

Supplementary Information

METHODS.....	2
PARTICIPANT RECRUITMENT.....	2
MISSING DATA.....	2
SUBSTANCE USE.....	2
RESULTS.....	3
SAMPLE CHARACTERISTICS.....	3
BSNIP-1.....	3
BNSIP2.....	4
EFIGURE 1: PREDICTED AND OBSERVED COMPOSITE SCORES.....	5
SHAP VALUES FROM LINEAR MODELS.....	6
EFIGURE 2: CONTRIBUTION OF PREDICTOR VARIABLES TO WITHIN DIAGNOSIS COGNITIVE PERFORMANCE INCLUDING EDUCATION AND SUBSTANCE USE.....	6
EFIGURE 3: CONTRIBUTION OF PREDICTOR VARIABLES TO WITHIN DIAGNOSIS COGNITIVE PERFORMANCE EXCLUDING EDUCATION AND SUBSTANCE USE.....	6
EFIGURE 4: CONTRIBUTION OF PREDICTOR VARIABLES TO COGNITIVE IMPAIRMENT INCLUDING EDUCATION AND SUBSTANCE USE.....	7
EFIGURE 5: CONTRIBUTION OF PREDICTOR VARIABLES TO COGNITIVE IMPAIRMENT EXCLUDING EDUCATION AND SUBSTANCE USE.....	8
EFIGURE 6: RELATIONSHIP BETWEEN MEDICATION DOSE AND COGNITIVE IMPAIRMENT.....	9
SHAP VALUES FOR UNCOLLATED DEMOGRAPHIC AND EDUCATIONAL PREDICTORS.....	10
EFIGURE 7 CONTRIBUTION OF PREDICTOR VARIABLES TO WITHIN DIAGNOSIS COGNITIVE PERFORMANCE INCLUDING EDUCATION AND SUBSTANCE USE.....	10
EFIGURE 8: CONTRIBUTION OF PREDICTOR VARIABLES TO WITHIN DIAGNOSIS COGNITIVE PERFORMANCE EXCLUDING EDUCATION AND SUBSTANCE USE.....	10
EFIGURE 9: CONTRIBUTION OF PREDICTOR VARIABLES TO COGNITIVE IMPAIRMENT INCLUDING EDUCATION AND SUBSTANCE USE.....	11
EFIGURE 10: CONTRIBUTION OF PREDICTOR VARIABLES TO COGNITIVE IMPAIRMENT EXCLUDING EDUCATION AND SUBSTANCE USE.....	12
DOMAIN SPECIFIC RELATIONSHIPS.....	13
EFIGURE 11: DIGIT SPAN.....	13
EFIGURE 12: VERBAL FLUENCY.....	14
EFIGURE 13: DIGIT SYMBOL CODING.....	15
EFIGURE 14: TOWER OF LONDON.....	16
EFIGURE 15: TRAIL MAKING.....	17
EFIGURE 16: VERBAL MEMORY.....	18
EFIGURE 17: CONTRIBUTION OF PREDICTOR VARIABLES TO COGNITIVE IMPAIRMENT WHEN PANSS SCORES INCLUDED	19
REFERENCES.....	20

Methods

Participant Recruitment

B-SNIP included sites in Baltimore, Chicago, Dallas, Detroit, Boston, and Hartford, Conn. All sites had experience with the recruitment, and study of psychotic patients. The sites used identical diagnostic and clinical assessment techniques, and similar approaches to recruitment.

The recruitment of probands in BSNIP was carried out by dimension (psychosis). all potential participants with a history of psychosis and suspected schizophrenia, schizoaffective disorder, or bipolar disorder were recruited. All probands in the study were clinically stable and in a non-acute symptom state. Full details are provided in previous reports describing the cohort.^{1,2}

The study protocol was approved by the institutional review board at each local site. After complete description of the study to the volunteers, written informed consent was obtained. Information regarding race/ethnicity was based on self-report with study defined options and gathered given the differences previously observed between groups in terms of both prevalence and outcomes.

Missing data

The BSNIP1 cohort did not contain details of childhood adversity, substance use, or parental age at birth. For the patient groups and controls, within group imputation was performed (mean for continuous, mode for categorical). For relatives this was not possible given all values were absent, for the linear model the overall sample mean was used, whereas for XGBoost this was left as N/A.

Substance Use

Substance use was assessed in the course of a standard clinical interview for each of the following substances individuals were asked if they had ever used the substance: use of cannabis, stimulants, cocaine, opiates, hallucinogens, and other substances. If lifetime use was endorsed participants were then asked about frequency both before and after the age of 18. These were each coded as follows for use both before and after age 18 (except for cannabis which was coded as use before 15, between ages of 15-18, and use after 18):

1 = None; 2 = 1-4 times; 3 = 5-10 times; 4 = 11-50 times; 5 = 51-100 times; 6 = >100 times; 9 = Unknown

This gave a total of 13 variables for substance use that were entered into the model (i.e. 2 variables for each of the categories listed above other than cannabis which had 3 variables).

Results

Sample characteristics

BSNIP-1

	Healthy Control	Schizophrenia	Schizophrenia Relative	Bipolar W Psychosis	Bipolar Relative	Schizoaffective Disorder	Schizoaffective Relative
N	345	316	323	251	291	191	209
Male (%)	154 (44.9)	217 (68.7)	101 (31.3)	92 (36.7)	100 (34.4)	77 (40.8)	63 (30.6)
Asian ethnicity (%)	16 (4.9)	7 (2.5)	3 (1.2)	7 (2.8)	6 (2.1)	1 (1.0)	2 (1.0)
Black or African American ethnicity(%)	94 (27.5)	139 (44.3)	123 (38.1)	44 (17.9)	48 (16.5)	77 (40.8)	70 (33.5)
Hispanic ethnicity (%)	30 (8.7)	22 (7.0)	25 (8.0)	20 (8.0)	18 (6.5)	17 (9.4)	20 (9.6)
White ethnicity (%)	192 (55.9)	134 (42.7)	164 (50.8)	174 (69.7)	213 (73.5)	85 (45.0)	112 (53.6)
Olanzapine Equivalents (mg)	0.0 (0.0)	20.7 (19.5)	2.3 (10.6)	11.6 (16.4)	0.8 (4.0)	17.4 (22.1)	1.5 (5.7)
Age (years)	37.2 (13.0)	35.7 (13.0)	42.9 (15.0)	36.0 (13.0)	40.4 (16.0)	36.7 (12.0)	39.6 (16.0)
Age of mother at the time of birth (years)	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Age of father at the time of birth (years)	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Father's Occupation ^b	3.4 (1.7)	4.0 (1.8)	3.9 (1.9)	3.3 (1.8)	3.5 (1.8)	4.1 (1.9)	3.8 (1.9)
Mother's Occupation ^b	4.0 (1.8)	4.2 (1.8)	4.6 (1.9)	4.0 (1.8)	4.0 (1.8)	4.3 (1.9)	4.2 (1.9)
Father Educational Scale ^b	3.1 (1.6)	3.3 (1.5)	3.5 (1.6)	3.1 (1.6)	3.2 (1.6)	3.5 (1.7)	3.5 (1.6)
Mother Educational Scale ^b	3.3 (1.5)	3.3 (1.3)	3.7 (1.5)	3.0 (1.3)	3.3 (1.4)	3.5 (1.4)	3.5 (1.5)
Subject Educational Scale ^b	2.6 (1.1)	3.6 (1.2)	3.1 (1.2)	2.9 (1.1)	2.8 (1.3)	3.5 (1.1)	3.1 (1.4)
Word Reading Score ^d	103.4 (13.7)	93.8 (16.5)	99.8 (53.4)	101.2 (14.5)	102.9 (14.0)	95.8 (14.7)	98.5 (16.0)
Emotional Abuse ^a	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Emotional Neglect ^a	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Physical Abuse ^a	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Physical Neglect ^a	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Sexual Abuse ^a	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Use of cannabis after age 18 ^c	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Use of cannabis between age 15 and 18 ^c	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Use of cannabis before age 15 ^c	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Use of other substances before age 18 ^c	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Use of other substances after age 18 ^c	N/A	N/A	N/A	N/A	N/A	N/A	N/A

^a Abuse and neglect scores are from the childhood trauma questionnaire²¹

^b The Hollingshead Socioeconomic Rating Scale is used to classify educational achievement ranging from 1 (completed graduate degree) to 7 (less than seven years of schooling); and parental occupation ranging from 1(major professional) to 8 (unskilled labouring)²²

^c The cannabis/substance use scores were derived from a standardised clinical interview and range from 1(no use) to 6 (over 100 times)

^d Word reading score is the WRAT-4 standardised score

BNSIP2

	Healthy Control	Schizophrenia	Bipolar With Psychosis	Schizoaffective Disorder
N	495	393	206	350
Male (%)	196 (39.6)	240 (61.1)	79 (38.8)	156 (44.6)
Asian ethnicity (%)	46 (9.3)	20 (5.3)	4 (2.4)	8 (2.3)
Black or African American ethnicity(%)	148 (30.1)	196 (49.9)	36 (17.5)	129 (37.1)
Hispanic ethnicity (%)	66 (13.5)	42 (10.7)	30 (14.6)	47 (13.7)
White ethnicity (%)	209 (42.4)	113 (29.0)	124 (60.2)	136 (39.1)
Olanzapine Equivalents (mg)	0.0 (0.6)	12.9 (15.2)	7.3 (10.8)	11.4 (13.4)
Age (years)	34.4 (12.0)	39.3 (12.0)	34.9 (11.0)	40.0 (12.0)
Age of mother at the time of birth (years)	27.4 (6.6)	25.8 (6.7)	28.0 (6.4)	25.8 (6.3)
Age of father at the time of birth (years)	30.6 (7.3)	30.3 (8.5)	30.4 (7.2)	29.6 (8.0)
Father's Occupation ^b	3.6 (1.8)	4.3 (1.8)	3.4 (1.9)	4.1 (1.9)
Mother's Occupation ^b	3.6 (1.7)	4.3 (1.9)	3.4 (1.6)	4.1 (1.8)
Father Educational Scale ^b	3.0 (1.5)	3.5 (1.7)	2.9 (1.7)	3.5 (1.7)
Mother Educational Scale ^b	3.0 (1.4)	3.5 (1.6)	2.9 (1.5)	3.5 (1.6)
Subject Educational Scale ^b	2.4 (1.0)	3.5 (1.2)	2.7 (1.0)	3.4 (1.1)
Word Reading Score ^d	102.8 (43.0)	92.4 (14.9)	102.3 (14.2)	93.5 (14.2)
Emotional Abuse ^a	7.5 (3.6)	11.4 (5.5)	12.6 (5.8)	13.6 (6.1)
Emotional Neglect ^a	8.5 (4.1)	12.3 (5.3)	12.1 (5.2)	13.4 (5.3)
Physical Abuse ^a	6.8 (2.8)	9.2 (4.7)	8.6 (4.8)	10.6 (5.6)
Physical Neglect ^a	6.4 (2.1)	9.2 (4.0)	8.5 (4.1)	9.2 (4.0)
Sexual Abuse ^a	6.2 (3.4)	8.5 (5.5)	8.9 (6.6)	11.1 (7.4)
Use of cannabis after age 18 ^c	1.2 (1.6)	2.2 (2.2)	2.2 (2.1)	2.2 (2.2)
Use of cannabis between age 15 and 18 ^c	0.7 (0.9)	1.2 (1.5)	1.3 (1.5)	1.4 (1.6)
Use of cannabis before age 15 ^c	1.9 (2.1)	2.8 (2.4)	3.3 (2.3)	3.1 (2.4)
Use of other substances before age 18 ^c	0.2 (0.3)	0.4 (0.6)	0.7 (0.8)	0.5 (0.7)
Use of other substances after age 18 ^c	0.4 (0.7)	0.8 (1.0)	1.3 (1.2)	1.1 (1.1)

eTable1 – Study specific clinical and demographic details

^a Abuse and neglect scores are from the childhood trauma questionnaire²¹

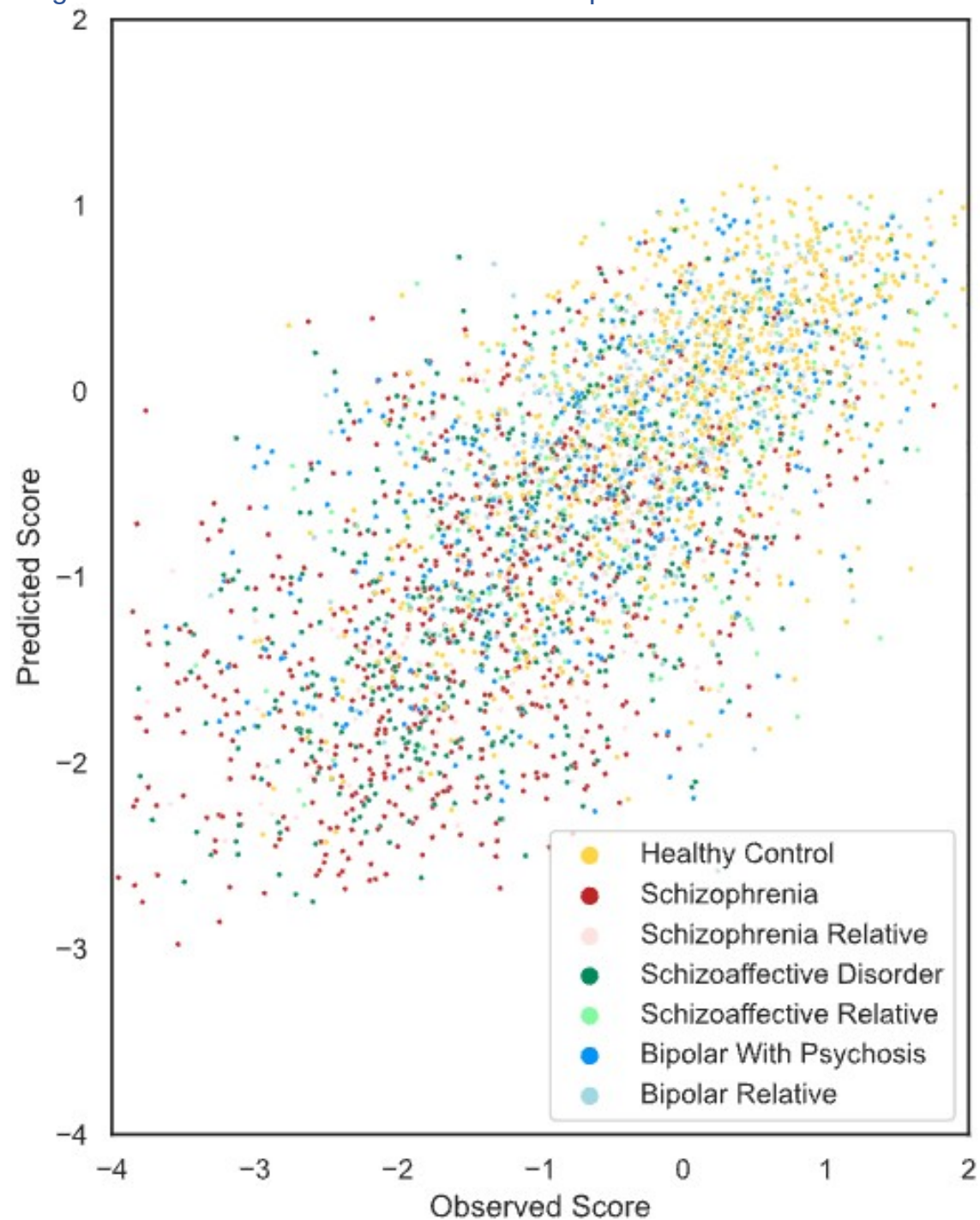
^b The Hollingshead Socioeconomic Rating Scale is used to classify educational achievement ranging from 1 (completed graduate degree) to 7 (less than seven years of schooling); and parental occupation ranging from 1 (major professional) to 8 (unskilled labouring)²²

^c The cannabis/substance use scores were derived from a standardised clinical interview and range from 1 (no use) to 6 (over 100 times)

^d Word reading score is the WRAT-4 standardised score

Significant differences ($p < 0.05$) were observed between BSNIP1 and BSNIP2 cohorts for olanzapine equivalents (all patient groups), age (controls, schizophrenia, and schizoaffective disorder groups), mothers occupation (controls and bipolar with psychosis groups), father's occupation (schizophrenia group), mothers educational scale (control group), subject educational scale (controls and bipolar with psychosis groups), word reading score (control group), sex (schizophrenia group), and ethnicity/race (controls and schizophrenia groups).

eFigure 1: Predicted and Observed Composite scores

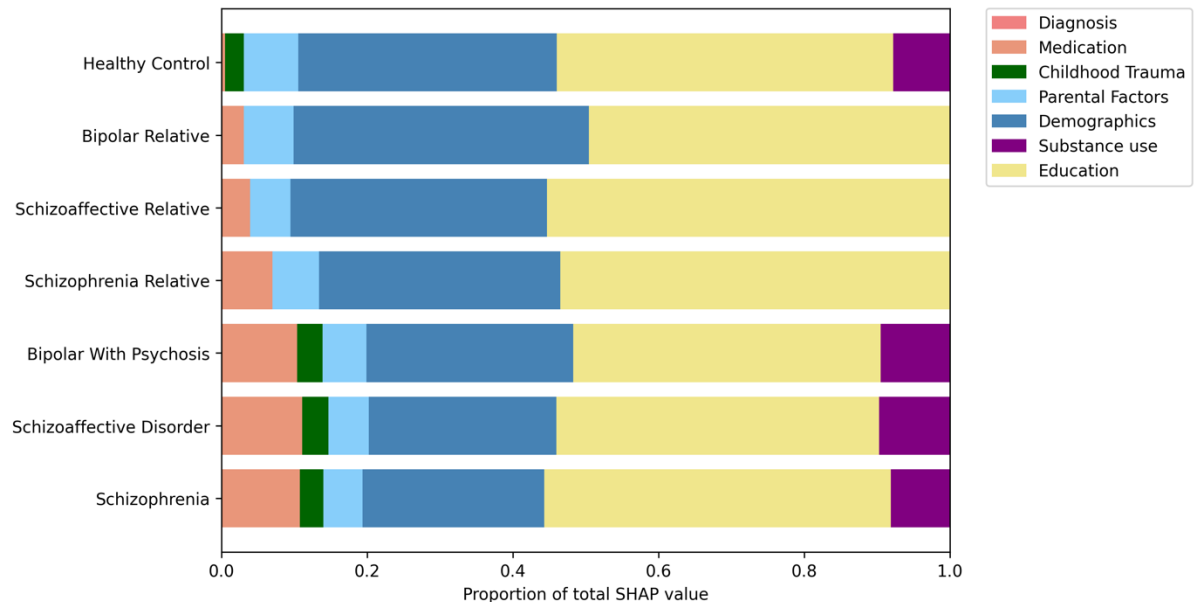


Out of sample correlation between predicted and observed composite cognitive score

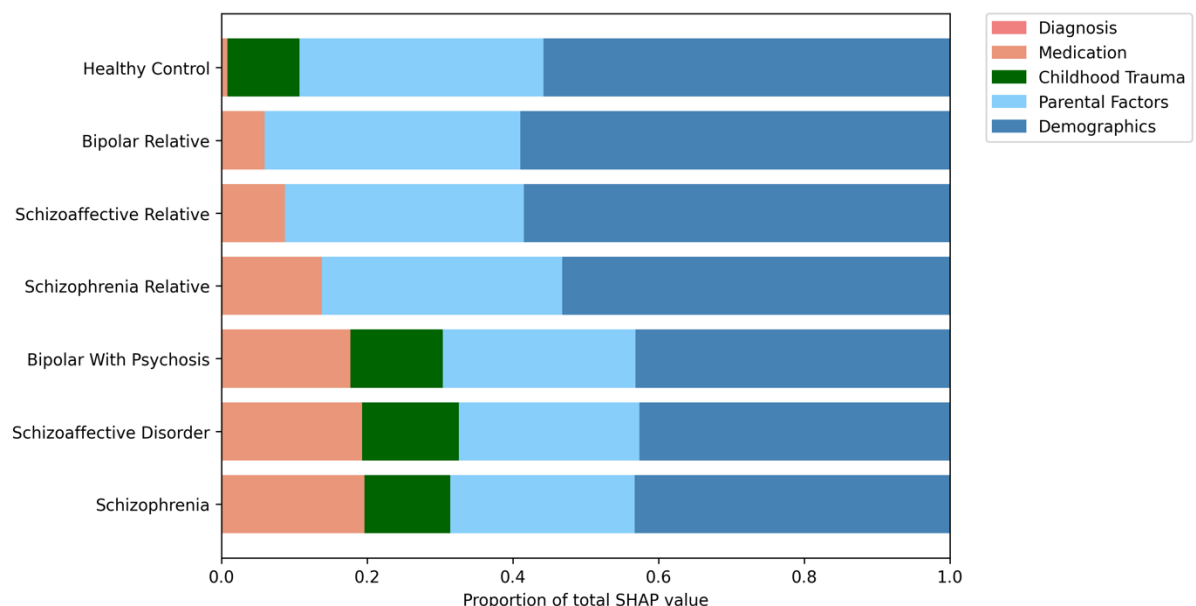
SHAP Values from linear models

Overall model performance was similar to that observed with XGBoost ($R^2=0.52$ ($SD=0.06$))

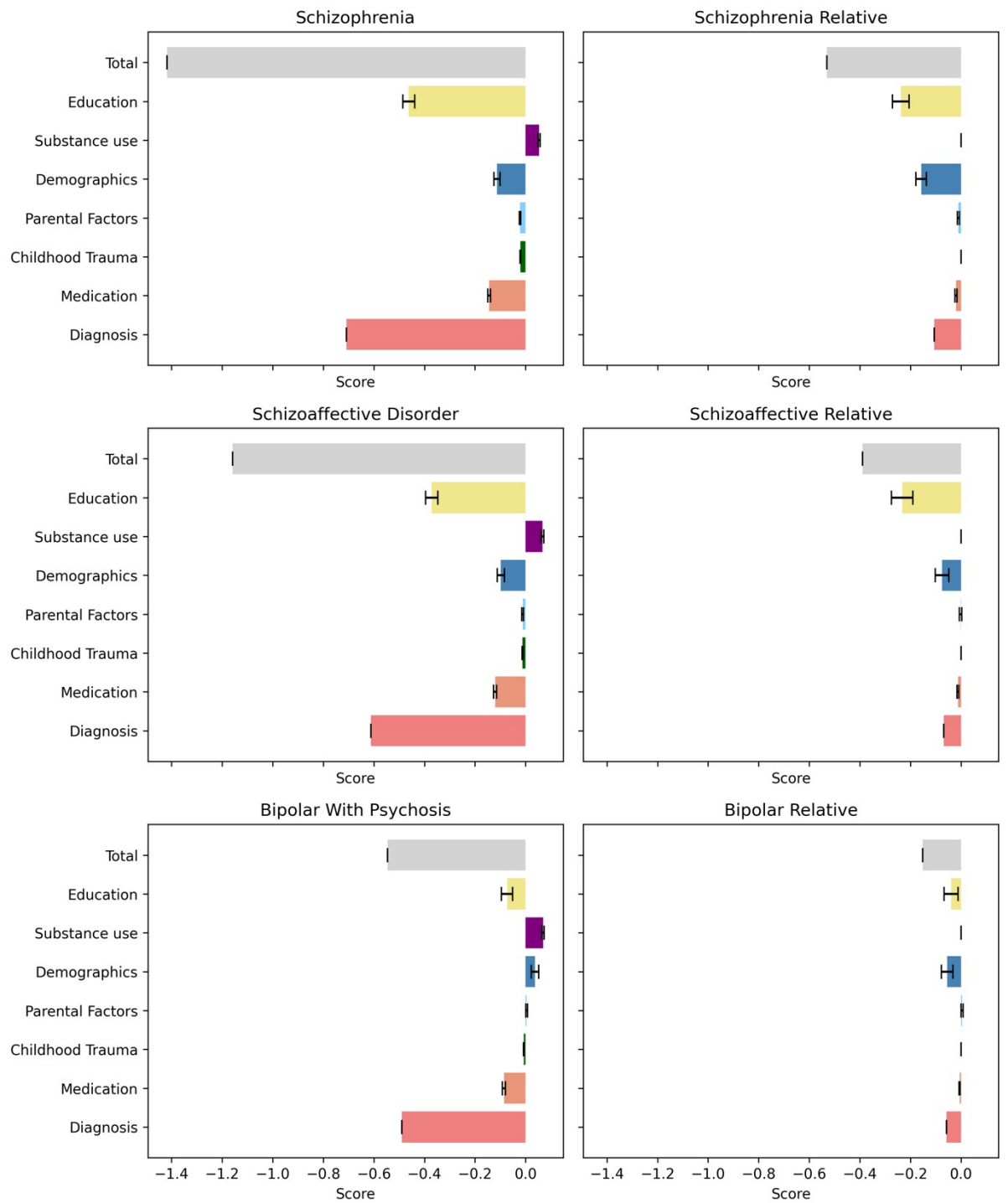
eFigure 2: Contribution of predictor variables to within diagnosis cognitive performance including education and substance use



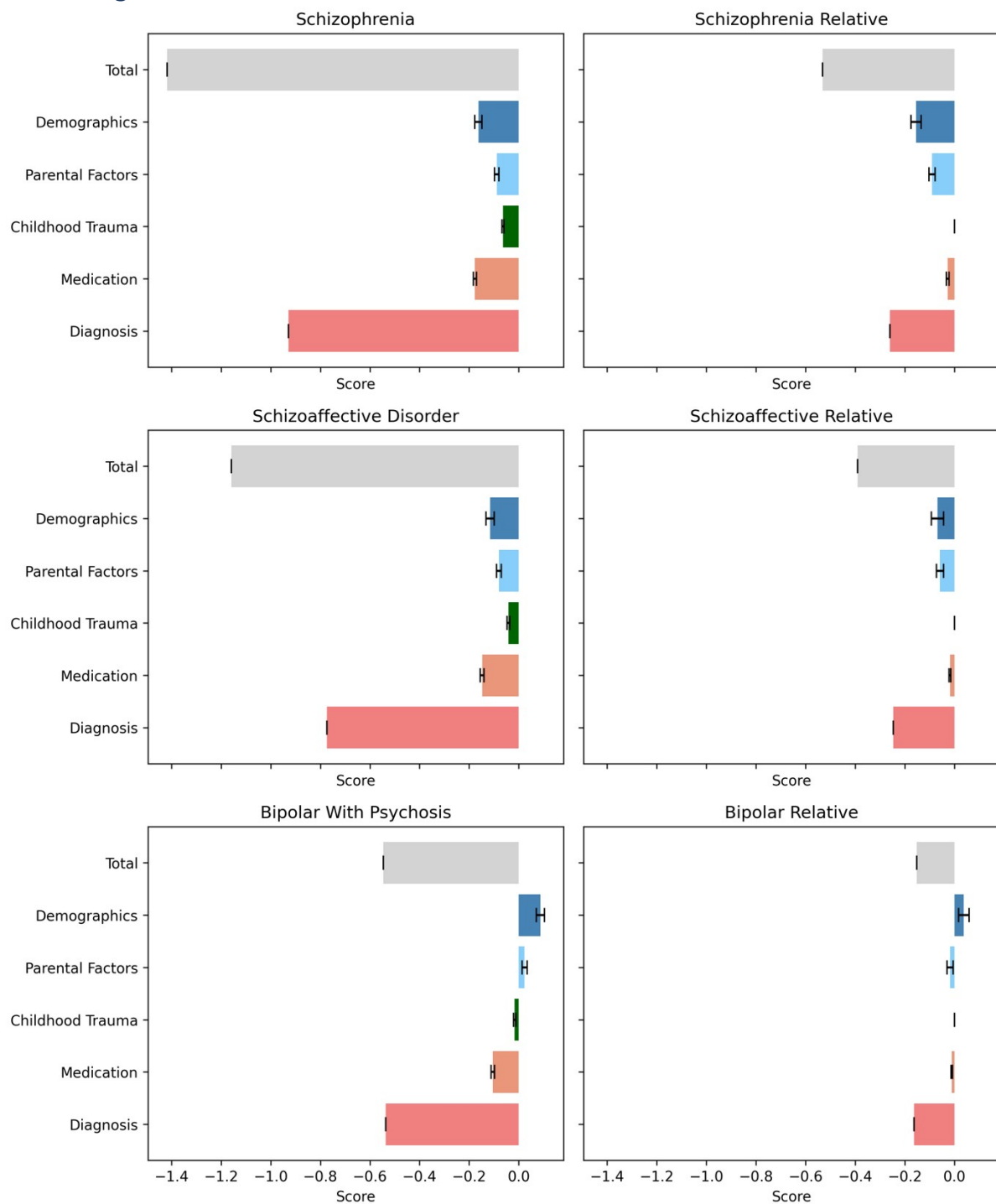
eFigure 3: Contribution of predictor variables to within diagnosis cognitive performance excluding education and substance use



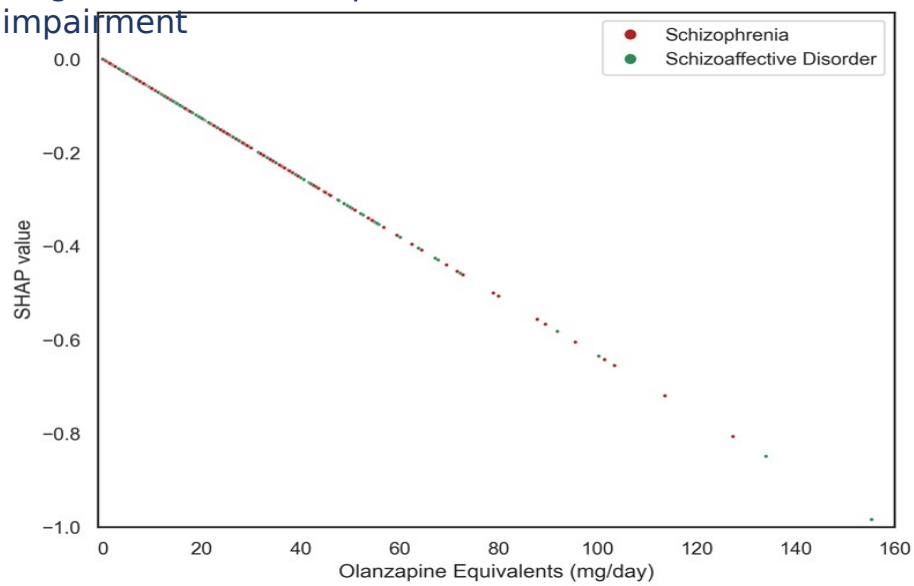
eFigure 4: Contribution of predictor variables to cognitive impairment including education and substance use



eFigure 5: Contribution of predictor variables to cognitive impairment excluding education and substance use

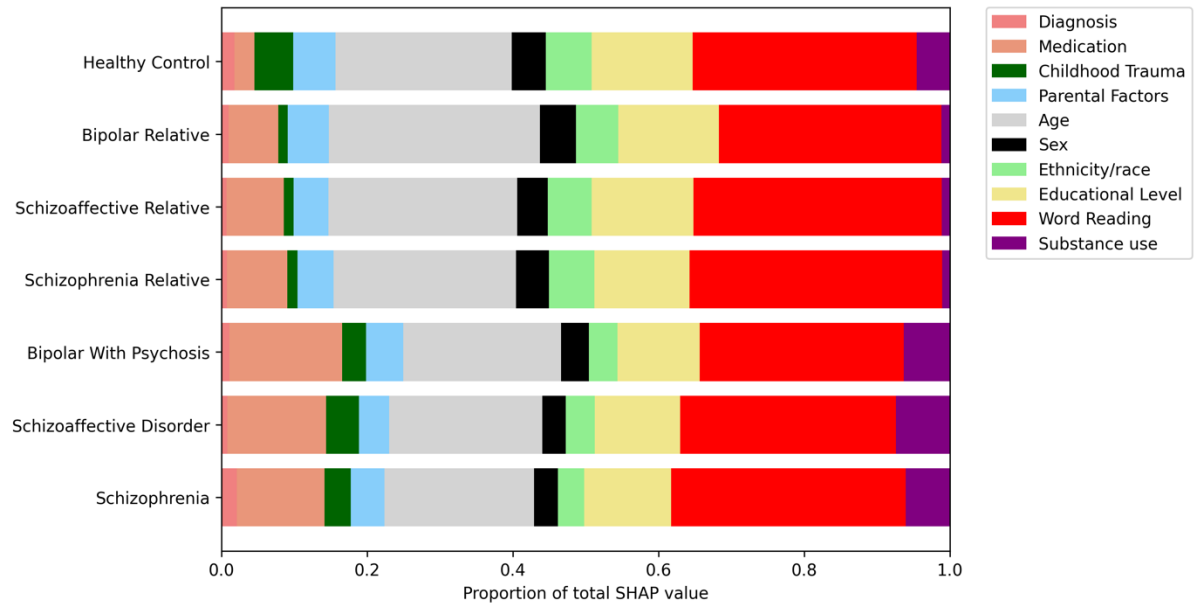


eFigure 6: Relationship between medication dose and cognitive impairment

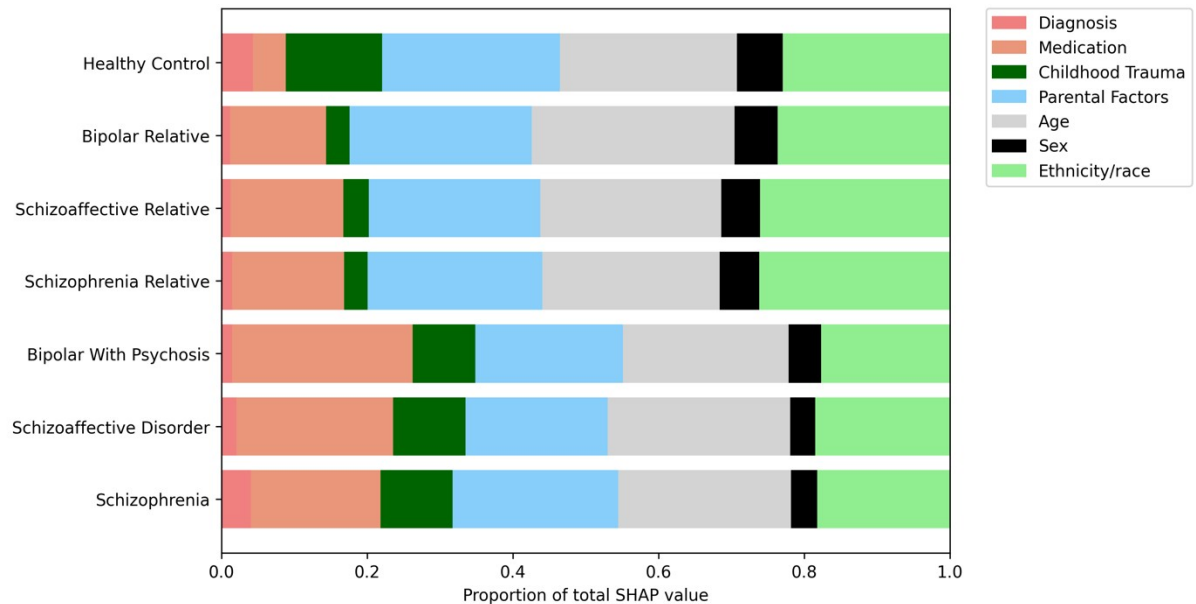


SHAP Values for uncollated demographic and educational predictors

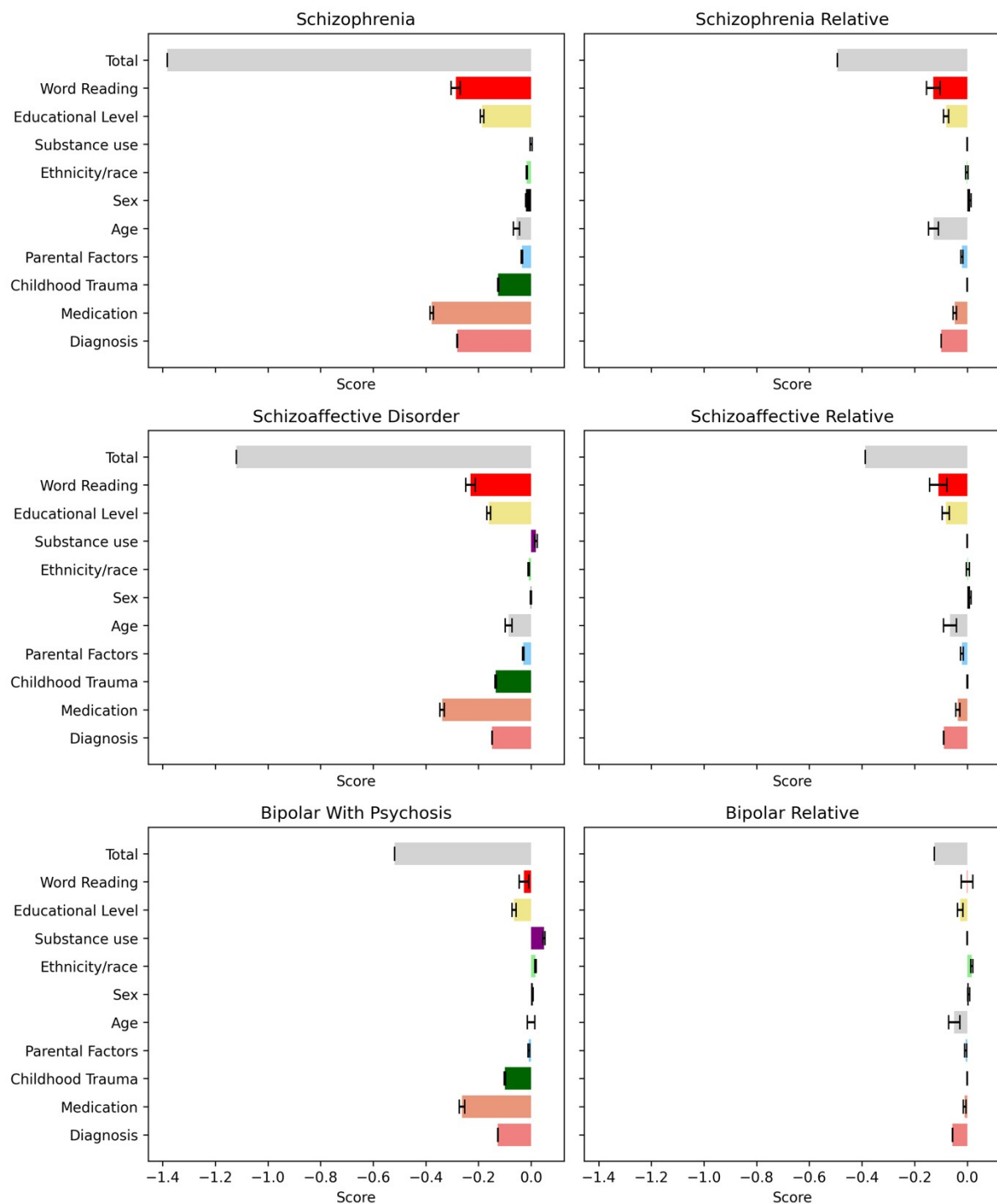
eFigure 7 Contribution of predictor variables to within diagnosis cognitive performance including education and substance use



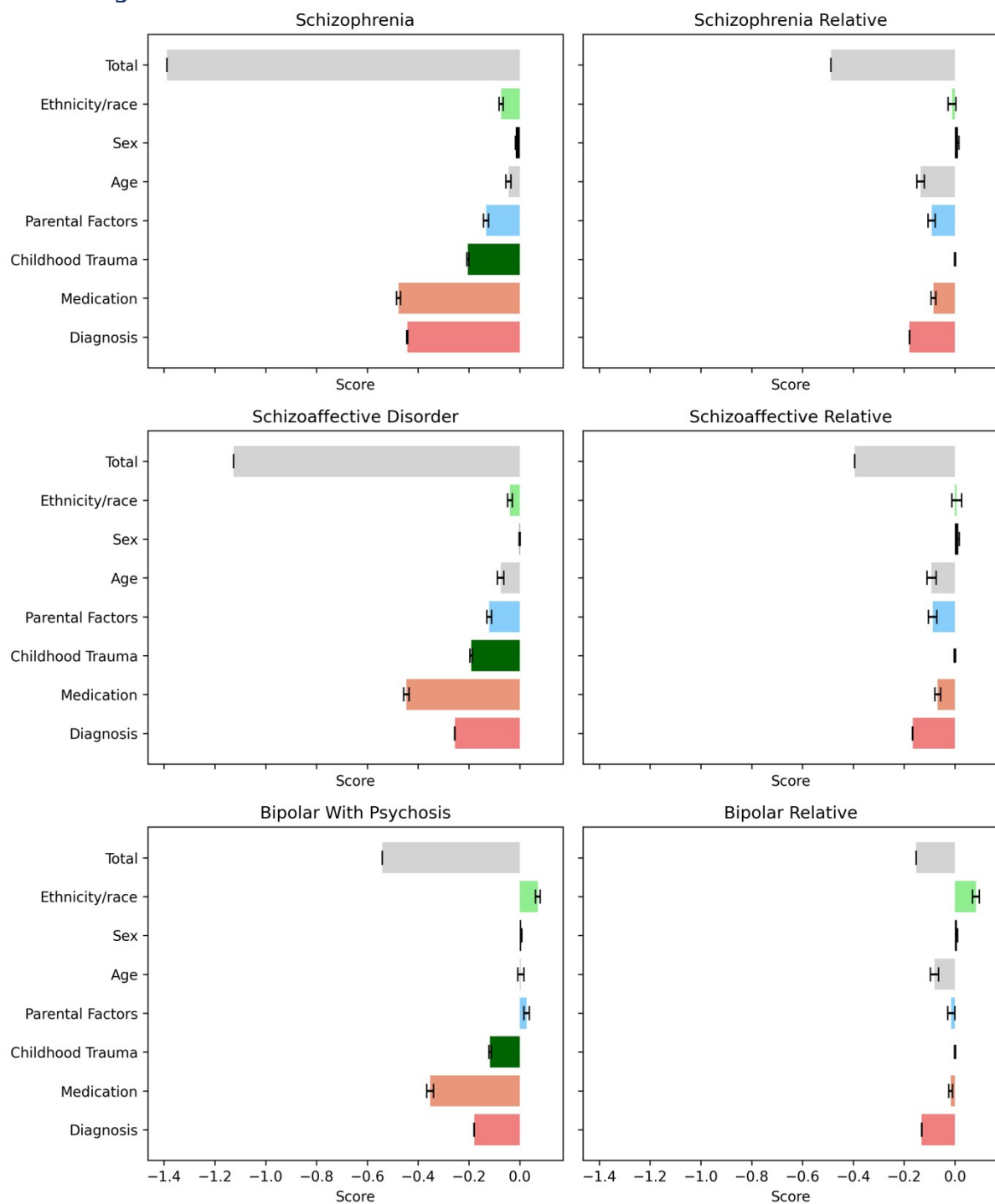
eFigure 8: Contribution of predictor variables to within diagnosis cognitive performance excluding education and substance use



eFigure 9: Contribution of predictor variables to cognitive impairment including education and substance use



eFigure 10: Contribution of predictor variables to cognitive impairment excluding education and substance use

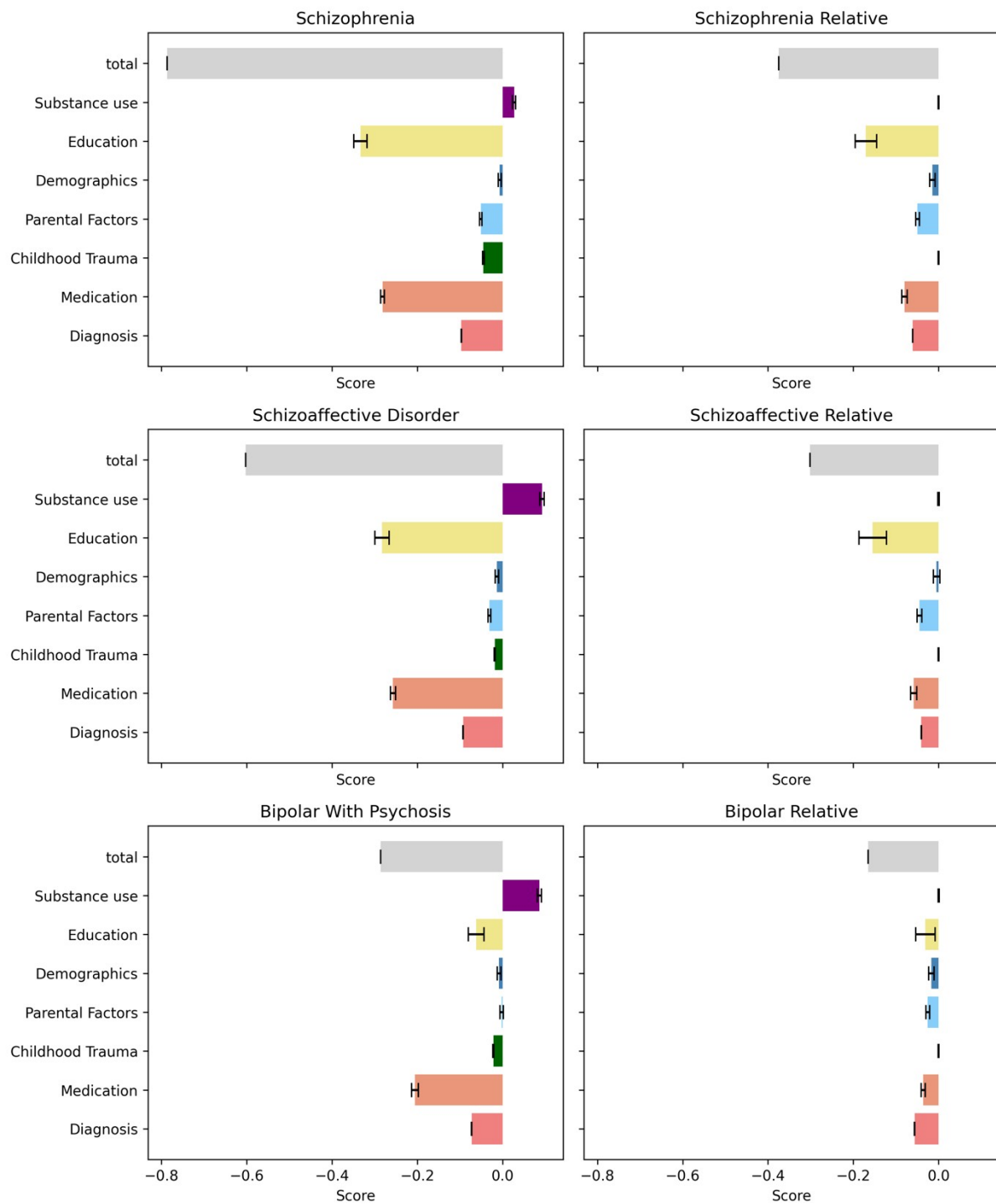


Domain Specific Relationships

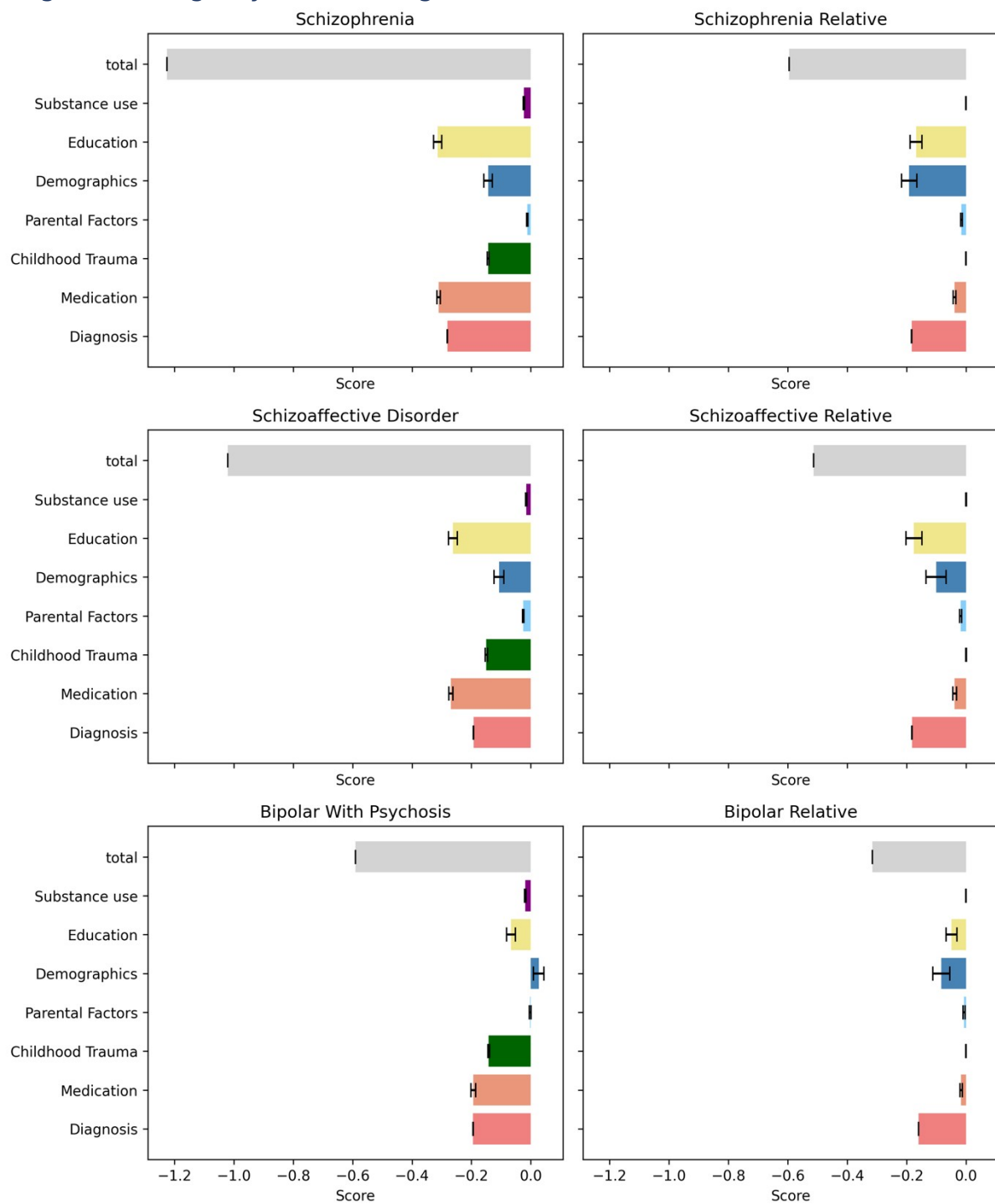
eFigure 11: Digit Span

eFigure 12: Verbal Fluency

Deconstructing Cognitive impairment in Psychosis: A Machine Learning Approach: Supplement

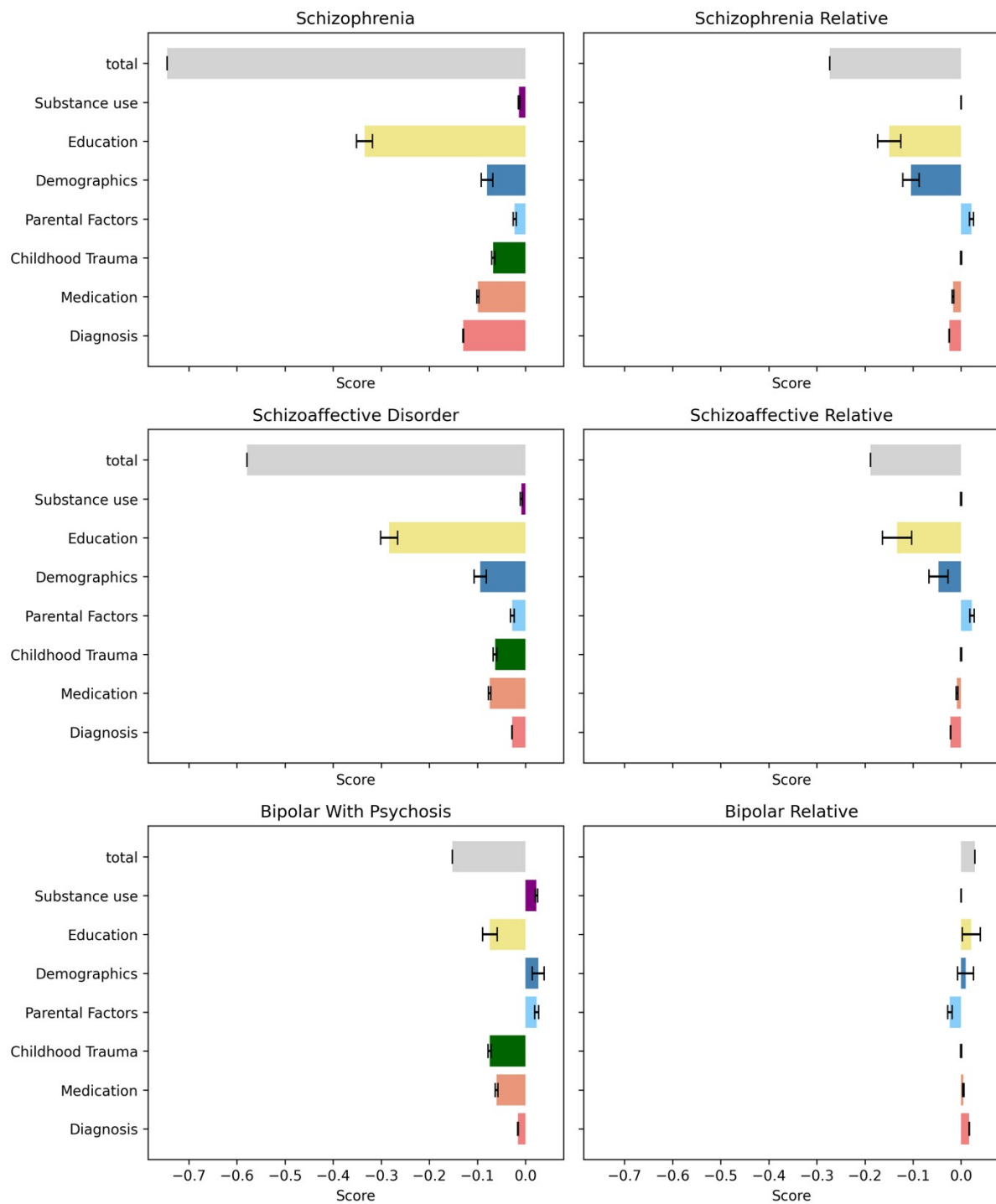


eFigure 13: Digit Symbol Coding



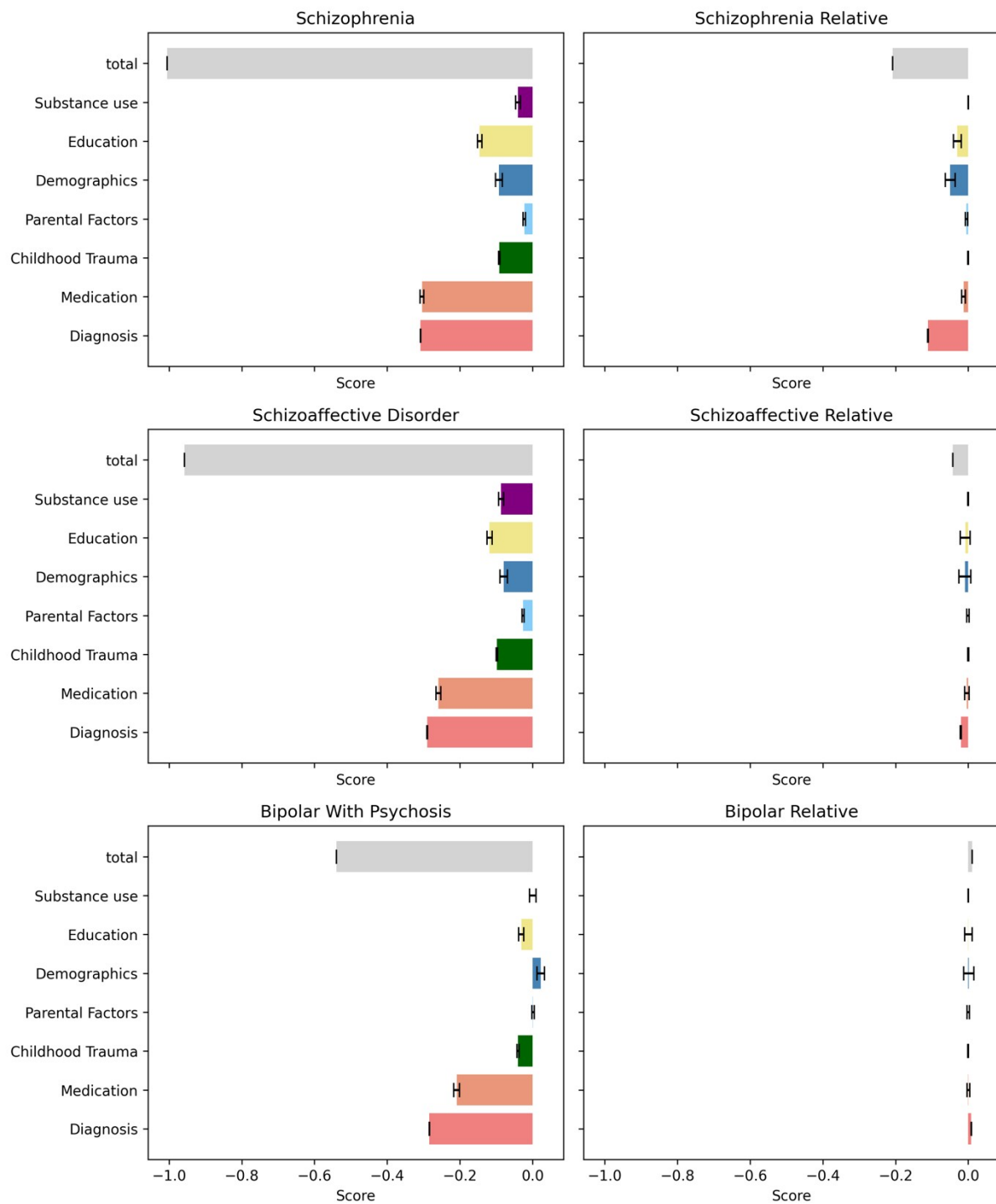
eFigure 14: Tower of London

Deconstructing Cognitive impairment in Psychosis: A Machine Learning Approach: Supplement



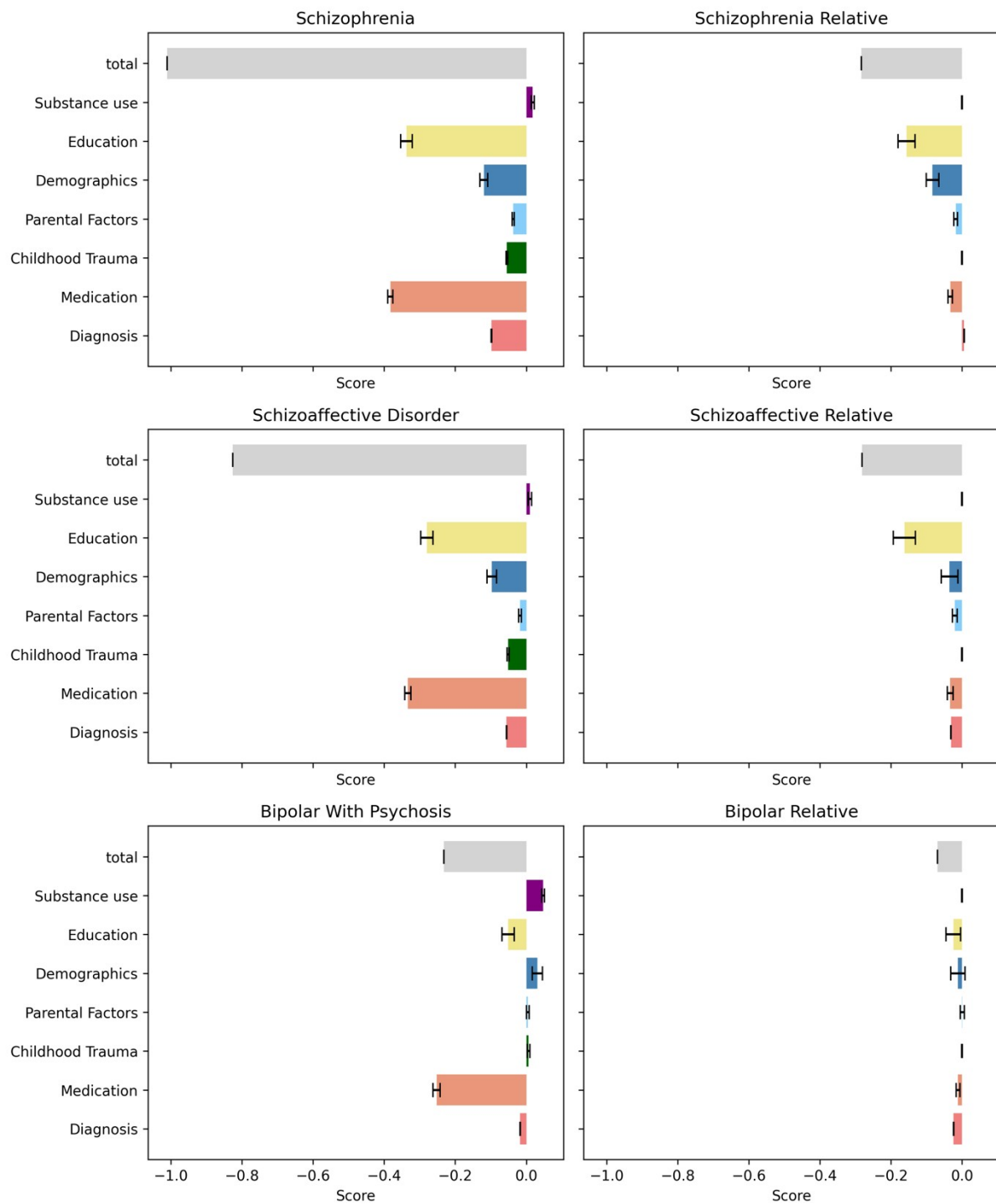
eFigure 15: Trail Making

Deconstructing Cognitive impairment in Psychosis: A Machine Learning Approach: Supplement

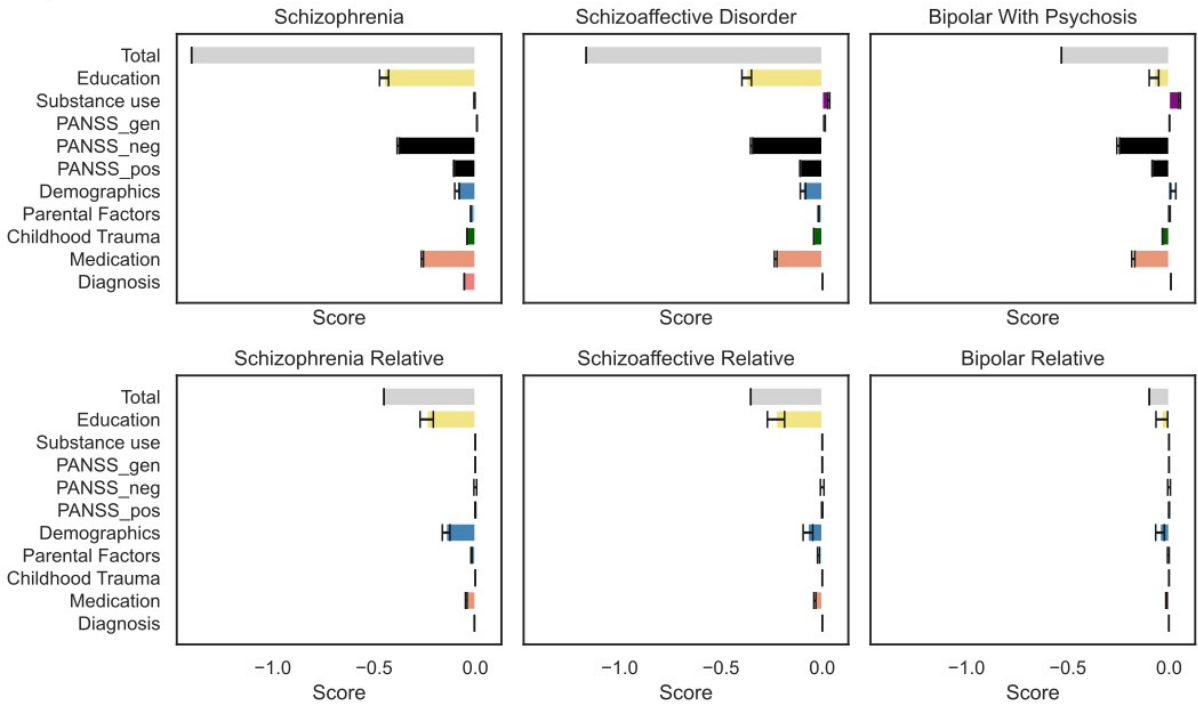


eFigure 16: Verbal Memory

Deconstructing Cognitive impairment in Psychosis: A Machine Learning Approach: Supplement



eFigure 17: Contribution of predictor variables to cognitive impairment when PANSS scores included



References

1. Tamminga CA, Ivleva EI, Keshavan MS, et al. Clinical phenotypes of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *Am J Psychiatry*. 2013;170(11):1263-1274.
2. Clementz BA, Parker DA, Trotti RL, et al. Psychosis Biotypes: Replication and validation from the B-SNIP consortium. *Schizophr Bull*. 2022;48(1):56-68.

