Development of Impulse Control Circuitry in Children of Alcoholics

Supplementary Information

Participants

Mean participant age at each scan is shown in Table S1. Independent samples *t*-tests revealed no significant differences between FH- and FH+ groups at each scan (scan 1: $F = 0.36$, $p = 0.78$; scan 2: $F = 0.01$; $p = 0.82$; scan 3: $F = 0.26$, $p = 0.35$; scan 4: $F = 0.25$, $p = 0.62$).

Table S1. Participant age (in years) by scan.

	All Participants		FH-		FH+	
	n	**Mean ± SD**	***n***	**Mean ± SD**	***n***	**Mean ± SD**
Age at scan 1	73	10.28 ±1.16	30	10.32 ±1.09	43	10.24 ±1.21
Age at scan 2	73	12.25 ±1.28	30	12.29 ±1.29	43	12.22 ±1.29
Age at scan 3	39	14.27 ±1.10	19	14.10 ±1.01	20	14.44 ±1.19
Age at scan 4	14	15.20 ±0.87	6	15.06 ±0.92	8	15.31 ±0.89

Participants were divided into two groups: those with at least one parent who had an alcohol use disorder (AUD) diagnosis during the child's lifetime (FH+; $n = 43$), and those with no parental history of an AUD (FH-; $n = 30$). Family history was determined only from the aspect of the child's lifetime in order to maximize the environmental burden of familial risk. Eighteen FH- participants had parents reporting a previous AUD occurring 2 or more years prior to the participant's birth; 7 had a parent diagnosed with alcohol abuse 2 to 17 years (9.7 years on average) prior to the child's birth; 4 had a parent diagnosed with alcohol dependence without physical dependence 3-14 years (7 years on average) prior to the child's birth; and 7 had a parent diagnosed with alcohol dependence with physical dependence 2-15 years (6 years on average) prior to the child's birth.

Our analyses did not include any subjects with incomplete family history information or with a parental AUD less than 2 years prior to but not after the child's birth. AUD diagnosis was calculated by a clinical psychologist based on Diagnostic Interview Schedule – Version 4 (1-3), Health History, and Drinking & Drug Use answers. As both Health History and Drinking & Drug Use are specific to the Michigan Longitudinal Study, details on both are outlined below.

Drinking and Drug History. This questionnaire incorporates items from the 1978 NIDA Survey (4), the American Drinking Practices Survey (5), and the VA Medical Center Research Questionnaire for Alcoholics (6). All of the items have been extensively used in a variety of survey and clinical settings. They provide data on quantity, frequency and variability of alcohol consumption, frequency of drug use, and multiple questions on consequences and troubles related to the use of these substances. Items have been carefully reviewed to yield information sufficient to provide diagnoses according to DSM-IV and Research Diagnostic Criteria.

Health History Questionnaire. This extensive self-administered history questionnaire was developed by the Rutgers Longitudinal Study (7) to assess health and illness status in 15 areas, including alcohol and drug use. The mother's form contains questions regarding the target child's birth and early developmental history. Also embedded within the instrument are all items from the short form of the Michigan Alcoholism Screening Test (8,9).

fMRI Task

The go/no-go task in the current study used both the timing (stimulus and fixation duration) and event distributions outlined in Durston *et al.* (10). The proportion of go trials was 75% of total trials across all 5 runs (185 out of 245 total trials), while no-go trials made up 25% of the total trial number across all 5 runs (60 out of 245 total trials) (Table S2). Additionally, each no-go trial could be preceded by 1, 3, or 5 go trials (Table S3) for a total of 20 each across all 5 runs.

Table S2. Proportion of no-go to go trials for each run and across the entire experiment.

	Run 1	**Run 2**	**Run 3**	**Run 4**	**Run 5**	**Total**
No-Go	11	13	12	12	12	60
Go	38	36	37	37	37	185
Total	49	49	49	49	49	245

Table S3. The number of go trials that preceded a no-go trial for each run and across the entire experiment.

Number of Go Trials	Run 1	Run 2	Run 3	Run 4	Run 5	Total
1	1	5	4	4	6	20
3	7	2	4	4	3	20
5	3	5	4	4	4	20

Participant head motion was corrected using FSL 5.0.2.2 (FMRIB, Oxford, UK), and runs exceeding 3 mm translation or rotation in any direction were excluded. The percent and number of runs excluded are presented below (Table S4).

Table S4. Percent and number of runs excluded due to motion per group, by scan. Runs excluded are from the number of participants (*n*) in the left-hand column of each group.

	FH-			FH+		
	n	**Percent Excluded**	**Runs Excluded/Total**	***n***	**Percent Excluded**	**Runs Excluded/Total**
Scan 1	4	4.7%	7/150	12	9.8%	21/215
Scan 2	3	2.0%	3/150	9	5.1%	11/215
Scan 3	0	0%	0/95	2	2.0%	2/100
Scan 4	1	3.0%	1/30	1	2.5%	1/40

Performance Results

Performance measures that fell outside of the mean ± 3 SD were removed (Hits: FH+ 3 data points, FH- 2; HitRT: FH+ 2, FH- 1). Table S5 shows task performance by scan. Independent samples *t*-tests revealed significant differences between FH- and FH+ groups for Hit reaction time (HitRT) and False Alarms (FA) at scan 1, for Hit accuracy and FA at scan 3, and for FA at scan 4. However, when running a linear mixed model looking at differences across age between the two groups using gender as a covariate, no interactions were found. Only a main effect of age was found for Hit accuracy and HitRT, and a main effect of gender was found for FA (see main manuscript).

Table S5. Task performance by scan.

Task Performance	All Participants	FH-	FH+
	Mean±SD	Mean ±SD	Mean ±SD
Overall mean			
Hit accuracy (%)	95.83 ±5.12	95.79 ±5.40	95.87 ±4.92
Hit reaction time (ms)	469.99 ±78.42	467.46 ±81.86	471.85 ±76.04
False alarm rate (%)	42.18 ±19.43	45.99 ±19.43	39.32 ±19.01
Scan 1			
Hit accuracy (%)	93.48 ±8.75	94.15 ±7.67	93.00 ±9.51
Hit reaction time (ms)	516.66 ±142.81	499.35 ±110.30	529.02 ±162.28
False alarm rate (%)	48.23 ±18.56	52.02 ±18.58	45.51 ±18.28
Scan 2			
Hit accuracy (%)	95.33 ±6.03	94.97 ±7.13	95.58 ±5.20
Hit reaction time (ms)	472.32 ±83.44	477.42 ±91.97	468.76 ±77.86
False alarm rate (%)	40.74 ±19.95	41.70 ±20.32	40.07 ±19.90
Scan 3			
Hit accuracy (%)	96.98 ±6.36	96.09 ±8.71	97.83 ±2.71
Hit reaction time (ms)	437.27 ±45.21	434.32 ±49.85	440.08 ±41.42
False alarm rate (%)	36.17 ±17.98	42.15 ±19.56	30.50 ±14.65
Scan 4			
Hit accuracy (%)	94.52 ±10.99	95.23 ±7.32	93.99 ±13.61
Hit reaction time (ms)	426.15 ±42.41	429.58 ±43.27	423.58 ±44.56
False alarm rate (%)	35.36 ±17.82	49.45 ±13.09	24.80 ±13.02

Because a main effect of gender was found for FA, we split participants by gender to look at mean FA for all participants across all scans, and then also by group across all scans. For all participants, the mean (±SD) FA was 32.5% ± 1.8 for females, and 46.1% ± 1.9 for males. For the FH- group, the mean was 36.8% ± 2.0 for females and 47.8% ± 1.9 for males. For the FH+ group, the mean was 31.1% ± 1.7 for females and 44.4% ± 1.9 for males. This pattern, where females made less false alarms than males, was visible at each scan (scan 1: female mean 41.9% ± 2.1, male mean 51.2% ± 1.7; scan 2: female mean 28.0% ± 1.4, male mean 46.2% ± 2.0; scan 3: female mean 27.7% ± 1.3, male mean 39.5% ± 1.9; scan 4: female mean 23.3% ±1.5, male mean 38.7% ± 1.8).

fMRI Results

Longitudinal Results with Participants Reporting Substance Use Removed. Five participants reported alcohol or drug use at the time of their first scan, and an additional eight reported alcohol or drug use at some point after their first scan (n = 13; Table 1). Linear mixed model analyses were rerun with these 13 individuals removed; however, results did not differ from the original results using all participants as there was a significant interaction between family history and age-centered for the right caudate (F = 13.47, p = 0.001), right middle cingulate (F = 5.74, p = 0.018), and the right middle frontal gyrus (F = 14.49, p = 0.001).

Longitudinal Results for Correct Reject vs. Baseline. For contrast interpretation, we also analyzed correct reject trials versus baseline for the three regions of interest (ROIs) that were significant in the main contrast of correct reject versus correct go trials. In SPM, ROIs for the right caudate, middle cingulate, and middle frontal gyrus were created based on the coordinates reported in Table 3. These ROIs were applied to correct reject versus baseline trials, and activation values were extracted and imported into SPSS. A linear mixed model analysis was performed with scan as a repeated measure, subject as a random factor, family history as a fixed factor, and gender as a covariate. A significant group by age interaction was found in the caudate (F = 5.98, p = 0.016) and middle frontal gyrus (F = 7.07, p = 0.009), but not in the middle cingulate (F = 2.90, p = 0.091). Post-hoc analyses revealed significant activation decreases with age in the FH- group in the caudate (FH-: F = 8.56, p = 0.006; FH+: F = 1.31, p = 0.254) and middle frontal gyrus (FH-: F = 9.83, p = 0.004; FH+: F = 1.31, p = 0.257), consistent with the results for correct reject versus go trials reported in the manuscript. These changes can be seen in Figure S2.

Intraclass-correlation Coefficient Analysis. Intra-class correlation coefficient (ICC) analysis was used to assess test-retest reliability of fMRI data in the go/no-go task. Here, we computed ICC maps using the toolbox and methods provided by Caceres *et al.* (11). The activation network was defined using task activation from scan 1 for all participants for the correct no-go versus correct go contrast (also used in the fMRI longitudinal analysis; see main manuscript). This activation network included 8 regions of interest that were significant at a cluster-corrected false discovery rate threshold of $p < 0.05$, 35 voxel extent. The following are the regions of interest: 1) right and 2) left supramarginal gyrus, 3) right insula, 4) left inferior frontal gyrus, 5) left superior temporal gyrus, 6) right precuneus, 7) right inferior temporal gyrus,

and 8) right putamen. Using the SPM ICC toolbox, ICC maps were created, and intra-subject reliability values (threshold $t = 3.2$) were calculated.

ICC between scans 1 and 2 for all participants in the current study ($n = 73$) was 0.03 for the activation network. This value is below the acceptable reliability minimum of 0.4 (e.g., 12-15). This was not unexpected, given that developmental changes are occurring within the activation network in the 1-3 years between scans 1 and 2. In a study comparing test-retest reliability across approximately 3 years in fMRI data, children demonstrated poor ICC values, while adolescents and adults showed fair to good values with respect to performance monitoring networks (16). Furthermore, test-retest reliability fMRI studies tend to focus on healthy young adult and adult populations; therefore, the expected range of ICC values pertaining to children and development is not well defined.

To test the reliability of our fMRI data without development as a confound, we ran an ICC analysis only on children who performed the go/no-go task in the scanner a second time within 1 month after an initial scan ($n = 5$). These second scans were performed specifically to assess test-retest reliability of blood-oxygen-level dependent (BOLD) activation; thus, these data were not included in the report of developmental effects, which is the focus of this manuscript. For these participants, the test-retest intra-subject reliability for the network was 0.5. This value is considered to be an acceptable level of reliability, as reported by Eaton *et al.* (15) and Aron *et al.* (12), based on guidelines originally proposed by Cicchetti (13,14). Therefore, we are confident of the reliability of the BOLD data reported here and its utility for investigating developmental changes in impulse control circuitry.

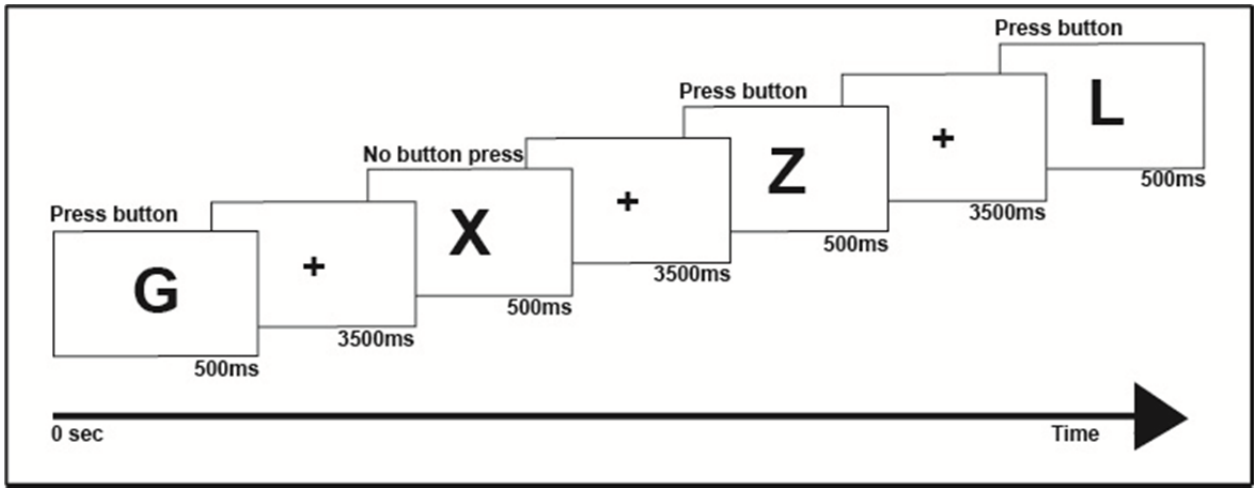

Figure S1. Schematic of Go/No-Go task. Participants were instructed to press the button to target stimuli (go trials; all letters except X) and make no response to infrequent non-target stimuli (no-go trials; letter X).

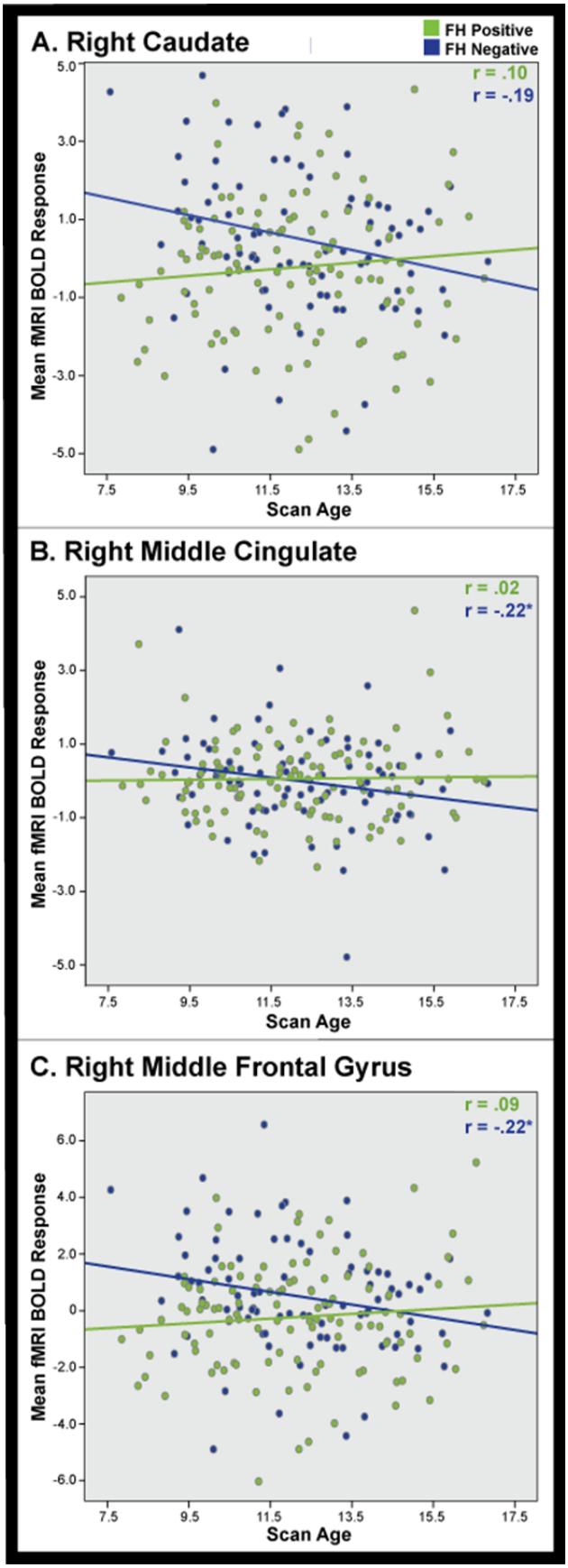

Figure S2. Correct reject versus baseline scatterplots for FH+ (green) and FH- (blue) groups illustrating changes in mean fMRI activation across time. Regions of interest (right caudate, right middle cingulate, and right middle frontal gyrus) were determined in whole-brain longitudinal analysis by the main contrast of interest. * indicates $p < 0.05$.

Supplemental References

1. Robins L, Helzer J, Croughan J, Ratcliff KS (1980): *The NIMH Diagnostic Interview Schedule: Its history, characteristics and validity*. St. Louis, MO: Washington University School of Medicine.
2. Robins LM, Helzer JH, Croughan J, Ratcliff KS (1981): National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. *Arch Gen Psychiat* 38:381-389.
3. Robins LN, Cottler LB, Bucholz KK, Compton WM, North CS, Rourke KM (2000): *Diagnostic Interview Schedule for the DSM-IV (DIS-IV).* St. Louis, MO: Washington University School of Medicine.
4. Johnston LD, Bachman JG, O'Malley PM (1979): *Drugs and the class of 1978: Behaviors, attitudes, and recent national trends (DPEW Publication No. ADM 79-877).* Washington, DC: National Institute on Drug Abuse, Division of Research, U.S. Department of Health, Education, and Welfare.
5. Cahalan D, Cisin IH, Crossley HM (1969): *American drinking practices: A national study of drinking behavior and attitudes (Monograph No. 6).* New Brunswick, NJ: Publications Division, Rutgers Center of Alcohol Studies.
6. Schuckit MA (1978): *Research questionnaire*. San Diego, CA: Alcoholism Treatment Program, V. A. Medical Center, University of California, San Diego.
7. Carpenter JA, Lester D (1980): *The etiology of alcoholism: A longitudinal study*. New Brunswick, NJ: Rutgers Center for Alcohol Studies.
8. Selzer M (1971): The Michigan Alcohol Screening Test (MAST): The quest for a new diagnostic instrument. *Am J Psychiat* 127:1653-1658.
9. Selzer ML (1975): A self-administered short Michigan Alcoholism Screening Test (SMAST). *J Stud Alcohol* 36:117-126.
10. Durston S, Thomas KM, Yang Y, Ulug AM, Zimmerman RD, Casey BJ (2002): A neural basis for the development of inhibitory control. *Developmental Sci* 5:F9–16.
11. Caceres A, Hall DL, Zelaya FO, Williams SCR, Mehta MA (2009): Measuring fMRI reliability with the intra-class correlation coefficient. *Neuroimage* 45:758-768.
12. Aron AR, Gluck MA, Poldrack MB (2006): Long-term test-retest reliability of functional MRI in a classification learning task. *Neuroimage* 29:1000-1006.
13. Cicchetti DV, Sparrow SA (1981): Developing criteria for establishing inter-rater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic* 86:127-137.
14. Cicchetti DV (2001): The precision of reliability and validity estimates revisited: distinguishing between clinical and statistical significance of sample size requirements. *J Clin Exp Neuropsychol* 23:695-700.
15. Eaton KP, Szaflarski JP, Altaye M, Ball AL, Kissela BM, Banks C, *et al*. (2008): Reliability of fMRI for studies of language in post-stroke aphasia subjects. *Neuroimage* 41:311-322.
16. Koolschijn PC, Schel MA, de Rooij M, Rombouts SA, Crone EA (2011) A three-year longitudinal functional magnetic resonance imaging study of performance monitoring and test-retest reliability from childhood to early adulthood. *J Neurosci* 31:4204-4212.