Supporting Information

Garrett et al. 10.1073/pnas.1504090112

SI Methods

Participants and Procedure. Our initial sample consisted of 75 adults, 12 of whom were removed from the final sample due to various data artifacts in the functional images [e.g., ghosting, excessive motion (>4-mm displacement)] or due to incomplete behavioral data.

Blood pressure and pulse were taken immediately before and 30 min after pill ingestion, and immediately before and after the scanning session. Before AMPH ingestion, anyone with blood pressure exceeding 140/90 was not allowed to continue with the study. Following AMPH ingestion, participants whose blood pressure was less than 80/60, exceeded 180/110, or whose pulse exceeded 100 beats per min, would have been precluded from entering the scanner; none in the current sample met any of these exclusionary criteria.

Data Preparation. Due to a technical problem with our response box during collection of behavioral data in the scanner, trials were missing at random for some subjects. In such cases, because we could not distinguish between missingness due to this technical problem and due to valid missing responses, we chose to only examine blocks of trials with complete behavioral and corresponding brain data. Subjects in the final sample had to have a minimum of 15 complete behavioral trials within-condition (fixation, 1-, 2-, and 3-back) and within-session (when they received either placebo or AMPH).

Cognition. *N*-back performance in the scanner was initially assessed by accuracy (percentage of correct responses, including both hits and correct rejections), and by reaction time means (RT_{mean}) and SDs (RT_{SD}). However, an initial examination of behavioral distributions by age group, drug condition, and task load revealed normal distributions for RT_{mean} and RT_{SD}, but extremely left-skewed distributions for 1- and 2-back Accuracy in younger adults, and 1-back Accuracy in older adults (one-sample Kolmogorov–Smirnov "tests for normality," $P < 0.05$ for all), representing ceiling effects (average young adult, 1-back = 99%, 2-back = 98%; older adult, 1-back = 97%). For this reason, and to maintain an equivalent amount of data for each behavioral measure in our models, all analyses throughout the current study focused only on RT_{mean} and RT_{SD}. Notably, targeted subsequent examinations of possible links between 3-back–only Accuracy and SD_{BOLD} revealed no reliable effects, thus further supporting the focus on RT_{mean} and RT_{SD} in the current study.

Extended fMRI Preprocessing. Beyond standard preprocessing steps (*Methods*), we subsequently examined all functional volumes for artifacts via independent component analysis (ICA) within-run, within-person, as implemented in FSL/MELODIC (1). Noise components were targeted according to several key criteria: (*i*) Spiking (components dominated by abrupt time series spikes ≥6 SDs); (*ii*) motion [prominent edge or "ringing" effects, sometimes (but not always) accompanied by large time series spikes]; (*iii*) susceptibility and flow artifacts (prominent air–tissue boundary or sinus activation; typically represents cardio/respiratory effects); (*iv*) white matter (WM) and ventricle activation (2); (*v*) low-frequency signal drift (3); (*vi*) high power in high-frequency ranges unlikely to represent neural activity (≥75% of total spectral power present above 0.13 Hz;); and (*vii*) spatial distribution ["spotty" or "speckled" spatial pattern that appears scattered randomly across ≥25% of the brain, with few if any clusters with ≥10 contiguous voxels (at 4-mm^3 voxel size)]. Examples of these various components we typically deem to be noise can be found in the supplementary materials of the study by Garrett et al. (4). By default, we use a conservative set of rejection criteria; if manual classification decisions are difficult due to the co-occurrence of apparent "signal" and "noise" in a single component, we typically elect to keep such components. Two independent raters of noise components were used (D.D.G. and J.M.); >90% interrater reliability was required on separate data before denoising decisions were made on the current data. Components identified as artifacts were then regressed from corresponding fMRI runs using the FSL regfilt command. The use of ICA denoising had dramatic effects in our past research, effectively removing 50% of the variance still present after traditional preprocessing steps, while simultaneously doubling the predictive power of BOLD signal variability (5). Thus, calculating BOLD signal variance from relatively artifact-free BOLD time series permits the examination of what is more likely meaningful neural variability.

As a final step to control for the possibility that AMPH may impact BOLD signal variability only by boosting the general physiological responsiveness of the participants, we applied the latest version of single-subject level PHYsiological correction using Canonical Autocorrelation Analysis (PHYCAA+; ref. 6). PHYCAA+ is an automated algorithm that (*i*) down-weights voxel variance in probable nonneuronal tissue, and (*ii*) identifies the multivariate physiological noise subspace in gray matter that is linked to nonneuronal tissue. PHYCAA+ thus estimates physiological noise directly from BOLD time series (precluding the need for external measures of heartbeat and respiration) and has recently been shown to outperform various alternative physiological denoising techniques, such as RETROICOR (7), RVHR (8), and COMPCOR (9), at improving both the prediction and reproducibility of resulting activation maps.

Partial Least-Squares Modeling Details: Analysis of Relations Between Brain Signal Variability (SD_{BOLD}) and AMPH (Pre/Post), Age Group (Younger/Older Adults), and Task Condition (Fixation/1/2/3-Back). To examine multivariate relations between SD_{BOLD}, AMPH, age group, and task condition during *n*-back, we used a Task partial least-squares (PLS) analysis (10, 11). Task PLS begins by calculating a between-subject covariance matrix (COV) between experimental conditions/groups and each voxel's SD_{BOLD}. COV is then decomposed using singular value decomposition (SVD):

$$SVD_{COV} = USV'. \quad \textbf{[S1]}$$

This decomposition produces a left singular vector of experimental condition/group weights (U), a right singular vector of brain voxel weights (V), and a diagonal matrix of singular values (S). This analysis produces orthogonal latent variables (LVs) that optimally represent relations between experimental conditions/groups and voxelwise SD_{BOLD} values. Each LV contains a spatial activity pattern depicting the brain regions that show the strongest relation to condition/group contrasts identified by the LV. Each voxel weight (in V) is proportional to the covariance between voxel SD_{BOLD} and the condition/group contrast. In the current study, we had a 2 (AMPH) × 2 (Age Group) × 4 (Task Condition) design, yielding a total of 16 possible latent dimensions (i.e., singular values in S) that could be estimated.

To obtain a summary measure of each participant's expression of a particular LV's spatial pattern, we calculated within-person "brain scores" by multiplying each voxel (*i*)'s weight (V) from

each LV (j) (produced from the SVD in Eq. **S1**) by voxel (i)'s SD_{BOLD} value, for each condition/group (k) within person (l), and summing over all (n) brain voxels:

$$\sum_{i=1}^{n} V_{ij} SD_{BOLDikl}. \quad \textbf{[S2]}$$

This is exactly equivalent to the multiplication of V by a subject's vector of SD_{BOLD} values for all voxels. Significance of detected relations between multivariate spatial patterns and conditions/groups was assessed using 1,000 permutation tests of the singular value corresponding to each LV. A subsequent bootstrapping procedure revealed the robustness of voxel saliences across 1,000 bootstrapped resamples of the data (12). By dividing each voxel's mean salience by its bootstrapped SE, we obtained "bootstrap ratios" (BSRs) as normalized estimates of robustness. We thresholded BSRs at a conservative value of ±4.25, which exceeds a 99.99% confidence interval.

To restrict all multivariate analyses to gray matter (GM), we masked our functional data with the GM tissue prior provided in FSL, thresholded at probability >0.37. We localized thresholded regions from all PLS model output by submitting resulting Montreal Neurological Institute (MNI) coordinates to the Anatomy Toolbox (version 1.8) in SPM8, which applies probabilistic algorithms to determine the cytoarchitectonic labeling of MNI coordinates (13, 14).

Mixed Model Details.

Modeling unique relations between SD_{BOLD}, AMPH, Age Group, and Task. The PLS models we specified in the current paper resulted in multivariate, latent-level relations between BOLD, AMPH, Age Group, and Task. Due to their multivariate nature, these models do not explicitly specify the unique importance of each effect in the solution. Subsequent mixed models can help parse these various effects, while flexibly and properly accounting for degrees of freedom and model covariances at within- (i.e., AMPH, Task) and between-subject (i.e., Age Group) levels (15). The PLS model above produced eight brain scores per person [AMPH (placebo vs. drug) × Task (fixation, 1/2/3-back)], and these scores became the dependent variables of interest in the mixed models. We then fit a model of the following form:

$$\begin{aligned} SD_{BOLD}BrainScore_{ijk} = {} & \beta_0 + \beta(AMPH_{jk}) + \beta(Task_{ijk}) \\ & + \beta(Age\ group_k) + \beta(AMPH_{jk} \times Task_{ijk}) \\ & + \beta(AMPH_{jk} \times Age\ group_k) \\ & + \beta(Task_{ijk} \times Age\ group_k) \\ & + \beta(AMPH_{jk} \times Task_{ijk} \times Age\ group_k) \\ & + e_{0ijk}. \end{aligned} \quad \textbf{[S3]}$$

Here, the PLS brain score for each task condition (i), AMPH condition (j), and participant (k) is modeled as a function of a model intercept (β_0), main effects for AMPH, Task, and Age group, all interactions, and a residual error term (e_{0ijk}). We did not include random intercepts and slopes in the final model; when included, model convergence was typically not achieved, likely as a result of modest sample size. All mixed models were run using the Mixed Models module in SPSS 22 (IBM).

Modeling relations between SD_{BOLD} and cognitive performance. Another goal in the current study was to link AMPH-related changes in SD_{BOLD} to AMPH-related changes in behavior during n-back (i.e., "change–change" relations). However, average levels of SD_{BOLD} and behavior may also be uniquely linked (i.e., "level–level" relations). In the current study, key "within" effects represent AMPH-placebo differences, and "between" effects represent the average of AMPH and placebo. Initially, for behavior (each for RT_{mean} and RT_{SD}) and SD_{BOLD} (from our PLS brain scores noted above), a total of six scores were available per subject [AMPH (placebo, drug) × Task (1/2/3-back)]. These scores permitted the creation of three change scores (AMPH-placebo) and three average scores [(AMPH+placebo)/2], one for each n-back condition, per variable of interest (RT_{mean}, RT_{SD}, and SD_{BOLD}) and subject. We could then fit separate change–change and level–level models.

For the AMPH-related change–change model, we built upon a general model of the following form:

$$Performance\_within_{ij} = \beta_0 + \beta(SD_{BOLD}\_within_{ij}) + \beta_n \ldots e_{0ij}. \quad \textbf{[S4]}$$

Here, we model within-person AMPH-related cognitive performance for each n-back task condition (i) and participant (j) as a function of a model intercept (β_0), the AMPH-related SD_{BOLD} brain score ($SD_{BOLD}\_within_{ij}$), other variables of interest, and residual error, e_{0ij}.

Subsequently, our level–level models were run as follows:

$$Performance\_between_{ij} = \beta_0 + \beta(SD_{BOLD}\_between_{ij}) + \beta_n \ldots e_{0ij}. \quad \textbf{[S5]}$$

Between-person cognitive performance for each n-back task condition (i) and participant (j) was modeled as a function of a model intercept (β_0), the average SD_{BOLD} brain score across placebo and AMPH conditions ($SD_{BOLD}\_between_{ij}$), other variables of interest, and residual error, e_{0ij}.

Various specific models used in the current study are outlined in *Results*, and full model results are contained in Table S2. Random intercepts and slopes were again not modeled due to lack of model convergence. For all mixed models outlined above, we chose compound symmetry (CS) as the covariance structure given that Akaike information criteria fits were typically significantly better than for the default diagonal covariance structure, and required fewer parameters to estimate. We also compared CS to the most bias-free available covariance structure (i.e., "unstructured" covariance), and due to a greatly increased number of estimated parameters (e.g., in Table S2, model 1, estimated parameters would have ballooned from 18 to 53) and our modest sample size, we elected to keep CS as our covariance structure. All models were fit using full-information maximum-likelihood estimation.

Model outliers were determined by calculating Cook's distance, which reflects the extent to which model residuals would change if a particular data point (in multivariate space) were excluded from the model. Larger Cook's distance values indicate more influential data points. Single observations exceeding a threshold of 2.5 SDs from the distribution of all multivariate observations were deemed overly influential on model parameters. In our sample, out of 496 total observations (62 subjects × 4 task conditions × 2 drug conditions), only two observations were outliers according to Cook's distance. However, mixed models permit missingness without subject-level listwise deletion, maximizing remaining available explanatory power and parameter robustness from the remaining 494 observations.

$Mean_{BOLD}$ Effects. We sought also to compare SD_{BOLD} results to a typical mean-based measure of BOLD activity ($mean_{BOLD}$). To calculate $mean_{BOLD}$ for each experimental condition, we first expressed each signal value as a percent change from the average of the last four scans from the previous block, and then calculated a mean percent change within each block and averaged across all blocks for a given condition (a typical method in the PLS data analysis framework). This effectively acts as an explicit

high-pass filter over the data. We then reran relevant PLS and mixed models described above, while using $mean_{BOLD}$ measures.

Physiological Component Derivation. Due to high covariance between pulse and systolic and diastolic blood pressure (BP) measurements, we used principal-component analysis (PCA) to derive a physiological component structure. BP and pulse were collected immediately before and immediately after each scanning session, allowing us to derive within-person (i.e., AMPH-related changes) and between-person (i.e., average across session) levels of these measures (see above for logic). In turn, we ran separate PCAs to create a separate component structure at each level. We found a single component representing the within-person level (eigenvalue, 1.88; systolic BP loading, 0.91; diastolic BP loading, 0.86; pulse loading, 0.56), and a single component representing the between-person level (eigenvalue, 1.88; systolic BP loading, 0.85; diastolic BP loading, 0.90; pulse loading, 0.58). Resulting components were then entered into our mixed models to test whether physiological measures would account for our primary model effects (*Results* in main text and Table S2, model 2).

SI Results

Multivariate Model Linking SD_{BOLD} to Age Group and Task Within the Placebo Condition Only. The current BOLD signal variability literature often finds higher cortical signal variability in younger, higher performers (16). However, sometimes (17, 18), but not always (19), subcortical regions exhibit the inverse effect such that SD_{BOLD} is higher in older, poorer performers. Our PLS model of SD_{BOLD} in the current study (clearly dominated by the AMPH effect on SD_{BOLD} in older adults; Fig. 1 and Fig. S2) showed no evidence of inverse effects; all regions were red/yellow, showing higher SD_{BOLD} in younger compared with older adults at placebo, and a large increase in SD_{BOLD} in older adults on AMPH. To verify that the dominant AMPH effect in older adults was not skewing our results away from possible inverse cortical–subcortical age effects off drug, we reran our PLS model with only placebo data. This ensures the model is more similar in form to previous BOLD signal variability studies (especially to ref. 19, which focused on SD_{BOLD} modulations across age groups and task conditions) and ensures that AMPH cannot impact the derivation of any latent dimensions via PLS. These placebo-based results corresponded with the overall model shown in Fig. 1. We found a single robust LV (cross-block covariance, 49.66%; permuted $P = 0.007$) showing that younger adults were again more variable than older adults on every experimental condition (red/yellow regions in Fig. S1), and we found no regions in which older adults were more variable than younger adults. As in our current Fig. 1 and Fig. S2, several cortical and DA-typical subcortical regions (putamen, caudate) remained present. Table S1, model 3, includes peak locations, bootstrap ratio values, and cluster sizes.

Mixed Models of SD_{BOLD} Are Not Driven by Physiological Artifacts. We took several precautions at the level of image preprocessing to ensure that the present SD_{BOLD} results are unlikely to be accounted for by physiological noise (see *SI Methods* for details). First, our ICA denoising procedure explicitly targets and removes cardiac and respiratory components from the fMRI data. Second, we subsequently applied PHYCAA+ (6) as an iterative procedure to further classify and remove nonneural signal sources remaining after ICA denoising. Finally, we estimated within- and between-person physiological components (derived from systolic/diastolic blood pressure and pulse measures collected immediately before and immediately after each scanning session; see *SI Methods* for details) and added these components (and relevant interactions) to the terms in our initial mixed model (Table S2, model 1). Results revealed no reliable effects of physiological components (Table S2, model 2), and critically, our key AMPH and AMPH × Age Group effects noted above were preserved. Thus, in combination with our use of ICA denoising and PHYCAA+, we find no evidence that our primary SD_{BOLD} results (Table S2, model 1) are attributable to blood pressure or pulse-related artifacts at within- or between-person levels.

Behavior-Only Models. For RT_{mean}, we noted significant effects of AMPH, Age Group, AMPH × Session Order, Task × Age Group, Task × Session Order, and a key AMPH × Age Group × Session Order effect (see Table S2, model 7, for full model results, and Table S3 for descriptive statistics). The three-way interaction is plotted in Fig. S5, *Left*. Inverse placebo-AMPH slopes exists for different session orders in both younger and older adults. Although slope inversion is common to both age groups, the inversion is much more pronounced for YA [AMPH × Session Order interaction, $F_{(1,200)} = 223.08$, $P = 2.23 \times 10^{-34}$] than for OA [AMPH × Session Order interaction, $F_{(1,106.15)} = 9.62$, $P = 0.002$]. Although Task × Age Group and Task × Session Order effects were also present, Task did not interact with AMPH. In combination with a lack of Task effects on SD_{BOLD} (independent of, or when interacting with, AMPH; Table S2, model 1), this supports the focus on AMPH as the primary within-person dimension linking SD_{BOLD} and WM performance in the current study.

For RT_{SD}, we found robust effects of Task, Age Group, AMPH × Session Order, Task × Age Group, and AMPH × Age Group × Session Order. This three-way interaction is plotted in Fig. S5, *Right*. Here, a strong placebo-AMPH inversion effect existed for different session orders in YA [AMPH × Session Order interaction, $F_{(1,200)} = 63.18$, $P = 1.36 \times 10^{-13}$], but far more subtly for OA [AMPH × Session Order interaction, $F_{(1,110)} = 1.27$, $P = 0.26$]. Similarly to RT_{mean}, we found no interaction between Task and AMPH. See Table S2, model 8, for full model results, and Table S3 for descriptive results.

Mixed Models of SD_{BOLD} Are Not Affected by Session Order Effects. In our mixed model of SD_{BOLD} (AMPH × Task × Age Group; Table S2, model 1), we showed significant effects of AMPH and AMPH × Age Group. Subsequent models predicting reaction time (RT_{mean} and RT_{SD}) also showed significant effects of Session Order (Table S2, models 3 and 4). To assuage concerns that our original SD_{BOLD} model results may also reflect Session Order, we reran model 1 in Table S2 as an AMPH × Task × Age Group × Session Order mixed factorial model predicting SD_{BOLD}. We found no effects of Session Order (or any interaction including Session Order) on SD_{BOLD} (all values of $P > 0.53$; Table S2, model 10). Notably, our key AMPH and AMPH × Age Group effects in the original Table S2, model 1, were fully maintained, with nearly identical F statistics as before.

1. Beckmann CF, Smith SM (2004) Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging* 23(2):137–152.
2. Birn RM (2012) The role of physiological noise in resting-state functional connectivity. *Neuroimage* 62(2):864–870.
3. Smith AM, et al. (1999) Investigation of low frequency drift in fMRI signal. *Neuroimage* 9(5):526–533.
4. Garrett DD, McIntosh AR, Grady CL (2014) Brain signal variability is parametrically modifiable. *Cereb Cortex* 24(11):2931–2940.
5. Garrett DD, Kovacevic N, McIntosh AR, Grady CL (2010) Blood oxygen level-dependent signal variability is more than just noise. *J Neurosci* 30(14):4914–4921.
6. Churchill NW, Strother SC (2013) PHYCAA+: An optimized, adaptive procedure for measuring and controlling physiological noise in BOLD fMRI. *Neuroimage* 82:306–325.
7. Glover GH, Li T-Q, Ress D (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med* 44(1):162–167.
8. Chang C, Glover GH (2009) Effects of model-based physiological noise correction on default mode network anti-correlations and correlations. *Neuroimage* 47(4):1448–1459.
9. Behzadi Y, Restom K, Liau J, Liu TT (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37(1):90–101.
10. McIntosh AR, Bookstein FL, Haxby JV, Grady CL (1996) Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage* 3(3 Pt 1):143–157.
11. Krishnan A, Williams LJ, McIntosh AR, Abdi H (2011) Partial least squares (PLS) methods for neuroimaging: A tutorial and review. *Neuroimage* 56(2):455–475.
12. Efron B, Tibshirani R (1993) *An Introduction to the Bootstrap* (Chapman & Hall/CRC, Boca Raton, FL).

13. Eickhoff SB, et al. (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25(4):1325–1335.
14. Eickhoff SB, et al. (2007) Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage* 36(3):511–521.
15. Singer JD, Willett JB (2003) *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence* (Oxford Univ Press, New York).
16. Garrett DD, et al. (2013) Moment-to-moment brain signal variability: A next frontier in human brain mapping? *Neurosci Biobehav Rev* 37(4):610–624.
17. Garrett DD, Kovacevic N, McIntosh AR, Grady CL (2011) The importance of being variable. *J Neurosci* 31(12):4496–4503.
18. Samanez-Larkin GR, Kuhnen CM, Yoo DJ, Knutson B (2010) Variability in nucleus accumbens activity mediates age-related suboptimal financial risk taking. *J Neurosci* 30(4):1426–1434.
19. Garrett DD, Kovacevic N, McIntosh AR, Grady CL (2013) The modulation of BOLD variability between cognitive states varies by age and processing speed. *Cereb Cortex* 23(3):684–693.

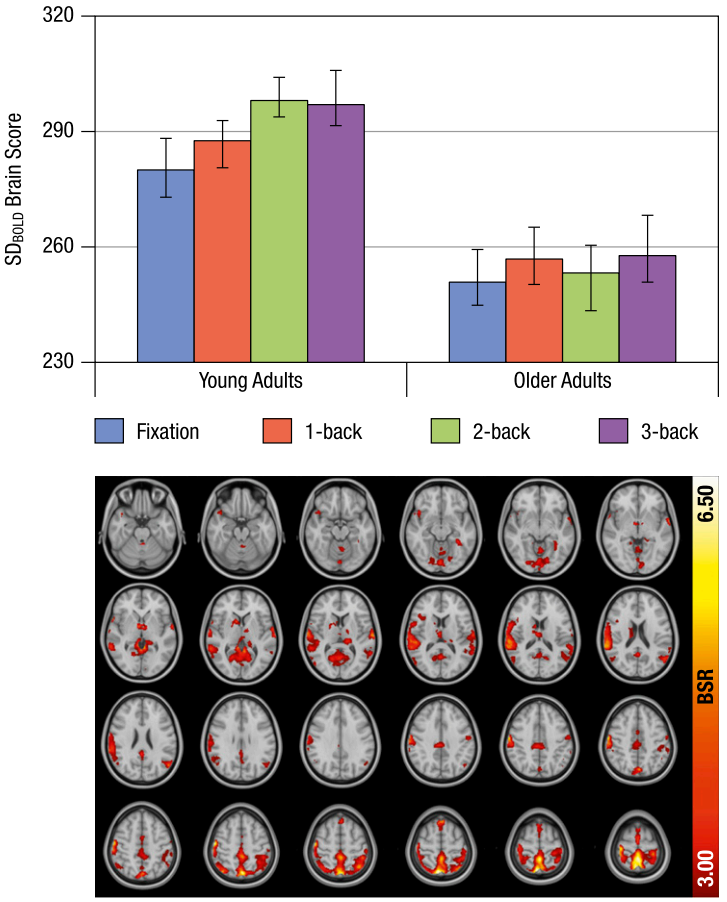

Fig. S1. Brain scores and spatial pattern from PLS model of relation between SD_{BOLD}, Age Group, and Task Condition at placebo only. Higher brain scores reflect higher SD_{BOLD} in red/yellow regions. Error bars represent bootstrapped 95% confidence intervals (1,000× with replacement). Every other 2-mm slice in *Z* direction is plotted, from $Z = -24$ to $Z = 68$. Images are plotted in neurological orientation (left is *Left*). BSR, bootstrap ratio. BSR level set to ±3.00 due to reduced data (i.e., only placebo data) compared with the model noted in Fig. 1.

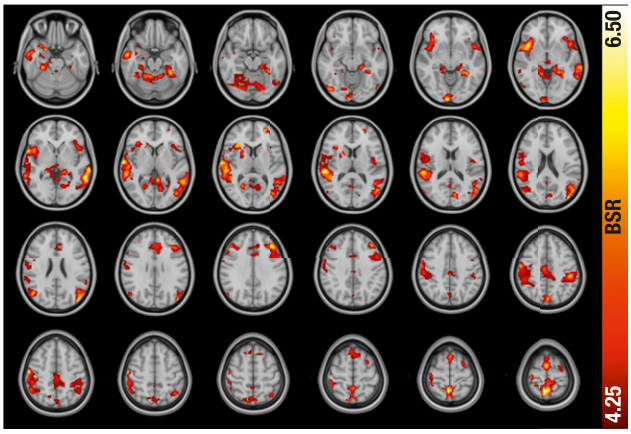

Fig. S2. Full axial view of SD_{BOLD} PLS spatial pattern (first latent variable). Every other 2-mm slice in *Z* direction is plotted, from $Z = -24$ to $Z = 68$. Images are plotted in neurological orientation (left is *Left*). BSR, bootstrap ratio.

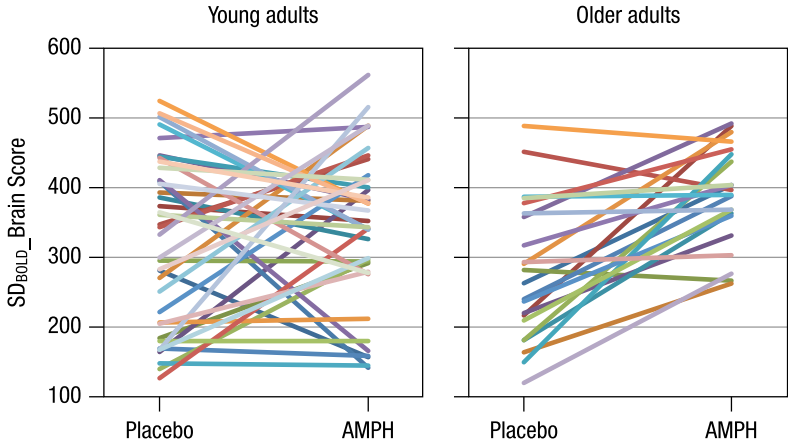

Fig. S3. Individual slopes representing placebo-AMPH shifts in SD_{BOLD} for young and older adults from an example *n*-back condition (2-back).

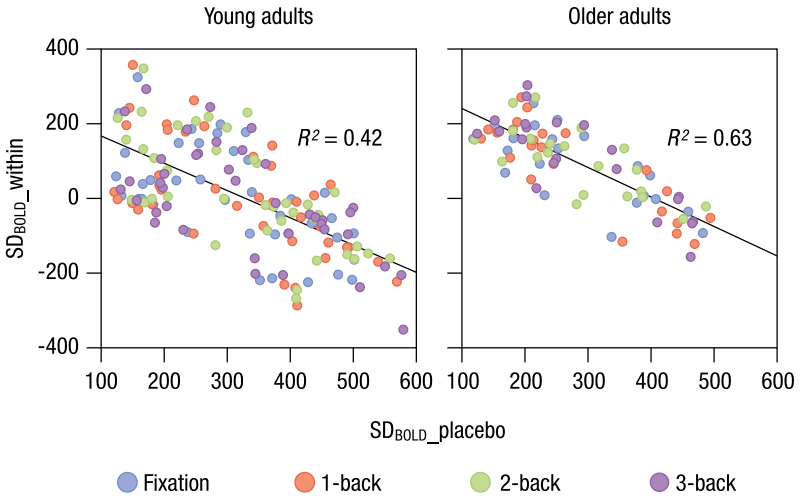

Fig. S4. Baseline (placebo) SD_{BOLD} negatively correlates with AMPH-related change in SD_{BOLD}. SD_{BOLD}_within (AMPH SD_{BOLD} − placebo SD_{BOLD}). No reliable task condition or session order differences in slope existed for either age group.

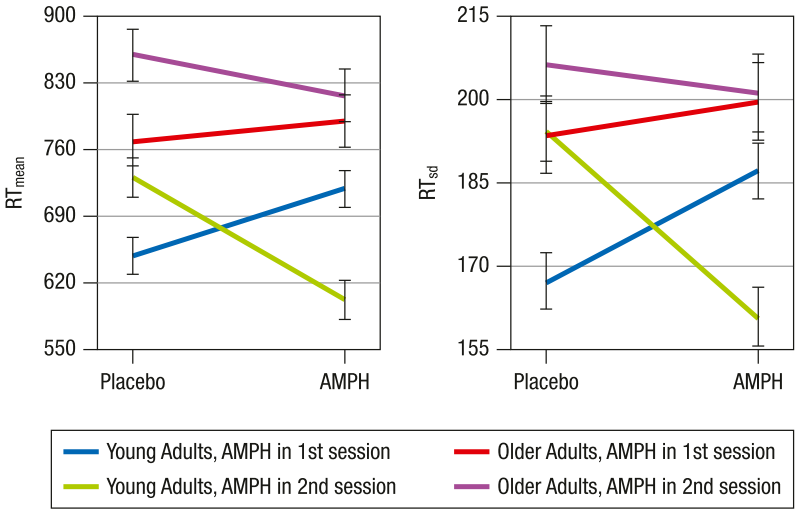

Fig. S5. Average *n*-back RT_{mean} (*Left*) and RT_{SD} (*Right*), by age group and drug session order. Plotted values represent estimated marginal means from the AMPH × Age Group × Session Order interaction noted in Table S2 (models 7 and 8 for RT_{mean} and RT_{SD}, respectively). Error bars represent ±1 SE.

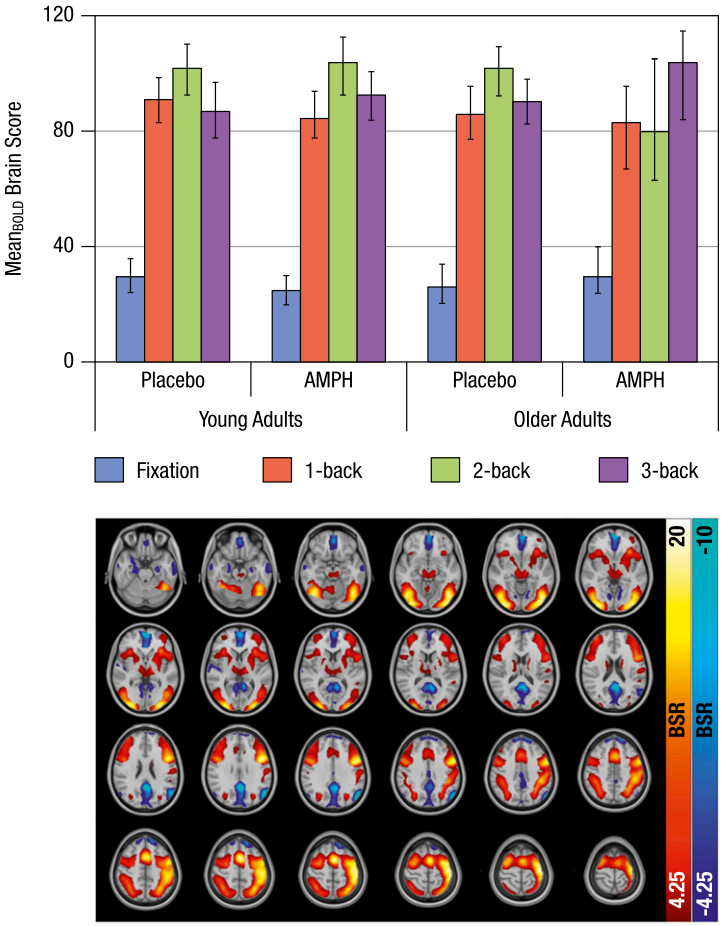

Fig. S6. Brain scores and spatial pattern from PLS model of relation between Mean$_{BOLD}$, Age Group, AMPH, and Task Condition. Higher brain scores reflect higher mean$_{BOLD}$ in red/yellow regions and lower mean$_{BOLD}$ in blue regions. Error bars represent bootstrapped 95% confidence intervals (1,000× with replacement). Every other 2-mm slice in *Z* direction is plotted, from $Z = -24$ to $Z = 68$. Images are plotted in neurological orientation (left is *Left*). AMPH, amphetamine; BSR, bootstrap ratio.

Table S1. PLS model peak activations, bootstrap ratios, and cluster sizes

Model	Region	Hem	MNI coordinates			BSR	Cluster size, voxels
			X	*Y*	*Z*		
1	SMA	L	−2	−8	74	7.18	4,783
SD_{BOLD}	Middle temporal gyrus	R	68	−32	−2	6.15	2,235
(Placebo and AMPH)	Superior temporal gyrus	L	−56	−30	12	6.11	3,935
	Temporal pole	L	−24	6	−28	6.06	165
	Middle frontal gyrus	R	40	34	34	5.84	751
	Calcarine gyrus	L	−8	−98	−8	5.83	273
	Inferior occipital gyrus	L	−42	−76	−14	5.8	1,182
	Superior frontal gyrus	R	24	54	8	5.8	90
	Parahippocampal gyrus	R	32	−44	−6	5.77	2,014
	Middle temporal gyrus	L	−56	−4	−22	5.65	301
	Postcentral gyrus	R	48	−28	44	5.48	678
	Fusiform gyrus	L	−28	−32	−24	5.45	169
	Angular gyrus	L	−48	−66	24	5.43	312
	Putamen	L	−22	10	10	5.32	133
	Superior frontal gyrus	R	28	6	62	5.27	121
	Calcarine gyrus	R	8	−76	16	5.21	110
	Insula lobe	R	46	16	−4	5.17	515
	Calcarine gyrus	L	−22	−60	6	5.15	108
	Middle cingulate cortex	L	−6	−14	44	5.04	494
	Middle frontal gyrus	L	−44	32	34	5.02	194
	Anterior cingulate cortex	R	4	30	28	4.92	425
	Insula lobe	R	40	−8	12	4.72	95
2	Angular gyrus	L	−54	−66	30	10.9	711
$Mean_{BOLD}$	Posterior cingulate cortex	L	−2	−54	24	10.41	4,572
(Placebo and AMPH)	Superior medial gyrus	L	0	60	0	10.39	2,480
	Superior frontal gyrus	L	−14	52	42	7.43	993
	Parahippocampal gyrus	L	−24	−14	−26	6.45	141
	Middle temporal gyrus	L	−64	−12	−20	6.08	396
	Middle occipital gyrus	R	52	−70	28	5.77	93
	Parahippocampal gyrus	R	22	−16	−26	5.48	328
	Middle temporal gyrus	R	62	−10	−16	5.11	110
	Superior temporal gyrus	R	66	−18	6	4.87	166
	Precentral gyrus	L	−42	−18	62	−18.69	25,253
	Inferior occipital gyrus	L	−38	−84	−8	−17.32	4,274
	Inferior occipital gyrus	R	34	−90	−10	−17.17	4,293
	Precentral gyrus	R	46	6	30	−11.74	3,498
	Superior parietal lobule	R	36	−56	54	−10.46	3,772
	Insula lobe	R	34	22	0	−9.33	2,428
3	Precuneus	L	−8	−48	76	−7.29	9,426
SD_{BOLD}	Cerebellar vermis	R	6	−48	2	−5.21	2,135
(Placebo only)	Thalamus	L	−6	−6	−2	−4.97	175
	Superior temporal gyrus	L	−62	−38	20	−4.9	1,678
	Superior temporal gyrus	R	64	−16	8	−4.78	671
	Middle temporal gyrus	R	54	−66	22	−4.63	512
	Thalamus	L	−4	−14	14	−4.33	373
	Calcarine gyrus	L	−2	−84	−12	−4.11	615
	Temporal pole	L	−42	16	−22	−4.04	149
	Posterior cingulate	L	−2	−46	24	−4	137
	Fusiform gyrus	R	34	−42	−12	−3.96	136
	Postcentral gyrus	R	58	−10	40	−3.62	114
	Supplementary motor area	L	0	0	48	−3.6	83
	Inferior frontal gyrus	L	−42	12	18	−3.52	90

Note: BOLD, blood oxygen level dependent; BSR, bootstrap ratio (model salience/bootstrapped SE); Hem, hemisphere; MNI, Montreal Neurological Institute.

Table S2. Mixed-model results

Model	Dependent variable	Predictor	*df1*	*df2*	*F*	*P*
1	SD_{BOLD}	AMPH	1	429.97	52.80	**1.75×10^{-12}**
		Task	3	430.46	1.08	0.36
		Age Group	1	62.06	0.08	0.79
		AMPH × Task	3	429.97	0.16	0.92
		AMPH × Age Group	1	429.97	42.15	**2.33×10^{-10}**
		Task × Age Group	3	430.46	0.29	0.83
		AMPH × Task × Age Group	3	429.97	0.19	0.91
2	SD_{BOLD}	AMPH	1	429.99	30.80	**5.00×10^{-08}**
		Task	3	430.47	1.09	0.35
		Age Group	1	61.94	3.14	0.08
		AMPH × Task	3	429.99	0.14	0.94
		AMPH × Age Group	1	429.99	12.31	**4.98×10^{-04}**
		Task × Age Group	3	430.47	0.30	0.83
		AMPH × Task × Age Group	3	429.99	0.19	0.91
		AMPH × Physio_within	2	107.31	2.44	0.09
		AMPH × Age Group × Physio_within	2	107.31	1.85	0.16
		Physio_between	1	62.24	2.84	0.10
		AMPH × Physio_between	1	429.99	0.01	0.95
		Age Group × Physio_between	1	62.24	0.41	0.53
		AMPH × Age Group × Physio_between	1	429.99	1.26	0.26
3	RT_{mean}_within	Age Group	1	70.03	2.24	0.14
		Session Order	1	70.03	109.60	**5.71×10^{-16}**
		SD_{BOLD}_within	1	95.57	0.39	0.54
		Age Group × Session Order	1	70.03	7.89	**0.01**
		Age Group × SD_{BOLD}_within	1	95.57	1.80	0.18
		Session Order × SD_{BOLD}_within	1	95.57	11.22	**0.001**
		Age Group × Session Order × SD_{BOLD}_within	1	95.57	4.72	**0.03**
4	RT_{SD}_within	Age Group	1	67.36	5.01	**0.03**
		Session Order	1	67.36	40.61	**1.94×10^{-08}**
		SD_{BOLD}_within	1	81.97	2.08	0.15
		Age Group × Session Order	1	67.36	1.38	0.25
		Age Group × SD_{BOLD}_within	1	81.97	4.57	**0.04**
		Session Order × SD_{BOLD}_within	1	81.97	5.96	**0.02**
		Age Group × Session Order × SD_{BOLD}_within	1	81.97	7.31	**0.01**
5	RT_{mean}_between	Age Group	1	93.67	9.52	**0.003**
		Session Order	1	93.67	0.07	0.79
		SD_{BOLD}_between	1	93.67	1.76	0.19
		Age Group × Session Order	1	93.67	2.74	0.10
		Age Group × SD_{BOLD}_between	1	93.67	2.87	0.09
		Session Order × SD_{BOLD}_between	1	93.67	0.00	1.00
		Age Group × Session Order × SD_{BOLD}_between	1	93.67	1.70	0.20
6	RT_{SD}_between	Age Group	1	78.04	4.99	**0.03**
		Session Order	1	78.04	0.18	0.67
		SD_{BOLD}_between	1	76.16	0.61	0.44
		Age Group × Session Order	1	78.04	0.01	0.92
		Age Group × SD_{BOLD}_between	1	76.16	2.15	0.15
		Session Order × SD_{BOLD}_between	1	76.16	0.34	0.57
		Age Group × Session Order × SD_{BOLD}_between	1	76.16	0.03	0.85
7	RT_{mean}	AMPH	1	306.11	11.74	**0.001**
		Task	2	306.67	223.20	**1.90×10^{-60}**
		Age Group	1	62.23	33.67	**2.37×10^{-7}**
		Session Order	1	62.23	0.81	0.37
		AMPH × Task	2	306.11	0.49	0.61
		AMPH × Age Group	1	306.11	2.50	0.12
		AMPH × Session Order	1	306.11	125.19	**1.36×10^{-24}**
		Task × Age Group	2	306.67	13.03	**4.00×10^{-6}**
		Task × Session Order	2	306.67	4.65	**0.01**
		Age Group × Session Order	1	62.23	2.66	0.11
		AMPH × Task × Age Group	2	306.11	0.68	0.51
		AMPH × Task × Session Order	2	306.11	0.21	0.81
		AMPH × Age Group × Session Order	1	306.11	34.96	**9.00×10^{-9}**
		Task × Age Group × Session Order	2	306.67	1.26	0.29
		AMPH × Task × Age Group × Session Order	2	306.11	1.26	0.28

Table S2. Cont.

Model	Dependent variable	Predictor	*df1*	*df2*	*F*	*P*
8	RT_{SD}	AMPH	1	310.00	1.19	0.28
		Task	2	310.00	334.67	**3.69×10^{-78}**
		Age Group	1	62.00	17.69	**8.50×10^{-5}**
		Session Order	1	62.00	0.47	0.50
		AMPH × Task	2	310.00	0.01	0.99
		AMPH × Age Group	1	310.00	1.56	0.21
		AMPH × Session Order	1	310.00	30.61	**6.73×10^{-8}**
		Task × Age Group	2	310.00	20.40	**4.74×10^{-9}**
		Task × Session Order	2	310.00	0.32	0.73
		Age Group × Session Order	1	62.00	0.39	0.54
		AMPH × Task × Age Group	2	310.00	0.69	0.50
		AMPH × Task × Session Order	2	310.00	1.37	0.26
		AMPH × Age Group × Session Order	1	310.00	13.05	**3.54×10^{-4}**
		Task × Age Group × Session Order	2	310.00	0.83	0.44
		AMPH × Task × Age Group × Session Order	2	310.00	0.58	0.56
9	$Mean_{BOLD}$	AMPH	1	434.00	0.24	0.63
		Task	3	434.00	104.69	**5.32×10^{-51}**
		Age Group	1	62.00	0.41	0.53
		AMPH × Task	3	434.00	1.08	0.36
		AMPH × Age Group	1	434.00	0.04	0.84
		Task × Age Group	3	434.00	0.93	0.43
		AMPH × Task × Age Group	3	434.00	0.72	0.54
10	SD_{BOLD}	AMPH	1	429.96	53.46	**1.30×10^{-12}**
		Task	3	430.46	1.10	0.35
		Age Group	1	62.07	0.09	0.77
		Session Order	1	62.07	0.02	0.90
		AMPH × Task	3	429.96	0.14	0.94
		AMPH × Age Group	1	429.96	41.90	**2.63×10^{-10}**
		AMPH × Session Order	1	429.96	0.39	0.53
		Task × Age Group	3	430.46	0.30	0.83
		Task × Session Order	3	430.46	0.30	0.83
		Age Group × Session Order	1	62.07	0.32	0.57
		AMPH × Task × Age Group	3	429.96	0.20	0.90
		AMPH × Task × Session Order	3	429.96	0.29	0.83
		AMPH × Age Group × Session Order	1	429.96	0.35	0.55
		Task × Age Group × Session Order	3	430.46	0.03	0.99
		AMPH × Task × Age Group × Session Order	3	429.96	0.38	0.77

Note: AMPH, amphetamine; RT_{mean}_within, (AMPH RT_{mean} – placebo RT_{mean}); RT_{SD}_within, (AMPH RT_{SD} – placebo RT_{SD}); SD_{BOLD}_within, (AMPH SD_{BOLD} – placebo SD_{BOLD}). Significant effects are highlighted in bold.

Table S3. Descriptive behavioral statistics

		Young adults				Older adults			
		AMPH in session 1		AMPH in session 2		AMPH in session 1		AMPH in session 2	
Measure	Task condition	Placebo Mean (SD)	AMPH Mean (SD)	Placebo Mean (SD)	AMPH Mean (SD)	Placebo Mean (SD)	AMPH Mean (SD)	Placebo Mean (SD)	AMPH Mean (SD)
RT_{mean}	1-back	561.20 (102.25)	613.54 (116.36)	636.90 (107.67)	518.65 (72.93)	665.26 (109.32)	709.69 (133.05)	795.29 (89.55)	757.92 (66.67)
	2-back	649.91 (117.89)	729.34 (119.91)	763.17 (109.11)	613.26 (82.16)	820.12 (111.29)	825.26 (93.77)	898.44 (113.03)	852.00 (99.37)
	3-back	734.20 (102.32)	813.15 (113.83)	792.56 (109.27)	675.49 (95.30)	823.84 (88.26)	834.06 (98.53)	885.18 (76.52)	844.90 (80.75)
RT_{SD}	1-back	125.82 (24.59)	114.33 (24.70)	139.13 (34.88)	105.11 (26.11)	160.41 (39.21)	149.34 (39.35)	163.15 (37.89)	163.34 (38.66)
	2-back	191.09 (40.12)	158.47 (32.11)	202.70 (33.68)	160.09 (31.68)	212.28 (15.51)	207.03 (26.04)	219.20 (31.61)	208.89 (21.62)
	3-back	244.64 (39.80)	229.39 (30.94)	241.14 (31.44)	217.68 (38.16)	226.53 (39.78)	224.73 (42.20)	236.73 (28.78)	231.49 (26.70)

Note: AMPH, amphetamine.