Supplemental Material

Detailed fNIRS Processing and Analysis Workflows

Pre-GLM Processing. fNIRS data were processed using the open-source, HOMER2 software (Huppert et al., 2009). Processing workflows were similar to previous reports (Jahani et al., 2017), but are also briefly detailed here. Raw NIRS signals were converted to optical density and an initial filter was applied to remove channels with SNRs < 1.75. Motion artifacts were corrected on a channel-by-channel basis using the spline-interpolation method (Scholkmann et al., 2010; p = .99, frame size = 10 s; see Brigadoi et al., 2013). A low-pass filter with a cut-off frequency of .5 Hz was then applied to remove, high-frequency noise. Optical density signals were then converted to oxy- and deoxyhemoglobin concentrations using the modified Beer-Lambert law with a partial path length factor of 6 (Boas et al., 2004).

Aggregate (Block) GLM and channel filtering. Unreliable and noisy channels were filtered (Powell et al., 2018). Channels were removed that did not show positive and reliable (r>.30 correlation between hemodynamic response across runs) task-versus-rest activation across the two task runs. These filtering decisions were based upon aggregate, block-averaged, task-versus-rest hemoglobin responses. Thus, filtering decisions did not take into account list size or WM phase conditions and did not bias hemoglobin estimates for any of these conditions described in the main text (i.e., list size or P-R differences). Estimates of hemoglobin changes were extracted from the average of all channels surviving filtering described in the main text. Task-versus-rest effects were modeled using least-squares GLMs which yield the weights of consecutive Gaussian basis functions (t_{step} = .02 s; t_{range} = 0-25 s). Third-order drifts and short-separation channel signals were also added as nuisance factors in these GLMs (same as Main Text processing). The sum of oxyhemoglobin concentrations during task-versus-rest periods (i.e., area-under-the-curve) estimated aggregate (positive or negative) brain activation.

Proof-of-Concept (Block) GLM. The same block-modeling approach was applied for the proof-of-principle analyses, as those used for channel filtering (results presented in Main Text Figure 2D). However, here, stimulus timings were added for the three different WM list size conditions (i.e., 2, K, K+1). These WM list size models were used to test whether a positive relationship existed between dorsolateral PFC oxyhemoglobin concentrations and WM demand when a priori calibrated list sizes increased (Braver et al., 1997; Callicott et al., 1999; Finn et al., 2017; Rypma et al., 1999; Satterthewaite et al., 2013; Schneider-Garces et al., 2009; Van

Snellenberg et al., 2015; Main Text Figure 2D). This test served as a proof-of-principle to demonstrate that brain activation (e.g., aggregate increases in oxyhemoglobin concentrations) could be significantly modulated by subject-specific increases in memory demand. We hypothesized increases in aggregate (area-under-the-curve) dorsolateral PFC oxyhemoglobin concentrations with increases from subcapacity to capacity, and from capacity to supracapacity list sizes. These hypotheses were supported (Main Text Figure 2D).

Additional Deoxyhemoglobin Analyses

The main text included correlations between P-R differences and *K* using oxyhemoglobin concentrations. Oxyhemoglobin estimates are the most common fNIRS dependent variable, due to their higher signal-to-noise ratios (SNRs) relative to deoxyhemoglobin estimates, which may also be recovered using fNIRS. For instance, median SNR during WM task activation was lower for deoxyhemoglobin estimates (*Median* = 2.07 [*MAD* = .75]), relative to oxyhemoglobin estimates (*Median* = 2.21 [*MAD* = .99]). To provide complete results, we detail analyses for deoxyhemoglobin estimates.

Main Text Correlations with *K* Using Deoxyhemoglobin. No significant relationships were observed between *K* and P-R differences in deoxyhemoglobin concentrations—*K*+1 letters condition: Spearman's ρ=-.47, p = .089; *K* letters condition: ρ=-.13, p = .651; 2 letters condition: ρ=.13, p = .646.

Supplemental Table 1. MNI Channel Coordinates

Channel	Type	x	y	z	MNI Label
A	Standard	32	45	18	Right Middle Frontal
B	Standard	27	56	34	Right Middle Frontal
C	Standard	24	33	35	Right Superior Frontal
D	Short	24	36	29	Right Superior Frontal
E	Standard	38	31	32	Right Inferior Frontal
F	Standard	22	46	23	Right Superior Frontal
G	Standard	16	46	27	Right Anterior Cingulum
H	Standard	-23	58	36	Left Superior Frontal
I	Standard	-30	43	14	Left Inferior Frontal
J	Standard	-33	30	22	Left Inferior Frontal
K	Standard	-27	41	39	Left Middle Frontal
L	Short	-27	46	29	Left Middle Frontal
M	Standard	-16	41	29	Left Superior Frontal
N	Standard	-37	37	20	Left Middle Frontal

Channel Montreal Neurological Institute (MNI) coordinates and anatomical labels were extracted from Atlas Viewer (Aastad et al., 2015) using a 3D digitizer described by Tsuzuki and colleagues (2007). Type references standard (30 mm) or short (7.5 mm) source-detector distances.

Supplemental References

Aasted, C.M. et al. (2015). Anatomical guidance for functional near-infrared spectroscopy: AtlasViewer tutorial. *Neurophotonics, 2*(2), 020801.

Boas, D.A. et al. (2004). Diffuse optical imaging of brain activation: Approaches to optimizing image sensitivity, resolution, and accuracy. *NeuroImage, 23*, S275-S288.

Braver, T.S. et al. (1997). A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage, 5*, 49-62.

Brigadoi, S. et al. (2013). Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data. *NeuroImage, 85*(1), 181-191.

Callicott, J.H. et al. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral Cor, 9*(1), 20-26.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci, 24*(1), 87-114.

Finn, A.S. et al. (2017). Functional brain organization of working memory in adolescents varies in relation to family income and academic achievement. *Dev Sci, 20*(5), e12450.

Hochberg, Y., Benjamini, Y. (1990). More powerful procedures for multiple significance testing. *Statistics in Medicine, 9*(7), 811-818.

Hubbard, N.A. et al. (2014). Central executive dysfunction and deferred prefrontal processing in veterans with Gulf War Illness. *Clin Psychol Sci,* 2(3), 319-327.

Huppert, T.J. et al. (2009). HomER: A reviewer of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl Opt, 48*(10), D280-298.

Jahani, S. et al. (2017). fNIRS can robustly measure brain activity during memory encoding and retrieval in healthy subjects. *Sci Reports, 7*(1), 9533.

Powell, L.J. et al. (2018). Using individual functional channels of interest to study cortical development with fNIRS. *Dev Sci, 21*(4), e12595.

Rypma, B. et al. (1999). Load-dependent roles of frontal brain regions in the maintenance of working memory. *NeuroImage, 9*(2), 216-226.

Satterthwaite, T.D. et al. (2013). Functional maturation of the executive system during adolescence. *J Neurosci, 33*(41), 16249-16261.

Scholkmann, F. et al. (2010). How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation. *Physiol Measur, 31*(5), 649-662.

Tsuzuki, D. et al. (2007). Virtual spatial registration of stand-alone fNIRS data to MNI space. *NeuroImage, 34*(4), 1506-1518.

Van Snellenberg, J.X. et al. (2015). Dynamic shifts in brain network activation during supracapacity working memory task performance. *Hum Brain Mapp, 36*(4), 1245-1264.