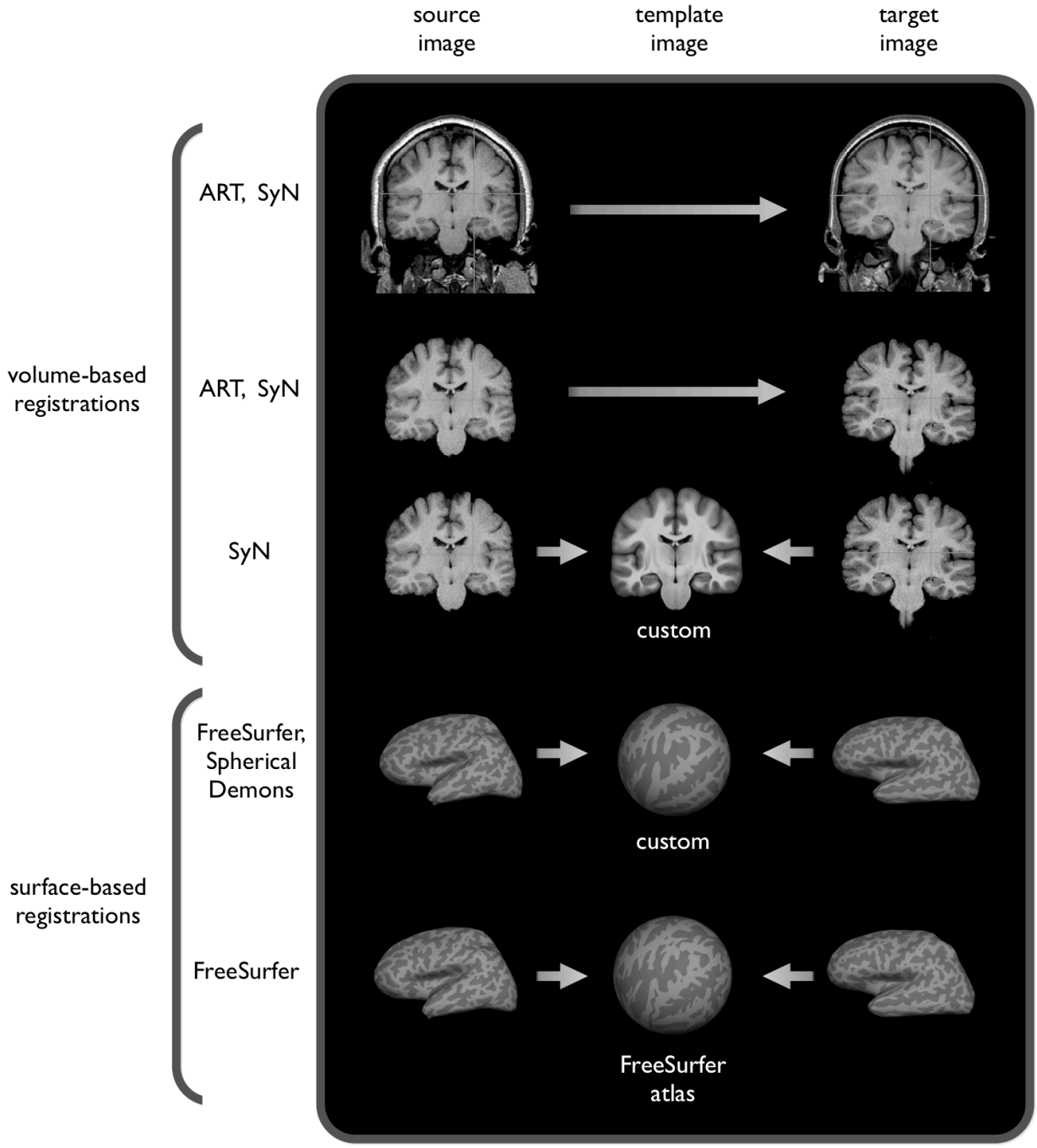
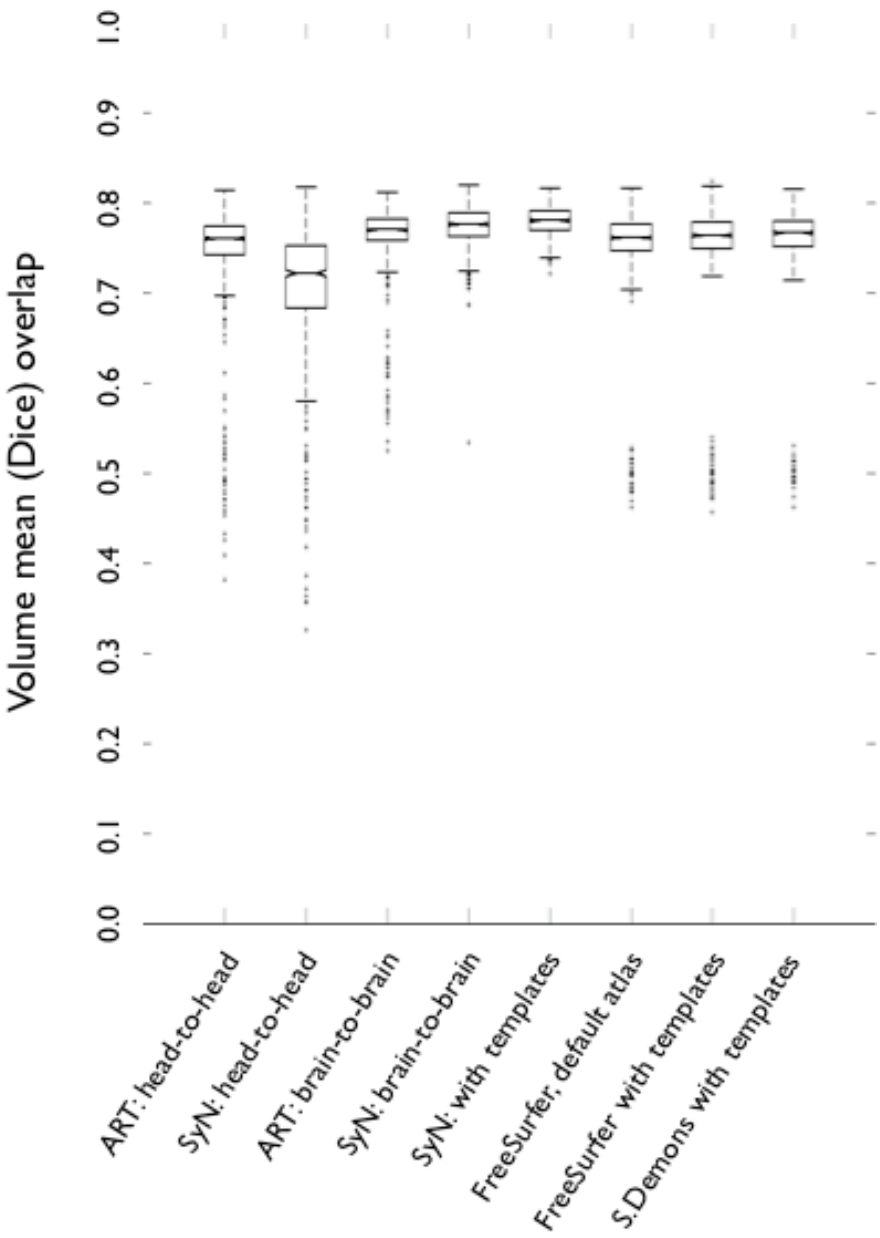
Supplementary Material for
"Evaluation of volume-based and surface-based brain image registration methods"
Arno Klein, Satrajit S. Ghosh, Brian Avants, B.T.T. Yeo, Bruce Fischl, Babak Ardekani,
James C. Gee, J.J. Mann, Ramin V. Parsey. (2010): http://www.mindboggle.info/papers

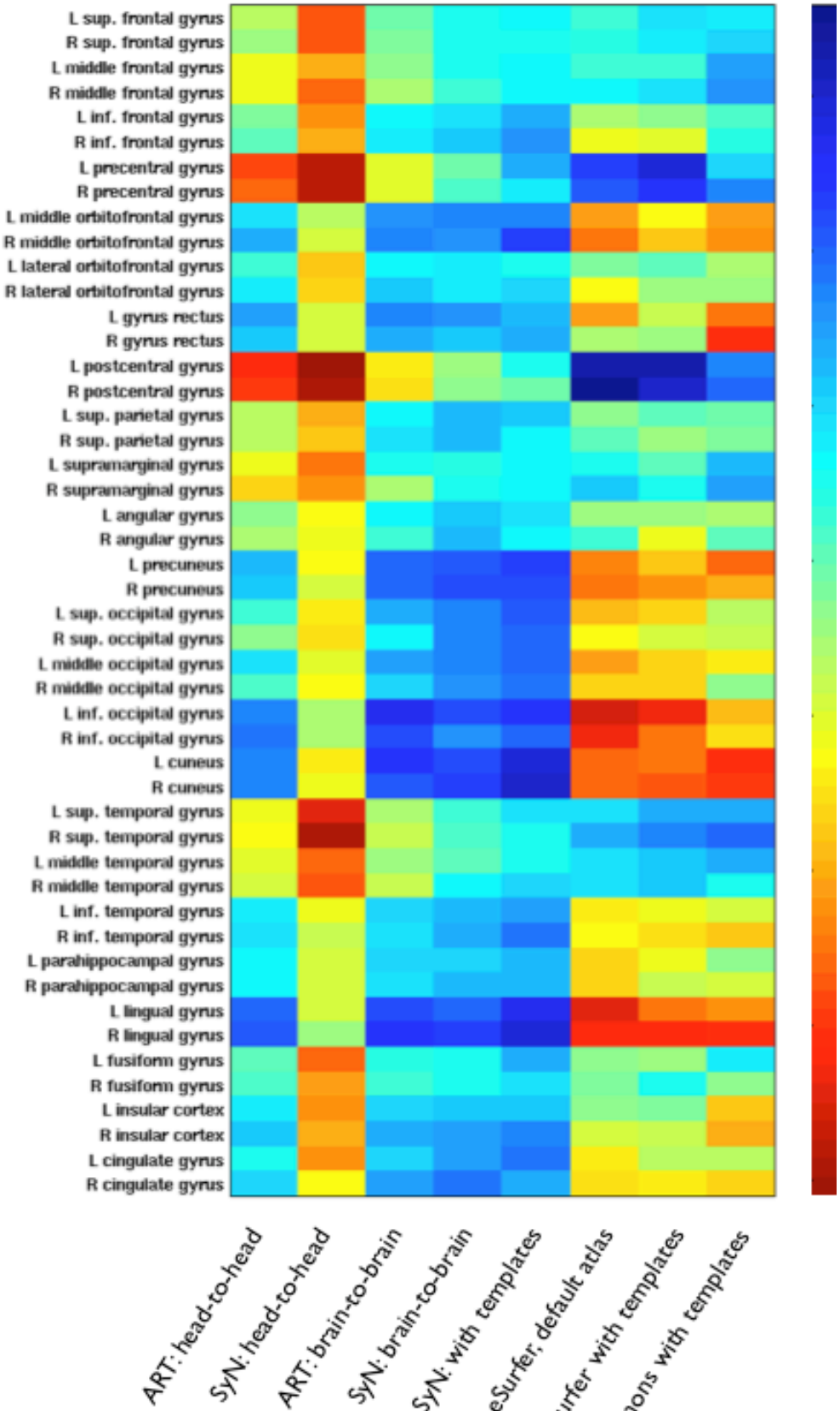


Supplementary Figure 1. Volume-based and surface-based registration methods
This schematic diagram shows an overview of the different registration methods in this study. The top half consist of volume-based registration methods (ART and SyN) and the bottom half of surface-based registration methods (FreeSurfer and Spherical Demons). The left column shows the source image, the right column shows the target image, and the middle column shows the intermediate registration template, if one was used. Each of the custom templates was created using the same registration method that was evaluated using that template (for example, FreeSurfer was used to create the custom template for use with FreeSurfer).



Supplementary Figure 2. Volume overlaps averaged across all regions (Test 1 with LPBA40 data) These box and whisker plots show the volume (Dice) overlap values between warped source label volumes and target label volumes, averaged first across all 56 regions then across all 760 registration pairs. Each target label volume was either the LPBA40 target (left graph) or the FreeSurfer-generated target mask (right graph). Registrations were performed using volume-based methods and surface-based methods. The performance of each registration method is represented by a box and has lines at the lower quartile, median, and upper quartile values; whiskers extend from each end of the box to the most extreme values within 1.5 times the interquartile range from the box. Outliers (+) have values beyond the ends of the whiskers; almost all of the outliers can be attributed to one brain image when it was used as a registration source. The volume-based and surface-based results should be considered separately, because for Test 1, the surface-based methods underwent a double resampling to take volume labels, register them as surfaces, and evaluate them again as volumes (see text and Figure 5). The surface-based methods are included to demonstrate the attempt made to compare volume- and surface-based registration methods using different target masks.

LPBA40 labels



Supplementary Figure 3. Test 1 indifference-zone ranking of the registration methods by region
This figure uses a color scale that reflects the relative performance of the registration methods (with blue indicating higher accuracy). Each colored rectangle represents the average score for a given method (column) for a given region (row), averaged over 760 registrations. The scores indicate the overlap accuracy of the method relative to each of the other methods, and were determined by indifference-zone ranking, which tests practical significance rather than statistical significance. For example, if the overlap is calculated for two different registration methods and these two values are within delta of one another (referred to as an "indifference zone" (Bechhofer 1954)), they are considered equal. The delta must correspond to a practical difference in registration. If we model a region as a cube, then a single-voxel offset along the normal to one of its faces means an offset of one-sixth of its surface. We therefore set delta to one-sixth of a target region's surface, assuming the surface to be that of a cube ($6\times edge^2-12\times edge$, where *edge* = the edge length of a cube with the volume of the target region).

LPBA40 labels	SyN overlap	FreeSurfer overlap	S.Demons overlap	Δoverlap	Resampling error non-overlap
L sup. frontal gyrus	79.55	81.45	81.26	1.90	5.31
R sup. frontal gyrus	82.56	81.28	81.34	1.22	6.10
L middle frontal gyrus	81.35	78.83	79.96	1.39	4.47
R middle frontal gyrus	78.65	80.91	81.90	3.25	4.12
L inf. frontal gyrus	79.91	73.01	73.99	5.92	5.42
R inf. frontal gyrus	76.65	71.42	73.24	3.41	5.16
L precentral gyrus	80.60	76.87	73.26	3.73	7.12
R precentral gyrus	71.64	74.19	72.45	2.55	7.31
L middle orbitofrontal gyrus	79.65	75.38	73.32	4.27	6.68
R middle orbitofrontal gyrus	73.78	74.19	73.23	0.41	6.87
L lateral orbitofrontal gyrus	62.28	58.12	57.36	4.16	11.27
R lateral orbitofrontal gyrus	56.32	52.27	52.65	3.67	11.19
L gyrus rectus	41.87	58.73	53.61	16.86	25.27
R gyrus rectus	34.89	58.51	51.21	23.62	31.01
L postcentral gyrus	74.51	75.63	72.21	1.12	8.77
R postcentral gyrus	64.87	74.04	71.58	9.17	9.37
L sup. parietal gyrus	68.75	68.25	67.98	0.50	9.40
R sup. parietal gyrus	59.02	67.74	68.04	9.02	8.91
L supramarginal gyrus	72.12	67.93	69.74	2.38	7.06
R supramarginal gyrus	69.69	67.95	69.50	0.19	6.99
L angular gyrus	63.17	59.68	59.77	3.40	9.85
R angular gyrus	64.69	61.52	63.62	1.07	7.74
L precuneus	56.99	54.35	52.48	2.64	12.30
R precuneus	68.31	53.25	53.35	14.96	14.67
L sup. occipital gyrus	60.27	53.01	55.04	5.23	11.53
R sup. occipital gyrus	52.05	51.64	52.31	0.26	10.87
L middle occipital gyrus	72.32	67.68	68.35	3.97	7.36
R middle occipital gyrus	70.51	66.43	68.90	1.61	7.27
L inf. occipital gyrus	70.31	55.32	60.50	9.81	6.68
R inf. occipital gyrus	70.88	61.17	64.48	6.40	6.66
L cuneus	46.20	44.68	42.25	1.52	20.01
R cuneus	59.71	45.49	44.93	14.22	23.3
L sup. temporal gyrus	86.69	82.05	81.83	4.64	4.23
R sup. temporal gyrus	84.92	83.67	83.88	1.04	4.20
L middle temporal gyrus	78.15	72.35	72.58	5.57	6.00
R middle temporal gyrus	78.22	73.92	73.54	4.30	5.58
L inf. temporal gyrus	68.28	61.65	63.31	4.97	7.43
R inf. temporal gyrus	68.22	63.70	62.91	4.52	9.55
L parahippocampal gyrus	76.04	72.15	72.99	3.05	9.50
R parahippocampal gyrus	74.79	70.95	70.36	3.84	9.53
L lingual gyrus	68.91	63.16	63.93	4.98	8.92
R lingual gyrus	71.83	61.99	62.09	9.74	10.95
L fusiform gyrus	75.40	71.14	72.89	2.51	7.32
R fusiform gyrus	73.18	72.25	70.37	0.93	7.55
L insular cortex	89.55	86.88	83.30	2.67	7.87
R insular cortex	89.37	83.96	81.86	5.41	8.18
L cingulate gyrus	68.83	66.73	66.35	2.10	12.43

Supplementary Table 1: Test 1 volume overlap results and resampling test results (LPBA40 data)
For each LPBA40 label (left column), the results for Test 1 (V-V) (volume labels and evaluation) are listed for SyN, FreeSurfer, and Spherical Demons as (percent) overlap averaged across 760 custom-template-based registrations. To the right of these columns is Δoverlap, the maximum absolute difference between SyN overlap and the overlap of the surface registration methods. The rightmost column lists the resampling test results (LPBA40 volume labels resampled to surface then back to volume and compared against the original volume labels) as non-overlap (100 - overlap). For every label, Δoverlap is less than non-overlap (by at least one standard deviation of the resampling value), which means that the resampling error is too great here to distinguish between the performance of volume and surface registration methods.

FS40 labels	SyN		FreeSurfer		S.Demons				Resampling error	
	overlap	distance	overlap	distance	overlap	distance	Δoverlap	Δdistance	non-overlap	distance
Sup. temporal sulcus	67.02	3.98	65.59	4.03	69.34	3.73	2.32	0.25	24.11	3.02
Caudal ant. cingulate	68.62	3.36	82.61	1.80	77.25	2.33	13.99	1.56	22.87	2.74
Caudal middle frontal	63.12	6.72	78.10	4.26	73.26	5.60	14.98	2.46	13.90	3.12
Cuneus	62.28	5.22	77.83	3.79	71.79	4.80	15.55	1.43	16.01	2.97
Entorhinal	67.53	2.52	68.49	2.75	65.30	2.82	2.23	0.30	24.03	2.03
Fusiform	76.76	4.41	80.93	3.62	79.77	3.99	4.17	0.79	15.59	3.22
Inferior parietal	73.07	7.68	82.48	5.79	83.45	5.54	10.38	2.14	11.24	4.36
Inferior temporal	71.23	4.96	82.74	3.37	80.54	3.77	11.51	1.60	14.19	2.85
Isthmus cingulate	75.44	2.11	80.98	1.55	78.10	1.85	5.54	0.56	20.76	1.87
Lateral occipital	69.51	7.69	81.13	5.78	82.04	5.55	12.53	2.14	14.48	4.92
Lateral orbitofrontal	87.77	2.88	91.32	2.00	87.72	2.71	3.55	0.88	10.33	2.68
Lingual	76.14	5.32	88.22	2.92	85.75	3.50	12.08	2.40	10.97	2.93
Medial orbitofrontal	82.29	3.43	89.78	1.94	85.04	2.91	7.49	1.48	11.46	2.67
Middle temporal	75.28	4.54	83.23	3.23	80.68	3.63	7.95	1.31	13.17	2.56
Parahippocampal	77.11	2.28	78.65	2.25	78.86	2.23	1.75	0.05	15.81	1.67
Paracentral	78.99	3.86	89.03	2.11	83.39	3.13	10.04	1.75	9.54	1.90
Pars opercularis	68.69	4.48	69.76	4.54	64.66	5.13	4.03	0.65	20.12	3.06
Pars orbitalis	72.16	3.29	76.74	2.97	75.11	3.28	4.58	0.31	20.06	2.69
Pars triangularis	66.04	5.06	74.38	4.25	76.11	4.11	10.07	0.95	20.32	3.40
Pericalcarine	52.54	5.72	79.15	2.65	69.03	3.94	***26.61**	***3.06**	14.28	2.04
Postcentral	75.19	5.39	85.67	2.94	78.39	4.67	10.48	2.45	9.70	2.26
Posterior cingulate	78.49	2.89	85.56	1.97	83.81	2.20	7.07	0.92	18.26	2.67
Precentral	77.73	4.94	86.75	3.17	81.13	4.74	9.02	1.77	9.18	2.59
Precuneus	79.17	4.81	85.99	3.28	85.11	3.60	6.82	1.53	9.67	2.64
Rostral ant. cingulate	64.97	3.20	78.94	1.93	65.39	3.15	13.97	1.27	20.45	1.92
Rostral middle frontal	71.08	7.51	80.36	5.86	81.90	5.74	10.82	1.77	13.66	4.95
Sup. frontal	82.98	6.32	90.35	3.59	88.19	4.46	7.37	2.73	7.53	3.03
Sup. parietal	68.40	7.09	86.41	3.78	83.24	4.80	***18.01**	3.31	10.52	3.09
Sup. temporal	88.19	2.85	91.24	2.02	88.27	2.72	3.05	0.83	7.78	2.07
Supramarginal	74.73	6.00	82.65	4.50	82.61	4.68	7.92	1.51	10.96	3.13
Frontal pole	49.69	3.94	66.53	2.54	60.53	2.97	16.84	1.40	33.71	2.35
Temporal pole	74.95	2.34	78.97	1.91	73.78	2.39	4.02	0.43	20.25	1.95
Transverse temporal	69.65	2.21	83.29	1.31	70.67	2.34	13.64	0.90	20.11	1.59
Insula	88.29	2.36	91.76	1.59	88.84	2.16	3.47	0.77	8.74	2.12

Supplementary Table 2: Test 2 surface overlap, distance results and resampling test results (FS40 data) For each FS40 label (left column, averaged across left and right hemispheres), the results for Test 2 (S-S) (surface labels and evaluation) are listed for SyN, FreeSurfer, and Spherical Demons as (percent) overlap averaged across 760 custom-template-based registrations. To the right of these columns is Δoverlap and Δdistance, the maximum absolute difference between overlap or distance calculated for SyN and for the surface registration methods. The rightmost columns list the resampling test results (FS40 surface labels resampled to volume then back to surface and compared against the original surface labels) as non-overlap (100 - overlap). For almost every label, Δoverlap and Δdistance are less than the resampling non-overlap and distance, which means that the resampling error is too great for these regions to distinguish between the performance of volume and surface registration methods. *****An asterisk is placed next to the only exceptions, the pericalcarine and superior parietal regions, whose values are greater by at least one standard deviation of the resampling values.

FS40 labels	SyN		FreeSurfer		S.Demons	
	overlap μ	SD	overlap μ	SD	overlap μ	SD
Sup. temporal sulcus	50.12	11.71	60.89	13.41	60.02	12.42
Caudal ant. cingulate	65.58	8.11	85.05	5.26	81.84	6.08
Caudal middle frontal	63.93	9.18	79.5	10.94	79.84	7.51
Cuneus	56.33	6.91	75.94	8.19	71.9	9.31
Entorhinal	67.23	6.96	77.09	7.08	73.56	6.19
Fusiform	65.59	7.58	71.76	9.66	70.63	8.95
Inferior parietal	65.38	5.49	81.74	4.8	80.68	4.81
Inferior temporal	67.15	4.9	83.14	4.99	83.9	4.43
Isthmus cingulate	72.45	5.02	83.96	4.13	82.08	4.47
Lateral occipital	64.59	4.73	85.77	6.01	82.77	4.42
Lateral orbitofrontal	80.41	2.5	84.12	4.44	84.22	3.16
Lingual	67.82	7.5	90.7	2.05	87.13	2.81
Medial orbitofrontal	77.07	4.12	89.13	3.36	88.02	2.91
Middle temporal	70.89	3.95	89.57	3.13	85.62	3.32
Parahippocampal	74.58	5.57	85.26	3.5	83.12	3.48
Paracentral	70	6.12	79.34	7.03	80.38	7.76
Pars opercularis	62.07	10.48	85.47	6.39	83.98	4.75
Pars orbitalis	64.92	6.42	73.69	12.9	73.44	11.22
Pars triangularis	60.99	7.91	76.11	7.39	74.15	6.77
Pericalcarine	44.83	8.48	74.07	7.77	75.65	7.75
Postcentral	66.8	9.06	78.87	5.73	69.38	7.77
Posterior cingulate	74.45	4.02	82.58	14.45	83.11	7.45
Precentral	70.94	8.73	88.12	3.99	86.6	3.11
Precuneus	73.43	3.95	86.48	13.98	87.13	6.13
Rostral ant. cingulate	68	8.34	87.88	3.64	87.67	2.64
Rostral middle frontal	69.1	4.64	82.58	7.3	73.47	8.58
Sup. frontal	79.35	3.36	80.54	5	82.68	4.32
Sup. parietal	63.25	5.06	90.31	2.81	90.19	2.47
Sup. temporal	81	2.66	84.88	4.85	85.11	3.08
Supramarginal	68.58	5.32	91.44	2.12	89.74	1.7
Frontal pole	54.58	11.96	84.16	5.33	84.93	4.04
Temporal pole	70.23	6.6	71.27	8.63	64.34	8.98
Transverse temporal	70.2	7.02	79.18	5.12	76.07	6.17

Supplementary Table 3: Test 3 volume overlap results (FS40 data)
For each FS40 label (left column, averaged across left and right hemispheres), the results for Test 3 (surface labels, volume evaluation) are listed for SyN, FreeSurfer, and Spherical Demons as percent overlap averaged across 760 custom-template-based registrations.

LPBA40 labels	SyN		FreeSurfer		S.Demons	
	overlap μ	SD	overlap μ	SD	overlap μ	SD
L sup. frontal gyrus	79.55	11.88	81.45	14.83	81.26	14.72
R sup. frontal gyrus	82.56	10.38	81.28	13.98	81.34	13.88
L middle frontal gyrus	81.35	8.9	78.83	14.4	79.96	13.64
R middle frontal gyrus	78.65	6.73	80.91	7.72	81.9	6.12
L inf. frontal gyrus	79.91	5.47	73.01	7.18	73.99	6.48
R inf. frontal gyrus	76.65	6.79	71.42	7.49	73.24	6.67
L precentral gyrus	80.6	6.56	76.87	12.34	73.26	13.61
R precentral gyrus	71.64	8.19	74.19	11.03	72.45	11.53
middle orbitofrontal gyrus	79.65	7.9	75.38	8.45	73.32	9.16
middle orbitofrontal gyrus	73.78	9.41	74.19	8.78	73.23	9.29
lateral orbitofrontal gyrus	62.28	14.17	58.12	14.66	57.36	14.63
lateral orbitofrontal gyrus	56.32	16.38	52.27	16.56	52.65	15.77
L gyrus rectus	41.87	18.97	58.73	15.21	53.61	17.23
R gyrus rectus	34.89	15.34	58.51	18.04	51.21	18.01
L postcentral gyrus	74.51	8.24	75.63	13.54	72.21	14.27
R postcentral gyrus	64.87	11.01	74.04	12.34	71.58	12.97
L sup. parietal gyrus	68.75	13.91	68.25	17.57	67.98	17.23
R sup. parietal gyrus	59.02	12.68	67.74	16.5	68.04	16.41
L supramarginal gyrus	72.12	9.74	67.93	10.28	69.74	9.9
R supramarginal gyrus	69.69	10.22	67.95	11.01	69.5	10.82
L angular gyrus	63.17	15.88	59.68	16.51	59.77	16.25
R angular gyrus	64.69	12.44	61.52	13.14	63.62	12
L precuneus	56.99	16.41	54.35	19.56	52.48	19.47
R precuneus	68.31	14.06	53.25	21	53.35	21.64
L sup. occipital gyrus	60.27	13.71	53.01	13.78	55.04	13.37
R sup. occipital gyrus	52.05	13.83	51.64	14.77	52.31	14.66
L middle occipital gyrus	72.32	8.41	67.68	8.14	68.35	8.1
R middle occipital gyrus	70.51	9.22	66.43	9.31	68.9	8.28
L inf. occipital gyrus	70.31	9.86	55.32	12.5	60.5	12.58
R inf. occipital gyrus	70.88	8.75	61.17	10.81	64.48	9.72
L cuneus	46.2	19.84	44.68	20.74	42.25	19.08
R cuneus	59.71	18.8	45.49	16.56	44.93	16.19
L sup. temporal gyrus	86.69	5.26	82.05	6.62	81.83	6.47
R sup. temporal gyrus	84.92	4.58	83.67	5.07	83.88	4.61
L middle temporal gyrus	78.15	6.58	72.35	7.64	72.58	7.17
R middle temporal gyrus	78.22	6.22	73.92	6.43	73.54	6.23
L inf. temporal gyrus	68.28	9.23	61.65	11.32	63.31	9
R inf. temporal gyrus	68.22	11.87	63.7	11.1	62.91	10.99
L parahippocampal gyrus	76.04	7.74	72.15	7.35	72.99	7.59
R parahippocampal gyrus	74.79	8.07	70.95	7.72	70.36	7.36
L lingual gyrus	68.91	10.19	63.16	11.28	63.93	10.89
R lingual gyrus	71.83	10.07	61.99	10.84	62.09	10.89
L fusiform gyrus	75.4	7.21	71.14	8.4	72.89	7.26
R fusiform gyrus	73.18	8.18	72.25	8.29	70.37	8.59
L insular cortex	89.55	3.17	86.88	3.04	83.3	6.03
R insular cortex	89.37	4.36	83.96	4.87	81.86	6.29
L cingulate gyrus	68.83	12.16	66.73	15.42	66.35	15.25

Supplementary Table 4: Test 4 surface overlap results (LPBA40 data)
For each LPBA40 label (left column), the results for Test 4 (volume labels, surface evaluation) are listed for SyN, FreeSurfer, and Spherical Demons as percent overlap averaged across 760 custom-template-based registrations.

FS40 subjects	hemisphere	FWHM
1 – 20	left	2.94
1 – 20	right	2.91
21 – 40	left	2.97
21 – 40	right	3.03
1 – 40	left	3.01
1 – 40	right	3.04

Supplementary Table 5: FreeSurfer template smoothness test (FS40 data)
Smoothness was measured on the sphere for each hemisphere of each of the custom FreeSurfer templates (FS40 subjects 1-20 and 21-40) and for all (1-40) FS40 subjects, corresponding to the default FreeSurfer template. After sampling the curvature files for a group of subjects to an atlas space (FreeSurfer command: “mris_preproc --target fsaverage --meas curv...”), the FWHM was computed on the spherical surface (using FreeSurfer's “mris_fwhm” command).