
SPECIAL SERIES

Ecological Validity and Neuroanatomical Correlates of the NIH EXAMINER Executive Composite Score

Katherine L. Possin,¹ Amanda K. LaMarre,¹ Kristie A. Wood,¹ Dan M. Mungas,² AND Joel H. Kramer¹

¹Department of Neurology, University of California – San Francisco, San Francisco, California

²Department of Neurology, University of California – Davis, Davis, California

(RECEIVED January 21, 2013; FINAL REVISION May 14, 2013; ACCEPTED May 15, 2013; FIRST PUBLISHED ONLINE June 14, 2013)

Abstract

Executive functions refer to a constellation of higher-level cognitive abilities that enable goal-oriented behavior. The NIH EXAMINER battery was designed to assess executive functions comprehensively and efficiently. Performance can be summarized by a single score, the “Executive Composite,” which combines measures of inhibition, set-shifting, fluency, and working memory. We evaluated the ecological validity of the Executive Composite in a sample of 225 mixed neurological patients and controls using the Frontal Systems Behavior Scale (FrSBe), an informant-based measure of real-world executive behavior. In addition, we investigated the neuroanatomical correlates of the Executive Composite using voxel-based morphometry in a sample of 37 participants diagnosed with dementia, mild cognitive impairment, or as neurologically healthy. The Executive Composite accounted for 28% of the variance in Frontal Systems Behavior Scale scores beyond age. Even after including two widely used executive function tests (Trails B and Stroop) as covariates, the Executive Composite remained a significant predictor of real-world behavior. Anatomically, poorer scores on the Executive Composite were associated with smaller right and left dorsolateral prefrontal volumes, brain regions critical for good executive control. Taken together, these results suggest that the Executive Composite measures important aspects of executive function not captured by standard measures and reflects the integrity of frontal systems. (*JINS*, 2014, 20, 20–28)

Keywords: Executive function, Magnetic resonance imaging, Validity of results, Neurological disorders, Prefrontal cortex, Neurodegenerative disorders

INTRODUCTION

Executive functions refer to a constellation of cognitive abilities that allow us to engage in goal-oriented behavior. They include the ability to generate and organize thought, to think flexibly, to mentally manipulate information, to self-monitor, and to adjust behavior as appropriate to the current context (Cummings & Miller, 2007; Lezak, Howieson, Bigler, & Tranel, 2012). Deficits in executive functions are prominent in a multitude of neurological conditions, including Alzheimer’s disease (Bondi et al., 2008), frontotemporal dementia (Possin et al., 2013; Torralva et al., 2007), Parkinson’s disease (Muslimovic, Post, Speelman, & Schmand, 2005), Huntington’s disease (Peinemann et al., 2005), multiple sclerosis (Parmenter, Shucard, & Shucard, 2007), traumatic brain

injury (Caeyenberghs et al., 2012; McDonald, Flashman, & Saykin, 2002), brain tumors (Correa, 2010), and vascular disease of the brain (Hayes, Donnellan, & Stokes, 2012; Sachdev et al., 2004). Executive dysfunction can devastate a patient’s level of functioning (Farias et al., 2009), predicting functional decline beyond memory or global cognition (Boyle, Paul, Moser, & Cohen, 2004; Cahn-Weiner et al., 2007). Executive functions rely heavily on the integrity of the frontal lobes, which represent over 30% of the cortical surface of the brain and play a major role in the organization of behavior and cognition (Clark, Cools, & Robbins, 2004; Diamond, 2002; Rolls, 2004; Stuss & Levine, 2002). The intactness of posterior cortical structures such as the parietal lobes (Champod & Petrides, 2010; Yin et al., 2012), deep gray matter (e.g., dorsal head of the caudate nucleus and the thalamus; Little et al., 2010; Provost, Petrides, & Monchi, 2010; Schmahmann & Pandya, 2008; Van der Werf et al., 2003), and the cerebellum (Stoodley & Schmahmann, 2010) are also necessary for good executive control.

Correspondence and reprint requests to: Katherine L. Possin, 675 Nelson Rising Lane, Ste 190, MC 1207, San Francisco, CA 94158. E-mail: kpossin@memory.ucsf.edu

Despite the high frequency of executive dysfunction in neurological disorders and its impact on daily living, executive deficits are rarely measured comprehensively or consistently. This is often true of clinical trials, which vary in their selection of executive measures or do not include them at all, even when the treatment targets cognitive disorders with prominent executive dysfunction (Brown et al., 2003; Coles et al., 2012; Vercelletto et al., 2011; Willmott & Ponsford, 2009). For example, using the search term: frontotemporal dementia on www.clinicaltrials.gov brings back 14 studies that investigate therapeutic pharmaceutical agents for the treatment of frontotemporal dementia. Of those 14, 10 include standardized neuropsychological measures of cognition, while only four include specific measures of executive function, with inconsistent selection (Clinicaltrials.gov Identifier: NCT00545974; NCT00127114; NCT00604591; NCT00187525). There remains a compelling need to have a battery of executive function tests that can be routinely integrated into neurobehavioral research and that will reliably and validly measure constructs that clinical investigators agree are important. The NIH EXAMINER battery was developed to address this need.

The NIH EXAMINER generates 11 indices to capture important aspects of executive functions including inhibition, set-shifting, working memory, and fluency. Verbal, visual, and spatial stimuli are used. The “Executive Composite” is calculated to summarize performance across the 11 indices. It has excellent reliability (test–retest = .93; Kramer et al., this issue) and provides a comprehensive estimate of an individual’s executive functions that is not specific to a single test or type of stimuli. Although sometimes measures of discreet executive functions are indicated, the Executive Composite could be useful for research and clinical applications when a global and reliable estimate of executive function is needed, for example, in clinical trials.

We investigated the validity of the NIH EXAMINER Executive Composite score *via* two studies. In Study 1, we investigated the ecological validity of the Executive Composite with an informant-based measure of real-world executive behavior, the Frontal Systems Behavior Scale™ (FrSBe) (Grace & Malloy, 2001), in a sample of patients with a variety of neurological disorders and neurologically healthy controls. The FrSBe has been validated as a measure of executive behavior by several studies (Basso et al., 2008; Chiaravalloti & DeLuca, 2003; Hellmuth et al., 2012; Lane-Brown & Tate, 2009; Malloy & Grace, 2005; Malloy, Tremont, Grace, & Frakey, 2007; Velligan, Ritch, Sui, DiCocco, & Huntzinger, 2002). The Composite was considered to have good ecological validity if it was a significant predictor of the FrSBe in regression models. In Study 2, we investigated the neuroanatomical correlates of the Executive Composite using voxel-based morphometry in a sample of patients with neurodegenerative disease and control participants to determine if the Composite correlated with brain volumes in regions known to be important for executive control.

METHOD

Study 1: Ecological Validity

Participants

All participants from the NIH EXAMINER validation sample who were administered the Trail Making Test (Reitan, 1955), the Stroop Test (Kramer et al., 2003), and the FrSBe, and who were 18 years of age or older, were included in “Study 1: Concurrent Validity.” The sample included 136 patients with neurological disorders and 89 neurologically healthy controls (HC). Fifty-four patients were diagnosed with a neurodegenerative disorder by a board-certified neurologist using probable research criteria, as follows: Alzheimer’s disease ($N = 24$; AD) (McKhann et al., 2011), behavioral variant frontotemporal dementia ($N = 17$; bvFTD) (Rascovsky et al., 2011), Parkinson’s disease ($N = 7$; PD), progressive supranuclear palsy ($N = 3$; PSP) (Litvan et al., 1996), and Huntington’s disease ($N = 3$; HD). The HD patients showed unequivocal motor signs with genetic confirmation of the disease. Thirty-one patients were diagnosed with mild cognitive impairment (Winblad et al., 2004). Thirty suffered a focal lesion due to ischemic stroke, tumor, or focal injury at least 3 months before participation, and 15 suffered a moderate to severe traumatic brain injury as defined by a Glasgow Coma Scale <12 at least 6 months before testing. Five patients were diagnosed with multiple sclerosis (Polman et al., 2005). On average, the patients and the controls were similar in age (patients: 65 ± 18 , HC: 64 ± 19), $p = .66$, and years of education (patients: 15.4 ± 2.7 , HC: 15.6 ± 3.0), $p = .67$; however, there were more females in the control group (56%) than the patient group (41%), $p = .03$ (Table 1). Participants were evaluated at six separate sites: Mayo Clinic Alzheimer’s Disease Research Center ($N = 85$), University of California – San Francisco Memory and Aging Center ($N = 68$), University of Iowa ($N = 38$), University of Texas Southwestern Medical Center ($N = 29$), University of California – Davis ($N = 4$), and University of Colorado – Denver ($N = 1$). Neurologically healthy controls underwent neurological and cognitive screening to verify health status. Patients and controls were excluded if they had current major psychiatric illness or substance abuse disorder, ongoing cancer treatment, known HIV, or history of metabolic abnormalities, major systemic medical illness, traumatic brain injury with >30 min loss of consciousness, seizure disorder, or diagnosis of developmental learning disability. For both Study 1 and Study 2, written informed consent was obtained from each participant or their legal guardian (with assent from the participant) before testing. The study was approved by the committees on human research of the respective institutions and was completed in accordance with the Helsinki Declaration.

Executive Function Assessment

Participants were administered the NIH EXAMINER battery in a quiet room using a standard 15.4-inch Dell Latitude D830 laptop. The tests and the methods for generating the

Table 1. Study 1: Demographic characteristics and executive scores by diagnostic group

Diagnosis	<i>N</i>	Age	Ed	% Male	EC	FrsBe	Trails B	Stroop
HC	89	64.2 (19.2)	15.6 (3.1)	44	.99 (.61)	69.5 (18.4)	62.3 (34.6)	53.2 (15.5)
MCI	31	74.2 (10.7)	15.7 (2.3)	52	.36 (.49)	98.1 (36.3)	102.3 (64.0)	39.3 (11.2)
Lesion	30	63.3 (11.0)	14.4 (2.8)	43	.33 (.55)	92.2 (24.7)	88.8 (42.0)	48.1 (14.1)
AD	24	79.7 (9.5)	16.0 (2.6)	79	-.07 (.71)	93.5 (27.1)	163.1 (99.2)	26.3 (14.2)
bvFTD	17	66.7 (10.7)	16.0 (2.1)	65	-.43 (.98)	137.1 (41.9)	147.8 (94.9)	31.3 (17.7)
TBI	16	31.2 (9.0)	13.6 (2.6)	56	.61 (.88)	113.5 (26.3)	61.9 (37.1)	59.9 (16.5)
PD	7	70.4 (7.8)	15.6 (2.2)	71	.18 (.70)	95.1 (14.6)	114.4 (83.3)	41.7 (10.5)
MS	5	49.6 (10.7)	16.8 (2.3)	80	.59 (.31)	104.0 (33.7)	69.2 (22.9)	54.4 (19.3)
PSP	3	74.7 (5.5)	18.3 (5.1)	100	-1.11 (.87)	128.7 (26.8)	233.0 (116.0)	16.3 (10.7)
HD	3	55.0 (10.8)	16.7 (3.1)	0	-.26 (.31)	127.7 (54.3)	112.7 (62.4)	36.0 (11.5)

Note. Values represent mean (standard deviation).

Ed = Education; EC = Executive Composite; FrsBe = Frontal Systems Behavior Scale; HC = healthy control; MCI = mild cognitive impairment; Lesion = focal lesion due to stroke; AD = Alzheimer's disease; bvFTD = behavioral variant frontotemporal dementia; TBI = traumatic brain injury; PD = Parkinson's disease; MS = multiple sclerosis; PSP = progressive supranuclear palsy; HD = Huntington's disease.

Executive Composite are described in Kramer et al. (this issue). The Trail Making Test and the Stroop test are widely used executive measures (Rabin, Barr, & Burton, 2005) and were administered during the same session. Completion time on Trails B and number correct on the Stroop Interference condition were used to operationalize executive performance on these tests. To evaluate for real-world executive behavior, the FrsBe was completed by an informant who knew the participant well, in most cases the spouse or other close relative, either in-person during the visit or at home and then sent back by mail. Total raw scores were used to operationalize real-world executive behavior.

Data analysis

Statistical analyses were performed using PASW 17.0 for Windows (SPSS Inc., Chicago, IL). Using hierarchical regression, we evaluated how well the Executive Composite predicts real-world executive behavior after accounting for the variance explained by age and gender. Next, we performed a second regression to evaluate whether the Executive Composite predicts real-world executive behavior after accounting for variance predicted by Trails B, Stroop, age, and gender. To address the possible confounding effects of motor dysfunction, the regressions were repeated

after removing patients diagnosed with Parkinson's disease, multiple sclerosis, progressive supranuclear palsy, or Huntington's disease. All *p* values less than .05 were considered to be significant.

Study 2: Neuroanatomical Correlates

Participants

MRI data were available on a sample of 18 patients and 19 HCs evaluated at the UCSF Neuroscience Imaging Center. The patients included eight patients diagnosed with bvFTD, six patients diagnosed with PSP, three patients diagnosed with AD, and one patient diagnosed with MCI using the same research criteria as in Study 1. The combined patients and the controls did not differ significantly in age, gender, or education, all *p* values > .05 (Table 2).

An additional sample of 50 healthy age-matched normal controls was used to create a template for inter-subject normalization and voxel-based morphometry.

Neuroimaging data acquisition

MRI scans were obtained on all Study 2 participants using a 3.0 Tesla Siemens (Siemens, Iselin, NJ) TIM Trio scanner

Table 2. Study 2: Demographic characteristics, Mini Mental State Exam scores, Clinical Dementia Rating Scale scores, and Executive Composite scores by diagnostic group

Diagnosis	<i>N</i>	Age	Education	% Male	MMSE	CDR-G	CDR-Box	EC
HC	19	68.1 (9.7)	15.9 (1.9)	47	29.7 (0.5)	0.0 (0.0)	29.7 (0.5)	1.18 (0.4)
bvFTD	8	63.2 (4.7)	15.3 (2.1)	38	26.3 (2.9)	1.1 (0.4)	26.3 (2.9)	-0.06 (0.6)
PSP	6	66.6 (5.4)	14.7 (1.8)	83	27.5 (1.2)	1.0 (0.5)	27.5 (1.2)	-1.12 (0.5)
AD	3	68.3 (7.7)	14.0 (2)	67	25.7 (3.5)	1.0 (0.0)	25.7 (3.5)	-0.01 (0.5)
MCI	1	64.4	14.0	0	28	0.5	28	0.32

Note. Values represent mean (standard deviation).

MMSE = Mini Mental State Examination; CDR-G = Clinical Dementia Rating Scale Global Score; CDR-Box = Clinical Dementia Rating Scale Box Score; EC = Executive Composite; HC = healthy control; bvFTD = behavioral variant frontotemporal dementia; AD = Alzheimer's disease; PSP = progressive supranuclear palsy; MCI = mild cognitive impairment.

equipped with a 12-channel head coil. Whole brain images were acquired using volumetric magnetization prepared rapid gradient-echo sequence (MPRAGE; TR/TE/TI = 2300/2.98/900 ms; $\alpha = 9^\circ$). The field of view was 240×256 mm, with 1×1 mm in-plane resolution and 1 mm slice thickness.

Voxel-based morphometry

We investigated the neuroanatomical correlates of the Executive Composite. Structural T1 images were corrected for bias field, segmented into gray matter, white matter, and cerebrospinal fluid, and initially normalized into Montreal Neurological Institute (MNI) space using the Unified Segmentation procedure (Ashburner & Friston, 2005) implemented in SPM 5 (Friston, Ashburner, Kiebel, Nichols, & Penny, 2007), running under MATLAB 8.0. More anatomically precise inter-subject registration was then performed with the DARTEL toolbox (Ashburner, 2007) by warping each subject's image to a template created from the 50 normal control subjects. Modulated gray and white matter probability maps were scaled by Jacobians, smoothed to a Gaussian kernel of 8 mm full-width at half maximum, then summed together to obtain a map of brain parenchyma (Wilson et al., 2010).

The Executive Composite was correlated voxel-wise with the summed gray and white matter probabilities for the 37 subjects as a single group using the voxel-based lesion-symptom mapping toolbox (Bates et al., 2003); neuroling.arizona.edu/resources.html). Statistical maps were thresholded at voxelwise $p < .001$ and then corrected for multiple comparisons by permutation analysis. Statistical maps were calculated for 1000 random assignments of normally distributed behavioral scores to patients, with the maximum cluster size recorded each time. The fifth percentile maximum cluster size was 454 mm^3 , so applying this as the minimum cluster size ensured corrected significance of $p < .05$. Age and total intracranial volume were included as nuisance covariates in all VBM analyses.

Two secondary VBM analyses were conducted with all subjects except those with bvFTD, and then with all subjects except those with PSP. The purpose of these analyses was to determine whether our primary VBM findings were specific to either of our largest diagnostic groups; that is, if the findings were no longer significant with either bvFTD or PSP removed. This was important because PSP patients display motor dysfunction and because bvFTD and PSP are each associated with characteristic patterns of atrophy. BvFTD causes atrophy in medial and orbital prefrontal cortex that extends to lateral prefrontal cortex with disease progression; early targets also include insula, medial temporal, and striatal-thalamic structures (Seeley et al., 2008). PSP causes atrophy of supplementary motor area, premotor cortex, caudate, thalamus, anterior insula, and brainstem (Whitwell et al., 2012). By emphasizing only findings that are significant in both of these secondary VBMs, we reduce the impact of any one diagnostic group. We restricted our secondary analyses within a mask of our full sample results thresholded

Table 3. Regression results predicting FrSBe scores from age, gender, and the Executive Composite

Predictor variables	β	95% CI for β	p Value
Age	-.16	(-.36, -.10)	.005
Gender	-.11	(-.24, .02)	.052
Executive Composite	-.54	(-.67, -.41)	<.001

FrSBe = Frontal Systems Behavior Scale; CI = confidence interval.

at voxelwise $p < .05$ with permutation correction and thresholded clusterwise at $p < .05$. For each of these two analyses within this mask, statistical maps were thresholded at voxelwise $p < .001$ with permutation correction, and then a clusterwise threshold of $p < .05$ was applied, which was 495 voxels for the analysis without the bvFTD subjects and 499 voxels without the PSP subjects.

RESULTS

Study 1: Ecological Validity

The regression model with age, gender, and the Executive Composite entered as predictors significantly predicted real-world executive behavior, $F(3,221) = 33.43$, $p < .001$, (Table 3). Using hierarchical regression, the Executive Composite accounted for 28% of the variance in FrSBe scores beyond age and gender, $p < .001$. When Trails B and Stroop scores were entered simultaneously with age and gender, the Executive Composite remained a significant predictor, $p < .001$, predicting an additional 4% of the variance after accounting for the other covariates (Table 4). Trails B was also a significant unique predictor, $p = .001$, but the Stroop did not predict significant unique variance, $p = .44$.

These regressions were repeated after removing patients with motor dysfunction from the sample. The Executive Composite accounted for 27% of the variance in FrSBe scores beyond age and gender, $p < .001$, and 5% of the variance beyond Trails B, Stroop, age, and gender, $p < .001$.

Study 2: Neuroanatomical Correlates

Lower Executive Composite scores were associated with smaller brain volumes covering a large extent of frontal regions in a lateral and medial pattern. The largest and

Table 4. Regression results predicting FrSBe scores from age, gender, Trails B, Stroop, and the Executive Composite

Predictor variables	β	95% CI for β	p Value
Age	-.23	(-.36, -.10)	.001
Gender	-.08	(-.21, .06)	.176
Trails B	.27	(.13, .40)	.001
Stroop	-.07	(-.20, .06)	.443
Executive Composite	-.32	(-.45, -.19)	<.001

FrSBe = Frontal Systems Behavior Scale; CI = confidence interval.

Table 5. Neuroanatomical correlates of the Executive Composite in the full sample

Region	MNI coordinates			Max T
	x	y	z	
Cluster 1 (15,583 mm ³ ; $p = .003$, corrected)	-6	4	31	5.24
Right frontal regions				
Middle frontal gyrus	34	6	40	4.45
Superior frontal gyrus	29	-3	68	4.34
Precentral gyrus	38	2	44	4.11
Supplementary motor area	10	-1	54	3.69
Mid cingulate	10	13	42	3.67
Inferior frontal gyrus	47	15	37	3.66
Anterior cingulate	8	36	24	3.53
Left frontal regions				
Superior frontal gyrus	-12	13	55	5.08
Supplementary motor area	-11	15	56	5.01
Inferior frontal gyrus	-34	17	34	4.37
Middle frontal gyrus	-38	17	35	4.12
Precentral gyrus	-39	4	34	4.10
Mid cingulate	-13	-11	48	3.94
Rolandic operculum	-54	4	4	3.66
Anterior cingulate	-11	30	30	3.44
Left parietal regions				
Left postcentral gyrus	-61	-4	28	4.38
Left temporal regions				
Left superior temporal gyrus	-50	-23	6	4.40
Subcortical and brainstem				
Brainstem	-4	-20	-10	5.24
Left thalamus	-14	-13	0	4.19
Right thalamus	14	-15	-1	3.56
Cluster 2 (1,095 mm ³ ; $p = .02$, corrected)	22	-47	-31	4.08
Right cerebellum	22	-44	-30	4.08
Cluster 3 (975 mm ³ ; $p = .03$, corrected)	-19	-49	-34	4.28
Left cerebellum	-20	-44	-32	4.28

Note. Results were thresholded at voxelwise $p < .001$ and corrected for multiple comparisons by permutation analysis. Cluster size was thresholded at $p < .05$; only clusters greater than 454 mm³ were considered significant. MNI coordinates for the clusters are the center of the mass, whereas MNI coordinates for the regions are the local maxima.
MNI = Montreal Neurological Institute.

most significant cluster included the bilateral superior, middle, and inferior frontal gyri, the supplementary motor area, the precentral gyrus, and the anterior and middle cingulate. This cluster extended to the left rolandic operculum, the left postcentral gyrus, the left superior temporal gyrus, the brainstem, and the bilateral anteroventrolateral thalamus. Smaller volumes of the right and left cerebellum were also associated with lower Executive Composite scores (Table 5).

To determine whether the neuroanatomical correlates were specific to bvFTD or PSP, the analyses were repeated twice, once with the bvFTD subjects removed (Table 6) and once with the PSP subjects removed (Table 7). Regions that were significant across all three analyses were one cluster in the right middle and superior frontal gyri that extended into the inferior frontal gyrus and middle cingulate, and one cluster in the left middle frontal gyrus that extended into the inferior frontal gyrus. This overlap of the three analyses is depicted in Figure 1.

DISCUSSION

We investigated the concurrent validity and neuroanatomical correlates of the NIH EXAMINER Executive Composite score. The Executive Composite was a robust predictor of real-world executive behavior, even after accounting for performance on the commonly used executive function tests Trails B and Stroop. Poor performance on the Executive Composite correlated with atrophy in brain regions important for executive control, including the right middle and superior frontal gyri and the left middle frontal gyri.

The Executive Composite is measured using objective testing, and the FrSBe is an informant-based rating scale. These methods have different limitations as measures of behavior. Objective testing captures the patient's behavior during one structured visit and may not capture intermittent behaviors or behaviors that emerge in less structured environments. Informant scales can address these limitations, but are dependent on the quality of the informant's report.

Table 6. Neuroanatomical correlates of the Executive Composite without the bvFTD subjects

Region	MNI coordinates			Max T
	x	y	z	
Cluster 1 (11,373 mm ³ ; <i>p</i> < .01)	-6	3	38	5.03
Right frontal regions				
Middle frontal gyrus	35	9	42	4.88
Precentral gyrus	42	9	47	4.85
Superior frontal gyrus	28	-3	68	4.20
Supplementary motor area	1	-1	56	4.41
Mid cingulate	13	-7	45	3.96
Inferior frontal gyrus	32	9	35	4.66
Left frontal regions				
Anterior cingulate	-12	31	30	3.81
Mid cingulate	-11	1	46	4.83
Precentral gyrus	-36	3	34	4.08
Middle frontal gyrus	-30	17	35	4.22
Supplementary motor area	-10	13	48	4.91
Superior frontal gyrus	-14	9	49	4.71
Rolandic operculum	-61	0	10	4.39
Inferior frontal gyrus	-61	9	9	3.61
Left parietal regions				
Postcentral gyrus	-62	-5	26	5.03
Left temporal regions				
Superior temporal gyrus	-50	-24	9	4.37
Cluster 2 (1,544 mm ³ ; <i>p</i> = .02)	-4	-18	-8	5.00
Brainstem	-2	-22	-10	5.00
Right thalamus	15	-12	0	3.46
Left thalamus	-13	-13	-1	4.01
Cluster 3 (913 mm ³ ; <i>p</i> = .03)	24	-44	-32	4.27
Right cerebellum	26	-42	-30	4.27
Cluster 4 (672 mm ³ ; <i>p</i> = .04)	-22	-43	-34	4.81
Left cerebellum	-28	-36	-40	4.81

Note. Results were thresholded at voxelwise *p* < .001 and corrected for multiple comparisons by permutation analysis. Cluster size was thresholded at *p* < .05; only clusters greater than 495 mm³ were considered significant. MNI coordinates for the clusters are the center of the mass, whereas MNI coordinates for the regions are the local maxima. MNI = Montreal Neurological Institute; bvFTD = behavioral variant frontotemporal dementia.

Because the Executive Composite and the FrSBe use distinct methods, their shared variance can be attributed to a shared construct. The unique contribution of the Executive Composite to FrSBe after accounting for performance on Trails B and Stroop indicates that the Composite is sensitive to components of executive function not tapped by these classic measures. Additional research will be needed to evaluate the Executive Composite's validity beyond other executive measures, such as measures of planning.

We did not control for speed in the Executive Composite, Trails B, or Stroop because speed is closely related to executive functions (Salthouse, 2005). The purpose of this study was to validate a measure that is optimally sensitive to executive functions, and controlling for speed would have removed meaningful variance. The analyses were repeated after removing patients diagnosed with disorders that cause motor dysfunction, and a similar pattern of results was found,

Table 7. Neuroanatomical correlates of the Executive Composite without the PSP subjects

Region	MNI coordinates			
	x	y	z	Max T
Cluster 1 (2,574 mm ³ ; <i>p</i> < .01)	24	31	35	5.36
Right frontal				
Middle frontal gyrus	36	36	44	5.36
Superior frontal gyrus	22	33	39	4.80
Anterior cingulate	6	37	25	4.30
Inferior frontal gyrus	34	19	31	3.77
Mid cingulate	9	26	34	3.61
Cluster 2 (1,049 mm ³ ; <i>p</i> = .02)	-30	32	29	4.11
Left frontal				
Inferior frontal gyrus, pars triangularis	-38	16	30	4.11
Middle frontal gyrus	-27	29	38	3.91
Superior frontal gyrus	-24	32	36	3.90
Cluster 3 (769 mm ³ ; <i>p</i> = .03)	-26	8	-36	4.04
Left temporal				
Temporal pole	-22	6	-36	4.01
Fusiform	-22	4	-41	4.01
Inferior temporal gyrus	-27	4	-42	3.90
Left frontal				
Inferior frontal gyrus, pars orbitalis	-31	27	-14	3.48

Note. Results were thresholded at voxelwise *p* < .001 and corrected for multiple comparisons by permutation analysis. Cluster size was thresholded at *p* < .05; only clusters greater than 495 mm³ were considered significant. MNI coordinates for the clusters are the center of the mass, whereas MNI coordinates for the regions are the local maxima. MNI = Montreal Neurological Institute; PSP = progressive supranuclear palsy.

suggesting that the relationship to real-world executive behavior was not driven by motor dysfunction.

The Executive Composite scores correlated with brain volume in regions important for executive control. In the full sample, the largest extent of the correlations was in the lateral and medial aspects of the frontal lobes bilaterally. Thalamic volumes corresponding to the ventral anterior and ventrolateral nuclei were also significant; these are regions important for higher cognitive and motor functions *via* circuits with the frontal lobes (Royall et al., 2002). Correlations with the supplementary motor area, the precentral gyrus, the midbrain and dorsal pons, and the cerebellum might reflect the speeded measures included in the composite score. The extent of correlations with the cerebellum subsumed subregions important for cognitive, sensorimotor, and limbic functions (Stoodley & Schmahmann, 2010). Volumes of the left superior temporal and the left postcentral gyri were the only temporal and parietal regions to reach significance. The limited extent of correlations within posterior regions could be explained by the multi-faceted nature of the Executive Composite. That is, posterior regions important for fundamental cognitive skills such as processing verbal or visual information were central to some tasks but not others, and so the impact of these brain regions on the multi-domain composite was diminished. These findings suggest that the multi-domain Executive Composite may emphasize higher-level processes not specific to any one domain. It is also likely that the lack of significant correlations

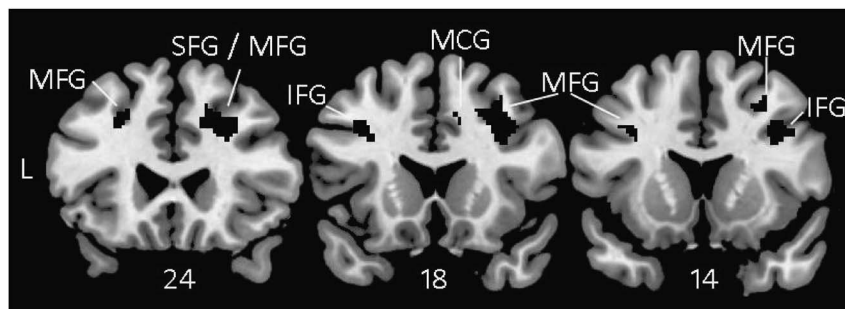


Fig. 1. Voxel-based morphometry showing brain regions where atrophy correlated with lower scores on the Executive Composite consistently across all three analyses: Full sample (Table 5), No bvFTD (Table 6), and No PSP (Table 7). These results are not specific to any one diagnostic group. All results were thresholded voxelwise at $p < .001$ and corrected for multiple comparisons using permutation analysis with $p < .05$. MFG = middle frontal gyrus; SFG = superior frontal gyrus; IFG = inferior frontal gyrus; MCG = middle cingulate gyrus; PSP = progressive supranuclear palsy.

with posterior regions reflects the atrophy patterns of the patient sample. Atrophy patterns in PSP and bvFTD are frontally predominant (Seeley et al., 2008; Whitwell et al., 2012). If we had included a sample of subjects with greater posterior disease, our findings may have been different. The extent to which the present findings generalize to other patient samples will need to be tested in future research.

To minimize the specific impact of bvFTD or PSP atrophy patterns on the neuroanatomical correlates, we repeated the voxel-based morphometry analyses once with the bvFTD patients removed and once with the PSP patients removed. Voxels that were significant in the full sample and in these subsidiary analyses were primarily in the right middle and superior frontal gyri and the left middle frontal gyri, but also in the right and left inferior frontal gyri and the right middle cingulate (Figure 1). We are most confident that these regions were important for Executive Composite performance irrespective of diagnosis. Prefrontal lesions have been associated with real world executive behavior decrements as measured by the FrSBe (Robinson, Calamia, Glascher, Bruss, & Tranel, this issue), suggesting a common neural mechanism for executive impairments on testing and in everyday life.

The NIH EXAMINER Executive Composite is a single score that represents performance across 11 indices of executive functions including measures of inhibition, set-shifting, fluency, and working memory that rely to varying degrees on verbal, spatial, or visual stimuli. In this study, we found that the Executive Composite is a robust predictor of executive behavior in the real world and correlates with atrophy in frontal brain regions important for executive control. The Executive Composite is not the appropriate choice for all clinical and research questions about executive functions, for example, when time is limited (it is based on a 30-min test battery) and a shorter test like Trails B is adequate. Furthermore, differential diagnosis and inferences about affected brain function is often better guided by the separate evaluation of discreet executive functions. It is, however, an appropriate choice when a single comprehensive score with high reliability, sensitivity to real-world executive

dysfunction, linear measurement across the ability spectrum, or multiple alternate forms is needed.

ACKNOWLEDGMENTS

This work was supported by NINDS (J.K., HHSN271200623661C); NIA (K.P., K23AG037566), (J.K., P01AG019724 and P50AG023501); and by the Hellman Family Foundation (K.P.). We thank Stephen Wilson, Ph.D., for his guidance on the voxel-based morphometry methods, and our research participants for their generous time and efforts. We have no conflicts of interest to report.

REFERENCES

- Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage*, *38*(1), 95–113.
- Ashburner, J., & Friston, K.J. (2005). Unified segmentation. *Neuroimage*, *26*(3), 839–851.
- Basso, M.R., Shields, I.S., Lowery, N., Ghormley, C., Combs, D., Arnett, P.A., & Johnson, J. (2008). Self-reported executive dysfunction, neuropsychological impairment, and functional outcomes in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *30*(8), 920–930.
- Bates, E., Wilson, S.M., Saygin, A.P., Dick, F., Sereno, M.I., Knight, R.T., & Dronkers, N.F. (2003). Voxel-based lesion-symptom mapping. *Nature Neuroscience*, *6*(5), 448–450.
- Bondi, M.W., Jak, A.J., Delano-Wood, L., Jacobson, M.W., Delis, D.C., & Salmon, D.P. (2008). Neuropsychological contributions to the early identification of Alzheimer's disease. *Neuropsychology Review*, *18*(1), 73–90.
- Boyle, P.A., Paul, R.H., Moser, D.J., & Cohen, R.A. (2004). Executive impairments predict functional declines in vascular dementia. *Clinical Neuropsychologist*, *18*(1), 75–82.
- Brown, P.D., Buckner, J.C., O'Fallon, J.R., Iturria, N.L., Brown, C.A., O'Neill, B.P., ... Shaw, E.G. (2003). Effects of radiotherapy on cognitive function in patients with low-grade glioma measured by the folstein mini-mental state examination. *Journal of Clinical Oncology*, *21*(13), 2519–2524.
- Caeyenberghs, K., Leemans, A., Leunissen, I., Gooijers, J., Michiels, K., Sunaert, S., & Swinnen, S.P. (2012). Altered structural networks and executive deficits in traumatic brain injury patients. *Brain Structure & Function*. [Epub ahead of

- print]. Retrieved from <http://link.springer.com/article/10.1007/s00429-012-0494-2/fulltext.html>.
- Cahn-Weiner, D.A., Farias, S.T., Julian, L., Harvey, D.J., Kramer, J.H., Reed, B.R., ... Chui, H. (2007). Cognitive and neuroimaging predictors of instrumental activities of daily living. *Journal of the International Neuropsychological Society*, *13*(5), 747–757.
- Chamod, A.S., & Petrides, M. (2010). Dissociation within the frontoparietal network in verbal working memory: A parametric functional magnetic resonance imaging study. *Journal of Neuroscience*, *30*(10), 3849–3856.
- Chiaravalloti, N.D., & DeLuca, J. (2003). Assessing the behavioral consequences of multiple sclerosis: An application of the Frontal Systems Behavior Scale (FrSBe). *Cognitive and Behavioral Neurology*, *16*(1), 54–67.
- Clark, L., Cools, R., & Robbins, T.W. (2004). The neuropsychology of ventral prefrontal cortex: Decision-making and reversal learning. *Brain and Cognition*, *55*(1), 41–53.
- Coles, A.J., Twyman, C.L., Arnold, D.L., Cohen, J.A., Confavreux, C., Fox, E.J., ... Compston, D.A. (2012). Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: A randomised controlled phase 3 trial. *Lancet*, *380*(9856), 1829–1839.
- Correa, D.D. (2010). Neurocognitive function in brain tumors. *Current Neurology and Neuroscience Reports*, *10*(3), 232–239.
- Cummings, J., & Miller, B.L. (2007). Conceptual and clinical aspects of the frontal lobes. In B.L. Miller & J.L. Cummings (Eds.), *The human frontal lobes* (pp. 12–24). New York: The Guilford Press.
- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D.T. Stuss, & R.T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). London, UK: Oxford University Press.
- Farias, S.T., Cahn-Weiner, D.A., Harvey, D.J., Reed, B.R., Mungas, D., Kramer, J.H., & Chui, H. (2009). Longitudinal changes in memory and executive functioning are associated with longitudinal change in instrumental activities of daily living in older adults. *Clinical Neuropsychologist*, *23*(3), 446–461.
- Friston, K.J., Ashburner, J., Kiebel, S., Nichols, T., & Penny, W. (2007). *Statistical parametric mapping*. London: Academic Press.
- Grace, J., & Malloy, P.F. (2001). *Frontal systems behavior scale (FrSBe): Professional manual*. Lutz, FL: Psychological Assessment Resources.
- Hayes, S., Donnellan, C., & Stokes, E. (2012). Associations between executive function and physical function poststroke: A pilot study. *Physiotherapy*, *99*(2), 165–171.
- Hellmuth, J., Mirsky, J., Heuer, H.W., Matlin, A., Jafari, A., Garbutt, S., ... Boxer, A.L. (2012). Multicenter validation of a bedside antisaccade task as a measure of executive function. *Neurology*, *78*(23), 1824–1831.
- Kramer, J.H., Jurik, J., Sha, S.J., Rankin, K.P., Rosen, H.J., Johnson, J.K., & Miller, B.L. (2003). Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer disease. *Cognitive and Behavioral Neurology*, *16*(4), 211–218.
- Kramer, J.H., Mungas, D., Possin, K.L., Rankin, K.P., Boxer, A.L., Rosen, H.J., ... Windmeyer, M. (2013). NIH EXAMINER: Conceptualization and development of an executive function battery. *Journal of the International Neuropsychological Society*, *19*(3), 000–000.
- Lane-Brown, A.T., & Tate, R.L. (2009). Measuring apathy after traumatic brain injury: Psychometric properties of the Apathy Evaluation Scale and the Frontal Systems Behavior Scale. *Brain Injury*, *23*(13–14), 999–1007.
- Lezak, M.D., Howieson, D.B., Bigler, E.D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York: Oxford University Press.
- Little, D.M., Kraus, M.F., Joseph, J., Geary, E.K., Susmaras, T., Zhou, X.J., ... Gorelick, P.B. (2010). Thalamic integrity underlies executive dysfunction in traumatic brain injury. *Neurology*, *74*(7), 558–564.
- Litvan, I., Agid, Y., Calne, D., Campbell, G., Dubois, B., Duvoisin, R.C., ... Zee, D.S. (1996). Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): Report of the NINDS-SPSP international workshop. *Neurology*, *47*(1), 1–9.
- Malloy, P., & Grace, J. (2005). A review of rating scales for measuring behavior change due to frontal systems damage. *Cognitive and Behavioral Neurology*, *18*(1), 18–27.
- Malloy, P., Tremont, G., Grace, J., & Frakey, L. (2007). The Frontal Systems Behavior Scale discriminates frontotemporal dementia from Alzheimer's disease. *Alzheimer's & Dementia: the Journal of the Alzheimer's Association*, *3*(3), 200–203.
- McDonald, B.C., Flashman, L.A., & Saykin, A.J. (2002). Executive dysfunction following traumatic brain injury: Neural substrates and treatment strategies. *Neurorehabilitation*, *17*(4), 333–344.
- McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Jr., Kawas, C.H., ... Phelps, C.H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*(3), 263–269.
- Muslimovic, D., Post, B., Speelman, J.D., & Schmand, B. (2005). Cognitive profile of patients with newly diagnosed Parkinson disease. *Neurology*, *65*(8), 1239–1245.
- Parmenter, B.A., Shucard, J.L., & Shucard, D.W. (2007). Information processing deficits in multiple sclerosis: A matter of complexity. *Journal of the International Neuropsychological Society*, *13*(3), 417–423.
- Peinemann, A., Schuller, S., Pohl, C., Jahn, T., Weindl, A., & Kassubek, J. (2005). Executive dysfunction in early stages of Huntington's disease is associated with striatal and insular atrophy: A neuropsychological and voxel-based morphometric study. *Journal of the Neurological Sciences*, *239*(1), 11–19.
- Polman, C.H., Reingold, S.C., Edan, G., Filippi, M., Hartung, H.P., Kappos, L., ... Wolinsky, J.S. (2005). Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Annals of Neurology*, *58*(6), 840–846.
- Possin, K.L., Feigenbaum, D., Rankin, K.P., Smith, G.E., Boxer, A.L., Wood, K., ... Kramer, J.H. (2013). Dissociable executive functions in behavioral variant frontotemporal and Alzheimer's dementias. *Neurology*, [Epub ahead of print]. Retrieved from <http://www.neurology.org/content/early/2013/05/08/WNL.0b013e318296e940.abstract>
- Provost, J.S., Petrides, M., & Monchi, O. (2010). Dissociating the role of the caudate nucleus and dorsolateral prefrontal cortex in the monitoring of events within human working memory. *European Journal of Neuroscience*, *32*(5), 873–880.
- Rabin, L.A., Barr, W.B., & Burton, L.A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, *20*(1), 33–65.
- Rascovsky, K., Hodges, J.R., Knopman, D., Mendez, M.F., Kramer, J.H., Neuhaus, J., ... Miller, B.L. (2011). Sensitivity

- of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, 134(Pt 9), 2456–2477.
- Reitan, R.M. (1955). The relation of the trail making test to organic brain damage. *Journal of Consulting Psychology*, 19(5), 393–394.
- Robinson, H., Calamia, M., Glascher, J., Bruss, J., & Tranel, D. (2013). Neuroanatomical correlates of executive functions: A neuropsychological approach using the EXAMINER battery. *Journal of the International Neuropsychological Society*, 19(3), 000–000.
- Rolls, E.T. (2004). The functions of the orbitofrontal cortex. *Brain and Cognition*, 55(1), 11–29.
- Royall, D.R., Lauterbach, E.C., Cummings, J.L., Reeve, A., Rummans, T.A., Kaufer, D.I., ... Coffey, C.E. (2002). Executive control function: A review of its promise and challenges for clinical research. A report from the Committee on Research of the American Neuropsychiatric Association. *Journal of Neuropsychiatry and Clinical Neurosciences*, 14(4), 377–405.
- Sachdev, P.S., Brodaty, H., Valenzuela, M.J., Lorentz, L., Looi, J.C., Wen, W., & Zagami, A.S. (2004). The neuropsychological profile of vascular cognitive impairment in stroke and TIA patients. *Neurology*, 62(6), 912–919.
- Salthouse, T.A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology*, 19(4), 532–545.
- Schmahmann, J.D., & Pandya, D.N. (2008). Disconnection syndromes of basal ganglia, thalamus, and cerebrotocerebellar systems. *Cortex*, 44(8), 1037–1066.
- Seeley, W.W., Crawford, R., Rascovsky, K., Kramer, J.H., Weiner, M., Miller, B.L., & Gorno-Tempini, M.L. (2008). Frontal paralimbic network atrophy in very mild behavioral variant frontotemporal dementia. *Archives of Neurology*, 65(2), 249–255.
- Stoodley, C.J., & Schmahmann, J.D. (2010). Evidence for topographic organization in the cerebellum of motor control versus cognitive and affective processing. *Cortex*, 46(7), 831–844.
- Stuss, D.T., & Levine, B. (2002). Adult clinical neuropsychology: Lessons from studies of the frontal lobes. *Annual Review of Psychology*, 53, 401–433.
- Torralva, T., Kipps, C.M., Hodges, J.R., Clark, L., Bekinschtein, T., Roca, M., ... Manes, F. (2007). The relationship between affective decision-making and theory of mind in the frontal variant of fronto-temporal dementia. *Neuropsychologia*, 45(2), 342–349.
- Van der Werf, Y.D., Scheltens, P., Lindeboom, J., Witter, M.P., Uylings, H.B., & Jolles, J. (2003). Deficits of memory, executive functioning and attention following infarction in the thalamus; a study of 22 cases with localised lesions. *Neuropsychologia*, 41(10), 1330–1344.
- Velligan, D.I., Ritch, J.L., Sui, D., DiCocco, M., & Huntzinger, C.D. (2002). Frontal Systems Behavior Scale in schizophrenia: Relationships with psychiatric symptomatology, cognition and adaptive function. *Psychiatry Research*, 113(3), 227–236.
- Vercelletto, M., Boutoleau-Bretonniere, C., Volteau, C., Puel, M., Auriacombe, S., Sarazin, M., ... Lacomblez, L. (2011). Memantine in behavioral variant frontotemporal dementia: Negative results. *Journal of Alzheimer's Disease*, 23(4), 749–759.
- Whitwell, J.L., Duffy, J.R., Strand, E.A., Machulda, M.M., Senjem, M.L., Gunter, J.L., ... Josephs, K.A. (2012). Neuroimaging comparison of primary progressive apraxia of speech and progressive supranuclear palsy. *European Journal of Neurology*, 20(4), 629–637.
- Willmott, C., & Ponsford, J. (2009). Efficacy of methylphenidate in the rehabilitation of attention following traumatic brain injury: A randomised, crossover, double blind, placebo controlled inpatient trial. *Journal of Neurology, Neurosurgery, and Psychiatry*, 80(5), 552–557.
- Wilson, S.M., Henry, M.L., Besbris, M., Ogar, J.M., Dronkers, N.F., Jarrold, W., ... Gorno-Tempini, M.L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, 133(Pt 7), 2069–2088.
- Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.O., ... Petersen, R.C. (2004). Mild cognitive impairment—beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. *Journal of Internal Medicine*, 256(3), 240–246.
- Yin, X., Zhao, L., Xu, J., Evans, A.C., Fan, L., Ge, H., ... Liu, S. (2012). Anatomical substrates of the alerting, orienting and executive control components of attention: Focus on the posterior parietal lobe. *PLoS One*, 7(11), e50590.