Supplemental Materials for: Evidence accumulation and associated error-related brain activity as computationally-informed prospective predictors of substance use in emerging adulthood

Alexander Weigard, Ph.D.[1], Sarah J. Brislin, Ph.D.[1], Lora M. Cope, Ph.D.[1], Jillian E. Hardee, Ph.D.[1], Meghan E. Martz, Ph.D.[1], Alexander Ly, Ph.D.[2,3], Robert A. Zucker, Ph.D.[1], Chandra Sripada. M.D., Ph.D.[1], & Mary M. Heitzeg, Ph.D.[1]

[1]Department of Psychiatry, University of Michigan

[2]Department of Psychological Methods, University of Amsterdam

[3]Machine Learning Group, Centrum Wiskunde & Informatica

Go/No-Go Task

Participants completed an event-related go/no-go task[1–3] during fMRI data collection in which they were presented with a string of letters (white on black background) that indicated whether they should respond (any letter other than "X"; 75% of trials) or inhibit their response ("X"; 25% of trials). Letters were presented for 500ms (3500ms interstimulus fixation interval) during 5 imaging runs of 49 trials each (245 trials total; 60 "X" trials).

MRI Data Acquisition Parameters

A high-resolution T1-weighted anatomical image was obtained using the following parameters: three-dimensional spoiled gradient-recalled echo, TR=25ms, minimum TE, FOV=25cm, 256x256 matrix, slice thickness=1.4mm. During runs of the go/no-go task, whole brain T2*-weighted functional images were acquired using a single-shot spiral in-out sequence[4] with the following parameters: TR=2000ms, TE=30ms, flip angle=90°, FOV=200mm, 29 axial slices, 64×64 matrix, in-plane resolution=3.12mm×3.12mm, and slice thickness=4mm. All scans were conducted with the same 3.0 T GE Signa scanner.

Pre-Processing and Single-Subject fMRI Analyses

Functional images were reconstructed using an iterative algorithm[5] and entered into the following pre-processing steps: 1) motion correction with realignment using FSL 5.0.2.2 tools (FMRIB, Oxford, United Kingdom), 2) spatial normalization to standard space as defined by the Montreal Neurological Institute template using Statistical Parametric Mapping 8 (SPM8: Wellcome Institute of Cognitive Neurology, London, United Kingdom) and using normalization of the T1-weighted anatomical image for guidance, 3) resampling to 2x2x2mm voxels in SPM8, and 4) spatial smoothing with a 6mm full-width half-maximum Gaussian kernel. Functional runs

were excluded from further analysis if they exceeded 3 mm translation or 3° rotation in any direction during the run.

A general linear model was fit in SPM8 to individual subjects' fMRI time series data with three regressors convolved with the hemodynamic response function: 1) correct "go" trials, 2) successful inhibition (SI) trials, in which participants withheld their response to a "no-go" stimuli, and 3) failed inhibition (FI) trials, in which participants made a response following "no-go" stimuli. Motion parameters from earlier realignment and average white matter signal intensity for each volume were also included as nuisance regressors. Individual statistical maps for the primary "error monitoring" contrast of interest (FI > correct "go") were generated for later analyses.

Correlations Between Raw and Composite Substance Use Measures

We investigated simple correlations between all individual and composite substance use outcome measures of interest from ages 22-26 in order to 1) evaluate whether they were moderately correlated with each other, as would be expected given prior research, and 2) ensure that our substance use composite (SUC) measure was well-representative of the use of all three substances. We also did the same with measures of prior cumulative use of all three substances from age 17 and the prior substance use composite (preSUC) that was utilized as a covariate in the primary analyses. Supplemental Table 1 displays Bayesian estimates of correlation coefficients between all of these variables. Inspection of these values indicates that, as expected, measures of average use of all three substances from ages 22-26 are moderately correlated with one another, and measures of cumulative use by age 17 also show strong interrelationships. Furthermore, both the SUC and preSUC show strong, and roughly equal, correlations with the three substance use measures that each composite measure was derived from, suggesting that

these composites provide representative indices of the use of common substances for specific developmental periods.

Sensitivity Analyses

To assess whether our findings were generally robust to the inclusion of prior substance use and other covariates in our prediction models, we conducted two sensitivity analyses using the same frequentist and Bayesian methods that were utilized in the primary analyses. First, we conducted prediction analyses which used measures of cumulative use of individual substances by age 17 (alcohol, marijuana, cigarettes) as covariates, in place of the preSUC, in order to evaluate whether our findings would hold when prior use of these substances was accounted for individually (Supplemental Tables 2 and 3). Next, we conducted prediction analyses without any of our previous covariates to evaluate whether our results were still robust even when models did not account for other relevant risk factors (Supplemental Tables 4 and 5).

Results of the first sensitivity analysis are highly similar to the results of our primary analyses reported in the manuscript; both predictors of interest, *v.avg* and PC1, show statistically significant relationships with the SUC outcome in frequentist tests, and Bayesian model comparison indicates substantial evidence for the inclusion of both predictors of interest in the model. Results of our second sensitivity analysis, without any covariates, were also similar, although there was slightly less evidence for the inclusion of the neural-level measure (PC1). PC1 was only a significant predictor in frequentist tests when *v.avg* was not also included in the model, and evidence for the inclusion of PC1 in Bayesian models was weaker than in the primary analyses. However, the best-fitting model was still one which contained both predictors of interest, rather than *v.avg* only. Taken together, results from these sensitivity analyses suggest that our primary results are generally robust to the inclusion, vs. exclusion, of covariates in our

regression models, and to alterations in the measurement of the prior substance use covariates, specifically.

Supplemental Tables and Figures

Supplemental Table 1. Correlations between measures of the use of individual substances, both averaged over ages 22-26 and cumulative use (Cu.) by age 17, as well as with the age 22-26 average substance use composite (SUC) and age 17 prior substance use composite (preSUC). Large-font numbers indicate the median of the Bayesian posterior distribution of the correlation coefficient, representing the most likely correlation value, while smaller-font numbers in italics indicate the 95% credible intervals of the posterior distribution, which represent the upper and lower bounds of the range in which there is a .95 probability that the correlation coefficient falls. DV = annual volume of alcoholic drinks (standard beverages); MF = annual marijuana use frequency (days of use); CF = annual cigarette use frequency (days of use)

	DV (22-26)	MF (22-26)	CF (22-26)	SUC (22-26)	Cu. DV (17)	Cu. MF (17)	Cu. CF (17)
DV (22-26)	—						
	—						
	—						
MF (22-26)	0.24	—					
	0.41	—					
	0.05	—					
CF (22-26)	0.34	0.29	—				
	0.49	*0.45*	—				
	0.16	*0.10*	—				
SUC (22-26)	0.73	0.70	0.75	—			
	0.80	*0.78*	*0.82*	—			
	0.62	*0.59*	*0.64*	—			
Cu. DV (17)	0.60	0.28	0.26	0.53	—		
	0.71	*0.45*	*0.43*	*0.65*	—		
	0.46	*0.10*	*0.08*	*0.37*	—		
Cu. MF (17)	0.40	0.39	0.24	0.47	0.71	—	
	0.55	*0.53*	*0.41*	*0.60*	*0.79*	—	
	0.23	*0.21*	*0.05*	*0.30*	*0.60*	—	
Cu. CF (17)	0.29	0.24	0.33	0.39	0.70	0.64	—
	0.45	*0.40*	*0.49*	*0.54*	*0.79*	*0.73*	—
	0.11	*0.05*	*0.15*	*0.22*	*0.59*	*0.50*	—
preSUC (17)	0.49	0.34	0.31	0.52	0.91	0.88	0.88
	0.61	*0.49*	*0.47*	*0.64*	*0.94*	*0.92*	*0.91*
	0.32	*0.16*	*0.13*	*0.36*	*0.86*	*0.83*	*0.82*

Supplemental Table 2. Results from our first sensitivity analysis, which included measures of prior use of individual substances as covariates rather than a prior use composite, involving frequentist regressions that predicted values of the age 22-26 substance use composite (SUC) with models that included 1) *v.avg*, 2) PC1, and 3) both predictors of interest, along with covariates. **Bolded** *p*-values survive false discovery rate correction for multiple comparisons within families defined by the individual regression models. Overall variance explained by each model (R^2) is displayed in parentheses. FH-AUD = family history of alcohol use disorder (either parent); Cu. DV = cumulative drink volume at age 17; Cu. MJ = cumulative marijuana use at age 17; Cu. CF – cumulative cigarette use at age 17

Model (R^2)		**Unstandardized**	**Standard Error**	**Standardized**	***t***	***p***
v.avg	(Intercept)	0.400	0.197			
(.385)	Sex	-0.255	0.119	-0.173	-2.150	0.034
	FH-AUD	0.061	0.134	0.037	0.456	0.649
	Cu. DV (17)	3.412e -4	1.216e -4	0.359	2.805	**0.006**
	Cu. MJ (17)	0.001	6.519e -4	0.203	1.730	0.087
	Cu. CF (17)	2.418e -6	2.460e -4	0.001	0.010	0.992
	v.avg	-0.202	0.068	-0.240	-2.963	**0.004**
PC1	(Intercept)	-0.124	0.127			
(.404)	Sex	-0.211	0.118	-0.143	-1.787	0.077
	FH-AUD	0.076	0.132	0.046	0.574	0.567
	Cu. DV (17)	3.788e -4	1.203e -4	0.399	3.150	**0.002**
	Cu. MJ (17)	0.001	6.417e -4	0.201	1.742	0.085
	Cu. CF (17)	1.602e -5	2.415e -4	0.008	0.066	0.947
	PC1	-0.096	0.027	-0.280	-3.493	**< .001**
v.avg	(Intercept)	0.238	0.198			
and	Sex	-0.205	0.115	-0.139	-1.775	0.079
PC1	FH-AUD	0.124	0.131	0.076	0.946	0.347
(.435)	Cu. DV (17)	3.743e -4	1.176e -4	0.394	3.182	**0.002**
	Cu. MJ (17)	0.001	6.277e -4	0.209	1.846	0.068
	Cu. CF (17)	-3.660e -5	2.372e -4	-0.018	-0.154	0.878
	v.avg	-0.159	0.067	-0.188	-2.354	**0.021**
	PC1	-0.082	0.028	-0.239	-2.971	**0.004**

Supplemental Table 3. Results from our first sensitivity analysis, which included measures of prior use of individual substances as covariates rather than a prior use composite, involving Bayesian regressions in which all possible models involving the predictors of interest, drift rate (*v.avg*) and error-related activation (PC1), were compared to a "null" model that included only the covariates of sex, family history of alcohol use disorder (FH-AUD) and separate measures of prior alcohol (DV), marijuana (MJ) and cigarette use (CF). In the "Model Comparison" section, P(M) is prior probability of the model, P(M|data) is the posterior probability of the model after seeing the data, BF_{10} is the Bayes factor comparing the model to the "null" model, and BF_M is a Bayes factor comparing the model to all other models from the analysis. The "Posterior Summaries" section reports the model-averaged mean, standard deviation (SD) and 95% credible intervals of posterior samples for coefficients of each predictor of interest, as well as inclusion probabilities obtained from model averaging; P(inc) is the prior probability of including each predictor, P(inc|data) is the posterior probability of including each predictor, and BF_{inc} is a Bayes factor for the change from prior to posterior inclusion odds for the predictor after seeing the data.

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	R^2
"Null" model (incl. Sex, FH-AUD, Cu. DV, Cu. MJ, Cu. CF)	0.250	0.004	0.012	1.000	0.330
v.avg + PC1	0.250	0.724	7.873	180.487	0.435
PC1	0.250	0.219	0.840	54.506	0.404
v.avg	0.250	0.053	0.169	13.271	0.385

Posterior Summaries of Coefficients

						95% Credible Interval	
Coefficient	**Mean**	**SD**	**P(inc)**	**P(inc|data)**	**BF_{inc}**	**Lower**	**Upper**
v.avg	-0.144	0.064	0.500	0.777	3.491	-0.272	-0.017
PC1	-0.074	0.026	0.500	0.943	16.467	-0.126	-0.022

Supplemental Table 4. Results from our second sensitivity analysis, which included no covariates for other substance use risk factors, involving frequentist regressions that predicted values of the age 22-26 substance use composite (SUC) with models that included 1) *v.avg*, 2) PC1, and 3) both predictors of interest. **Bolded** *p*-values survive false discovery rate correction for multiple comparisons within families defined by the individual regression models. Overall variance explained by each model (R^2) is displayed in parentheses.

Model (R^2)		**Unstandardized**	**Standard Error**	**Standardized**	***t***	***p***
v.avg	(Intercept)	0.582	0.210			
(.076)	*v.avg*	-0.233	0.080	-0.276	-2.931	**0.004**
PC1	(Intercept)	0.020	0.069			
(.059)	PC1	-0.083	0.033	-0.243	-2.554	**0.012**
v.avg	(Intercept)	0.501	0.211			
+ **PC1**	*v.avg*	-0.195	0.081	-0.230	-2.403	**0.018**
(.109)	PC1	-0.064	0.033	-0.186	-1.944	0.055

Supplemental Table 5. Results from our second sensitivity analysis, which included no covariates for other substance use risk factors, involving Bayesian regression analyses in which all possible models involving the predictors of interest, drift rate (*v.avg*) and error-related activation (PC1), were compared to a "null" model that included only the regression intercept parameter. In the "Model Comparison" section, P(M) is prior probability of the model, P(M|data) is the posterior probability of the model after seeing the data, BF_{10} is the Bayes factor comparing the model to the "null" model, and BF_M is a Bayes factor comparing the model to all other models from the analysis. The "Posterior Summaries" section reports the model-averaged mean, standard deviation (SD) and 95% credible intervals of posterior samples for coefficients of each predictor of interest, as well as inclusion probabilities obtained from model averaging; P(inc) is the prior probability of including each predictor, P(inc|data) is the posterior probability of including each predictor, and BF_{inc} is a Bayes factor for the change from prior to posterior inclusion odds for the predictor after seeing the data.

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	R^2
"Null" model (incl. Sex, FH-AUD, PSUC)	0.250	0.04	0.124	1.000	0.000
v.avg + PC1	0.250	0.472	2.680	11.903	0.109
v.avg	0.250	0.346	1.584	8.717	0.076
PC1	0.250	0.143	0.500	3.605	0.059

Posterior Summaries of Coefficients

						95% Credible Interval	
Coefficient	Mean	SD	P(inc)	P(inc|data)	BF_{inc}	Lower	Upper
v.avg	-0.174	0.077	0.500	0.817	4.478	-0.326	-0.022
PC1	-0.057	0.031	0.500	0.615	1.596	-0.119	0.004

Supplemental References

1. Weigard, A. et al. Cognitive modeling informs interpretation of go/no-go task-related neural activations and their links to externalizing psychopathology. http://biorxiv.org/lookup/doi/10.1101/614420 (2019) doi:10.1101/614420.
2. Heitzeg, M. M. et al. Left middle frontal gyrus response to inhibitory errors in children prospectively predicts early problem substance use. Drug Alcohol Depend. **141**, 51–57 (2014).
3. Durston, S., Thomas, K. M., Worden, M. S., Yang, Y. & Casey, B. J. The Effect of Preceding Context on Inhibition: An Event-Related fMRI Study. NeuroImage **16**, 449–453 (2002).
4. Glover, G. H. & Law, C. S. Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. Magn. Reson. Med. **46**, 515–522 (2001).
5. Fessler, J. A., Sangwoo Lee, Olafsson, V. T., Shi, H. R. & Noll, D. C. Toeplitz-based iterative image reconstruction for MRI with correction for magnetic field inhomogeneity. IEEE Trans. Signal Process. **53**, 3393–3402 (2005).