SUPPLEMENTARY MATERIAL.

Statistical Analyses.

1) **Pattern Matrix and Factor Loadings for the 4-factor and for the 5-factor solutions from the Exploratory Factor Analyses in Healthy Controls.**

For Healthy Controls, we conducted an Exploratory Factor Analysis (EFA) with OBLIMIN rotation. The Principal Component Analysis (PCA) and Scree Plot showed the presence of 5 components with Eigen Values > 1, although the Eigenvalue for 5 was 1.014. The 5-factor solution separated out the 5 theoretical domains used to design the assessments, with auditory and visual assessments highly loading onto their specific construct and not on the others. In the 4-factor solution, performances on the auditory and visual speed of processing and executive functioning tasks loaded onto the same component.

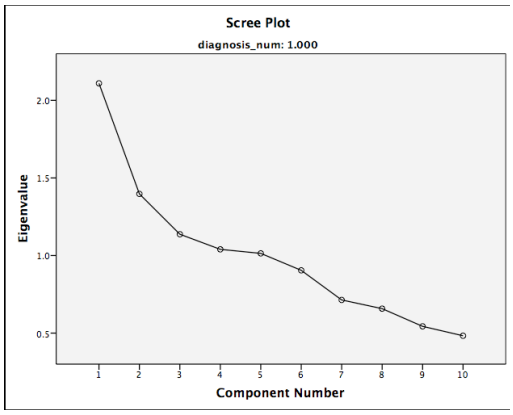
Total Variance Explained[a]

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings[b]
	Total	% of Variance	Cumulative %	Total
1	2.110	21.096	21.096	1.459
2	1.397	13.970	35.067	1.498
3	1.137	11.372	46.438	1.196
4	1.040	10.398	56.836	1.350
5	1.014	10.138	66.975	1.575
6	.905	9.046	76.020	
7	.714	7.137	83.158	
8	.658	6.577	89.734	
9	.544	5.436	95.170	
10	.483	4.830	100.000	

Extraction Method: Principal Component Analysis.

a. diagnosis_num = 1.000

b. When components are correlated, sums of squared loadings cannot be added to obtain a total variance.



Pattern Matrix[a,b]

	Component			
	1	2	3	4
BHQ_SoundSweeps_threshold_flipped	.747	-.017	.083	.210
BHQ_VisualSweeps_threshold_flipped	.638	-.217	.231	-.105
BHQ_SustainedAuditoryAttention_sdRT_target_flipped	-.055	.835	.038	.024
BHQ_SustainedVisualAttention_sdRT_target_flipped	.128	.698	.190	-.025
BHQ_AuditoryTaskSwitcher_threshold_flipped	.592	.210	-.087	-.099
BHQ_TaskSwitcherDouble_threshold_flipped	.532	.111	-.374	-.085
BHQ_Auditory Associates_reversal_numItems	-.042	.346	-.160	-.580
BHQ_Visual Associates_reversal_numItems	-.023	-.158	.144	-.880
BHQ_VoiceChoice_speedthreshold_flipped	.238	-.011	.635	-.257
BHQ_Emotion Motion_threshold_speed	-.087	.240	.714	.083

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

a. diagnosis_num = 1.000

b. Rotation converged in 7 iterations.

Pattern Matrix[a,b]

	Component				
	1	2	3	4	5
BHQ_SoundSweeps_threshold_flipped	.129	.079	-.221	.224	-.835
BHQ_VisualSweeps_threshold_flipped	.337	-.273	.149	-.124	-.496
BHQ_SustainedAuditoryAttention_sdRT_target_flipped	-.057	.884	.017	.074	-.066
BHQ_SustainedVisualAttention_sdRT_target_flipped	.307	.564	.411	-.003	.061
BHQ_AuditoryTaskSwitcher_threshold_flipped	.588	.123	.030	-.064	-.224
BHQ_TaskSwitcherDouble_threshold_flipped	.810	-.036	-.059	-.034	.092
BHQ_Auditory Associates_reversal_numItems	.025	.468	-.292	-.526	-.007
BHQ_Visual Associates_reversal_numItems	.091	-.139	.054	-.887	.051
BHQ_VoiceChoice_speedthreshold_flipped	-.309	.076	.244	-.301	-.654
BHQ_Emotion Motion_threshold_speed	-.055	.042	.032	.014	-.873

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

a. diagnosis_num = 1.000

b. Rotation converged in 18 iterations.

In Schizophrenia patients, the PCA and scree plot for the EFA with OBLIMIN rotation identified two possible components. However, the Eigen Value for cCmponent 1 was 4.007 and the Eigenvalue for Component 2 was 1.115, suggesting the presence of a preponderant latent variable, putatively a common generalized deficit, and possibly of a small second factor.

2) Criteria to compare Confirmatory Factor Analyses (CFA) models

We compared CFA factor models using the following parsimony-adjusted information criterion-based fit indices: the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and sample size adjusted BIC (SABIC) - for which the model with lower values is favored, χ 2 divided by degrees of freedom (χ2/df) - for which lower values are preferred, the root mean square error of approximation (RMSEA) - for which models with values < .06 and closer to 0 are preferred, and comparative fit index (CFI) – for which models with values > .95 and closer to 1are preferred (Hu and Bentler, 1999).[1]
Models were also evaluated for statistical fit with a chi-square test, comparing the proposed model to the null model. A significant p-value indicates inadequate fit, as in this case a significant p-value would indicate that the null model fit better than the proposed model.
Finally, nested models were evaluated using the chi-square difference test (Δχ2), which determines whether the model fit is significantly changed as a result of adding or removing free parameters from the model (Bollen, 1989).[2] Specifically, the Δχ2 explicitly tests whether the change in χ2 value is meaningful considering the change in df, and a significant Δχ2 statistic favors the model with the lower chi-square value.

1. Hu, L. & Bentler, P. M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* **6,** 1–55 (1999).
2. Bollen, K. A. A New Incremental Fit Index for General Structural Equation Models. *Sociol. Methods Res.* **17,** 303–316 (1989).

---

1. Description of Assessments

Domain of Perception

The **Sound Sweeps** assessment is a time-order judgment task of a sequence of two frequency-modulated (FM) sound sweeps, and it measures auditory psychophysical efficiency, indicative of both auditory perceptual and attentional operations. Users are asked to identify the direction of tonal change in a sequence of two successive FM sound sweeps presented for a short duration, as either “up” (from a lower to a higher pitch) or “down” (from a higher to a lower pitch). Users are asked to indicate whether the two sweeps were upward-downward, upward-upward, downward-downward, or downward-upward, by clicking up and down arrow buttons appearing on the screen. Wrong answers on the first sweep automatically progress the task to the next trial. Auditory processing speed is determined using a single staircase method, a 2-down-1-up procedure for determining sensitivity measures (i.e., estimating threshold; Levitt, 1971). The staircase algorithm adaptively modifies the interstimulus interval (ISI) between the two FM sound sweeps and the sweep duration, which is held equal to the ISI, as the performance changes trial by trial. The length of the sound sweeps starts at 251ms, with a maximal length of 1000ms. The staircase procedure terminates after 40 trials. The resulting threshold is **the logarithm** with **base 10 of** number of seconds of ISI (and sweep duration) at which the subject correctly performs 70.7% of trials, allowing for a measure of auditory psychophysical threshold under moderate perceptual challenge. Therefore, a more positive threshold is indicative of a greater speed of processing *deficit* in the auditory modality.
https://research.brainhq.com/#assessment/1E

The **Visual Sweeps** assessment is a time-order judgment task of a sequence of two Gabor Patches, and it measures visual psychophysical efficiency, indicative of both visual perceptual and attentional operations. In the assessment, users are asked to identify the spatial direction of change in a sequence of two successive FM visual sweeps, as either “inward” (from a lower to a higher spatial frequency) or “outward” (from a higher to a lower spatial frequency). Two response buttons showing inward and outward directions then appear on the screen. Users are asked to hit the response buttons in the correct order. Wrong answers on the first sweep automatically progress the task to the next trial. Visual processing speed is determined using a single staircase method (2-down-1-up). The staircase algorithm adaptively modifies ISI between the two FM visual sweeps and the sweep duration, which is held equal to the ISI, as the performance changes trial by trial. The length of the sweeps starts at a length of 200ms, with a maximal length of 1000ms. The staircase procedure terminates after 40 trials. The resulting threshold is **the logarithm** with **base 10 of** number of milliseconds of ISI (and

sweep duration) at which the subject correctly performs 70.7% of trials, allowing for a measure of visual psychophysical threshold under moderate perceptual challenge. Therefore, a more positive threshold is indicative of a greater speed of processing *deficit* in the visual modality.
https://research.brainhq.com/#assessment/41

Domain of Attention

The **Sustained Auditory Attention** assessment is a Test Of Variables of Attention (T.O.V.A.) - a continuous performance (go-no-go) test. In the assessment, users are asked to press the spacebar as fast as possible whenever a target (a 750Hz sound) is presented, and to withhold the response whenever a foil (a 440Hz sound) is presented. Targets are presented 33.3% of the trials, and foils are presented 66.6% of the trials. Each stimulus is presented for 1000ms, followed by a response interval of 1200ms. The assessment consists of 200 trials and is non-adaptive. The main outcome is the standard deviation of the Reaction Time (RT) for accurate target trials. Participants with irregular RTs are hypothesized to have more attentional lapses. Therefore, a more positive threshold is indicative of a greater attentional deficit in the auditory modality.
https://research.brainhq.com/#assessment/39/

The **Sustained Visual Attention** assessment is a T.O.V.A., a continuous performance (go-no-go) test where users are asked to press the spacebar as fast as possible whenever a target is presented (star at the top half of the screen), and to withhold the response whenever a foil (star at the bottom half of the screen) is presented. Targets are presented 33.3% of the trials, and foils are presented 66.6% of the trials. Each stimulus is presented for 1000ms, followed by a response interval of 1200ms. The assessment consists of 200 trials and is non-adaptive. The main outcome is the standard deviation of the RT for accurate target trials. Participants with irregular RTs are hypothesized to have more attentional lapses. Therefore, a more positive threshold is indicative of a greater attentional deficit in the visual modality.
https://research.brainhq.com/#assessment/3A

Domain of Executive Functioning

The **Auditory Task Switcher** assessment is a task switching task that measures the set shifting process in the auditory modality for the domain of executive functioning. In each trial of the assessment, a number is played. Users must decide, as quickly as possible, whether the number is odd or even, or whether it is spoken by a male or female voice, depending on the instruction written on the screen. The decision rule (odd/even vs. male female) changes pseudo-randomly from trial to trial, with a 50% change of occurring at the start of each trial. The length of the response window changes adaptively from trial to trial using a 4-up-2-down procedure, ranging between 1276ms (best) to 4096ms (worst). The response window begins when the stimulus sound begins. The staircase procedure terminates after 84 trials. The resulting threshold is **the logarithm** with **base 2** of the length of the response window in milliseconds at which the subject correctly performs 67.8% of trials, allowing for a measure of switch cost. Therefore, a more positive threshold is indicative of a greater set shifting deficit in the auditory modality.
https://research.brainhq.com/#assessment/1F

The **Task Switcher:Double** assessment is a task switching task that measures the set shifting process in the visual modality of the domain of executive functioning. In each trial of the assessment, two shapes are shown on the screen. Users must decide, as quickly as possible, whether the two shapes match in shape, or whether they match in color, depending on the instruction written on the screen. If they match, the user should hit the left arrow, otherwise hit the right arrow. The decision rule changes pseudo-randomly from trial to trial, with equal distribution of the four trial types: same-rule/same-response, same-rule/switch-response, switch-rule/switch-response. The length of the response window changes adaptively from trial to trial using a 4-up-2-down procedure, ranging between 1276ms (best) to 4096ms (worst). The response window begins when the visual stimuli are presented. The staircase procedure terminates after 85 trials. The resulting threshold is **the logarithm** with **base 2** of the length of the response window in milliseconds at which the subject correctly performs 67.8% of trials, allowing for a measure of switch cost. Therefore, a more positive threshold is indicative of a greater set shifting deficit in the visual modality.
https://research.brainhq.com/#assessment/22

Domain of Learning

The **Auditory Associates** assessment is a paired associate learning task that measures response learning and bond formation in the auditory modality. In each trial of the task, a ring of items is presented on the screen and users hear in sequence syllables associated with each item. Each syllable sound is presented for a fixed duration of 700 ms. Users must memorize the location of each heard syllable. During the response phase, users must click in order on the items on the screen that match the train of heard syllables. The number of syllables to memorize on each trial ranges from 2-8. The adaptive algorithm requires the user to respond accurately to three consecutive trials with the same number of items, before the number of items increments by one. In case of an incorrect trial, the number of items automatically decrements by one. The procedure terminates after 27 trials. The resulting threshold is the mean number of items recalled. Therefore, a smaller threshold is indicative of a greater deficit in learning in the auditory modality.
https://research.brainhq.com/#assessment/2E

The **Visual Associates** assessment is a paired associate learning task that measures response learning and bond formation in the visual modality. In each trial of the task, a ring of items is presented on the screen and users are asked to memorize the items and their locations. The number of items to memorize on each trail ranges from 2-8 and the duration of each stimulus ranges from 1-3 seconds. During the response phase, users are shown the same items, one at a time, and asked to click in order on the location of the screen where the item was originally presented. The initial stimulus duration is 3 seconds. For each correct trial, the stimulus duration decrements by one second until it has reached 1, at which point the number of items increments by one and the stimulus duration resets to 3 seconds. For incorrect trials, the number of items decrements by one and the stimulus duration is reset to 3. The adaptive procedure terminates after 27 trials. The resulting threshold is the mean number of items recalled. Therefore, a smaller threshold is indicative of a greater deficit in learning in the visual modality.
https://research.brainhq.com/#assessment/2F

Domain of Socio-Affective Processing

The **Voice Choice** assessment is a prosody detection task that measures emotion recognition in the auditory modality. In each trial of the task, a sentence of neutral content (e.g. “today is Tuesday”) is spoken by male and female speakers to convey 1 of 5 different emotions (neutral, happy, sad, angry or afraid). The gender of the speaker (male vs female) changes pseudo-randomly from trial to trial, with a 50% change of occurring at the start of each trial. During the response phase (which appears 500 ms after stimulus offset) three emotion labels appear, and users have to select the label that matches the prosody of the sentence played during the stimulus phase. The length of the sentence changes adaptively between 3 potential lengths: length 1 for short sentences (1-2 words), length 2 for medium sentences (3-4 words), and length 3 for long sentences (5-7 words). The threshold is determined using a single staircase (2-up-1-down) procedure, which terminates after 20 trials. Under the assumption that it is easier to identify prosody from longer sentences, a more positive threshold is indicative of a greater deficit in emotion recognition in the auditory modality.
https://research.brainhq.com/#assessment/30

The **Emotion Motion assessment** is a facial emotion recognition task, and it measures emotion recognition in the visual modality. In each trial of the task, a short movie clip is presented, showing an actor or actress displaying one of 7 potential emotions (neutral, angry, afraid, surprised, disgusted, sad and happy).
During the response phase (which appears 500 ms after stimulus offset), three emotion labels appear, and users have to select the label that matches the emotion presented in the clip. The speed of stimulus presentation changes adaptively from 1 (slowest speed, longest duration of the clip) to 10 (fastest speed, shortest duration of the clip), with a starting speed of 3, using a 2-up-1-down staircase procedure that terminates after 15 trials.
Under the assumption that it is easier to recognize emotions from longer clips, a lower threshold is indicative of a greater deficit in emotion recognition in the visual modality.
https://research.brainhq.com/#assessment/24

3. Correlation Matrix between ONAs T subscores and MCCB T subscores in the SZ sample

Spearman's rho		MCCB Speed of Processing	MCCB Attention	MCCB Executive Functioning	MCCB Learning	MCCB Social Cognition
BHQ Speed Of Processing	Correlation Coefficient	.536**	.309*	.267*	.320*	.428**
	Sig. (2-tailed)	0	0.019	0.039	0.012	0
BHQ Attention	Correlation Coefficient	.258*	.574**	0.239	.302**	.309*
	Sig. (2-tailed)	0.048	0	0.074	0.01	0.019
BHQ Executive Functioning	Correlation Coefficient	.294*	0.134	.396**	.312*	.325*
	Sig. (2-tailed)	0.022	0.322	0.002	0.017	0.011
BHQ Learning	Correlation Coefficient	.309*	.294*	.258*	.418**	0.202
	Sig. (2-tailed)	0.019	0.023	0.048	0.001	0.136
BHQ Socio-Affective Processing	Correlation Coefficient	.294*	.353**	0.222	.294*	0.134
	Sig. (2-tailed)	0.023	0.007	0.088	0.023	0.322