. Please find our point-by-point response to the remaining comments from Reviewer 2 below. We have also updated our main text and supplementary information accordingly.

The paper is now much improved, with many clarifications and additional analyses.
I have, however, still a few questions/concerns:

1. For the in-person study, the reported results on sense of control seem to be from a one-tailed t-test (t(40)=1.93, p= 0.031)? The two-tailed t-test was thus likely non-significant (p=0.062). I understand the temptation to do a one-tailed test, especially if the results go in the direction we expected, but I think that in this case the most appropriate would be to do a two-tailed test. The exception would have been if this particular directional hypothesis and analysis had been pre registered, which it wasn't. It is totally ok if the result is non-significant, it should just not be labeled as significant if it isn't, thus the text should be reworded accordingly.

Response: Thank you for this careful observation. We did not intend to implement a one-tailed t-test for this analysis. You are correct that a two-tailed t-test is more appropriate, and we have updated our results descriptions to reflect these changes. Briefly, now group difference in perceived controllability becomes non-significant for the in-person study but remains significant for the online study, albeit larger Cohen's d for the in-person than the online sample. We have updated all relevant descriptions and interpretations accordingly (while noting that ratings from three smoker subjects were missing). We have also added effect size measures to all outcomes of interest, which might provide more information to readers in terms of interpretation of the reported effects.

Main text, line 109:
"Although smokers reported a lower sense of control (52.40%±20.76) compared to non-smokers (65.91%±22.39; t(37)=-1.93, p= 0.062; Cohen's d= -0.63; Figure 2E), this difference was not statistically significant (potentially due to the small sample size; due to technical failures, sense of control ratings for three non-smoker subjects were also missing). Taken together, these model-agnostic behavioral results reveal impaired social controllability in smokers, primarily indexed by their reduced ability to raise offers in the Controllable condition."

Line 168:

"Finally, online smokers self-reported a significantly reduced sense of control than online non-smokers (smokers: 52.68%±34.46, non-smokers: 61.32%±34.63; p= 0.0442; Cohen's d= -0.25; Figure 4E), despite a smaller effect size compared to the in-person study. Taken together, these model-agnostic analyses of a much larger and variable online sample provided converging evidence that smokers showed impairments in their ability to exploit the controllability of their social interactions, indexed by reduced ability to raise offers in the Controllable condition. We also found that group differences in perceived controllability ratings diverged between in-person and online samples, reflecting the huge variability in subjective perception."

2. Similarly, are the p-value results related to the online sample one-tailed or two-tailed?

Response: The p-values related to the online sample are always two-tailed. For the bootstrapping method, each iteration performed a two-tailed t-test. We have clarified this in the *Statistics and Reproducibility* section.

Main text line 488: "... (iii) the two-tailed t-value of the difference between the two surrogate groups was calculated."

3. I appreciate the added analyses answering the other reviewers' comments, namely looking at potential effects of depression and/or impulsivity that could also help explain the results. I do not understand the authors' claim that "The significant intercept shows that on average healthy subjects have higher model-estimated controllability than smoker subjects". Shouldn't the effect of group (non-smokers vs. smokers) instead be captured by the "group" parameter also included in the general linear model? But maybe I am missing something…

Response: Thank you for pointing out this error. You are absolutely correct that the group term should indicate if the controllability measure differed between groups (in this case not in the GLM). The significant intercept, in fact, indicates that the value of delta, when all predictor variables are set to zero, is significantly non-zero. We have double checked relevant descriptions in the supplementary tables to ensure these results are correctly interpreted (Table S6-S8)

4. On that note, the group predictor in both the depression and impulsivity general linear models was actually non-significant, and the whole GLMs seem to explain very little of the total variance in delta. Overall, this suggests that group membership (being a smoker or non-smoker) explains very little of the variation in delta (mentally simulated influence of one's actions on future social outcomes). How do the authors interpret these findings?

Response: Thank you for raising this interesting point. We agree that overall, these regressions account for little variance in delta, suggesting that the particular combinations of

input variables were not great for accounting for the variances in delta. However, our intention with these GLMs was not to maximize the explained variance in delta by finding the most optimal combination of variables (e.g. including additional predictors, such as demographics or more clinical scores, could potentially increase the explained variance). Instead, our primary question was to see whether smokers and non-smokers differed on delta; and this added GLM was designed to rule out the possibility that individual negative affect or impulsivity measures were driving the group differences.

Inspired by your question, we also further explore how the different numeric value ranges of different input variables (e.g. 0~1 for group but 0~63 for BDI) might have affected the GLM results. Our original analysis did not normalize the values, and we have since conducted new GLMs by z-scoring BDI/BIS. Overall, the patterns still hold (see revised Table S6 & S7). However, the group term in the BDI GLM is now approaching significance in the GLM with group and BDI, a trend consistent with the direct t test of group means. The group term in the updated BIS GLM remains non-significant. In either case, these results might still not be directly comparable to the t test, as the variance explained is shared between multiple variables. We have also added brief discussions in the figure legends of updated Table S6 and S7.

Minor

1. I suggest replacing "health controls" with "non-smokers", as there may be healthy smokers and non-health non-smokers.

Response: We agree and have reworded all "healthy controls" as "non-smokers".

2. Could the effect sizes for each result be reported?

Response: Thank you for the suggestion. Effect sizes for all statistical tests are now reported as Cohen's d values.