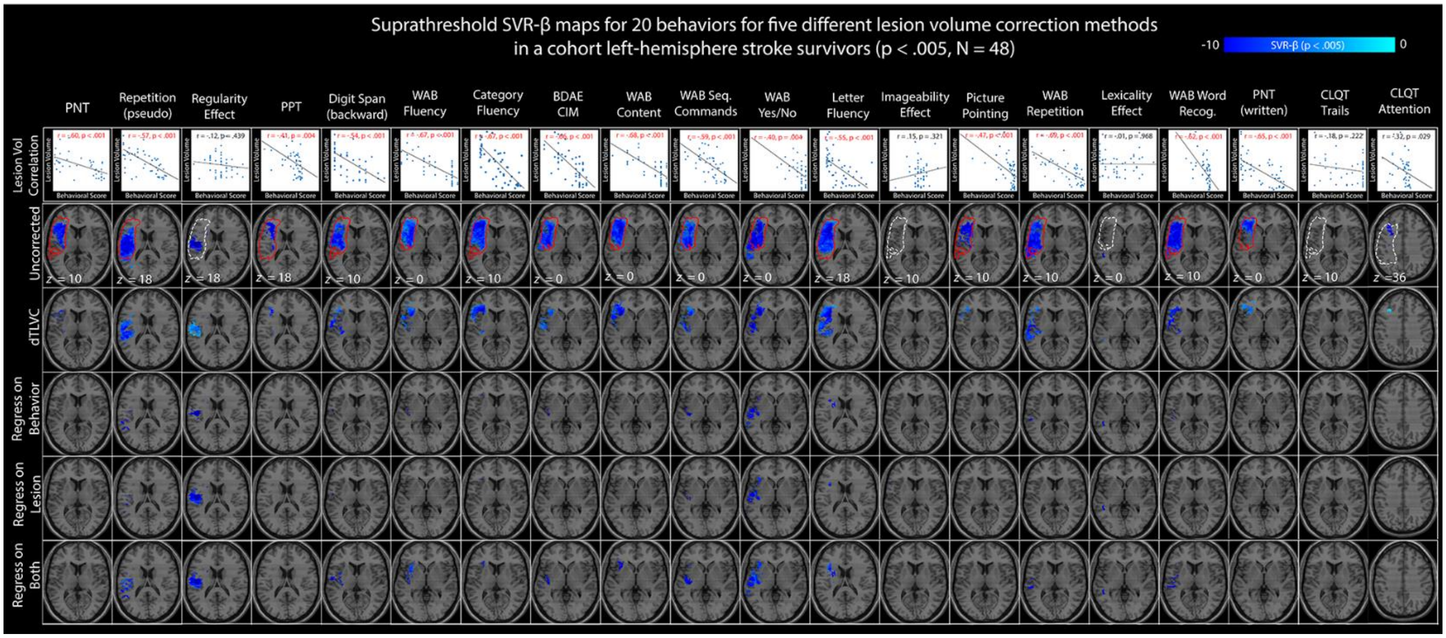
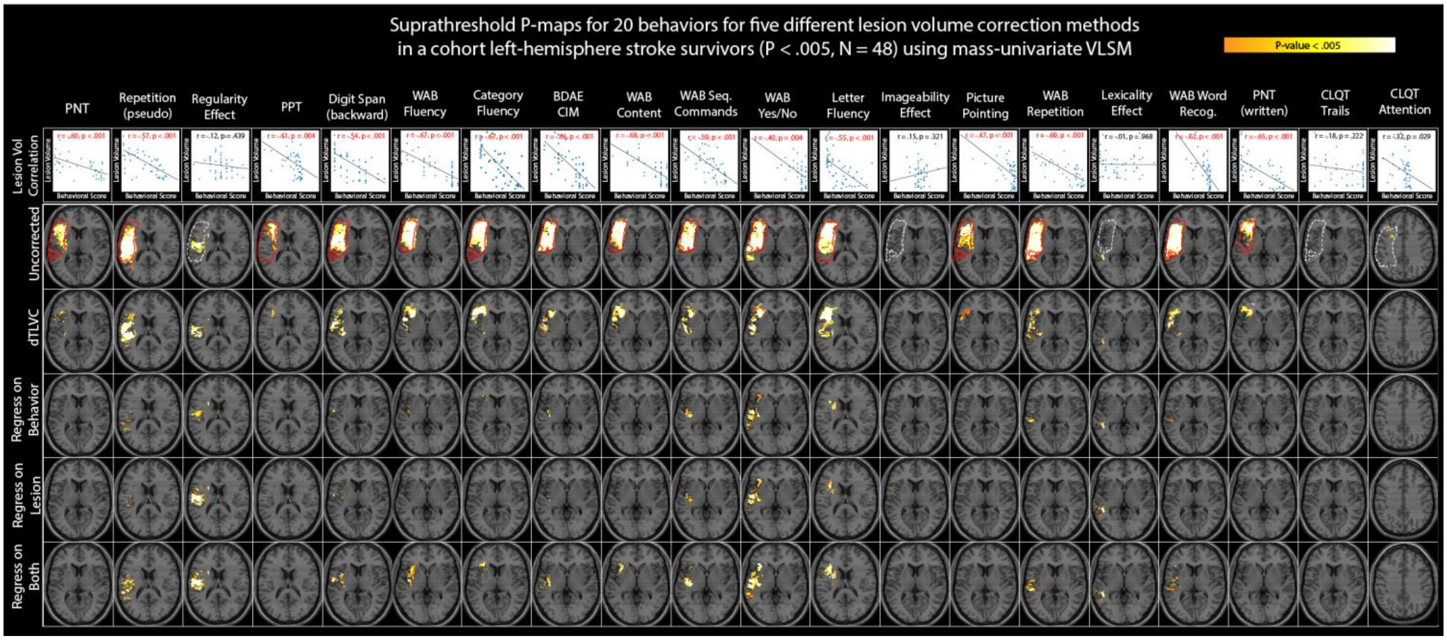
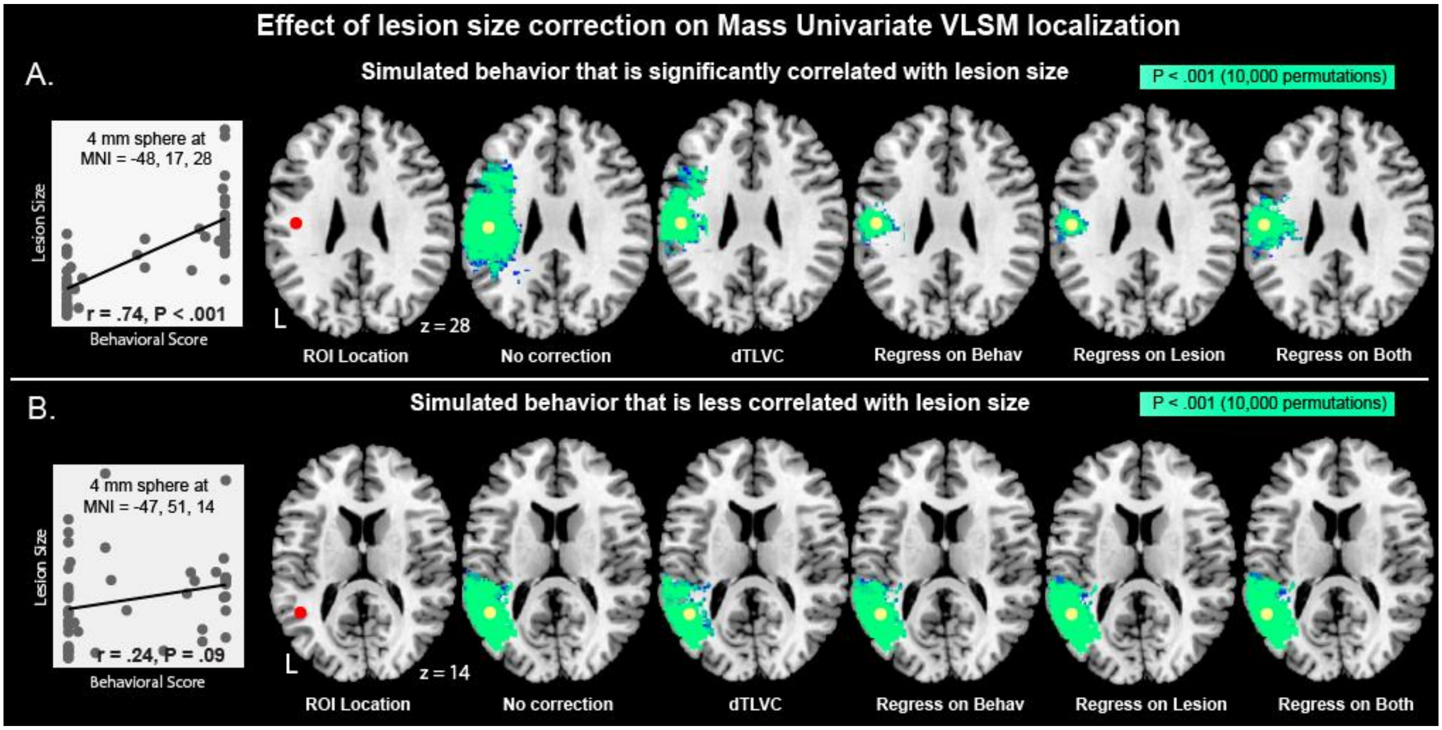


Supplemental Figure 1. This figure plots the correlation between behavioral performance and lesion volume for 20 behaviors measured in a cohort of left-hemisphere stroke survivors ($N = 48$), with significant correlations in red. Below each plotted correlation, thumbnails are shown of SVR-LSM results using five different lesion volume correction methods. Values displayed are suprathreshold SVR-β values thresholded voxelwise at $p < .005$. Columns correspond to behaviors and rows correspond to lesion volume correction methods. MNI coordinate in the *z* plane is shown in white for each column. The lesion volume correction methods displayed include no correction, dTLVC, regression of lesion volume out of behavior, regression of lesion volume out of voxelwise lesion statuses, and regression of lesion volume out of both behavior and lesion data. The top row of lesion-symptom maps, which are uncorrected for lesion volume, also depicts regions from **Figure 3d** for behaviors significantly correlated ($p < .05$) with lesion volume (red outline) and not significantly correlated with lesion volume (dotted white outline). Behaviors include Philadelphia Naming Test score (PNT), pseudoword repetition, regularity effect, Pyramids and Palm Trees (PPT), backward digit span, Western Aphasia Battery (WAB) Fluency score, category fluency, score on complex ideational material (CIM) from the Boston Diagnostic Aphasia Examination (BDAE), WAB Content score, WAB Sequential Commands score, WAB Yes/No score, letter fluency, imageability effect, picture pointing score, WAB Repetition score, lexicality effect, WAB Word Recognition score, written PNT score, and the Trails and Attention subscores on the Cognitive Linguistics Quick Test (CLQT).



Supplemental Figure 2. This figure replicates the results from **Supplemental Figure 1**, which used SVR-LSM, with traditional mass-univariate lesion-symptom mapping. This figure demonstrates that the concern about lesion volume as a confound is not limited to multivariate lesion-symptom mapping. In this figure, P values are derived from beta values from voxelwise linear regression (see **Supplemental Methods**). Statistical maps are corrected at voxelwise *p* < .005, based on 10,000 permutations. Each of the 20 columns corresponds to one behavior and each and five rows correspond to a lesion volume correction methods, as in **Supplemental Figure 1**.



Supplemental Figure 3. This figure replicates the results from **Figure 6**, which used SVR-LSM, with traditional mass-univariate lesion-symptom mapping methods. Effect of lesion volume correction on mass-univariate VLSM localization for (A) a simulated behavior significantly correlated with lesion volume and (B) a simulated behavior less correlated with lesion volume. Behaviors were simulated by summing the number of lesioned voxels within a 4mm sphere (red) placed based on its correlation with lesion volume. At the left of each pane is shown a scatterplot of lesion volume versus the simulated behavior across the patient sample. Voxels with statistically significant results at voxelwise $p < .001$ based on 10,000 permutations are shown in bright green.

Behavior under investigation	Lesion volume correction method				
	No correction	dTLVC	Regress on Lesion	Regress on Behavior	Regress on Both
Philadelphia Naming Test (Oral)	•	•	○	○	○
Philadelphia Naming Test (Written)	•	•	○	○	○
Category Fluency	•	•	○	○	•
Letter Fluency	•	•	○	○	•
Western Aphasia Battery Content	•	•	○	○	•
Western Aphasia Battery Fluency	•	•	○	○	•
Backward Digit Span	•	•	○	○	•
Western Aphasia Battery Repetition	•	•	○	○	•
Pseudoword Repetition	•	•	○	○	•
Concreteness Effect (concrete minus abstract)		○	○	○	○
Regularity Effect (regular minus exception)	•	•	•	○	•
Lexicality Effect (word minus pseudoword)	○	○	○	○	○
Pyramids and Palm Trees	•	○	○	○	○
Western Aphasia Battery Word Recognition	•	•	○	○	○
Western Aphasia Battery Yes/No	•	•	•	•	•
Picture Pointing	•	•	○	○	○
BDAE Complex Ideational Material	•	•	○	○	○
CLQT Attention Score	○	○		○	○
CLQT Symbol Trails Score	○	○	○	○	○
Western Aphasia Battery Sequential Commands	•	•	○	○	•

Supplemental Table 1. Table listing which of 20 behaviors were significant for each of five lesion volume correction methods using traditional, mass-univariate lesion-symptom mapping. A solid bullet (•) indicates that significant findings for a given combination of behavior and lesion volume correction method which survived voxelwise thresholding, $p < .005$ (one-tailed) and clusterwise correction, $p < .05$ based on 10,000 permutations. An empty bullet (○) indicates that the analysis resulted in voxels that survived voxelwise thresholding but did not survive cluster correction.

Supplemental Methods

Mass-univariate voxel-based lesion-symptom mapping was conducted to replicate some of our multivariate analyses in a traditional framework. For these analyses, lesion volume correction was applied using a nuisance model, just as with the multivariate analyses. The SVR-LSM toolbox was modified so that instead of estimating a brain-wide multivariate SVR models relating lesion location to a behavior in question, each voxel in the brain was submitted to a linear model predicting the values at that voxel across the patient sample from the behaviors across the patient sample. The beta coefficients from each linear model accumulated in a whole-brain image, which was retained for permutation testing. Each permutation was modeled in this way. All resulting statistic images were thresholded voxelwise at $p < .005$ or $p < .001$ (one-tailed) and by cluster extent at $p < .05$, based on 10,000 permutations (voxelwise and clusterwise thresholds applied to match the multivariate procedures).