**SUPPLEMENT 1**

Detailed exclusion criteria of the current study, age distribution, task visual presentation, data preprocessing and sensitivity analysis are described below.

**Recruitment/Diagnostic procedure of NeuroIMAGE project**

The diagnostic algorithm for ADHD included assessment with the Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS)1 and the Conners’ rating scale2,3. KSADS is a semi-structured diagnostic interview based on the DSM-IV criteria and symptoms. Elevated scores in the KSADS lead to additional assessment with the Conners' questionnaires. Children < 18 years were evaluated with the [Conners' Parent Rating Scales](http://addwarehouse.com/shopsite_sc/store/html/conners-parent-rating-scales-cprs-r-l.html) (CPRS-R:L) and the [Conners' Teacher Rating Scales](http://addwarehouse.com/shopsite_sc/store/html/conners-teacher-rating-scales-ctrs-r-l.html) (CTRS-R:L), while children ≥ 18 years were assessed with the CPRS-R:L and the Conners’ Adult ADHD Rating Scales (CAARS-S:S).

In NeuroIMAGE, ADHD patients were characterized by an estimated IQ level ≥ 70, a T-score ≥ 63 on at least one of the DSM-IV ADHD scales in at least one of the Conners’ rating scales (DSM Inattentive behavior, DSM Hyperactive/Impulsive behavior, DSM Total) and an onset age before 12 that met the DSM-IV criteria for impairment and pervasiveness. Moreover, healthy controls were characterized by an estimated IQ level ≥ 70, a T-score < 63 on all the DSM-IV ADHD scales and three or fewer symptoms in case of children and adolescents and two or fewer symptoms for adults as derived from the diagnostic algorithm that had no affected siblings. Participants that did not meet criteria for inclusion in the ADHD or the control group were excluded from subsequent analyses.

**Sample characteristics**

****

Figure S1. Age distribution in (i) effects of family history of SUD (FH)

****

Figure S2. Age distribution in (i) effects of substance misuse (SM)

**Data pre-processing**

SVs were calculated based on the proportion of delayed choices4. In detail, the highest possible SV was defined as the average of the large-delayed reward and the highest small-immediate reward (i.e. 4.5), whereas the lowest possible SV was the average of 0 and the lowest small-immediate reward (0.5). SV in each delay was then calculated as the number of large-delayed choices divided by the number of total choices, which was then multiplied with the highest possible SV and was finally added to the lowest possible SV.

SVs and delays were normalized, so that SVs were expressed as proportions of the large-delayed reward and delays as proportions of the maximum delay. The areas under the curve (AUCs) between two successive delays were subsequently computed. For this, we used the function (x2 – x1)\*[y1+y2)/2], where x1 and x2 were successive delays, and y1 and y2 were the SVs in these delays. Successive AUCs were then summed to produce the total AUC5, which we used for the analysis.

**Results**



Figure S3. Association between area under the curve (AUC) and in the unemotional scale of ICU in (A) ADHD & (B) Controls with negative family history of SUD (FH-; black squares & dashed regression line) & positive family history of SUD (FH+; blue circles & solid regression line); AUCs are before transformation to normality

**Sensitivity analysis**

*(i) Effects of Family history of SUD (FH)*

Results showed no main ADHD (p=.144) or FH effects (p=.792) or ADHD\*FH interaction effects (p=.1). There were significant FH\*callousness (p=.016) and FH\*unemotional interaction effects (p=.038) and trends for ADHD\*unemotional (p=.059) and ADHD\*FH\*unemotional interaction effects (=.099). We also found significant main effects for age (p=.041), scan-site (p=.011) and IQ (p=.004).

*(ii) Effects of substance misuse (SM)*

We found significant group differences (p=.045), with greater discounting for ADHD+SM compared to ADHD-only (p=.041), and significant age effects (p=.008).

**REFERENCES**

1. Kaufman, J. *et al.* Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version ( K-SADS-PL ): Initial Reliability and Validity Data. *J. Am. Acad. Child Adolesc. Psychiatry***36**, 980–988 (1997).

2. Conners, C. K., Sitarenios, G., Parker, J. D. A. & Epstein, J. N. Revision and Restandardization of the Conners Teacher Rating Scale ( CTRS-R ): Factor Structure , Reliability , and Criterion Validity. **26**, 279–291 (1998).

3. Conners, C. K., Erhardt, D., Epstein, J. N., Parker, J. D. A., Sitarenios, G., & Sparrow, E. Self-ratings of ADHD symptoms in adults I: Factor structure and normative data. **3**, 141–151 (1999).

4. Mies, G. W., Ma, I., de Water, E., Buitelaar, J. K. & Scheres, A. Waiting and working for rewards: Attention-Deficit/Hyperactivity Disorder is associated with steeper delay discounting linked to amygdala activation, but not with steeper effort discounting. *Cortex***106**, 164–173 (2018).

5. Scheres, A., Tontsch, C., Thoeny, A. L. & Sumiya, M. Temporal reward discounting in children, adolescents, and emerging adults during an experiential task. *Front. Psychol.***5**, 1–7 (2014).