SUPPLEMENT 1
Participant Age

Participants were 8-21 years old at time of recruitment, with a mean age of 14.23 years (SD=3.64 kurtosis and skewness were -1.14 and .016, respectively) at time of assessment. There were 2,358 participants between the ages of 8-10 (M=9.44; SD=.84), 2,135 participants between the ages of 11-13 (M=12.49; SD=.85), 2,323 participants between the ages of 14-16 years (M=15.48; SD .86), 2,162 participants between the ages of 17-19 years (M=18.34, SD=.8), and 434 participants were 20 years and older (M=20.72 SD=.5).

Executive Function (EF) Tasks From the Penn Computerized Neurocognitive Battery

Penn Continuous Performance task (PCPT). On this task, the stimuli consisted of 7-segment displays (e.g., similar to the display on a digital clock). Stimuli were presented on the screen for 300ms followed by a blank screen for 700ms. Total allowable response time per trial was 1000ms. The task was presented in two blocks with a total of 144 trials; targets were numbers during block one and letters for block two. Several behavioral measures were generated from the PCPT (response inhibition and attentional vigilance). Response inhibition represents the participant's overall ability to appropriately inhibit a motor response, and attentional vigilance represents the ability to focus attentional resources on specific stimuli. Of note, the PCPT attentional vigilance score loads highly on a general EF factor.[1]

Penn Conditional Exclusion Test (PCET). On this task, the stimuli consisted of shapes that could vary across several dimensions. These dimensions were either the shape itself (e.g., square, star), size (e.g., large, small), or the thickness of the lines that compromise the shape (e.g., thick, thin). On each trial, the participant was presented with four shapes and needed to figure which of the shapes was the "odd man out" (i.e., the dimension by which all but one of the stimuli matched). The rule (i.e., dimension by which the one stimuli "differed") changed after ten

consecutive correct trials. Participants were not informed of the rule switch and needed to learn the new rule through the feedback presented after every trial. The task consisted of a maximum of 144 trials, with a total of 48 possible trials for each new learning rule. A performance score was created to reflect overall correct responses and total rule learning by multiplying the overall correct responses by the number of rules learned (a value of 1 was added to number of rules learned to avoid multiplication by 0 when no rules were learned).

Penn Letter N-Back Test. On this task, stimuli were presented for 500ms, with an ISI of 2500ms. A total number of 90 trials were presented.

Signal Detection Measures. P_r reflects the ability to accurately detect whether the trial was a "go" or "no-go" and is conceptually similar to d', but does not require a correction in instances of small numbers (i.e., no errors of commission).[2] For B_r, negative values reflect a cautious response style (i.e., bias to withhold response); positive values reflect an impulsive response style (i.e., bias to execute response). P_r was calculated as the hit rate ([0.5+correct targets trials]/[1+total target trials]) minus the false alarm rate ([.05+ incorrect foil trials]/[1+total foil trials]). B_r was calculated with the formula (false alarm rate/[1- P_r])-0.05. Measures of P_r and B_r are mathematically independent.

Bifactor Model

Model Loadings and Variance Explained. The values for $Omega_H$, $Omega_{Anxious\text{-}Misery}$, $Omega_{Fear}$, $Omega_{Externalizing}$, and $Omega_{Psychosis}$ are 0.86, .04, .03, .03, and .02, respectively. The sum of squared loadings general psychopathology, mood, fear, externalizing, and psychosis are 40.9, 7.3, 6.9, 7.2, and 5.6, respectively. Table S2 presents the top 25 highest loading items on each of the clinical domains.

Associations Between Psychopathology and Signal Detection Variables. The models containing the signal-detection theory EF subcomponents of continuous performance task (CPT) discrimination accuracy, CPT response bias, and NBACK discrimination accuracy had the following fit indices: comparative fit index (CFI) = 0.96, 0.90, 0.94, and 0.90, respectively; Tucker-Lewis Index (TLI) = 0.96, 0.89, and 0.94, respectively; root mean-square error of approximation (RMSEA) = 0.027±0.001, 0.027±0.001, and 0.027±0.001, respectively. For the NBACK response bias model, a Bayes estimator was used to achieve convergence. As such, model fit indexes are not available for this model.

Results from these models (Table S2) revealed that for all clinical domains except for externalizing and fear domains, lower discrimination accuracy across the two tasks predicted higher symptoms. For the fear domain, higher discrimination accuracy predicted higher symptoms, but only on the N-Back task. Interestingly, for the response bias measures, the only significant main effect emerged in the fear domain. A negative response bias (i.e., cautious bias) on the PCPT predicted heightened fear symptoms.

References

1. Moore TM, Reise SP, Gur RE, Hakonarson H, Gur RC. Psychometric properties of the Penn Computerized Neurocognitive Battery. Neuropsychology [Internet]. 2015;29:235-46.
2. Snodgrass JG, Corwin J. Pragmatics of measuring recognition memory: Applications to dementia and amnesia. J Exp Psychol Gen. 1988;117:34-50.

Table S1. Factor Loadings From the General Execution Function Bifactor Model

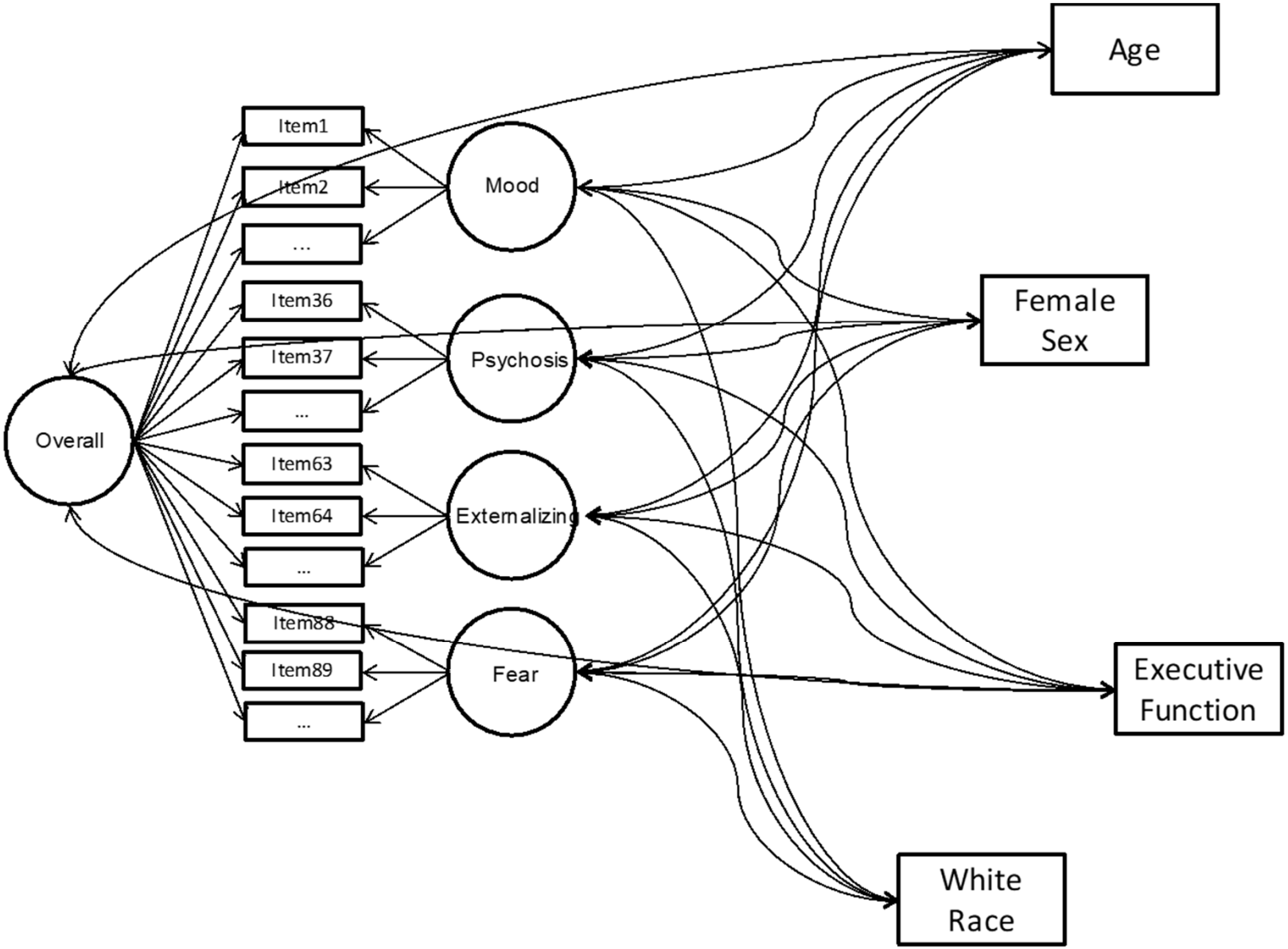
General Psychopathology		Anxious Misery		Fear		Externalizing		Psychosis	
Item	Loading	Item	Loading	Item	Loading	Item	Loading	Item	Loading
MAN	0.84	GAD	0.83	SOC	0.72	ADD	0.81	PSY	0.81
MAN	0.84	GAD	0.82	SOC	0.70	ADD	0.81	PSY	0.77
DEP	0.83	OCD	0.59	SOC	0.70	ADD	0.71	PSY	0.63
MAN	0.83	OCD	0.58	SOC	0.67	ADD	0.68	PSY	0.62
MAN	0.83	PAN	0.57	AGR	0.64	ADD	0.68	PSY	0.58
OCD	0.83	OCD	0.55	AGR	0.63	ODD	0.64	PSY	0.57
MAN	0.82	OCD	0.54	AGR	0.62	ADD	0.63	PSY	0.57
OCD	0.82	OCD	0.51	AGR	0.60	ODD	0.63	PSY	0.56
MAN	0.81	DEP	0.51	SOC	0.59	ADD	0.61	PSY	0.56
OCD	0.80	PAN	0.51	AGR	0.58	ADD	0.61	PSY	0.55
OCD	0.80	OCD	0.50	AGR	0.55	ODD	0.58	PSY	0.54
DEP	0.79	DEP	0.50	AGR	0.54	ADD	0.56	PSY	0.49
OCD	0.77	OCD	0.48	PHB	0.50	ODD	0.53	PSY	0.49
DEP	0.76	PAN	0.48	SEP	0.49	CD	0.52	PSY	0.46
MAN	0.74	OCD	0.46	PHB	0.47	ODD	0.50	PSY	0.46
PAN	0.73	OCD	0.45	SEP	0.43	CD	0.43	PSY	0.43
DEP	0.73	OCD	0.44	PHB	0.43	CD	0.42	PSY	0.35
PSY	0.72	OCD	0.44	PHB	0.42	CD	0.39	PSY	0.27
PSY	0.70	OCD	0.41	AGR	0.40	TX	0.38	PSY	0.23
PSY	0.70	OCD	0.41	PHB	0.40	CD	0.37	PSY	0.21
ODD	0.70	SUI	0.39	PHB	0.38	CD	0.36	PSY	0.06
OCD	0.70	OCD	0.39	SEP	0.38	CD	0.24	MAN	0.03
PSY	0.69	OCD	0.38	PHB	0.36	TX	0.21	MAN	-0.02
SUI	0.69	OCD	0.36	PHB	0.33	PSY	0.21	MAN	-0.03
OCD	0.69	PSY	0.36	SEP	0.30	CD	0.19	MAN	-0.04

Note. Loading refers to estimated factor loadings. ADD = attention hyperactivity; AGR = agoraphobia; CD = conduct; DEP = depression; GAD = generalized anxiety disorder; MAN = Mania; OCD = obsessive-compulsive disorder; ODD = oppositional defiant; PAN = panic; PHB = phobia; PSY = psychosis; SEP = separation; SOC = social anxiety; SUI = suicide; TX = treatment.

Table S2. Signal Detection Variable (Discrimination Accuracy and Response Bias) Predicting Clinical Domains

	Overall Psychopathology		Anxious Misery		Fear		Externalizing		Psychosis	
	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*
PCPT P_r										
P_r	**-0.110**	.000	**-0.249**	.000	-0.095	.002	-0.046	.105	**-0.325**	.000
Age	**0.294**	.000	**-0.205**	.000	**-0.183**	.000	**-0.154**	.000	**-0.252**	.000
Sex	**-0.112**	.000	**0.331**	.000	**0.328**	.000	**-0.059**	.000	**0.084**	.000
Race	**-0.229**	.000	**0.242**	.000	0.032	.024	0.013	.351	0.023	.131
$P_r \times$ Age	**0.518**	.000	**-0.402**	.000	**-0.515**	.000	**-0.433**	.000	**-0.229**	.000
$P_r \times$ Sex	**-0.430**	.000	**0.859**	.000	**0.654**	.000	**0.394**	.000	**0.571**	.000
PCPT B_r										
B_r	0.032	.013	0.010	.493	**-0.093**	.000	0.016	.212	-0.026	.062
Age	**0.274**	.000	**0.125**	.000	-0.036	.012	**-0.125**	.000	**-0.079**	.000
Sex	0.011	.403	**0.225**	.000	**0.222**	.000	**-0.170**	.000	**-0.057**	.000
Race	**-0.181**	.000	**0.190**	.000	**-0.094**	.000	**-0.071**	.000	**-0.065**	.000
$B_r \times$ Age	0.038	.002	0.013	.400	**-0.088**	.000	0.012	.368	-0.025	.066
$B_r \times$ Sex	0.028	.028	0.003	.837	**-0.061**	.000	0.032	.014	-0.039	.006
N-Back P_r										
P_r	**-0.181**	.001	**-1.025**[a]	.000	**0.697**	.000	-0.138	.035	**-0.465**	.000
Age	**0.288**	.000	**-0.228**	.000	0.015	.494	**-0.161**	.000	**-0.193**	.000
Sex	**-0.094**	.000	**0.357**	.000	**0.273**	.000	**-0.060**	.000	**0.083**	.000
Race	**-0.229**	.000	**0.149**	.000	0.044	.006	-0.003	.818	-0.021	.178
$P_r \times$ Age	**0.499**	.000	**0.409**	.000	**-1.162**[a]	.000	**-0.307**	.000	-0.021	.766
$P_r \times$ Sex	**-0.356**	.000	**0.863**	.000	**0.423**	.000	**0.380**	.000	**0.510**	.000
N-Back B_r										
B_r	0.135	.002	-0.032	.282	-0.051	.172	0.126	.008	0.068	.148
Age	**0.273**	.000	**0.147**	.000	-0.022	.075	**-0.121**	**.000**	**-0.044**	**.001**
Sex	**-0.064**	.000	**0.237**	.000	**0.219**	**.000**	**-0.137**	**.000**	-0.007	.308
Race	**-0.189**	.000	**0.174**	.000	**-0.062**	**.000**	**-0.057**	**.000**	**-0.112**	**.000**
$B_r \times$ Age	-0.132	.002	0.054	.154	0.036	.247	-0.085	.046	-0.059	.171
$B_r \times$ Sex	0.004	.403	-0.004	.414	-0.007	.346	-0.030	.043	-0.018	.193

Note: Boldface indicates significant effects ($p \leq .001$). PCPT = Penn Continuous Performance task.
[a]Though model estimation terminated normally, coefficients with absolute value >1.0 were possible due to non-positive-definite residual variance/covariance matrices.


Age
Item1
Item2
...
Mood
Item36
Item37
...
Psychosis
Female
Sex
Overall
Item63
Item64
...
Externalizing
Item88
Item89
...
Fear
Executive
Function
White
Race