Supporting Information

Kenett et al. 10.1073/pnas.1717362115

SI Network Construction

The semantic networks of the LSC and HSC groups were computed similar to the computational approach applied by Kenett et al. (1, 2). This approach consists of two parts. First, participants generate free association responses to cue words. The free association task is based on the method used in Rubinsten et al. (3), where participants are presented with a cue word and have 1 min to generate as many associative responses they could for that cue word. The LSC and HSC groups generated free associations to a list of 96 cue words. These cue words were based on fluency norms collected to a list of 36 categorical norms gathered by Henik and Kaplan (ref. 4; e.g., fruits, trees, and countries). The top four high-frequency words from each category were selected. These words were then tested for their degree of concreteness on a seven-point Likert scale, and only concrete words (average Likert score of three or higher) were selected. The final pool of cue words consisted of 96 words from 24 categories (Table S1). Thus, these four cue word per category are considered as the a priori components of the network.

The representations of the semantic networks of the LSC and HSC groups were based on a modified version of the method developed applied in ref. 1. Our revised method takes into account not only the correlation of associations, based on the overlap of associative features to a pair of cue words, but also the number of participants generating these overlapping associative features. The calculation of a link between two cue words (which represents the semantic similarity between them) is achieved in the following way. For each pair of cue words, we analyze only the associative responses generated to them (similar to the method used in ref. 1). For each of these matched associative responses, we sum the lower amount of participants generating them: $\mathrm{Link}(C_i, C_j) = \sum_{k=1}^{\mathrm{Associations}} \min(\#P\_C_iA(k), \#P\_C_jA(k))$, where Associations is the total number of associative responses given to cue words i and j, and $\#P\_C_{i/j}A(k)$ is the amount of participants in the sample generating the kth associative response to cue words i and j.

For example, for the pair of cue words tent and bus we examine the overlap of associative responses given to these two cue words (Table S2). A possible overlap of associative responses given to both the cue word tent and the cue word bus can be trip (given by 9 participants to tent and 22 participants to bus), big (given by 1 participant to tent and 5 participants to bus), and army (given by 2 participants to tent and 9 participants to bus). In the original method applied by ref. 1, Pearson's correlation is used to compute the correlation between these two cue words, based on the overlap of associative responses. In our revised method, each of the associative responses given to both cue words and the amount of participants generating these associative responses for both cue words is taken into account, in relation to all of the associative responses generated to each of the two cue words, to generate a link strength between the two cue words. Thus, our revised method takes into account the correlation of associations, as well as the amount of participants generating these overlapping associative features.

We compared the reliability of our revised method with the original method via three different analyses of the associative responses generated by both groups to the cue words tent and bus (Fig. S1). In the first analysis (original data), we simply compute the link correlations based on both approaches (Fig. S1*A*). In the second analysis (participant switching), we switch the number of participants generating the associative responses house, trip, and army. This switch was done so that the number of participants generating the same associative response is kept constant (Fig. S1*B*, highlighted in yellow). We expect that the link strength between the cue words bus and tent will remain the same after such a switch because the two cue words still have the same amount of similar associative responses. However, this switch has a strong effect on the calculation of the link strength between the cue words bus and tent based on the original method (0.64 vs. 0.49) but no effect on our revised method (Table S2). In the third case (participant amplification), we examined how increasing the amount of participants generating an associative response increases the strength of the link between two cue words. We multiplied by 5 the amount of participants generating the associative responses cockroach and game to the two cue words and computed the link strength with both methods (Fig. S1*C*, highlighted in yellow). We expect this manipulation to have a meaningful effect on the strength of the computed link. This analysis revealed that whereas the original method led to a small decrease (0.64 vs. 0.61), our revised method led to a meaningful increase (0.68 vs. 0.92) in the link strength (Table S2).

It is important to note that our approach is affected by the amount of responses, which can be either a result of the sample size or the amount of associations generated per cue word. In our study, the two groups were matched on sample size. To control for possible association fluency confounds (see also ref. 1), we compared the amount of associative responses generated to the cue words by both groups. This revealed that the HSC individuals generated a higher number of associative responses to the cue words (416.41, SD = 46.74) than LSC individuals (319.41, SD = 44.30). To control for this confound, we normalize the weights of the HSC network by a factor of 1.3 (the ratio between the average amount of responses of both groups) and compare the distribution of link strengths in both networks (Fig. S2).

To analytically examine the critical percolation threshold of the semantic networks, we compare the critical percolation threshold of both LSC and HSC networks to a random network with the same average degree. In a nonweighted random network, the average degree for percolation at criticality is one, $\langle k \rangle = 1$. We computed this theoretical threshold for each of the networks, which resulted in a threshold = 0.185 for both networks. As can be seen in Fig. 2*A*, the critical threshold in the empirical LSC network is about 0.153, and in the empirical HSC network is about 0.172. Thus, both empirical critical thresholds are below the theoretical critical threshold. This provides further analytical support for the effect of the community structure of the networks on their robustness.

SI Stability Analysis of Percolation Process

We examined the stability of the percolation analysis based on a correlation analysis. To conduct this analysis, we conducted 500 reiterations of the percolation analysis, where in each iteration a Gaussian noise was added to the network links with a mean value of zero and a varied standard variation, of 10^{-4} to 10^{-2}. This range is chosen to examine the effect of noise with one and even two orders of magnitude above the noise level expected if the weights of the links were randomly distributed (10^{-4}; *Materials and Methods*).

A percolation component number is defined as the percolation step number in which the component is disconnected from the giant component. To conduct the correlation analysis, we define a words vector, in which each cell represents a cue word (from 1 to 96) and contains the percolated component number of the word. The correlation analysis is conducted according to Pearson's correlation between the cue words vectors of the network before and after adding the noise. The correlation is calculated separately for each of the networks (LSC and HSC), for all amounts of noise added (Table S3). This analysis reveals significantly high correlation values for the different amounts of noise added, proving the stability of the percolation analysis.

SI Link Type Analysis

The link type analysis results in connectivity scores between each possible pair of components for both LSC (Table S4) and HSC (Table S5) networks separately and is used to construct Fig. 3 *A* and *B*. For this analysis, only components with more than three nodes were included.

SI LSC and HSC Percolated Components Overlap Analysis

We examined the matching between the percolated components of the two groups (Table S6). This is done by analyzing the matching nodes in a percolated component between both networks. For a specific percolated component in the LSC network we examine how much it overlaps with any percolated component in the HSC network. The components that were used for this analysis are only components with more than three nodes. Finally, we compute a matching percentage score for each percolated LSC component, which is calculated according to the number of overlapping nodes divided by the number of nodes composing that component. This analysis reveals that most of the components in LSC match a single or more components in the HSC network. The matching is about 87% (Table S7). For example, line 6 in Table S6 represents the sixth percolated component of the LSC network and its overlap with the HSC percolated components (Fig. 3). For this component, six nodes out of seven overlap between the components of the two groups; therefore, the overlap score for component 6 is 85%. As seen in the table, most of the components in LSC match a single or more components in the HSC network (87%). To conduct the matched connectivity link analysis, we merge small components in the HSC to match large components in the LSC network (Table S7).

1. Kenett YN, Anaki D, Faust M (2014) Investigating the structure of semantic networks in low and high creative persons. *Front Hum Neurosci* 8:407.
2. Kenett YN, Kenett DY, Ben-Jacob E, Faust M (2011) Global and local features of semantic networks: Evidence from the Hebrew mental lexicon. *PLoS One* 6:e23912.
3. Rubinsten O, Anaki D, Henik A, Drori S, Farn I (2005) Norms to free associations in Hebrew. *Word Norms in Hebrew*, eds Henik A, Rubinsten O, Anaki D (Ben-Gurion University, Beer-Sheva, Israel), pp 17–35.
4. Henik A, Kaplan L (2005) Content of categories: Findings regarding categories in Hebrew and comparison findings in the USA. *Word Norms in Hebrew*, eds Henik A, Rubinstein O, Anaki D (Ben-Gurion Univ, Beer-Sheva, Israel), pp 52–67.

Bus	Tent	Association
1	3	stone
4	14	house
1	2	mess
5	1	big
1	1	cockroach
1	7	rain
2	1	door
1	7	friend
2	1	soldier
5	1	window
9	22	trip
2	1	kid
1	5	fun
2	1	uncomfortable
1	2	water
2	1	route
1	2	journey
2	2	game
1	1	flashlight
2	9	army
1	1	yellow

A

Bus	Tent	Association
1	3	stone
14	4	house
1	2	mess
5	1	big
1	1	cockroach
1	7	rain
2	1	door
1	7	friend
2	1	soldier
5	1	window
22	9	trip
2	1	kid
1	5	fun
2	1	uncomfortable
1	2	water
2	1	route
1	2	journey
2	2	game
1	1	flashlight
9	2	army
1	1	yellow

B

Bus	Tent	Association
1	3	stone
4	14	house
1	2	mess
5	1	big
5	5	cockroach
1	7	rain
2	1	door
1	7	friend
2	1	soldier
5	1	window
9	22	trip
2	1	kid
1	5	fun
2	1	uncomfortable
1	2	water
2	1	route
1	2	journey
10	10	game
1	1	flashlight
2	9	army
1	1	yellow

C

Fig. S1. Link construction method comparison: (*A*) original data, (*B*) participant switching, and (*C*) participant amplification. All plots present the same associative responses generated to the cue words tent and bus, with the manipulations described above.

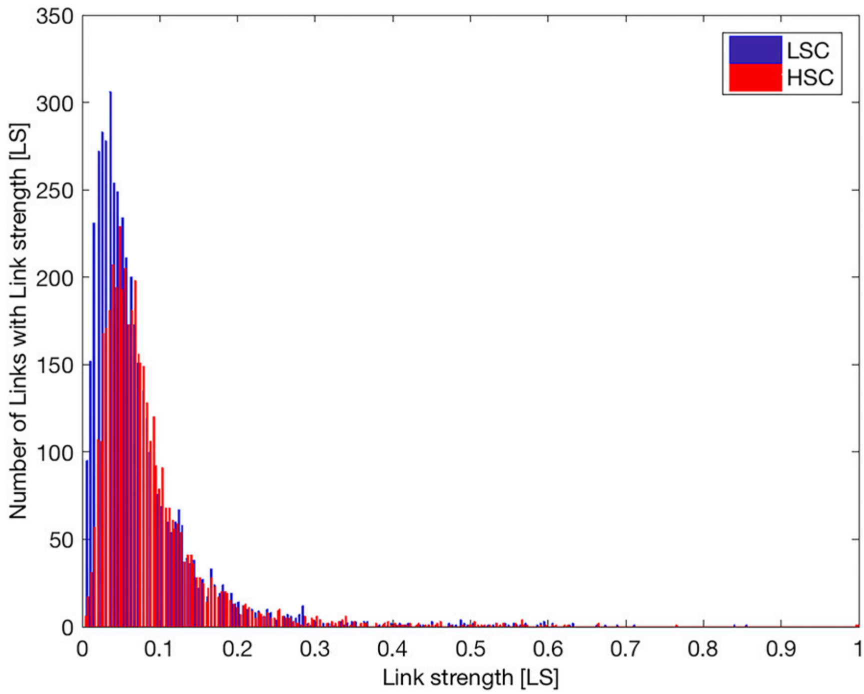

Fig. S2. Histogram of link strength distributions of the LSC (red) and HSC (blue) networks.

Table S1. Stimulus cue words used in the continuous free association paradigm, translated into English

Category	Cue words
Trees	Margosa (6), oak (9), eucalyptus (7), cypress (14)
Vegetables	Eggplant (35), cucumber (60), radish (82), potato (95)
Fabric	Cotton (42), flax (80)
Animals	Duck (13), stork (33), peacock (37), rooster (96)
Spices	Oregano (5), salt (59), sugar (68), paprika (79)
Furniture	Chair (45), counter (54), table (90), drawer (91)
Tools	Screwdriver (48), saw (61), nail (62), hammer (77)
Fruit	Pineapple (10), banana (12), pear (65), apple (94)
Apartment	Roof (17), window (29), stairs (50), elevator (64)
Animals #2	Giraffe (19), donkey (32), dog (43), mouse (75)
Housing	Tent (1), palace (11), apartment (22), shed (83)
Agricultural tools	Tractor (39), rake (49), shears (53), pitchfork (84)
Animals #3	Bee (20), fly (23), beetle (27), ant (66)
Sports	basketball (40)
Flowers	Iris (8), chrysanthemum (36), anemone (44), daffodil (67)
Metal	Iron (15), gold (24)
School	Notebook (55), book (72), newspaper (74)
Cooking	Fork (52), pan (56), knife (70), bowl (86)
Weapons	Bow (34), tank (38), rifle (88), cannon (93)
Cloth	Belt (25), shirt (26), hat (41), coat (63)
Instruments	Guitar (18), flute (30), piano (78), clarinet (85)
Dairy	Cheese (16), dairy (28), butter (31), cream (92)
Transportation	Bus (2), bicycle (4), car (58), train (89)
Marine travel	Ship (3), rowboat (69), boat (71), submarine (81)
Sea life	Dolphin (21), shark (46), whale (47), carp (87)
Office	Stapler (51), eraser (57), ruler (73), pencil (76)

Numbers presented in Fig. 1 are matched to the cue words in parentheses. Numbers of labels were assigned in an ascending fashion according to their spelling in Hebrew.

Table S2. Calculation of link strength between cue words tent and bus with original and revised methods, based on original data, participant switching, and participant amplification

Analysis	Original method	Revised method
Original data	0.64	0.68
Participant switching	0.49	0.68
Participant amplification	0.61	0.92

Link strength was normalized by 50 for comparison needs only for the combined and LS networks and has no meaning because it is a relative size.

Table S3. Stability test for HSC and LSC network

Noise SD	LSC	HSC
0.0001	1.000 (0)	0.999 (0.001)
0.001	0.992 (0.005)	0.997 (0.002)
0.01	0.920 (0.035)	0.949 (0.028)

The calculation of the correlation was done 500 times for different noises added to the network. Noise SD, SD of noise Gaussian added to link weights; LSC, low semantic creative individuals; HSC, high semantic creative individuals.

Table S4. LSC connectivity scores between the network percolated components

Component	1	2	3	4	5	6	7	8
1	0.11	0.05	0.05	0.05	0.05	0.07	0.05	0.04
2		0.13	0.09	0.05	0.05	0.06	0.05	0.04
3			0.07	0.05	0.04	0.05	0.04	0.04
4				0.04	0.05	0.06	0.04	0.04
5					0.04	0.05	0.04	0.04
6						0.05	0.05	0.04
7							0.05	0.05
8								0.04

Rows and columns indicate the different network percolated components. Values represent the connectivity score within and between components.

Table S5. HSC connectivity scores between the network percolated components

Component	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.08	0.07	0.05	0.05	0.04	0.05	0.06	0.06	0.04	0.07	0.04	0.06	0.05
2		0.15	0.10	0.05	0.05	0.07	0.06	0.07	0.05	0.06	0.04	0.06	0.05
3			0.05	0.05	0.05	0.06	0.06	0.05	0.04	0.05	0.04	0.05	0.05
4				0.05	0.05	0.05	0.04	0.06	0.04	0.04	0.05	0.05	0.05
5					0.05	0.05	0.04	0.06	0.04	0.05	0.05	0.04	0.05
6						0.04	0.05	0.05	0.04	0.05	0.05	0.05	0.06
7							0.06	0.06	0.04	0.04	0.04	0.06	0.06
8								0.06	0.06	0.06	0.05	0.07	0.05
9									0.07	0.06	0.06	0.05	0.06
10										0.05	0.05	0.05	0.05
11											0.04	0.04	0.05
12												0.05	0.06
13													0.06

Rows and columns indicate the different network percolated components. Values represent the connectivity score within and between components.

Table S6. Node overlap between LSC and HSC network percolated components

LSC/HSC	1	2	3	4	5	6	7	8	9	10	11	12	Match
1	4												100%
2	1	3											75%
3				7									100%
4		1		2			3						50%
5			4			6		8					100%
6		1			6								85%
7							1		17	7			96%
8					1		1	7			4	12	92%

Rows and columns denote LSC/HSC network percolated component numbers, according to the order in which they break during the percolation analysis.

Table S7. Matched components for the link connectivity analysis

LSC	HSC	Merged component
1	1	1
2	2	2
3	4	4
4	7	7
5	8	3, 6, 8
6	5	5
7	10	10, 11
8	13	9, 12, 13