Incentives Boost Model-based Control Across a Range of Severity on Several Psychiatric Constructs

Supplemental Information

Participants

Participants agreed to participate in a two-day study and were excluded from the second day for random responses or demonstrated non-adherence. They were 77.9% White/Caucasian, 6.1% Asian or Asian American, 0.5% Native American or Other, 9.4% African American or Black, and 6.1% Hispanic or Latino. Household income was 16.6% less than $20,000, 30.1% $20,000-$40,000, 23.6% $40,000-$60,000, 15.3% $60,000-$80,000, and 14.4% greater than $80,000. It has been shown that asking a commitment question can improve participant honesty (1), therefore we asked participants “It is important to us that you answer questions honestly. Will you answer the questions honestly?”.

On day one, participants completed a progressive-adaptive version of Raven’s progressive matrices IQ test (2), the novel two-step paradigm, and 9 clinical scales in randomized order. On day two, participants completed an additional behavioral task and 10 additional clinical scales in randomized order. The adaptive IQ test is comprised of five questions, progressing easier with incorrect responses. Therefore, decreasing IQ indicates random responding or non-adherence and these participants were excluded from analyses (N = 67). Participants were excluded for repeating the study or timing out on > 20% of task trials (N = 30). Lastly, subjects were excluded for non-estimable parameters in the computational model (N = 2) and drug or alcohol intoxication at the time of study completion (N = 3). The final sample varied slightly in size due to missing observations between N = 810 to N = 839.

Gaussian Random Walk

One alien planet reward distribution was initialized randomly between 0 and 4 points and the other between 5 and 9 points. The reward distributions for both planets drifted between 0 and 9 according to a Gaussian random walk with a standard deviation of 2. Each participant received a new randomly generated reward distribution.

Computational Model

We modeled the two-step task using a reinforcement learning model that captures the mixture of model-based and model-free control exhibited by humans (3–5). The model free algorithm is SARSA(λ) (6), which updates the state-action pair (*s,a*) at step *i* and trial *t* in the following manner:

$$Q_{MF}(s,a) \leftarrow Q_{MF}(s,a) + \alpha\delta_{i,t}e_{i,t}(s,a)$$

where the prediction error is defined as:

$$\delta_{i,t} = r_{i,t} + Q_{MF}(s_{i+1,t}, a_{i+1,t}) - Q_{MF}(s_{i,t}, a_{i,t}).$$

The learning rate a captures the rate at which an individual incorporates reward feedback (i.e. low a leads to slow incorporation and high a leads to quick incorporation) and $e_{i,t}(s,a)$ is an eligibility trace updated in the following way:

$$e_{i,t}(s,a) \leftarrow e_{i-1,t}(s_{i,t}, a_{i,t}) + 1.$$

The eligibility trace allows step 1 values to be partially updated following step 2 outcomes. Note that we could substitute this point estimation scheme with a Bayesian updating scheme, such as the hierarchical Gaussian filter (7) or a Kalman filter (8), though this would not affect the conclusions of our paper and therefore we opted for the simpler point estimation scheme.

The model-based strategy uses a learned transition function to compute values, rather than retrieving them from a cache:

$$Q_{MB}(s_A, a_j) = P(s_B|a_j)\, Q_{MF}\,(s_B) + P(s_C|a_j)\, Q_{MF}(s_C)$$

where $P(s_B|a_j)$ is the probability of transitioning to state S_B after choosing action a_j, and $Q_{MF}\,(s_B)$ and $Q_{MF}(s_C)$ are the immediate reward estimates at the second-step states (note that these are not dependent on actions since there is no second-step action.

The mixture of model-free and model-based control is governed by a weighting parameter w:

$$Q_{net}(s_A, a_j) = wQ_{MB}(s_A, a_j) + (1 - w)Q_{MF}(s_A, a_j)$$

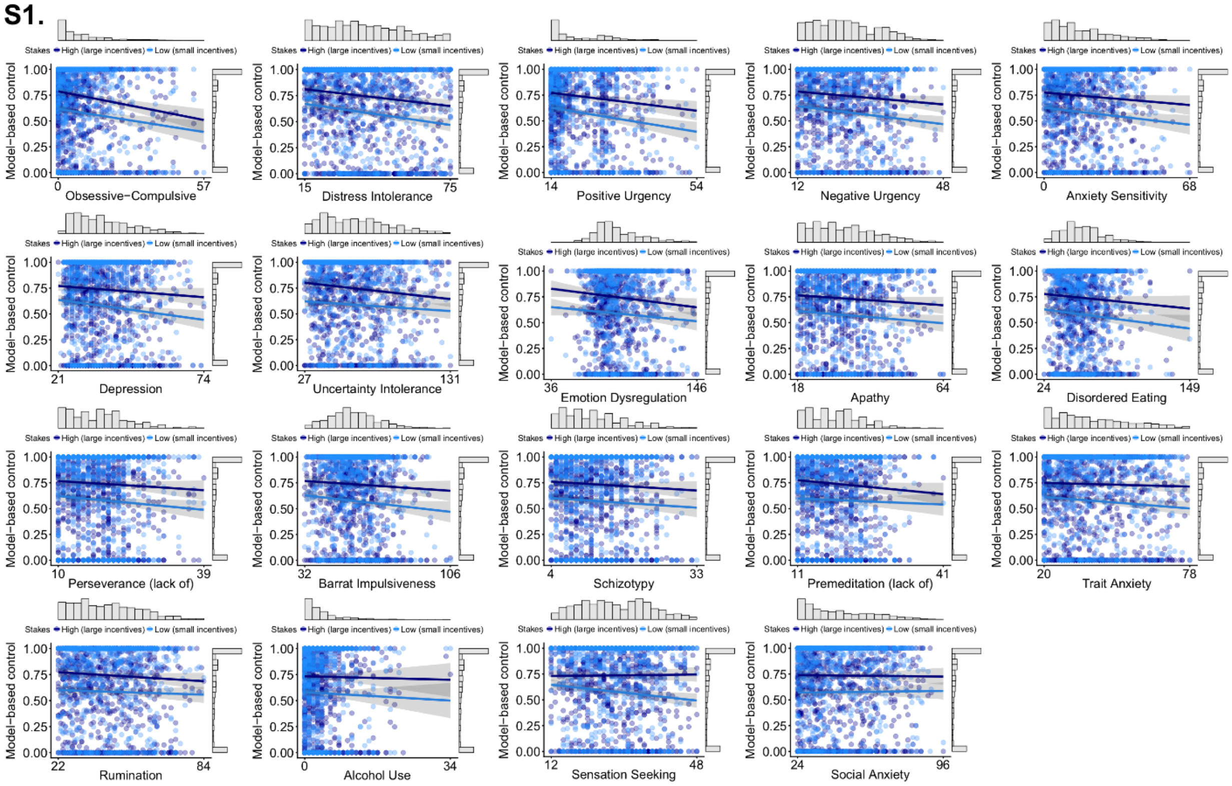
Importantly, *w* was fit independently for the low stakes and high stakes conditions to allow for comparison of model-based control across stakes. Kool and colleagues (3) compared various model formulations of the novel two-step and found that the current model outperformed other approaches which allowed other parameters to vary (e.g., choice inverse temperature). They showed that the change in model-based control due to stakes cannot be completely explained by other model parameters.

Finally, the probability of choosing action *a* given state *s* on a trial is an adapted form of Luce's choice rule:

$$P(s_{i,t}) = \frac{exp(\beta[Q_{net}(s_{i,t}, a) + \pi \cdot rep(a) + \rho \cdot resp(a)])}{\sum_{a'} exp(\beta[Q_{net}(s_{i,t}, a') + \pi \cdot rep(a') + \rho \cdot resp(a')])},$$

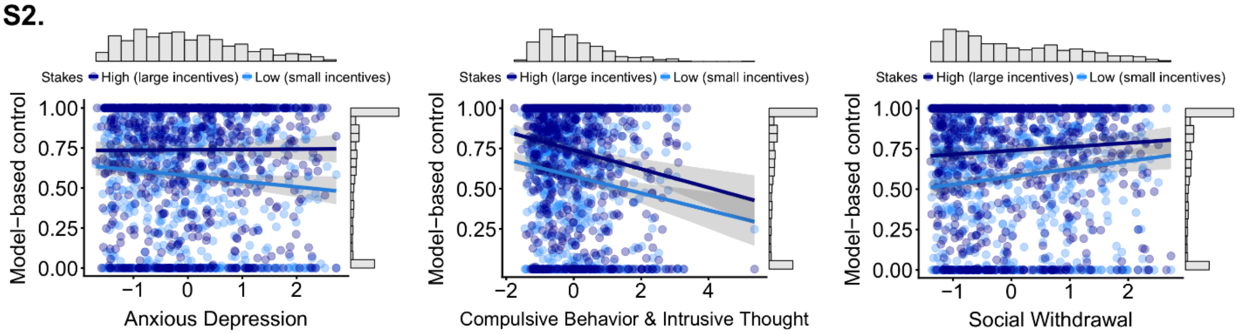
where we have included two parameters to capture additional variance: motor stickiness p and stimulus stickiness r (choosing the same stimulus twice independent of its value). b represents the inverse temperature parameter that controls choice stochasticity, with lower values promoting exploration and higher values promoting exploitation. Parameters were optimized using maximum a posteriori estimation with empirical priors (3, 9). See Kool et al. (2017) for additional details concerning task design and computational model-fitting.

S1.



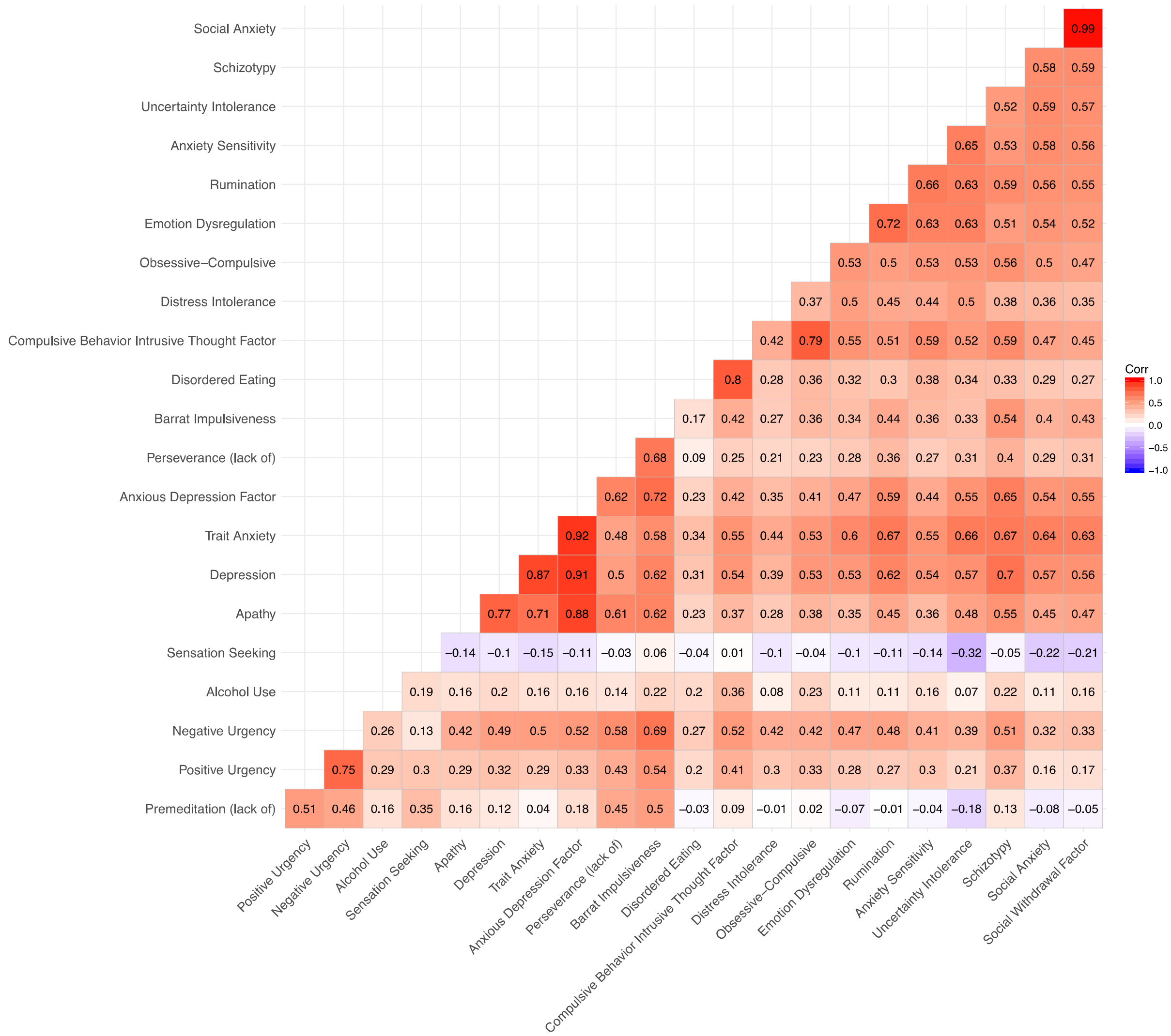
Supplementary Figure S1: Relationship between self-report measures and model-based control; regression lines with credible intervals. Marginal histograms show distribution of data. Self-report measure observed min and max on x-axis.

S2.

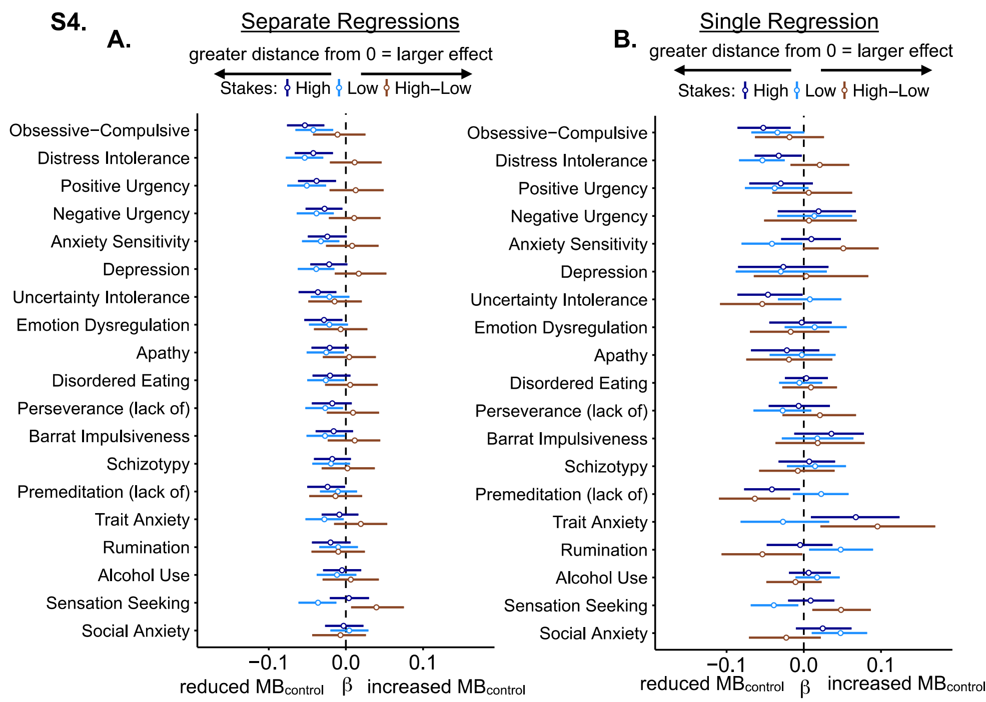


Supplementary Figure S2: Relationship between factors and model-based control; regression lines with credible intervals. Marginal histograms show distribution of data. Self-report measure observed min and max on x-axis.

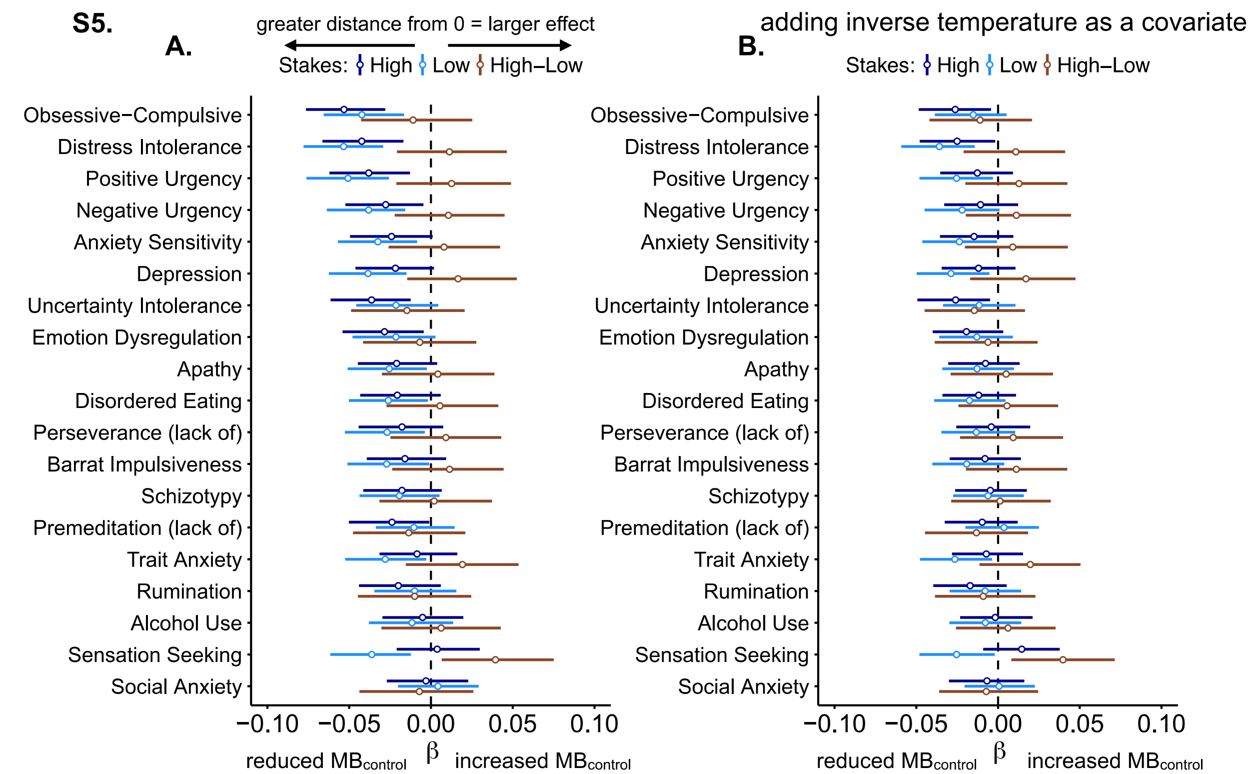
S3.



Supplementary Figure S3: Correlation matrix of all constructs.

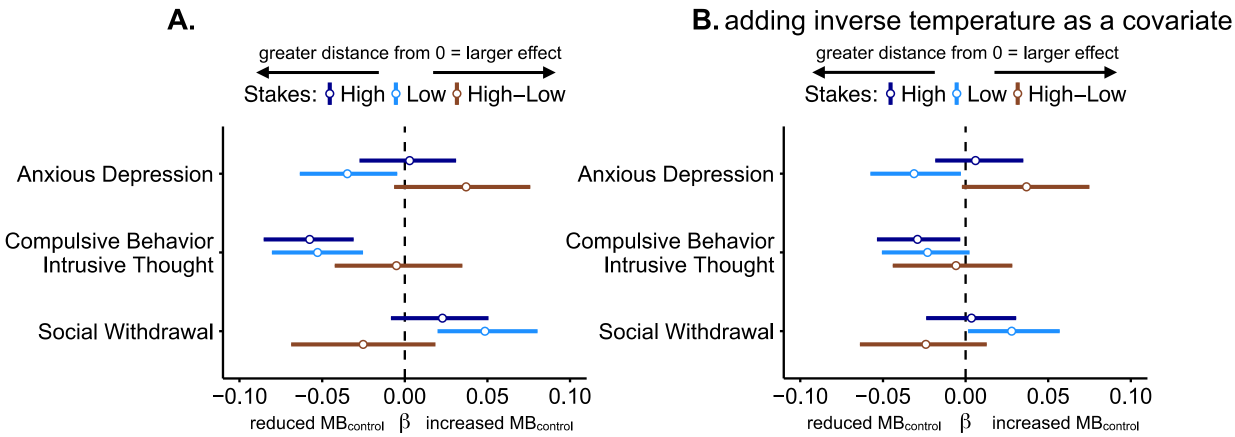


Supplementary Figure S4: Coefficient plots with credible intervals containing 95% of the posterior probability density around the mean. A. Each self-report measure entered into a separate regression covarying age, IQ, and gender (Figure 3A main text). B. All self-report measures entered simultaneously into a single regression.



Supplementary Figure S5: Coefficient plots with credible intervals containing 95% of the posterior probability density around the mean. A. Each self-report measure entered into a separate regression covarying age, IQ, and gender (Figure 3A main text). B. Added inverse temperature as a covariate to regression model for each self-report measure.

S6.



Supplementary Figure S6: Coefficient plots with credible intervals containing 95% of the posterior probability density around the mean. A. All factors entered into a single regression covarying age, IQ, and gender (Figure 3A main text). B. Added inverse temperature as a covariate to regression model.

Supplementary Table S1. Exact language from the cited source articles, and abbreviated definitions were used for the construct description. Factor loadings from the Gillan et al., 2016 supplement were used to generate factor scores. Note, Gillan and colleagues reversed scored Alcohol Use items: 2, 3, and 8; however, these are not reverse scored on the original measure or for our analyses.

Brief descriptions of constructs

Construct	Description
Obsessive-Compulsive[a]	OCD symptoms (checking, washing, obsessing, mental neutralizing ordering, hoarding, doubting) (10)
Distress Intolerance	Capacity to withstand and experience negative psychological states (11)
Positive Urgency	Tendency to act rashly or maladaptively in response to positive mood states (12)
Negative Urgency	Strong immediate need to avoid negative emotions or physical sensations (13)
Anxiety Sensitivity	Tendency to respond fearfully to physiological cues of anxiety (e.g. increased heart rate) (14)
Depression[a]	Affective (e.g. sad), physiological (e.g. sleep disturbance), and psychological (e.g. hopeless) symptoms (15)
Uncertainty Intolerance	Tendency to consider possibility of negative event unacceptable, regardless of likelihood (16)
Emotion Dysregulation	Lack of emotional arousal, awareness, understanding, acceptance, and perseverance when emotional (17)
Apathy[a]	Lack of motivation not due to diminished consciousness, cognitive impairment, or emotional distress (18)
Disordered Eating[a]	Dieting, bulimia and food preoccupation, and oral control (19)
Perseverance (lack of)	Difficulty maintaining attention and vulnerability to intrusive and interfering information (13)
Barratt Impulsiveness[a]	Dimensions of impulsivity (attentional, motor, non-planning impulsiveness) (20)
Schizotypy[a]	Unusual experiences, cognitive disorganization, introvertive anhedonia, and impulsive non-conformity (21)
Premeditation (lack of)	Difficulty considering the long-term consequences of actions (13)
Trait Anxiety[a]	Stable tendency to experience and attend to negative emotions (22)
Rumination	Thinking repetitively and passively about negative emotions, with a focus on distress (23)
Alcohol Use[a]	Hazardous and harmful alcohol consumption, drinking behavior, and alcohol related problems (24)
Sensation Seeking	Willingness to take risks (financial, legal, physical) for novel-intense experiences (13)
Social Anxiety[a]	Anxiety and avoidance of social situations likely to induce fear of evaluation (25)

[a.] scale included in Gillan et al. (2016)

Supplementary Table S2. Means and standard deviations all self-report measures.

Construct	Mean	(SD)
Alcohol Use	4.03	(4.7)
Anxiety Sensitivity	17.26	(13.57)
Apathy	32.56	(9.93)
Barratt Impulsiveness	57.05	(12.36)
Depression	37.49	(10.48)
Disordered Eating	56.61	(17.8)
Distress Intolerance	39.32	(15.79)
Emotion Dysregulation	86.6	(17.45)
Negative Urgency	23.98	(7.88)
Obsessive Compulsive	9.96	(10.81)
Perseverance (lack of)	18.22	(5.81)
Positive Urgency	22.52	(8.73)
Premeditation (lack of)	18.85	(5.42)
Rumination	42.75	(13.73)
Schizotypy	12.55	(5.98)
Sensation Seeking	27.78	(8.56)
Social Anxiety	43.97	(16.59)
Trait Anxiety	39.36	(14.54)
Uncertainty Intolerance	63.74	(23.66)

Supplemental References

1. Clifford S, Jerit J (2016): Cheating on Political Knowledge Questions in Online Surveys. *Public Opin Q*. 80: 858–887.

2. Raven J (2000): The Raven's Progressive Matrices: Change and Stability over Culture and Time. *Cogn Psychol*. 41: 1–48.

3. Kool W, Gershman SJ, Cushman F (2017): Cost-benefit arbitration between multiple reinforcement learning systems. *Psychol Sci*. 28: 1321–1333.

4. Kool W, Cushman FA, Gershman SJ (2016): When Does Model-Based Control Pay Off? *PLOS Comput Biol*. 12: 1–34.

5. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011): Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 69: 1204–1215.

6. Sutton RS, Barto AG, Book AB (1998): *Reinforcement learning: An Introduction*. Cambridge, MA: MIT Press.

7. Mathys C (2011): A Bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci*. 5: 1–20.

8. Gershman SJ (2015): A Unifying Probabilistic View of Associative Learning. *PLOS Comput Biol*. 1–20.

9. Gershman SJ (2016): Empirical priors for reinforcement learning models. *J Math Psychol*. 71: 1–6.

10. Foa EB, Huppert JD, Leiberg S, Langner R, Kichic R, Hajcak G, Salkovskis PM (2002): The obsessive-compulsive inventory: Development and validation of a short version. *Psychol Assess*. 14: 485–496.

11. Simons JS, Gaher RM (2005): The distress tolerance scale: Development and validation of a self-report measure. *Motiv Emot*. 29: 83–102.

12. Cyders MA, Smith GT, Spillane NS, Fischer S, Annus AM, Peterson C (2007): Integration of impulsivity and positive mood to predict risky behavior: Development and validation of a measure of positive urgency. *Psychol Assess*. 19: 107–118.

13. Whiteside SP, Lynam DR (2001): The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. *Pers Individ Dif*. 30: 669–689.

14. Taylor S, Zvolensky MJ, Cox BJ, Deacon B, Heimberg RG, Ledley DR, *et al.* (2007): Robust dimensions of anxiety sensitivity: Development and initial validation of the Anxiety Sensitivity Index-3. *Psychol Assess*. 19: 176–188.

15. Zung W, Durham N (1965): A Self-Rating Depression Scale. *Arch Gen Psychiatry*. 12: 63–70.

16. Carleton RN, Norton MAPJ, Asmundson JG (2007): Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *J Anxiety Disord*. 21: 105–117.

17. Gratz KL, Roemer L (2004): Multidimensional assessment of emotion regulation and dysregulation. *J Psychopathol Behav Assess*. 26: 41–54.

18. Marin R, Biedrzycki R, Firinciogullari S (1991): Reliability and validity of the apathy evaluation scale. *Psychiatry Res*. 38: 143–162.

19. Garner D, Olmsted MP, Bohr Y, Garfinkel P (1982): The eating attitudes test: psychometric features and clinical correlates. *Psychol Med*. 12: 871–878.

20. Patton JH, Stanford MS, Barratt ES (1995): Factor structure of the Barratt impulsiveness scale. *J Clin Psychol*. 51: 768–74.

21. Mason O, Linney Y, Claridge G (2005): Short scales for measuring schizotypy. *Schizophr Res*. 78: 293–296.

22. Spielberger CD, Gorsuch RL, Lushene PR, Vagg PR, Jacobs AG (1983): *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press.

23. Nolen-Hoeksema S, Morrow J (1991): A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta earthquake. *J Pers Soc Psychol*. 61: 115–121.

24. Saunders JB, Aasland OG, Babor TF, De La Fuente JR, Grant M (1993): Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption. II. *Addiction*. 88: 791–804.

25. Liebowitz MR (1987): Social phobia. *Mod Probl Pharmacopsychiatry*. 22: 141–173.