Supplemental Information

Methods:

Reward Task:

For guessing trials, participants pressed a button to guess whether the value of a card would be higher or lower than '5' (3000 msec). Next, participants viewed the actual value of the card (possible value of 1 to 9; 500 msec) and outcome feedback (Win: green upward-facing arrow; Loss: red downward-facing arrow; 500 msec). At the end of the trial, participants viewed a fixation cross (3000 msec). For control trials, participants pressed a button to the letter "X" (3000 msec) and then viewed an asterisk (500 msec), yellow circle (500 msec), and fixation cross (3000 msec). The entire task lasted approximately 6 minutes.

The task was a block design comprised of 3 win (80% win, 20% loss trials), 3 loss (80% loss, 20% win trials), and 3 control (no change in win/loss) blocks. Each control block consisted of 6 control trials; each guessing block consisted of 5 trials presented in an oddball format (Win block: win, win, win, loss, win; Loss block: loss, loss, win, loss, loss). The experimenter misled participants to believe that task performance determined outcome. Before the scan, participants practiced in an fMRI simulator. The importance of optimizing performance and remaining still was emphasized during both practice and the scan.

Gray matter regions:

Resulting gray matter variables included bilateral cortical thickness measures of: banks of the superior temporal sulcus, caudal middle frontal gyrus, caudal anterior cingulate gyrus, cuneus cortex, entorhinal gyrus, fusiform gyrus, inferior parietal cortex, inferior temporal cortex, isthmus of the anterior cingulate cortex, lateral occipital cortex, lateral orbitofrontal cortex, lingual gyrus, medial orbitofrontal cortex, middle temporal cortex, parahippocampal cortex, paracentral cortex, pars opercularis of the inferior frontal gyrus, pars orbitalis of the inferior frontal gyrus, pars triangularis of the inferior frontal gyrus, pricalcarine cortex, postcentral gyrus, posterior cingulate cortex, precentral cortex, precuneus, rostral anterior cingulate gyrus, rostral middle frontal cortex, superior frontal cortex, superior parietal cortex,

superior temporal cortex, supramarginal gyrus, frontal pole, temporal pole, transverse temporal gyrus, and insula.

Optimal lambda selection:

k-fold cross validation was uses to determine the mean squared error associated with each lambda (k=10). Optimal lambda selection was based on chi square comparisons of residual deviance for each model.

	Lambda	mean error	residual deviance	degrees of freedom	AIC
Lambda.min	0.06267	1.0914	26.315	60	52.315
Lambda.1se	0.10952	1.1787	55.151	67	67.151
optimal lambda	0.07548	1.1174	33.593	64	51.593

We chose to use the optimal lambda as our penalty variable for the following reasons. Lambda.min with 12 non-zero variables was not a significantly better model fit than optimal lambda with 8 variables (chi square = 7.278 with 4 d.f. $p > .1$) The sparsity principle tells us to use the more parsimonious model. The optimal lambda with 8 non-zero variables was a significantly better model fit than the lambda.1se with 5 non-zero variables (chi square=21.558 with 3 d.f. $p < .001$).

LASSO analysis including movement as a predictor of substance use

We performed an additional LASSO analysis including all of the variables from the main analysis with the addition of peak x,y,z movement. The results at lambda.1se showed that the same variables from the main analysis (left mPFC to win, left insula to loss, left caudal anterior cingulate thickness, parental

stress, antipsychotic medication, age, depression, and mania scores) plus peak movement predicted future substance use. In this model less movement during the reward task predicted future substance use. This finding suggests that future substance users may be concentrating and interesed in the potential to win or lose during the reward task which is perhaps reflected in greater stillness during the task.