Further Description of IMAGEN Study

The IMAGEN study is an ongoing European multi-center study on risk-taking behavior in teenagers. It is an integrated project funded by the European Commission in the 6th Framework Program "Life Science"(1), specifically designed to provide comprehensive behavioral, neuropsychological, genetic, functional and structural neuroimaging data related to behavioral disinhibition and reward processing in a representative sample of 2,000 young adolescents at 8 study-sites in Germany, England, France and Ireland. The recruitment strategy employed resulted in the recruitment of 40% of families who initially expressed interest in the project. Main reasons for exclusion were lack of availability of children and parents for a full assessment day at the research institute, age and contraindication for magnetic resonance imaging such as braces, premature birth, diseases of the central nervous system, brain trauma, or medication (including ADHD stimulant medication).

A total of 2232 participants across 8 European sites were recruited via high-schools in geographical areas with minimal ethnic diversity to maximize ethnic homogeneity as a prerequisite for future genome-wide association analyses. To obtain a diverse sample in terms of socio-economic status, emotional and cognitive development, private-, state-funded schools and special educational units were equally targeted within those areas. After data quality control, complete and reliable data sets for 1778 volunteers with an average age of 14.43 years (SD = 0.35) and an even gender ratio (n = 948 girls, i.e. 51 %) were included analyses. Table S1 gives an overview of the sample distribution across sites.

Of these 1778 adolescents, 4.4% (78) were identified as having a diagnosis of CD (37), ADHD (30) or both (11) according to the Development and Well-Being Assessment interview(19), 3.6% (65) reported problematic alcohol use, and 11.6% (165) reported drug use. At 16 years, 6.3% (76) were identified as having a diagnosis of CD (25), ADHD (31) or both (10), 18.0% (218) reported problematic alcohol use, and 27.1% (328) reported drug use.

TABLE S1. Distribution of Participants (N=1864) Across Sites

Variable			%	N
Study site	England	London	12	224
		Nottingham	14	261
	Ireland	Dublin	10	186
	Germany	Berlin	13	242
		Hamburg	14	261
		Mannheim	11	205
		Dresden	14	261
	France	Paris	12	224

Further Description of Measures and Assessment Protocol

As the Psytools program was run at the participant's home without direct supervision of the research team, the reliability of each individual's data were checked in a two-stage procedure. Before every task, adolescents were asked to report on the current testing context including questions about their attentional focus and the confidentiality of the setting. Automated flags highlighted potentially problematic testing situations and were followed-up by research assistants face-to-face with the volunteer in a confidential setting. Final reliability ratings were assigned which led to in- or exclusion of the data. Exclusion criteria were: During reaction-time tasks, the child indicated that he/ she was listening to music or was exposed to disturbing noise or their mean reaction time per block within a task was below 100ms or they pressed the same response key throughout the task, during the questionnaires they indicated to have been in a

hurry or somebody was watching or they indicated to have known or taken the sham drug Relevin. Also, data sets showing inconsistent responses across or extreme outliers within measures were excluded (see next section). The child self-report and the parent-report clinical screening and interview where administered under supervision at the research institutes. Participants were explained the confidential nature of the interview but answered the questions on their own to establish openness and confidentiality. Research assistants were available on demand if further support was required. Data were rated as unreliable in case the research assistant had reason to believe from their observation of the interview session that the volunteer did not answer the questions appropriately e.g. as a result of language or reading problems.

Substance use measures: The AUDIT was developed and validated by the World Health Organization to assist the brief assessment of alcohol use disorders and was specifically designed for international use. It exists in all three languages, and was validated on primary health care patients in six countries. For this study, the scale total for problematic or harmful alcohol use in the last year was used including items about feelings of guilt or remorse after drinking, being unable to remember what happened the night before because of drinking, being injured or having injured someone as a result of drinking and relevant others being concerned about their drinking and suggestions to cut down. The five response options range from 0 ("never") to 4 ("daily or almost daily).

The ESPAD items used in this study comprise inverted age of onset of drinking alcohol (coded as 16 years for non-drinkers) to indicate positive associations between higher levels of risky behavior and earlier onset, alcohol use frequency (number of lifetime occasions, seven response options from '0' to '40 or more') and quantity (number of drinks on a typical day when

drinking, five response options from '1' to '10 or more', automatically coded 0 for non-drinkers), lifetime number of occasions being drunk (seven response options from '0' to '40 or more') and binge drinking (defined as having five or more drinks in a row, six response options from '0' to '10 or more times') and severity of lifetime illicit drug use, a composite score calculated as the total of the responses to number of occasions (seven response options for each item from '0' to '40 or more') taking either of the following substances or groups of substances: marijuana or hashish, inhalants, tranquilizer or sedatives, amphetamines, LSD, magic mushrooms or hallucinogens (excluding LSD), crack, cocaine, heroin, narcotics, ecstasy, ketamine or phenylclinidine, GHB or liquid ecstasy, and anabolic steroids.

Externalizing problems: The SDQ is a brief behavioral screening questionnaire for 3-16 year olds which assesses emotional and conduct symptoms, hyperactivity/inattention, prosocial behavior, and peer relationship problems (five items each). It offers an algorithm combining parent and adolescent report to yield scores predicting the likelihood of psychiatric caseness based on prevalence overall and per symptom area (0 = unlikely, 1 = possible, 2 = probable; for details on the algorithm see http://www.sdqinfo.com/ea1.html) . The SDQ therefore provides a short screening for caseness and service use (2).

As a comprehensive measure of psychiatric symptoms and caseness, the DAWBA interview was assessed in adolescents and parents. Based on the combination of parent- and self-report, a prognosis for the likelihood of having a disorder is calculated. The generated band ranges from level 0 up to level 5 corresponding to the approximate prevalence rates in an epidemiological sample for disorder in question, ranging from less than 0.1% up to 70%.

Diagnostic criteria were based on the Diagnostic Statistical Manual, Version 4 for ADHD and CD.

Self-report bullying as a perpetrator was introduced as a further externalizing behavior to the models. The total of 4 items describing verbal (e.g. calling somebody mean names), passive (e.g. completely ignoring somebody) and physical (e.g. hitting somebody) peer bullying in the past 6 months was adapted from the Bully/Victim Questionnaire used in the Youth Risky Behavior Survey (Brener, Collins, Kann et al. 1995). Response options on number of occasions ranged from 1 ("none") to 5 ("several times a week"). Internal consistency in the current sample was good with a Cronbach's $\alpha = .81$.

Personality: The SURPS reliability and concurrent and predictive validity is now well established. Woicik and colleagues found a good internal reliability of the SURPS subscales with Cronbach's Alpha ranging from .7 to .8 in an adolescent sample (mean age: 15.7 years), as well as a good 2-month test-retest reliability with intra-class correlations ranging from .68 (for AS) to .88 (for SS) in a late adolescent sample (mean age: 18.8 years). In the current sample, the personality trait scales IMP and SS correlate positively ($r=.15$, $p=.002$).

Passive Avoidance Learning Paradigm (PALP) training: Three training blocks with numbers 1 and 2 representing the "correct" and the "wrong" number were used before each test block to demonstrate the association between their responses (hits and misses) and the outcome (winning or losing points): First, the subject was forced to respond to every trial, then to withhold the responses and finally to respond in the best possible way to gain the maximum amount of points. Following three training blocks test blocks each consisting of 10 test trials were presented, repeating the set of 8 two-digit numbers each time in a different random order. The stimuli set used for each condition was balanced across subjects. The number appeared on

the screen for 3 seconds during which the subject was asked to decide whether to respond by pressing the space bar or to withhold the response. Afterwards, the number disappeared, leaving an inter-trial-interval of 1 second before the next number was displayed. At the upper left corner of the computer screen, a running score was continuously displayed. Three conditions were applied: reward only, punishment only and reward-punishment combined. Written instructions informed the subjects that the score would influence a real reward at the end of task. However, this did not apply to French families as the local ethics committee did not allow for any monetary reward, but were still motivated to win points or avoid losing points.

Delay discounting: Delay discounting was assessed with the Kirby Delay Discounting Questionnaire(3). With this measure, delay discounting is determined from 27 hypothetical choice questions for either immediate or delayed money, with the delays ranging from 7 to 186 days. Participants were instructed that, although hypothetical, to make choices as though they are actually going to receive the money they choose. The Kirby was scored as described previously by Kirby et al.(3), with k values (an index of delay discounting) assigned according to choice patterns across the 27 items, with larger k values indicating greater delay discounting of value for the delayed options.

Neuroimaging Tasks

Two MRI sessions were run lasting 45 minutes each; each included a combination of structural and functional MR scans. Before each session, the volunteers familiarized themselves with the scanner and the tasks in a practice session. In the scanner, volunteers were provided with two response grips with one button for each index finger and a goggle system for visual stimulation, (NordicNeuroLab, Bergen, Norway). They also received a brief visual and verbal reminder of the instructions before each task. Further information about the functional magnetic

resonance imaging procedure with links to task specifications can be found online at http://www.imagen-europe.com/en/Publications_and_SOP.php. All images were acquired on 3T magnetic resonance scanners; an overview of the scanner specifications as well as the quality control and standardization procedure across sites can be found elsewhere(1).

Region of interest (ROI) data were generated using WFU Pickatlas(4) and associated anatomical atlases(5-7) with the exception of the ventral striatum (VS) mask on the MID task. The VS mask was hand-drawn on an average anatomical image that was normalized into MNI space with SPM8. The VS mask specified the ventral ($z < 0$) part of the caudate and nucleus accumbens region. The regressors modeling the experimental conditions were convolved using SPM's default hemodynamic response function. The estimated model parameter maps were linearly combined to yield contrast maps. The mean contrast value within each ROI was calculated for each subject. SPM8 (Wellcome Trust Centre for Neuroimaging) was used to preprocess and analyse fMRI data on this task. Single-subject echo-planar images (EPI) were initially co-registered with the T1 structural image. Functional images were realigned and resliced to the first volume.

During the Stop-Signal-Reaction-Time task(15) volunteers responded to regularly presented visual go stimuli (arrows pointing left or right) but were instructed to withhold their response when the go stimulus was followed unpredictably by a stop-signal (arrow pointing upwards). Stopping difficulty was manipulated across trials by varying the delay between the onset of the go arrow and the stop arrow (stop-signal delay) using a previously described tracking algorithm.(15) A block contained 400 go trials and 80 variable-delay stop trials with between three and seven go trials between two stop trials. Stimulus duration in go trials was 1,000 ms and varied in stop trials (0–900ms, 50-ms steps) in accordance with the tracking

algorithm (initial delay = 250 ms). For this task, the first level model on unsmoothed single subject data included the movement realignment regressors plus four task-specific regressors: (1) successful inhibitions, (2) errors of commission, (3) incorrect responses on Go trials, and (4) late responses on Go trials. The contrast images that were created for each participant and were used for the current analyses were the successful inhibitions and errors of commission.

For the MID task, the first level model on unsmoothed single subject data included six regressors for successful and six regressors for unsuccessful trials: (1) anticipation of large reward, (2) anticipation of small reward, (3) anticipation of no reward, (4) feedback large reward, (5) feedback small reward, (6) feedback no reward, yielding a total of 12 regressors. Trials in which subjects failed to respond were modelled as separate error trials. Movement parameters from the realignment procedure were included as covariates in the first level model for each subject. Contrast images of parameter estimates were created for each participant. The current analysis focused on the contrast anticipation of large award > anticipation of no reward. There was no punishment condition in this task.

Further Information on fMRI Tasks

SSRT task: Difficulty of stopping was manipulated across trials by varying the delay between the onset of the go arrow and the stop arrow (stop-signal delay, SSD).[56] The practice session outside the scanner was run on a computer and consisted of 60 trials that lasted about 2 minutes. Volunteers were asked to press a corresponding key in response to the go-stimuli, i.e. the left arrow key in response to a left pointing arrow and the right arrow key to a right pointing arrow. Volunteers were also instructed to try and withhold their response when an upwards arrow followed the go-stimuli, however, they were explicitly reminded to try and respond as fast as possible to the go stimuli. In the scanner, a block of about 16 minutes followed, containing

400 go trials and 80 variable delay stop trials with a minimum of 3 and a maximum of 7 go trials between two stop trials. The go stimulus duration in go trials was 1000ms and varied, depending on the SSD, on stop trials.

MID task: A modified version of the Monetary Incentive Delay (MID)[58] was used to assess brain response to reward anticipation, in which each trial included a reward anticipation phase, a reward response phase, a feedback phase and a fixation period. The current analysis will only focus on the anticipation phase, in which participants were presented with cues (that varied between 4 and 4.5 seconds) signaling the amount of reward that could be won on a given trial (large reward, small reward or no reward). Subjects could win points (10 points in the large reward condition and 2 in the small reward condition) by responding to a response cue. The time window in which responses were counted as "win" was adjusted dynamically during the course of the experiment according to subjects' performance, such that, on average, subjects won on 66% of all trials. The response and feedback phase had a total duration of 2 seconds. Points were then converted to sweet food snacks (M&Ms) following testing (5 points per M&M). In total, subjects completed 22 trials per condition, 66 trials in total. Four seconds of inter-trial fixation separated the trials. The current analysis will focus on 3 ROIs identified as specific regions of interest, where mean activity levels for each contrast were extracted during the reward anticipation phase involving the contrast of large reward anticipation – no reward anticipation. The ROIs (Right and Left) selected were: Ventral Striatum, Orbital Cortex and Inferior Frontal Gyrus. As BOLD response in Right and Left Ventral Striatum was highly correlated, they were averaged to create a bilateral ventral striatum response score.

References

1. Schumann G, Loth E, Banaschewski T, Barker G, Buechel C, Conrod PJ, Dalley JW, Flor H, Gallinat J, Garavan H, Heinz A, Itterman B, Lathrop M, Martinot J-L, Mann K, Poline JB, Robbins TW, Rietschel M, Reed L, Smolka M, Spanagel R, Stephens DN, Stroehle A, Paus T, IMAGEN consortium. The IMAGEN study: Reinforcement-related behaviour in normal brain function and psychopathology. Molecular Psychiatry. 2010;15:1128-1139.
2. Goodman R, Scott S. Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: is small beautiful? J Abnorm Child Psychol. 1999;27:17-24.
3. Kirby KN, Petry NM, Bickel WK. Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. Journal of Experimental Psychology: General. 1999;128:78–87.
4. Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 2003;19:1233-1239.
5. Lancaster JL, Summerlin JL, Rainey L, Freitas CS, Fox PT. The Talairach Daemon a database server for talairach atlas labels. NeuroImage. 1997;5.
6. Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT. Automated Talairach atlas labels for functional brain mapping. Hum Brain Mapp. 2000;10:120-131.
7. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage. 2002;15:273-289.

TABLE S2. Correlations Between Externalizing Problem Indicators at 14 Years (Baseline)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. CD screen SR	1.00													
2. ADHD screen SR	**0.39**	1.00												
3. ADHD band	**0.21**	**0.33**	1.00											
4. CD band	**0.38**	**0.21**	**0.36**	1.00										
5. CD screen PR	**0.34**	**0.22**	**0.40**	**0.44**	1.00									
6. ADHD screen PR	**0.27**	**0.43**	**0.61**	**0.31**	**0.45**	1.00								
7. Bullying	**0.23**	**0.12**	**0.13**	**0.13**	**0.12**	**0.11**	1.00							
8. Age drinking onset	**0.22**	**0.15**	0.03	**0.19**	**0.10**	**0.09**	**0.10**	1.00						
9. Drinking problems	**0.20**	**0.14**	**0.06**	**0.22**	**0.09**	**0.09**	**0.11**	**0.21**	1.00					
10. Number drugs	**0.20**	**0.11**	**0.07**	**0.23**	**0.08**	**0.08**	0.03	**0.22**	**0.23**	1.00				
11. Drunkenness	**0.21**	**0.12**	**0.07**	**0.24**	**0.09**	**0.12**	**0.07**	**0.36**	**0.47**	**0.32**	1.00			
12. Bingeing	**0.20**	**0.12**	0.04	**0.26**	**0.13**	**0.09**	0.05	**0.34**	**0.41**	**0.34**	**0.65**	1.00		
13. Drinking Q*F	**0.22**	**0.09**	0.02	**0.26**	**0.07**	0.04	**0.07**	**0.45**	**0.45**	**0.31**	**0.60**	**0.69**	1.00	
14. English	**0.06**	**0.15**	-0.01	-0.04	**-0.12**	0.05	**0.14**	0.00	**0.08**	0.02	**0.12**	**-0.06**	0.01	1.00
15. Gender	**0.08**	0.01	**0.17**	**0.08**	0.02	**0.17**	0.01	0.01	0.01	0.04	-0.01	-0.02	0.00	0.03

ADHD – Attention Deficit Hyperactivity Disorder; CD – conduct disorder symptoms; SP – self reported; PR – parent reported; Q*F – Quantity by Frequency; bold identifies correlations significant at the level of $p<.05$.

TABLE S3. Correlations Between Covariates and Externalizing Problem Indicators at 14 Years

	CD screen SR	ADHD screen SR	ADHD band	CD band	CD screen PR	ADHD screen PR	Bullying	Age drinking onset	Drinking problems	Number drugs	Drunkenness	Bingeing	Drinking Q*F
Impulsivity	**0.41**	**0.45**	**0.23**	**0.24**	**0.23**	**0.27**	**0.18**	**0.20**	**0.17**	**0.14**	**0.15**	**0.16**	**0.17**
Sensation-seeking	**0.12**	**0.09**	**0.08**	**0.07**	0.03	**0.07**	0.05	**0.14**	**0.08**	**0.12**	**0.11**	**0.11**	**0.11**
DD K	**0.11**	**0.10**	**0.10**	**0.12**	**0.07**	**0.12**	-0.03	0.05	**0.08**	**0.06**	**0.09**	**0.11**	**0.10**
Com Err	**0.07**	**0.10**	**0.11**	**0.08**	0.04	**0.14**	0.04	0.03	**0.07**	0.03	**0.07**	**0.06**	0.03
Verbal IQ	**-0.05**	**-0.12**	**-0.07**	**-0.07**	**-0.10**	**-0.21**	0.00	0.03	**-0.06**	-0.01	**-0.08**	**-0.08**	0.01
Spatial IQ	**-0.07**	**-0.09**	**-0.13**	**-0.11**	**-0.08**	**-0.22**	-0.01	-0.02	**-0.09**	**-0.06**	**-0.11**	**-0.10**	-0.05
SS Basal G	-0.01	-0.03	0.03	0.00	0.01	0.02	0.02	-0.03	0.03	**-0.06**	0.01	0.01	-0.03
SS Parietal	0.04	0.04	0.03	0.03	0.03	0.01	-0.02	0.01	**0.07**	-0.02	**0.08**	0.03	0.04
SS Orb-Fr	0.01	-0.01	0.01	0.02	0.00	0.01	-0.05	-0.02	0.01	-0.02	-0.01	0.00	0.03
SS Med Orb-Fr	0.02	0.05	-0.01	0.05	0.03	0.02	0.02	0.03	0.03	0.00	0.01	0.02	0.00
SS SN_STN	-0.05	-0.01	0.02	-0.05	-0.04	0.03	-0.03	-0.05	0.02	-0.03	0.03	0.02	0.01
SS R FRONT	-0.02	-0.02	-0.02	-0.03	-0.01	-0.02	0.01	-0.01	-0.04	-0.03	-0.01	-0.03	-0.03
SS pSMA PCG	**-0.07**	**-0.07**	-0.04	**-0.08**	-0.04	-0.04	-0.02	-0.04	**-0.10**	-0.05	**-0.07**	**-0.08**	**-0.08**
SF Bas G	0.05	**0.09**	0.01	0.01	0.00	0.06	0.05	0.03	0.02	0.01	0.04	0.04	0.01
SF Orb-Fr	0.02	0.02	-0.02	**0.06**	-0.02	-0.01	-0.03	0.02	-0.01	0.02	0.02	0.03	0.02
SF Med Orb-Fr	0.04	0.02	0.03	0.03	0.01	0.05	0.02	0.02	-0.03	0.00	-0.04	-0.02	-0.01
SF Parietal	-0.01	-0.01	-0.01	-0.04	0.01	0.00	0.03	-0.03	-0.02	-0.02	**-0.07**	-0.05	-0.02
SF BiFrontal	0.02	-0.01	0.03	0.00	0.04	0.02	0.02	0.01	-0.03	-0.03	0.00	-0.01	0.00
SF SN_STN	-0.01	0.01	-0.03	-0.01	-0.04	0.02	-0.05	0.01	0.03	0.01	0.04	0.04	0.02
MID LIFG	0.05	0.02	0.01	0.03	0.02	-0.02	-0.04	0.01	0.00	-0.05	-0.03	-0.01	-0.02
MID RIFG	0.05	0.02	0.03	0.03	0.02	0.00	0.00	0.03	0.00	-0.02	0.00	0.01	-0.01
MID BIVS	0.01	-0.02	0.02	-0.01	-0.02	**-0.08**	-0.02	-0.03	0.00	-0.03	-0.02	-0.01	-0.04
MID RORB	0.03	0.00	0.05	0.01	0.03	0.02	0.05	0.04	0.03	-0.03	0.02	0.03	0.01
MID LORB	0.01	-0.01	0.01	0.03	0.02	-0.01	-0.03	0.03	0.04	-0.01	0.02	0.02	0.02

ADHD – Attention Deficit Hyperactivity Disorder; CD – conduct disorder symptoms; SP – self reported; PR – parent reported; Q*F – Quantity by Frequency; DD – Delay Discounting; Com Err – Commision Errors (Go No-go); SS – Succesful Stop; SF – Stop Failure; G – Ganglia; Orb-Fr – Orbitofrontal; Med – Medial; SN – Substantia Nigra; STN – Sub Thalamic Nucleus; R Front – Right Frontal; pSMA – presupplementary Motor Area; PCG – Pre-Central Gyrus; PCC – Pre-Central Cortex; MID: Monetary Incentive Delay Task; LIFG – Left Inferior Frontal Gyrus; RIFG – Right Inferior Frontal Gyrus; BIVS – Bilateral Ventral Striatum; RORB – Right Orbito-frontal; LORB – Left Orbito-frontal; bold identifies correlations significant at the level of $p<.05$.

TABLE S4. Correlations Among Covariates

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. Impulsivity	1.00																						
2. Sensation-seeking	**0.16**	1.00																					
3. DD K	**0.11**	0.01	1.00																				
4. Com Err	**0.12**	-0.04	**0.12**	1.00																			
5. Verbal IQ	**-0.08**	0.03	**-0.12**	**-0.19**	1.00																		
6. Spatial IQ	**-0.08**	0.01	**-0.13**	**-0.26**	**0.44**	1.00																	
7. SS Basal G	-0.05	-0.01	-0.05	-0.05	-0.01	0.01	1.00																
8. SS Parietal	0.04	0.02	**0.06**	0.00	0.00	0.00	**0.19**	1.00															
9. SS Orb-Fr	-0.01	-0.03	-0.04	-0.01	0.02	0.05	**0.10**	**0.14**	1.00														
10. SS Med Orb-Fr	0.01	**0.06**	-0.05	-0.01	-0.01	0.01	**0.14**	**-0.11**	**0.07**	1.00													
11. SS SN_STN	-0.03	-0.01	-0.03	-0.01	0.01	0.05	**0.35**	**0.14**	0.03	-0.01	1.00												
12. SS R FRONT	-0.01	0.04	0.01	0.03	0.04	-0.04	**-0.40**	**-0.24**	**-0.20**	-0.04	**-0.14**	1.00											
13. SS pSMA PCG	-0.04	-0.01	**-0.06**	0.01	0.02	0.00	**-0.40**	**-0.40**	-0.01	**-0.09**	**-0.19**	**0.24**	1.00										
14. SF Bas G	0.02	0.02	-0.01	0.00	**-0.06**	-0.02	**0.43**	**0.06**	-0.03	**0.08**	**0.11**	**-0.14**	**-0.19**	1.00									
15. SF Orb-Fr	0.04	-0.05	0.02	0.03	-0.01	0.03	0.01	**0.09**	**0.25**	**0.07**	0.03	**-0.08**	-0.04	**0.14**	1.00								
16. SF Med Orb-Fr	-0.03	0.03	-0.04	**-0.06**	0.04	0.03	0.05	-0.01	-0.03	**0.22**	0.01	0.02	**-0.07**	**0.18**	-0.04	1.00							
17. SF Parietal	-0.01	**-0.06**	-0.04	**0.06**	-0.01	-0.02	**-0.13**	**-0.39**	-0.03	**0.15**	**-0.10**	**0.14**	**0.16**	**-0.22**	**-0.20**	0.01	1.00						
18. SF BiFrontal	0.00	0.01	0.03	0.07	0.03	-0.03	**-0.34**	**-0.09**	**-0.07**	**-0.09**	**-0.10**	**0.37**	**0.13**	**-0.49**	**-0.21**	**-0.25**	**0.22**	1.00					
19. SF SN_STN	0.00	0.03	-0.01	**-0.07**	-0.01	0.03	**0.16**	0.05	-0.02	**-0.06**	**0.35**	**-0.06**	**-0.07**	**0.35**	0.01	**0.08**	**-0.17**	-0.22	1.00				
20. MID LIFG	0.00	0.03	0.00	0.01	-0.03	-0.03	0.03	0.03	-0.02	-0.01	0.04	-0.03	-0.05	0.00	-0.04	-0.02	**0.07**	0.03	0.01	1.00			
21. MID RIFG	-0.01	0.05	0.02	-0.01	0.03	0.00	0.01	-0.01	-0.03	0.02	0.02	-0.03	0.00	0.00	-0.02	0.01	**0.07**	-0.03	-0.01	**0.63**	1.00		
22. MID BIVS	0.01	0.00	-0.03	**-0.06**	0.05	**0.08**	0.03	0.02	-0.01	-0.05	-0.02	-0.05	0.02	0.05	-0.02	0.00	0.04	**-0.06**	0.00	**0.40**	**0.46**	1.00	
23. MID RORB	0.01	0.04	0.02	0.02	**0.06**	0.02	0.04	-0.02	0.03	0.04	0.01	-0.05	0.02	**-0.06**	0.05	0.01	0.00	-0.03	**-0.07**	**0.27**	**0.50**	**0.22**	1.00
24. MID LORB	0.00	0.02	0.01	0.02	0.05	-0.03	0.04	0.03	0.02	-0.01	0.04	**-0.06**	**-0.06**	**-0.12**	-0.03	-0.05	**0.09**	**0.06**	**-0.06**	**0.54**	**0.32**	**0.25**	**0.45**

DD – Delay Discounting; Com Err – Commision Errors (Go No-go); SS – Succesful Stop; SF – Stop Failure; G – Ganglia; Orb-Fr – Orbitofrontal; Med – Medial; SN – Substantia Nigra; STN – Sub Thalamic Nucleus; R Front – Right Frontal; pSMA – presupplementary Motor Area; PCG – Pre-Central Gyrus; PCC – Pre-Central Cortex; MID: Monetary Incentive Delay Task; LIFG – Left Inferior Frontal Gyrus; RIFG – Right Inferior Frontal Gyrus; BIVS – Bilateral Ventral Striatum; RORB – Right Orbito-frontal; LORB – Left Orbito-frontal;bold identifies correlations significant at the level of $p<.05$.

TABLE S5. Correlations Between Externalising Problem Indicators at 16 Years (Follow-Up)

		1	2	3	4	5	6	7	8	9	10	11	12
1	CD screen PR	1.00											
2	ADHD screen PR	**0.46**	1.00										
3	CD screen SR	**0.34**	**0.31**	1.00									
4	ADHD screen SR	**0.18**	**0.36**	**0.43**	1.00								
5	ADHD band	**0.28**	**0.46**	**0.22**	**0.19**	1.00							
6	CD band	**0.35**	**0.24**	**0.28**	**0.14**	**0.32**	1.00						
7	Bullying	**0.11**	**0.10**	**0.14**	**0.14**	**0.08**	**0.07**	1.00					
8	Age drinking onset	**0.11**	**0.10**	**0.17**	**0.18**	0.06	**0.15**	0.05	1.00				
9	Drinking problems	**0.08**	**0.08**	**0.12**	**0.07**	0.00	**0.16**	**0.15**	**0.24**	1.00			
10	Number drugs	**0.15**	**0.11**	**0.18**	**0.06**	**0.13**	**0.22**	0.05	**0.25**	**0.32**	1.00		
11	Drunkenness	**0.12**	**0.10**	**0.18**	**0.12**	0.04	**0.23**	0.03	**0.32**	**0.54**	**0.49**	1.00	
12	Bingeing	**0.14**	**0.12**	**0.21**	**0.11**	**0.09**	**0.21**	0.00	**0.38**	**0.49**	**0.49**	**0.77**	1.00
13	Drinking Q*F	**0.14**	**0.12**	**0.19**	**0.12**	**0.08**	**0.23**	0.02	**0.39**	**0.49**	**0.50**	**0.78**	**0.78**

ADHD – Attention Deficit Hyperactivity Disorder; CD – conduct disorder symptoms; SP – self reported; PR – parent reported; Q*F – Quantity by Frequency; bold identifies correlations significant at the level of $p<.05$.

TABLE S6. Correlations Between Externalizing Problem Indicators at 16 Years (Columns) and 14 Years (Rows)

	Symptoms assessed at 16 years												
Symptoms at 14 years	CD screen PR	ADHD screen PR	CD screen SR	ADHD screen SR	ADHD band	CD band	Bullying	Age drinking onset	Drinking problems	Number drugs	Drunken-ness	Binge	Drinking Q*F
CD screen PR	**0.54**	**0.36**	**0.29**	**0.16**	**0.25**	**0.27**	**0.09**	**0.10**	0.05	0.06	**0.10**	**0.12**	**0.08**
ADHD screen PR	**0.28**	**0.63**	**0.24**	**0.34**	**0.38**	**0.21**	**0.11**	0.06	0.02	0.04	**0.09**	**0.10**	**0.08**
CD screen SR	**0.27**	**0.24**	**0.47**	**0.26**	**0.17**	**0.24**	**0.13**	**0.17**	**0.08**	**0.17**	**0.16**	**0.15**	**0.13**
ADHD screen SR	**0.14**	**0.35**	**0.25**	**0.52**	**0.21**	**0.11**	**0.07**	**0.13**	0.04	0.03	**0.09**	**0.09**	**0.07**
ADHD band	**0.26**	**0.48**	**0.17**	**0.18**	**0.51**	**0.21**	**0.11**	0.04	-0.01	**0.08**	0.06	**0.09**	0.06
CD band	**0.30**	**0.24**	**0.22**	**0.09**	**0.25**	**0.33**	0.05	**0.13**	**0.10**	**0.16**	**0.19**	**0.16**	**0.17**
Bullying	**0.10**	**0.15**	**0.18**	**0.12**	**0.07**	**0.09**	**0.35**	0.06	**0.07**	0.01	0.05	0.03	0.00
Age drinking onset	**0.11**	**0.08**	**0.16**	**0.12**	**0.07**	**0.13**	-0.01	**0.54**	**0.20**	**0.23**	**0.35**	**0.36**	**0.34**
Drinking problems	**0.08**	**0.10**	**0.14**	**0.09**	0.03	**0.14**	0.06	**0.19**	**0.18**	**0.21**	**0.28**	**0.20**	**0.25**
Number drugs	**0.07**	0.05	**0.16**	**0.10**	**0.08**	**0.08**	-0.02	**0.20**	**0.09**	**0.29**	**0.21**	**0.18**	**0.20**
Drunkenness	**0.08**	0.04	**0.11**	**0.08**	-0.01	**0.12**	-0.03	**0.25**	**0.21**	**0.29**	**0.46**	**0.33**	**0.34**
Bingeing	**0.10**	0.02	**0.14**	0.04	0.03	**0.17**	-0.05	**0.26**	**0.16**	**0.31**	**0.38**	**0.34**	**0.37**
Drinking Q*F	0.05	0.00	**0.11**	0.04	0.01	**0.11**	-0.01	**0.37**	**0.23**	**0.32**	**0.38**	**0.38**	**0.43**

ADHD – Attention Deficit Hyperactivity Disorder; CD – conduct disorder symptoms; SP – self reported; PR – parent reported; Q*F – Quantity by Frequency; bold identifies correlations significant at the level of p<.05.

TABLE S7. Correlations Between Externalizing Problem Indicators at 16 Years and All Covariates at 14 Years

	CD screen SR	ADHD screen SR	CD screen PR	ADHD screen PR	ADHD band	CD band	Bullying	Age drinking onset	Drinking problems	Number drugs	Drunken-ness	Bingeing	Drinking Q*F
English	0.04	**0.17**	**-0.11**	**0.10**	-0.04	**-0.07**	**0.09**	0.00	0.04	**-0.07**	0.06	**-0.08**	-0.03
Gender	0.05	**-0.07**	-0.02	**0.16**	**0.10**	0.00	-0.01	0.01	0.00	**0.10**	**0.10**	**0.10**	**0.13**
Impulsivity	**0.31**	**0.31**	**0.20**	**0.25**	**0.18**	**0.19**	**0.08**	**0.16**	**0.08**	**0.10**	**0.14**	**0.15**	**0.13**
Sensation-seeking	**0.14**	**0.09**	0.01	**0.10**	0.06	**0.07**	0.02	**0.12**	0.05	**0.10**	**0.13**	**0.14**	**0.15**
DD K	0.06	0.02	0.01	**0.09**	**0.10**	0.06	-0.03	0.02	-0.02	0.03	0.02	0.05	0.03
Com Err	0.03	0.04	**0.09**	**0.16**	**0.09**	0.06	0.00	-0.01	0.06	0.01	0.04	0.04	0.01
Verbal IQ	0.02	-0.05	-0.07	**-0.20**	0.00	-0.04	-0.01	0.05	0.03	**0.15**	0.04	0.04	0.06
Spatial IQ	**-0.10**	-0.04	-0.02	**-0.16**	**-0.08**	**-0.08**	0.04	0.01	0.02	-0.01	-0.03	**-0.07**	-0.02
SS Basal G	0.00	0.02	0.00	0.04	0.04	-0.03	0.05	-0.01	0.01	-0.01	0.02	0.00	0.02
SS Parietal	0.01	0.02	0.01	0.02	0.00	-0.04	-0.03	0.03	**0.08**	-0.01	0.05	0.01	0.04
SS Orb-Fr	-0.03	0.01	-0.04	-0.05	-0.01	0.04	-0.04	0.03	-0.01	0.01	0.00	0.01	0.01
SS Med Orb-Fr	0.03	-0.02	0.05	0.03	0.03	0.00	-0.03	-0.04	-0.03	0.00	0.05	0.00	0.00
SS SN_STN	-0.03	-0.02	-0.01	0.03	0.03	0.04	0.01	-0.05	-0.01	-0.03	-0.03	-0.04	-0.01
SS R FRONT	0.01	-0.02	-0.01	-0.06	-0.02	-0.03	-0.01	-0.01	-0.02	0.01	-0.01	0.00	-0.03
SS pSMA PCG	**-0.07**	-0.04	-0.04	**-0.10**	-0.04	0.03	-0.04	-0.05	-0.03	-0.05	**-0.09**	-0.06	**-0.08**
SF Bas G	0.04	**0.07**	0.03	**0.07**	0.02	-0.02	**0.07**	0.00	0.03	0.05	0.06	0.02	0.02
SF Orb-Fr	-0.03	-0.01	-0.05	-0.03	-0.01	-0.03	-0.05	0.00	-0.02	-0.01	0.00	0.01	-0.01
SF Med Orb-Fr	0.03	0.04	0.04	0.02	0.05	-0.02	0.01	0.04	0.00	0.04	0.06	0.04	0.06
SF Parietal	0.04	0.03	0.05	0.02	0.04	0.06	0.01	-0.04	-0.03	0.00	-0.01	-0.01	-0.04
SF BiFrontal	0.03	0.03	-0.01	-0.05	0.01	0.00	-0.06	0.04	0.00	-0.03	-0.03	0.01	-0.02
SF SN_STN	0.01	0.00	-0.02	0.00	-0.03	0.01	0.02	0.00	0.00	0.03	0.00	-0.02	0.02
MID LIFG	0.04	-0.01	-0.02	-0.04	0.04	0.04	0.03	-0.02	0.02	0.00	0.04	0.04	0.02
MID RIFG	0.05	-0.03	0.00	0.00	0.02	0.06	0.01	-0.04	0.04	0.02	0.04	0.03	0.04
MID BIVS	0.01	-0.06	0.00	-0.02	-0.02	-0.01	0.01	-0.03	0.03	-0.02	0.01	0.01	0.01
MID RORB	0.00	-0.03	0.02	**0.07**	**0.09**	0.01	0.00	0.02	0.01	0.01	0.04	0.03	0.03
MID LORB	0.03	-0.05	0.03	0.00	0.06	0.01	0.00	-0.02	0.05	0.00	0.05	0.05	0.04

ADHD – Attention Deficit Hyperactivity Disorder; CD – conduct disorder symptoms; SP – self reported; PR – parent reported; Q*F – Quantity by Frequency; DD – Delay Discounting; Com Err – Commision Errors (Go No-go); SS – Succesful Stop; SF – Stop Failure; G – Ganglia; Orb-Fr – Orbitofrontal; Med – Medial; SN – Substantia Nigra; STN – Sub Thalamic Nucleus; R Front – Right Frontal; pSMA – presupplementary Motor Area; PCG – Pre-Central Gyrus; PCC – Pre-Central Cortex; MID: Monetary Incentive Delay Task; LIFG – Left Inferior Frontal Gyrus; RIFG – Right Inferior Frontal Gyrus; BIVS – Bilateral Ventral Striatum; RORB – Right Orbito-frontal; LORB – Left Orbito-frontal; bold identifies correlations significant at the level of $p<.05$.

FIGURE S1. Regions of interest include activation from (A) the substantia nigra (blue) and subthalamic nucleus (red) and from (B) the pre-SMA (red) and precentral gyri (blue) when successfully inhibiting a motor response and from a grouping of bilateral frontal areas (C) including the anterior cingulate (red), anterior insulae (blue) and the inferior frontal gyrus (green) when making an error of commission. Activation during reward anticipation (D) was also calculated for the left orbitofrontal cortex (red) and left inferior frontal gyrus (blue).