Supplementary Information for

Emotion, Impatience and Addictive Behavior

Charles A. Dorison, Ke Wang, Vaughan W. Rees, Ichiro Kawachi, Keith M. M. Ericson, Jennifer S. Lerner

Corresponding author: Charles A. Dorison
Email: cdorison@g.harvard.edu

This PDF file includes:

Supplementary Text
Figure S1 to S3
Tables S1 to S8

Supplementary Text

Summary

For Study 1 (pages 2-8), we report the procedure to identify current smoking status, additional demographic information for all four waves of the survey, full details of all regressions (including an additional measure of effect size: odds ratio), analyses after controlling for other psychologically-meaningful variables, additional details of longitudinal analyses, and analyses with the sub-sample of non-depressed participants. For Study 2 (pages 9-11), we report emotion manipulation checks, a visual depiction of the main effect, pre-registered exclusion criteria, and results with exclusions. For Study 3 (pages 12-16), we report emotion manipulation checks, a visual depiction of the main effects, pre-registered exclusion criteria, results with exclusions, moderation analysis, results of a replication study, and full details on how we calculated the required rate of return (RRR). For Study 4 (pages 17-19), we report emotion manipulation checks, additional demographics on participants, full details of Bayesian analyses, and unforeseen complications during the course of the study. Finally, we also report additional information on use and validity of online samples (page 19).

Study 1

Procedure to identify current smoking status. Current smoking status was identified using responses to three questions (Fig. S1). First, participants were asked "At what age did you have your very first cigarette?" Second, if they did not answer "never had a cigarette" to the first question, they were then asked "Have you ever smoked cigarettes regularly -- that is, at least a few cigarettes every day?" Third, if they answered yes to the second question, they were finally asked "Do you smoke cigarettes regularly now?" Participants who answered yes to this third question were identified as current smokers.

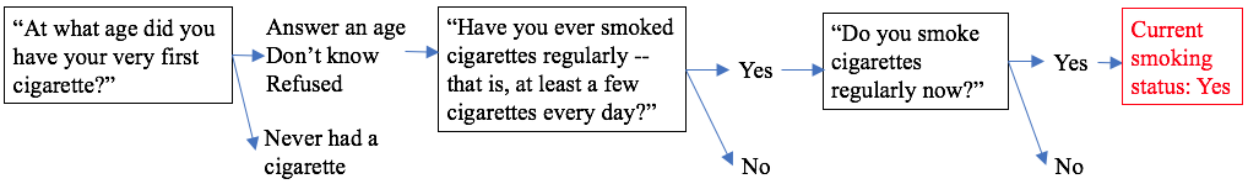

Fig. S1. Visual depiction of three questions used to identify current smoking status.

Additional demographic information for all four waves of the survey. Study 1 utilized data from four waves of the Midlife in the United States (MIDUS) survey, a nationally representative sample of the United States population aged 24-74 (1). Information about data collection and demographic variables for all four waves are summarized in Table S1.

Table S1. Demographic information across the four datasets. The datasets are nationally representative of the United States population aged 24-74.

Variables	MIDUS 1	MIDUS 2	MIDUS 3	MIDUS Refresher
Dates of collection	Jan 1995 - Jan 1996	Jan 2004 - Sep 2006	May 2013 – Nov 2014	Nov 2011 – Sep 2014
N	7,108	4,963	3,294	3,577
Females (%)	3632 (52%)	2647 (53%)	1810 (55%)	1856 (52%)
Age (SD)	46.38 (13.00)	55.43 (12.45)	63.64 (11.35)	50.51 (14.38)
Median, Household income, USD[a]	55,000	57,500	67,500	70,500
Education[b] (SD)	6.77 (2.49)	7.20 (2.52)	7.51 (2.51)	7.79 (2.49)
Major depressive disorder (%)	942 (13%)	522 (11%)	327 (10%)	476 (13%)
Smoke regularly (%)	1629 (23%)	768 (15%)	306 (9%)	471 (13%)

[a] If household income was less than $300,000, it was coded as a continuous variable. If it was greater than $300,000, it was coded as $300,000.

[b] Coding for education: No school/some grade school (1); Eighth grade/junior high school (2); Some high school (no diploma/no GED) (3); GED (4); Graduated from high school (5); 1 to 2 years of college, no degree yet (6); 3 or more years of college, no degree yet (7); Graduated from 2-year college or vocational school, or Associate's degree (8); Graduated from 4- or 5-year college or Bachelor's degree (9); Some graduate school (10); Master's degree (11); PH.D, ED.D, MD, DDS, LLB, LLD, JD, or other professional degree (12).

Two points merit note regarding the nationally-representative nature of this dataset. First, while broadly nationally-representative, the MIDUS 1 sample oversampled individuals from five metropolitan areas. Second, as seen in S1, women and high-income individuals were more likely to remain in the survey over the course of MIDUS 2 and MIDUS 3.

Full details of all regressions (including an additional measure of effect size: odds ratio)

Table S2. Regressions controlling for all negative emotions

	sadness		anger		fear		shame	
	beta (SE)	OR (95% CI)	beta (SE)	OR (95% CI)	beta (SE)	OR (95% CI)	beta (SE)	OR (95% CI)
MIDUS 1	.32*** (.04)	1.38 (1.28-1.49)						
MIDUS 2	.23** (.07)	1.26 (1.09-1.46)	.16* (.06)	1.17 (1.03-1.33)	.25** (.09)	1.29 (1.09-1.53)	-.03 (.08)	.97 (.83-1.14)
MIDUS 3	.29* (.11)	1.34 (1.07-1.66)	.10 (.10)	1.11 (.91-1.35)	.08 (.14)	1.08 (.82-1.41)	.16 (.12)	1.17 (.92-1.48)
MIDUS Refresher	.51*** (.09)	1.67 (1.39-2.00)	.13 (.08)	1.14 (.97-1.33)	-.06 (.11)	.94 (.75-1.16)	.01 (.10)	1.01 (.83-1.23)
Combine MIDUS 2, 3, and Refresher	.33*** (.05)	1.39 (1.26-1.54)	.14** (.05)	1.15 (1.05-1.26)	.12* (.06)	1.13 (1.00-1.27)	.02 .06)	1.02 (.91-1.14)

SE = standard error; OR = odds ratio; CI = confidence interval.
***$P < .001$, **$P < .01$, *$P < .05$.

Table S3. Regressions controlling for all negative emotions and demographics

	sadness		anger		fear		shame	
	beta (SE)	OR (95% CI)	beta (SE)	OR (95% CI)	beta (SE)	OR (95% CI)	beta (SE)	OR (95% CI)
MIDUS 1	.25*** (.04)	1.29 (1.19-1.39)						
MIDUS 2	.21** (.08)	1.23 (1.05-1.43)	.07(.07)	1.08 (.94-1.23)	.20* (.09)	1.22 (1.02-1.46)	.01 (.09)	1.01 (.85-1.20)
MIDUS 3	.25* (.12)	1.28 (1.01-1.61)	.09(.11)	1.09 (.89-1.34)	.03 (.14)	1.03 (.78-1.36)	.13 (.13)	1.13 (.88-1.45)
MIDUS Refresher	.44*** (.10)	1.55 (1.28-1.88)	.09(.09)	1.09 (.92-1.29)	-.08 (.12)	.93 (.74-1.16)	-.02 (.10)	.98 (.80-1.19)
Combine MIDUS 2, 3, and Refresher	.28*** (.05)	1.33 (1.20-1.48)	.09† (.05)	1.09 (1.00-1.20)	.08 (.06)	1.08 (.95-1.22)	.02 (.06)	1.02 (.91-1.14)

SE = standard error; OR = odds ratio; CI = confidence interval.

***$P < .001$, **$P < .01$, *$P < .05$, † $P < .10$

Analyses after controlling for other psychologically-meaningful variables

Even after controlling for sense of control, social integration, positive relations with others, and loneliness, sadness significantly predicted smoking status in 3 out of 4 waves, whereas no other negative emotion significantly predicted smoking status in any of the 4 waves.

Table S4. Regressions controlling for all negative emotions, demographics, and psychologically-meaningful variables

	sadness		anger		fear		shame	
	beta (SE)	OR (95% CI)	beta (SE)	OR (95% CI)	beta (SE)	OR (95% CI)	beta (SE)	OR (95% CI)
MIDUS 1	.28*** (.28)	1.32 (1.21-1.44)						
MIDUS 2	.20* (.08)	1.23 (1.04-1.45)	.05 (.07)	1.06 (.92-1.21)	.18† (.10)	1.20 (.99-1.44)	-.00 (.09)	.99 (.83-1.19)
MIDUS 3	.21 (.13)	1.23 (.95-1.57)	.07 (.11)	1.07 (.86-1.33)	.04 (.15)	1.04 (.78-1.40)	.13 (.13)	1.13 (.87-1.46)
MIDUS Refresher	.36*** (.11)	1.44 (1.17-1.77)	.06 (.09)	1.06 (.89-1.26)	-.16 (.12)	.85 (.67-1.08)	-.05 (.11)	.95 (.77-1.17)
Combine MIDUS 2, 3, and Refresher	.25*** (.06)	1.29 (1.15-1.44)	.07 (.05)	1.07 (.97-1.18)	.04 (.07)	1.05 (.92-1.19)	.00 (.06)	1.00 (.89-1.13)

SE = standard error; OR = odds ratio; CI = confidence interval.
***$P < .001$, **$P < .01$, *$P < .05$, † $P < .10$

Additional details of longitudinal analyses. We analyzed individuals at Time 1 who were non-smokers. Of the 3,640 individuals who responded at Time 2, 109 (3.0%) reported smoking at Time 2 (10 years later) but were self-reported non-smokers at Time 1. Of the 2,547 individuals who responded at Time 3, 54 (2.1%) reported smoking at Time 3 (20 years later) but were self-reported non-smokers at Time 1.

Analyses with the sub-sample of non-depressed participants. The frequency of participants with Major Depressive Disorder (MDD) ranged from 10% to 13% across waves (see Table S1). We conducted identical analyses to those reported in the manuscript on only the subset of participants who did not demonstrate MDD.

As in the manuscript, we organize the results around two questions concerning sadness and smoking.

Question 1: *Would sadness, but not other negative emotions, predict being a smoker in the present?*

To answer question one, we identified all negative emotions measured in the MIDUS datasets. The set included sadness, anger, shame, and fear in MIDUS 2, MIDUS 3, and the MIDUS Refresher (no negative emotions other than sadness were measured in MIDUS 1). We tested the hypothesis by regressing smoking status on all four negative emotions simultaneously.

Consistent with the ATF prediction that sadness is positively associated with smoking status, even after controlling for other negative emotions, sadness significantly predicted self-reported point prevalence of smoking in all three waves in which other negative emotions were collected (*b*'s = 20, .33, .53, all *z*'s > 2.24, all *p*'s < .05). No other emotion significantly predicted smoking status in any wave, with average betas comparatively small in combined-samples analyses (fear: .07; anger: .13; shame: .00).

Given evidence that income (e.g., 2), age (e.g., 3), and gender (e.g., 4) predict either smoking initiation, quitting, or both, it could be that sadness serves as a proxy for such demographic variables. We tested this question with simultaneous entry of all independent variables in logit regressions. Sadness significantly predicted smoking status after controlling for demographic variables and for all other negative emotions identified in previous analyses. Results for sadness were significant in MIDUS 1,

MIDUS 3, and MIDUS Refresher (b's = .21, .32, .42, z's > 2.37, all p's < .05) and marginally significant in MIDUS 2 (b = .17, z = 1.81, p = .071). Combined-samples analyses of the four datasets revealed strong support to conclude that the relationship between sadness and smoking holds across datasets, even after controlling for the influence of demographic variables (b = .28, z = 4.26, p < .001). Moreover, no other negative emotion significantly predicted smoking status in any wave, with average betas relatively small (fear: .06; anger: .08; shame: .01).

Question 2: *Among non-smokers at Time 1, could sadness predict their likelihood of smoking at Time 2 (10 years later) and Time 3 (20 years later)?*

We hypothesized that sadness felt at Time 1 among non-smokers would predict smoking 10 and 20 years later. We found directional, although not statistically significant, support for this hypothesis: sadness reported at Time 1 among non-smokers predicted smoking 10 years (b = .28, p = .074) and 20 years (b = .31, p = .172) later. It is worth noting that the sizes of the relationships were similar to those reported in Study 1; however, the smaller sample size may have led the analyses to be underpowered. Indeed, of the 3,259 non-depressed non-smokers who responded at Time 2 and 2,288 who responded at Time 3, 85 (i.e., 2.6%) were smokers at Time 2 (10 years later) and 45 (i.e., 2.0%) were smokers at Time 3 (20 years later).

Follow-up analyses confirmed the same key moderator of this relationship. Among non-depressed non-smokers who had smoked regularly, sadness at Time 1 significantly predicted smoking 10 years later (b = .38, p = .036) and directionally predicted smoking 20 years later (b = .38, p = .116). However, among non-depressed non-smokers who had never smoked regularly, sadness did not predict smoking 10 (b = -.55, p = .319) or 20 (b = -.77, p = .427) years later. Thus, sadness again appears to be associated with a persistence of smoking over time in people with a history of smoking, rather than initiation of smoking in people who had never smoked. This exactly mirrored the results reported with the full sample.

Study 2

Emotion manipulation checks. We tested whether the emotion manipulations were effective in both magnitude and specificity. In terms of magnitude, participants in the sadness (disgust) condition reported higher levels of sadness (disgust) and lower levels of neutrality than individuals in the neutral condition (all t's > 5.46, all p's < 0.001). In terms of specificity, participants in the sadness (disgust) condition reported feeling more sadness (disgust) than any other emotion (all t's > 8.01, all p's < 0.001).

Visual depiction of main effect. Results are depicted in Figure S2. The craving measure achieved a high level of reliability both before and after the emotion induction (alphas = .93 and .94, respectively).

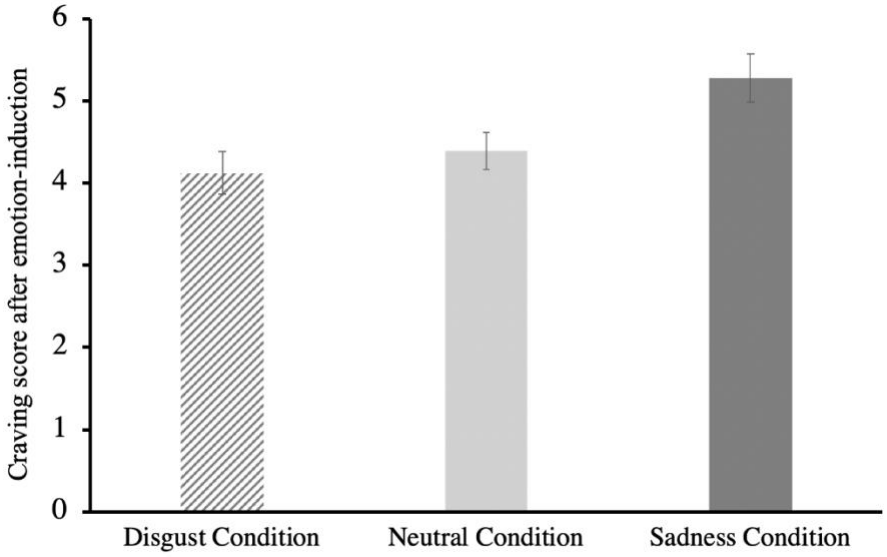

Fig. S2. Participants in the sadness condition reported higher levels of craving compared to individuals in the disgust and neutral conditions. Individuals in the disgust condition reported statistically equivalent levels of self-reported craving compared to individuals in the neutral condition. N = 342 unique participants. Error bars represent 1 standard error.

Pre-registered exclusion criteria. Because of the challenges inherent in inducing specific emotions via an online platform where participants are in uncontrolled, varied settings, we pre-registered two exclusion criteria that could help ensure a quality sample. Specifically, we pre-registered the ability to exclude participants for either (or both) of the two reasons listed below:

1. Failure to comply with instructions. This criterion may be operationalized in any of the following three ways: not answering all three attention/comprehension check questions correctly; reporting technical problems; and violating instructions (e.g., writing fewer than 10 words or writing nonsense responses in the writing task, reporting taking the study in more than one sitting, reporting being distracted by background noise, reporting smoking at any point during the survey, reporting doing something other than taking the survey at any point during the survey, admitting to not being a smoker, or admitting to smoking within the last hour).
2. Failure to pass the manipulation check. This criterion will be operationalized in the following way: showing no absolute increase in self-reported sadness (disgust) in the sadness (disgust) condition.

Results with exclusions.

It turned out that our concerns about inducing emotion online were unfounded. As described in greater detail below, regardless of whether we used either set of pre-registered exclusions or used the full sample, we observed the same pattern of results. We therefore chose to maximize statistical power and minimize sample modifications by reporting the full sample without exclusions in the main text of the manuscript.

Table S5. Results of Study 2 with different exclusion criteria

Exclusion	No exclusion (Reported in the main text)	Set 1: Full set of pre-registered exclusions	Set 2: Full set of pre-registered exclusions, minus exclusions based on emotion inductions
Results	First, we conducted a pairwise contrast between the sadness and neutral conditions. As predicted, we found evidence that sadness increased craving as compared to a neutral state (b = 0.58, se = 0.21, t = 2.82, p = 0.005, d = 0.29). Second, we conducted a pairwise contrast between disgust and neutral conditions. We found directional, but not statistically significant, evidence that disgust *decreased* craving as compared to a neutral state (b = -0.35, se = 0.22, t = -1.56, p = 0.12, d = -0.09). Finally, we conducted a pairwise contrast between the sadness and disgust conditions: sadness significantly increased craving as compared to disgust (b = 0.96, se = 0.25, t = 3.84, p < 0.001, d = 0.35).	First, we conducted a pairwise contrast between the sadness and neutral conditions. As predicted, we found evidence that sadness increased craving as compared to a neutral state (b = 0.70, se = 0.24, t = 2.93, p = 0.004, d = 0.30). Second, we conducted a pairwise contrast between disgust and neutral conditions. We found directional, but not statistically significant, evidence that disgust *decreased* craving as compared to a neutral state (b = -0.46, se = 0.25, t = -1.82, p = 0.07, d = -0.18). Finally, we conducted a pairwise contrast between the sadness and disgust conditions: sadness significantly increased craving as compared to disgust (b = 1.18, se = 0.29, t = 4.10, p < 0.001, d = 0.45).	First, we conducted a pairwise contrast between the sadness and neutral conditions. As predicted, we found evidence that sadness increased craving as compared to a neutral state (b = 0.57, se = 0.23, t = 2.50, p = 0.013, d = 0.21). Second, we conducted a pairwise contrast between disgust and neutral conditions. We found directional, but not statistically significant, evidence that disgust *decreased* craving as compared to a neutral state (b = -0.46, se = 0.25, t = -1.82, p = 0.10, d = -0.18). Finally, we conducted a pairwise contrast between the sadness and disgust conditions: sadness significantly increased craving as compared to disgust (b = 0.99, se = 0.27, t = 3.65, p < 0.001, d = 0.34).

Study 3

Emotion manipulation checks. We tested whether the emotion manipulation was effective in both magnitude and specificity. In terms of magnitude, participants in the sadness condition reported higher levels of sadness ($t(396) = -16.70$, $p < 0.001$) and lower levels of neutrality ($t(396) = 8.50$, $p < 0.001$) than individuals in the neutral condition. In terms of specificity, participants in the sadness condition reported feeling more sadness than any other emotion (all t's > 8.02, all p's < 0.001).

Visual depiction of the main effects.

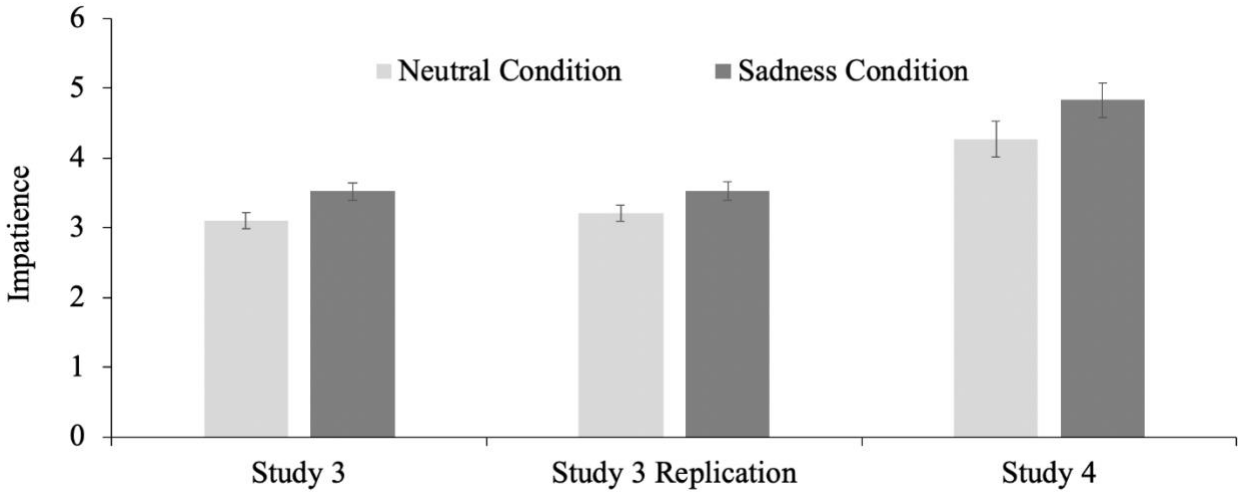

Fig. S3. Participants in the sadness condition made more impatient choices regarding cigarette puffs than did participants in the neutral condition. N = 725 unique participants across 5,622 total choice lists. Error bars represent 1 standard error adjusted for repeated measures.

Pre-registered exclusion criteria. Study 3 was conducted approximately one year apart from Study 2. The list below contains exactly the exclusions we pre-registered, which were similar in content to the list in Study 2 even though the presentation format differs.

1. Participants who do not answer all three video attention/comprehension check questions correctly.

2. Participants who report experiencing technical problems.

3. Participants who do not follow directions, either by writing fewer than 10 words in the writing prompt or by writing nonsense responses. 4. Participants who write hostile comments to the experimenter at the end of the survey.

5. Participants who report taking the study in more than one sitting, or report being distracted by background noise while taking this survey. 6. Participants who report smoking at any point during the survey.

7. Participants' responses that don't show monotonicity on a list-by-list basis (i.e., irrational flipping behavior).

8. Participants who report doing something other than taking the survey at any point during the survey (e.g., going to the bathroom, watching TV, chatting).

9. Participants who show no absolute increase in self-reported sadness of any size (i.e., the sum score of pre-sadness ≥ the sum score of post-sadness) in the sad condition.

Results with pre-registered exclusions

As was the case in Study 2, with two sets of exclusions, we observed the same pattern of results as those reported in the main text.

Table S6. Results of Study 3 with different exclusion criteria

Exclusion	No exclusion (Reported in the main text)	Set 1: Full set of pre-registered exclusions	Set 2: Full set of pre-registered exclusions, minus exclusions based on emotion inductions
Results	Smokers in the sad condition showed greater impatience for hypothetical cigarette puffs than did smokers in the neutral condition (b = 0.43, se = 0.18, t = 2.42, p = 0.016, d = 0.19; standard errors adjusted for repeated measures).	Smokers in the sad condition showed greater impatience for hypothetical cigarette puffs than did smokers in the neutral condition (b = 0.45, se = 0.22, t = 2.09, p = 0.038, d = 0.20; standard errors adjusted for repeated measures).	Smokers in the sad condition showed greater impatience for hypothetical cigarette puffs than did smokers in the neutral condition (b = 0.49, se = 0.22, t = 2.28, p = 0.024, d = 0.21; standard errors adjusted for repeated measures).

Moderation analysis. The effect of sadness on impatience choices was not moderated by most variables we tested (i.e., availability of an immediate option, depression, nicotine dependence) except household income. Without exclusion, however, household income significantly moderated the effect of sadness (b = -0.45, se = 0.17, t = -2.54, p = 0.011). The effect is primarily driven by participants whose household income is below or equal to the median (b = 0.83, se = 0.27, t = 3.05, p = 0.003, d = 0.35), rather than those whose household income is above the median (b = 0.04, se = 0.23, t = 0.17, p = 0.863, d = 0.02). After pre-registered exclusions, this interaction was unreliable (b = - 0.37, se = 0.22, t = -1.68, p = 0.095). After the full set of pre-registered exclusions minus the exclusions based on emotion inductions, this interaction was significant (b = - 0.44, se = 0.22, t = -2.03, p = 0.043). Our replication study, however, did not find the interaction significant (reported below). Thus, the effect of sadness on impatience choices was not consistently moderated by household income. The one instance of moderation may be Type I error, due to the high number of moderators tested. Alternatively, the lack of additional interactions could be Type II error, given that interactions require large sample sizes. Future research could explore these possibilities.

Results of a replication study. We conducted a replication study of Study 3 with an identical design aside from three minor methodological differences. First, for the wording on the choice lists, in the replication we asked questions in the form "X puffs in Y minutes" rather than "X puffs Y minutes from now" in Study 3. Second, we assessed the appraisal themes in a different way. Specifically, in the replication we included one reverse-coded item per appraisal theme whereas in Study 3 all items were scored in the same direction except for one reverse-coded item for valence. These appraisal themes did not significantly mediate the effect of sadness on impatience in either study. Finally, we included a measure of risk preference but did not include depression, nicotine dependence, subjective SES, or

childhood SES in the replication – none of these measures moderated the effect of sadness on impatience in Study 3.

In the replication, we recruited 362 smokers through Amazon's Mechanical Turk (217 male, 144 female, 1 non-binary/other, mean age = 34.36, age range = 18-72 years).

We first tested whether the emotion manipulation was effective in both magnitude and specificity with the full sample without exclusion. In terms of magnitude, we found that participants in the sadness condition reported higher levels of sadness (M = 4.30 vs. M = 1.35, $t(360) = -12.75$, $p < .001$) and lower levels of neutrality (M = 1.86 vs. M = 3.46, $t(360) = 7.84$, $p < .001$) than individuals in the neutral condition. In terms of specificity, we found that participants in the sadness condition reported feeling more sadness than any other emotion (all *t's* < - 4, all *p's* < .001).

We next turned to our primary confirmatory hypothesis: whether sadness had a causal effect on impatient choices. Without exclusion, induction of sadness in smokers produced greater impatience for hypothetical cigarette puffs than a neutral emotion condition, $b = 0.32$, $se = 0.18$, $t = 1.80$, $p = 0.073$, $d = 0.15$ (Figure 3, center panel). With exclusions, we obtained nearly identical results: smokers in the sadness condition tended to produce greater impatience for hypothetical cigarette puffs than did smokers in the neutral emotion condition, $b = 0.48$, $se = 0.22$, $t = 2.17$, $p = .031$, $d = 0.22$ (Set 1 exclusion as in Study 3) and $b = 0.44$, $se = 0.21$, $t = 2.08$, $p = .038$, $d = 0.20$ (Set 2 exclusion as in Study 3). The effect of sadness on impatience choices was not consistently moderated by any variables we tested (i.e., availability of an immediate option, household income, and the number of hours since last smoking).

To obtain a sense of the size of the sadness effect, we calculated a required rate of return (RRR) with the full sample. The RRR indicated the average increase in number of puffs per minute smokers required in order to wait for a delayed reward, where higher numbers indicated higher levels of impatience. Smokers in the neutral condition had an RRR of 7.3%, indicating that (on average) they required an increase in puffs of 7.3% per minute in order to wait for a delayed reward. Smokers in the sadness condition were more impatient: they had an RRR of 8.0%, indicating a 9% increase. Again,

sadness steepened their discount rate for cigarettes much like sadness has been shown to steepen discount rates for monetary reward (5).

Full details of how we calculated the required rate of return (RRR). We calculated required rate of return (RRR) in four steps. First, we identified the switching point where participants switched from preferring the earlier, smaller reward (t1, x1 vs. t2, x2A) to the larger, later reward (t1, x1 vs. t2, x2B). Second, we imputed an indifference point (t1, x1 vs. t2, x2C), where x2C = (x2A + x2B)/2. Third, we computed a discount factor between t1 and t2 where D(t1, t2) = x1/x2C. If the participant always chose the larger, later reward, their discount factor was set to 1 (i.e., no discounting). If the participant never chose the larger, later reward, we set x2C as if they had switched if there were one more line (e.g., if the last line was 9 puffs, we set x2C = 9.5). Last, we "minutized" (i.e., annualized, but over minutes instead of years) the discount factor to identify the RRR by setting -ln(RRR) = D(t1,t2)^(1/(t2-t1)). RRR thus served as a measure for the average increase in number of puffs per minute smokers required in order to wait for a delayed reward, where higher numbers indicated higher levels of impatience.

Table S7. RRR in different studies

Study	Study 3	Study 3 Replication	Study 4
RRR results (full sample)	Smokers in the neutral condition had an RRR of 6.9%, indicating that (on average) they required an increase in puffs of 6.9% per minute to wait for a delayed reward. Smokers in the sadness condition were more impatient: they had an RRR of 8.1%, indicating an 18% increase from smokers in the neutral control.	Smokers in the neutral condition had an RRR of 7.3%, indicating that (on average) they required an increase in puffs of 7.3% per minute in order to wait for a delayed reward. Smokers in the sadness condition were more impatient: they had an RRR of 8.0%, indicating a 9% increase from smokers in the neutral control.	Smokers in the neutral condition had an RRR of 9.1%, indicating that (on average) they required an increase in puffs of 9.1% per minute in order to wait for a delayed reward. Smokers in the sadness condition were more impatient: they had an RRR of 10.2%, indicating a 12.1% increase from smokers in the neutral control.

Study 4

Emotion manipulation checks. As in Studies 2-3, we tested whether the emotion manipulation was effective in both magnitude and specificity. In terms of magnitude, participants in the sadness condition reported higher levels of sadness ($t(156) = -7.94$, $p < 0.001$) and lower levels of neutrality ($t(156) = 3.73$, $p < 0.001$) than individuals in the neutral condition. In terms of specificity, participants in the sadness condition reported feeling more sadness than any other emotion (t's > 2.07, p's < 0.05), except gratitude ($t(81) = 0.52$, $p = 0.604$). Given the relatively low SES characteristics of our participants, it is possible that the relatively large sum participation fee in this study ($50) triggered high levels of gratitude (see information below on participant demographics).

Additional demographics of participants. Compared to participants in Study 3, participants in Study 4 had significantly lower education ($M_{study\ 3} = 3.25$, $SD_{study\ 3} = 1.58$ vs. $M_{study\ 2} = 4.07$, $SD_{study\ 2} = 1.35$, $t(554) = 6.11$, $p < .001$; 51% of participants' highest education was lower than college in Study 4) and less household income ($M_{study\ 3} = 3.81$, $SD_{study\ 3} = 2.92$ vs. $M_{study\ 2} = 5.76$, $SD_{study\ 2} = 2.87$, $t(554) = 7.20$, $p < .001$; The median response of entire household income in 2017 before taxes was between $20,000 to $29,999 in Study 4). Given the relatively low socioeconomic characteristics of participants in Study 4, it is possible that the relatively large sum participation fee in this study ($50) made gratitude as high as sadness for participants in the sadness condition.

Full details of Bayesian analyses. We examined the posterior probability distribution over parameter values and considered the relationship between the highest posterior density (HDP) interval estimates and a region of practical equivalence (ROPE) to the null value (6-7). We set a very weakly informative prior for the fixed effects (normal distribution with mean 0 and standard deviation 10), recommended by Stan Development Team (8; see an example of application in 9) and used the conventional ROPE ±0.1 for Cohen's d (7). We analyzed Studies 3-4 sequentially. Study 3 indicated that the HDP interval estimate for the standardized effect size of sadness is [0.03, 0.34], and the estimated posterior probability of the effect

being larger than the null ROPE is 87%. Using the posterior mode and standard deviation from Study 3 as priors for Study 3 Replication, the HDP interval estimates for the standardized effect size of sadness is [0.05, 0.27], and the estimated posterior probability of the effect being larger than the null ROPE is 88%. Using the posterior mode and standard deviation from Study 3 Replication as priors for Study 4, the updated HDP interval estimates for the standardized effect size of sadness is [0.07, 0.28], and the estimated posterior probability of the effect being larger than the null ROPE is 91%. Therefore, Study 4 strengthens the evidence from Study 3 and Study 3 Replication in supporting our hypothesis.

Unforeseen complications during the course of the study. We ran into many unforeseen complications during the course of the study. We detail these below. Due to this high number of unforeseen violations, we chose to take the most conservative approach and to include all participants in our analyses rather than make post-hoc choices about which participants to exclude.

Table S8. Unforeseen complications in Study 4.

Unforeseen complications	Explanations	Number of participants
Assumed being observed during the experiment/ Uncomfortable being observed smoking	One participant waved at the observation window and said hi to the (assumed) experimenter watching her. One participant expressed feeling uncomfortable with being observed taking the baseline puff and asked the experimenter to not watch.	2
Took more than one baseline puff	Some participants explained that their first breath did not get the smoke through the device.	3
Complained about not passing CO test/waiting	Three participants were annoyed and complained to the experimenter about not passing the CO test and/or waiting too long for their turn to begin the experiment after other participants' being late.	3
Used nicotine patch		1
Computer illiterate	Two participants did not know how to use a computer, including the mouse, keyboard, headphones, and how to answer the Qualtrics survey. They asked the experimenter to teach them. They filled the survey with trials and errors (e.g., mistakenly clicking the "print" button). They asked the experimenter to type for them in the writing task.	2

Checked phone	Although participants were asked to turn off their cell phone, three participants used their phone during the survey.	3
Might not be a daily smoker	One participant said that he did not smoke on the day prior. One participant reported no smoking for 7 days before the experiment but still said that he was a daily smoker when the experimenter asked.	2
Headphones were loose	Three participants buzzed the experimenter in the middle of the video and said that they didn't hear all the sound.	3
Smoked when were not supposed to during the study	One participant misunderstood the instruction and smoked during the course of the study.	1
Smoked without device after choice	Five participants smoked without notifying the experimenter to use the device after making choices.	5
Questionable comprehension	Four participants could not answer comprehension questions correctly even after the experimenter pointed out where they could find answers in the instructions. The experimenter had to tell them the right answer in order to proceed.	4
Weird behavior	One participant behaved weirdly during the survey. About three times, he stopped filling out the survey for about three minutes, shook his head, waved his arms in the air, and made some noise. When the experimenter asked, the participant said he was ok.	1

Additional information on use and validity of online samples.

A recent review revealed that as of 2015, over 1,200 experiments using mturk have been published every year (10), and thus a large empirical base exists for testing questions of validity of online samples. Indeed, PNAS itself offers many papers for examining the validity of online samples. In the first half of the year 2019, PNAS published 9 papers using mturk samples (11-20). Six of these papers used *exclusively* online samples (11-16). At the moment, the weight of evidence reveals that the results from studies with online samples largely mirror the results from matched studies with in-person samples (e.g., 21-27). It can also be helpful to compare results from mturk to other online samples such as Prolific (e.g., 28-29).

References

1. T. R. Barry. The Midlife in the United States (MIDUS) series: A national longitudinal study of health and well-being. *Open Health Data* **2**, e3 (2014).
2. R. Hiscock, L. Bauld, A. Amos, J. A. Fidler, M. Munafò, Socioeconomic status and smoking: A review. *Ann NY Acad Sci* **1248**, 107-123 (2012).
3. N. Breslau, E. L. Peterson, Smoking cessation in young adults: Age at initiation of cigarette smoking and other suspected influences. *Am J Public Health* **86**, 214-220 (1996).
4. E. M. Barbeau, N. Krieger, M. J. Soobader, Working class matters: Socioeconomic disadvantage, race/ethnicity, gender, and smoking in NHIS 2000. *Am J Public Health* **94**, 269-278 (2004).
5. J. S. Lerner, Y. Li, E. U. Weber, The financial costs of sadness. *Psychol Sci* **24**, 72-79 (2013).
6. J. K. Kruschke, Bayesian assessment of null values via parameter estimation and model comparison. *Perspect Psychol Sci* **6**, 299-312 (2011).
7. J. K. Kruschke, Rejecting or accepting parameter values in Bayesian estimation. *Adv Methods Pract Psychol Sci* **1**, 270-280 (2018).
8. Stan Development Team, Prior choice recommendations. Available at https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations (accessed 1 Sep 2019).
9. T. Strandberg, D. Sivén, L. Hall, P. Johansson, P. Pärnamets, False beliefs and confabulation can lead to lasting changes in political attitudes. *J Exp Psychol Gen* **147**, 1382-1399 (2018).
10. J. Bohannon, Mechanical Turk upends social sciences. *Science* **352**, 1263-1264 (2016).
11. J. Cone, K. Flaharty, M. J. Ferguson, Believability of evidence matters for correcting social impressions. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9802-9807 (2019).
12. L. Tieu, P. Schlenker, E. Chemla, Linguistic inferences without words. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9796-9801 (2019).
13. A. Z. Enkavi, I. W. Eisenberg, P. G. Bissett, G. L. Mazza, D. P. MacKinnon, L. A. Marsch, R. A. Poldrack, Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5472-5477 (2019).

14. H. Lu, Y. N. Wu, K. J. Holyoak, Emergence of analogy from relation learning. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4176-4181 (2019)

15. C. M. de Melo, S. Marsella, J. Gratch, Human cooperation when acting through autonomous machines. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3482-3487 (2019).

16. G. Pennycook, D. G. Rand, Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 2521-2526 (2019).

17. M. N. Meyer, P. R. Heck, G. S. Holtzman, S. M. Anderson, W. Cai, D. J.Watts, C. F. Chabris, Objecting to experiments that compare two unobjectionable policies or treatments. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 10723-10728 (2019).

18. B. L. Turner, E. M. Caruso, M. A. Dilich, N. J. Roese, Body camera footage leads to lower judgments of intent than dash camera footage. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 1201-1206 (2019).

19. A. Guterstam, H. H. Kean, T. W. Webb, F. S. Kean, M. S. Graziano, Implicit model of other people's visual attention as an invisible, force-carrying beam projecting from the eyes. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 328-333 (2019).

20. Z. Chen, D. Whitney, Tracking the affective state of unseen persons. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7559-7564 (2019).

21. J. J. Horton, D. G. Rand, R. J. Zeckhauser, The online laboratory: Conducting experiments in a real labor market. *Exp Econ* **14**, 399-425 (2011).

22. D. G. Rand, The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *J. Theor. Biol.* **299**, 172-179 (2012).

23. J. K. Goodman, C. E. Cryder, A. Cheema, Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *J Behav Decis Mak* **26**, 213-224 (2013).

24. G. Paolacci, J. Chandler, Inside the Turk: Understanding Mechanical Turk as a participant pool. *Curr Dir Psychol Sci* **23**, 184-188 (2014).

25. D. J. Hauser, N. Schwarz, Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behav Res Methods* **48**, 400-407 (2016).
26. S. Clifford, R. M. Jewell, P. D. Waggoner, Are samples drawn from Mechanical Turk valid for research on political ideology?. *Res & Polit* **2**, 1-9 (2015).
27. G. Paolacci, J. Chandler, P. G. Ipeirotis, Running experiments on amazon mechanical turk. *Judgm Decis Mak* **5**, 411-419 (2010).
28. S. Palan, C. Schitter, Prolific. ac—A subject pool for online experiments. *J Behav Exp Finance* **17**, 22-27 (2018).
29. E. Peer, L. Brandimarte, S. Samat, A. Acquisti, Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *J Exp Soc Psychol* **70**, 153-163 (2017).