

Figure 1: Results of the correlation between each subject's RDM and each networks' layer RDMs for two architectures: DeepCluster VGG16 modified for unsupervised learning and standard VGG16. The dashed lines represent the correlations with the random networks. The solid lines represents the correlations with the trained networks. The grey band represents the noise ceiling for the MRI data. Asterisks show in which layers one version of the network correlated with IT better than the other. a) Results for DeepCluster VGG16. b) Results for standard VGG16

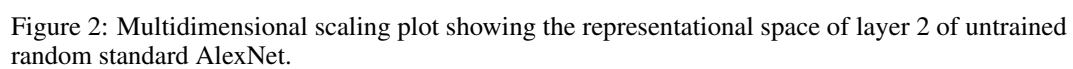


Table 1: Mediation model testing whether the relation between layer 2 of random DeepCluster and IT is mediated by semantic or perceptual features.

| path | coef | se | p value | CI[2.5%] | CI[97.5%] | Sig |
|----------------------|-----------|----------|---------|-----------|-----------|-----|
| category ~ X | -0.20 | 0.76 | 0.79 | -1.70 | 1.29 | No |
| animacy ~ X | 3.69 | 1.00 | 0.0002 | 1.73 | 5.66 | Yes |
| size ~ X | 1.29e+05 | 8.29+03 | 0.000 | 1.13e+05 | 1.45e+05 | Yes |
| thinness ~ X | 3.58e+01 | 0.96 | 0.000 | 3.39e+01 | 3.77e+01 | Yes |
| orientation ~ X | 1.68e+01 | 1.01 | 0.000 | 1.48e+01 | 1.88e+01 | Yes |
| silhouette ~ X | -1.05e+01 | 3.91e-01 | 0.000 | -1.13e+01 | -9.77 | Yes |
| contrast ~ X | 1.17e+05 | 3.52e+04 | 0.001 | 4.78e+04 | 1.86e+05 | Yes |
| hue ~ X | 3.13e+02 | 6.38e+01 | 0.000 | -4.38e+02 | -1.88e+02 | Yes |
| luridness ~ X | 2.21e+02 | 2.49e+01 | 0.000 | 1.72e+02 | 2.70e+02 | Yes |
| Y ~ category | 0.01 | 0.003 | 0.000 | 0.009 | 0.02 | Yes |
| Y ~ animacy | 0.06 | 0.002 | 0.000 | 0.05 | 0.006 | Yes |
| Y ~ size | -5.59e-07 | 2.54e-07 | 0.03 | -1.06e-06 | -6.14e-08 | Yes |
| Y ~ thinness | 0.03 | 0.002 | 0.000 | 0.02 | 0.03 | Yes |
| Y ~ orientation | 0.004 | 0.002 | 0.09 | -5.58e-04 | 0.008 | No |
| Y ~ silhouette | -0.07 | 0.006 | 0.000 | -0.08 | -0.06 | Yes |
| Y ~ contrast | 2.12e-07 | 5.02e-08 | 0.000 | 1.14e-07 | 3.11e-07 | Yes |
| Y ~ hue | 5.68e-05 | 2.78e-05 | 0.04 | 2.19e-06 | 1.11e-04 | Yes |
| Y ~ luridness | 3.83e-04 | 7.08e-05 | 0.000 | 2.45e-04 | 5.22e-04 | Yes |
| Total | 3.52 | 0.13 | 0.000 | 3.26 | 3.77 | Yes |
| Direct | 2.43 | 0.14 | 0.000 | 2.16 | 2.70 | Yes |
| Indirect category | -0.003 | 0.01 | 0.8 | -0.03 | 0.02 | No |
| Indirect animacy | 0.20 | 0.057 | 0.000 | 0.10 | 0.31 | Yes |
| Indirect size | -0.02 | 0.03 | 0.5 | -0.08 | 0.03 | No |
| Indirect thinness | 0.4 | 0.07 | 0.000 | 0.2 | 0.5 | Yes |
| Indirect orientation | 0.003 | 0.04 | 0.9 | -0.06 | 0.07 | No |
| Indirect silhouette | 0.50 | 0.65 | 0.000 | 0.38 | 0.62 | Yes |
| Indirect contrast | 0.02 | 0.008 | 0.000 | 0.006 | 0.03 | Yes |
| Indirect hue | -0.03 | 0.01 | 0.004 | -0.056 | -0.02 | Yes |
| Indirect luridness | 0.05 | 0.02 | 0.000 | 0.02 | 0.09 | Yes |

Table 2: Mediation model testing whether the relation between layer 7 of trained standard AlexNet and IT is mediated by semantic or perceptual features.

| path | coef | se | p value | CI[2.5%] | CI[97.5%] | Sig |
|----------------------|-----------|----------|---------|-----------|-----------|-----|
| category ~ X | 0.99 | 0.04 | 0.000 | 0.91 | 1.08 | Yes |
| animacy ~ X | 1.13 | 0.057 | 0.000 | 1.01 | 1.25 | Yes |
| size ~ X | 6.67e+03 | 4.92e+02 | 0.000 | 5.69 | 7.62e+03 | Yes |
| thinness ~ X | 0.05e-02 | 0.07e-02 | 0.45 | -0.08e-02 | 0.18 | No |
| orientation ~ X | 1.06 | 0.06 | 0.000 | 0.94 | 1.17 | Yes |
| silhouette ~ X | -0.57 | 0.02 | 0.000 | -0.62 | -0.53 | Yes |
| contrast ~ X | -5.57 | 2.08 | 0.79 | -4.62e+03 | 3.51e+03 | No |
| hue ~ X | 2.85e+01 | 3.74 | 0.000 | 2.12 | 3.58e+01 | Yes |
| luridness ~ X | 1.30e+01 | 1.47 | 0.000 | 1.01e+01 | 1.59e+01 | Yes |
| Y ~ category | 0.01 | 0.003 | 0.000 | 0.01 | 0.02 | Yes |
| Y ~ animacy | 0.06 | 0.002 | 0.000 | 0.05 | 0.06 | Yes |
| Y ~ size | -5.59e-07 | 2.54e-07 | 0.03 | -1.06e-06 | -6.14e-08 | Yes |
| Y ~ thinness | 0.03 | 0.002 | 0.000 | 0.02 | 0.03 | Yes |
| Y ~ orientation | 0.004 | 0.002 | 0.09 | -5.58e-04 | 0.008 | No |
| Y ~ silhouette | -0.07 | 0.006 | 0.000 | -0.08 | -0.06 | Yes |
| Y ~ contrast | 2.12e-07 | 5.02e-08 | 0.000 | 1.14e-07 | 3.12e-07 | Yes |
| Y ~ hue | 5.68e-05 | 2.79e-05 | 0.04 | 2.19e-06 | 1.11e-04 | Yes |
| Y ~ luridness | 3.83e-04 | 7.08e-05 | 0.000 | 2.45e-04 | 5.22e-04 | Yes |
| Total | 0.20 | 0.008 | 0.000 | 0.18 | 0.21 | Yes |
| Direct | 0.11 | 0.008 | 0.000 | 0.09 | 0.1 | Yes |
| Indirect category | 0.005 | 0.003 | 0.08 | -7.04e-05 | 0.01 | No |
| Indirect animacy | 0.06 | 0.003 | 0.000 | 0.05 | 0.07 | Yes |
| Indirect size | -0.008 | 0.002 | 0.000 | -0.01 | -0.005 | Yes |
| Indirect thinness | 0.002 | 0.002 | 0.40 | -0.002 | 0.005 | No |
| Indirect orientation | 0.002 | 0.002 | 0.42 | -0.003 | 0.006 | No |
| Indirect silhouette | 0.03 | 0.004 | 0.000 | 0.02 | 0.04 | Yes |
| Indirect contrast | -1.13e-04 | 4.07e-04 | 0.78 | -9.22e-04 | 6.30e-04 | No |
| Indirect hue | 5.16e-04 | 7.38e-04 | 0.53 | -7.77e-04 | 0.002 | No |
| Indirect luridness | 0.004 | 0.001 | 0.000 | 0.002 | 0.006 | Yes |