Table 1. List of 38 phenotypes in HCP-A used for prediction in ST. The information sections are in Lifespan HCP 2.0 Data Release Appendix 2: Details and References for Behavioral & Clinical Instruments.

Full Names	Info Section
Life satisfaction	3.2.b NIH Toolbox
Meaning and Purpose	3.2.c NIH Toolbox
Positive Affect	3.2.a NIH Toolbox
Sadness	3.1.a – NIH Toolbox
Fear-Affect	3.1.b NIH Toolbox - Fear-Affect Survey
Fear-Somatic	3.1.b NIH Toolbox - Fear-Somatic Survey
Anger-Affect	3.1.c NIH Toolbox - Anger-Affect Survey
Anger-Hostility	3.1.c NIH Toolbox - Anger-Hostility Survey
Anger-Aggression	3.1.c NIH Toolbox - Anger-Aggression Survey
Penn Emotion Recognition Test - anger	3.3 – Emotion Recognition
Penn Emotion Recognition Test - fear	3.3 – Emotion Recognition
Penn Emotion Recognition Test - no emotion	3.3 – Emotion Recognition
Penn Emotion Recognition Test - sad	3.3 – Emotion Recognition
Emotional Support	3.4 – Social Relationships
Instrumental Support	3.4 – Social Relationships
Friendship	3.4 – Social Relationships
Loneliness	3.4 – Social Relationships
Perceived Rejection	3.4 – Social Relationships
Perceived Hostility	3.4 – Social Relationships
Perceived Stress	3.4 – Social Relationships
Self-Efficacy	3.4 – Social Relationships
Agreeableness	3.8 – Personality - NEO-FFI Short Form
Openness	3.8 – Personality - NEO-FFI Short Form
Conscientiousness	3.8 – Personality - NEO-FFI Short Form
Neuroticism	3.8 – Personality - NEO-FFI Short Form
Extraversion	3.8 – Personality - NEO-FFI Short Form
Endurance	3.6 – Physical Functioning
Gait Speed	3.6 – Physical Functioning
Strength	3.6 – Physical Functioning
Picture Sequence Memory	2.1 – Cognition
Dimensional Change Card Sort	2.2 – Cognition
Flanker Task	2.3 – Cognition
Oral Reading Recognition	2.4 – Cognition
Picture Vocabulary	2.5 – Cognition
Processing Speed	2.6 – Cognition
Delay Discounting	2.11 – Self-regulation/Impulsivity - Delay Discounting
List Sorting	2.7 – Cognition
Pittsburgh Sleep Quality Index	3.5 – Sleep

Table 2. Hyperparameters for ST and TT stages. A learnable token with 1024 dimension was attached to the input, and we forwarded the output token for the last time point to either a linear layer with softmax (classification), or a linear layer solely (regression). All trainings were conducted on an Nvidia RTX 3090 GPU with 24GB of memory.

Hyperparameters	ST	TT	
		Classification	Regression
Batch Size	16	3/5/10	16
Learning Rate	$1e^{-4}$	$1e^{-3}$	$1e^{-3}$
Drop Out	0.1	0.2	0.2
Nonlinear in DIP	SiLU	SiLU	SiLU